TD

User Profiling with Feature Selection and Explainability

Essays on three case studies across different domains

DOCTORAL THESIS

Diogo Nuno Teixeira Freitas

DOCTOR DEGREE IN INFORMATICS ENGINEERING

SPECIALIZATION: ARTIFICIAL INTELLIGENCE

UNIVERSIDADE da MADEIRA

A Nossa Universidade

www.uma.pt

October | 2025

Fundação para a Ciência e a Tecnologia

Referência da bolsa:

2021.07966.BD (DOI: 10.54499/2021.07966.BD)

User Profiling with Feature Selection and Explainability

Essays on three case studies across different domains

DOCTORAL THESIS

Diogo Nuno Teixeira Freitas
DOCTOR DEGREE IN INFORMATICS ENGINEERING
SPECIALIZATION: ARTIFICIAL INTELLIGENCE

SUPERVISION
Fernando Manuel Rosmaninho Morgado Ferrão Dias

CO-SUPERVISION
Eduardo Leopoldo Fermé

User Profiling with Feature Selection and Explainability
Essays on three case studies across different domains

DOCTORAL THESIS

Diogo Nuno Teixeira Freitas
DOCTOR DEGREE IN INFORMATICS ENGINEERING
SPECIALIZATION: ARTIFICIAL INTELLIGENCE

EXAMINATION COMMITTEE

Chair: Mónica da Silva Cameirão, PhD
Assistant Professor, University of Madeira

Examiners: Cèsar Femi Ramirez, PhD
Full Professor, Valencia Polytechnic University

Marco Paulo Ferreirinha Garapa, PhD
Faculty Member of the Regional Secretary for Education, Science and Technology, assigned to the Faculty of Exact Sciences and Engineering, University of Madeira

Maria Vanina Martinez, PhD
Senior Research Scientist, Artificial Intelligence Research Institute

Eduardo Leopoldo Fermé, PhD
Full Professor, University of Madeira

To my family, for their endless love and support.

Acknowledgements

First, I would like to thank my advisors, Professor Eduardo Fermé and Professor Fernando Morgado-Dias, for their guidance and support throughout this work.

I am also grateful to my family (Mary, Duarte, and Afonso) for their constant presence during this journey and for their patience and understanding, especially in moments when I was not easy to approach.

My thanks also go to my PhD colleagues, who were always available to provide feedback on my work and presentations. I cannot enumerate all of them here, but they know who they are.

I would like to express my gratitude to Professors Cèsar Ferri and José Hernández-Orallo for welcoming me on three occasions at the Polytechnic University of Valencia during my research visits, and for always posing the most challenging questions that encouraged me to reflect and improve my work. I also thank Professor Santiago Budría for enabling my participation in several workshops across Spain, where I had the opportunity to present and discuss my research.

I would also like to express my acknowledgment to Dr. Maria Vanina Martínez and Professor Cèsar Ferri for their evaluation and constructive suggestions during the first assessment of this work, which contributed significantly to its improvement.

In addition, I extend my thanks to all my professors from the Faculty of Exact Sciences and Engineering. Without their teaching, I would not have the knowledge and methodologies that I have today. I am also thankful to the University of Madeira, the Interactive Technologies Institute, ARDITI, and all the non-teaching staff for providing an appropriate environment for research and work.

I wish to thank my students, whose questions and curiosity have made me a better teacher, and from whom I have also learned a lot.

I am equally grateful to the anonymous reviewers of my work, whose comments and suggestions helped me to improve the quality and clarity of my research.

This work was supported by the Portuguese Foundation for Science and Technology (FCT) under the grant number 2021.07966.BD.

"I'm a man on fire walking through your street
With one guitar and two dancing feet
Only one desire that's left in me
I want the whole damn world
To come dance with me"

— **Alex Ebert**, Man on Fire

Abstract

User profiling is the process of constructing a structured representation of the user within a system. This representation includes information such as preferences, behaviors, and characteristics. Based on the profile, the system can recommend services and products or, in this work, suggest actions. Machine learning methods are commonly used to this end, as they can identify complex patterns among large numbers of attributes.

However, not all attributes are relevant. High-dimensional datasets often contain irrelevant, redundant, or noisy features that obscure valuable patterns and reduce model accuracy. To address this, dimensionality reduction techniques—particularly feature selection—are essential. Equally important is the ability to explain a model's output, since understanding why a model produces a given outcome builds trust and clarifies which steps can change an undesirable situation.

This thesis applies feature selection, explainability, causal discovery, and machine teaching techniques to user profiling. The goal is to support decision-making by identifying the most relevant features, clarifying causal mechanisms, and ensuring that stakeholders understand why recommendations are made. Specifically, we investigate the mRMR (minimum-Redundancy-Maximum-Relevance) method for feature selection, examine explainability strategies such as feature importance analysis and counterfactuals, apply causal discovery to map cause-and-effect relationships, and use machine teaching to explore profile simplification.

We apply this approach in four domains: (i) Marine litter: developing static profiles to identify those who could benefit from literacy interventions; (ii) Football injuries: building predictive models based on player profile dynamics to forecast risk; (iii) Energy poverty: designing models, using counterfactuals, and applying causal discovery to understand health–poverty links; and (iv) Concept complexity: using machine teaching to study profile simplification.

These applications show how profiling can deliver targeted literacy interventions, prevent sports injuries, inform preventive policies in energy poverty, and improve the efficiency of user representations and concept learnability.

Keywords: User profiling, Profile dynamics, Explainable AI, Causal discovery, Machine teaching, Feature selection

Resumo

A criação de perfis de utilizador consiste na construção de uma representação estruturada do utilizador dentro de um sistema, incluindo informações sobre preferências, comportamentos e características. Com base neste perfil, o sistema pode recomendar serviços e produtos ou, no contexto deste trabalho, sugerir ações. Para tal, recorremos a métodos de aprendizagem automática, capazes de identificar padrões complexos em grandes conjuntos de atributos.

Contudo, nem todos os atributos são relevantes, sobretudo em conjuntos de dados de elevada dimensionalidade. Certas características podem ocultar padrões importantes e reduzir a precisão dos modelos. Para mitigar este problema, utilizamos técnicas de redução de dimensionalidade, em particular a seleção de características. É igualmente importante explicar os resultados dos modelos, pois compreender por que razão uma previsão é gerada aumenta a confiança e permite identificar ações para modificar situações indesejadas.

Este trabalho aplica técnicas de redução de dimensionalidade, explicabilidade, causalidade e ensino automático (machine teaching) à criação de perfis. O objetivo é apoiar a tomada de decisão, identificando atributos relevantes, clarificando mecanismos causais e assegurando que os intervenientes compreendem as recomendações. Investigamos o método mRMR (minimum-Redundancy-Maximum-Relevance), estratégias de explicabilidade como a importância de características (feature importance) e explicações contrafactuais, aplicamos descoberta causal para identificar relações de causa-efeito e recorremos ao ensino automático para simplificação de perfis.

A abordagem é aplicada em quatro domínios: (i) Lixo marinho: perfis estáticos para orientar intervenções de literacia ambiental; (ii) Lesões no futebol: modelos preditivos baseados na dinâmica do perfil do jogador; (iii) Pobreza energética: modelos preditivos, explicações contrafactuais e causalidade para compreender a relação entre saúde e pobreza; e (iv) Complexidade conceptual: uso de ensino automático para a simplificação de perfis.

Estes estudos mostram como a criação de perfis pode apoiar campanhas ambientais, prevenir lesões desportivas, fundamentar políticas preventivas contra a pobreza energética e melhorar a eficiência das representações de utilizadores.

Palavras-chave: Perfis de utilizador, Dinâmica de perfis, Inteligência artificial explicável, Descoberta causal, Ensino de máquina, Seleção de atributos

Contents

Appendices

List of Figures

Acronyms

***k*-NN**	*k*-Nearest Neighbors *(p. 19)*
AdaBoost	Adaptive Boosting *(pp. 64, 69, 70, 72, 73, 75–80, 82)*
AI	Artificial Intelligence *(pp. 32, 105, 106, 125, 152, 156)*
ANN	Artificial Neural Network *(pp. 89, 111)*
API	Application Programming Interface *(p. 145)*
AUC	Area Under the receiver operating characteristic Curve *(p. 80)*
CB	Content-Based *(pp. 27, 29)*
CDR	Cross-Domain Recommendation *(p. 27)*
CE	Collaborative Evolution *(p. 29)*
CF	Collaborative Filtering *(pp. 27–29)*
CNN	Convolutional Neural Network *(pp. 17, 20)*
DT	Decision Tree *(pp. 41, 51, 52)*
FNN	Feedforward Neural Network *(pp. 64, 69, 70, 72, 73, 75, 76, 78, 79, 82)*
FPR	False Positive Rate *(p. 70)*
GDPR	General Data Protection Regulation *(pp. 12, 34)*
GFE	General Knowledge Enhanced Framework for Explainable Sequential Recommendation *(p. 31)*
GMEAN	Geometric Mean *(pp. 73–80)*
GOP	Generalized Ordered Probit *(pp. 127, 128)*
GPS	Global Positioning Systems *(pp. 63, 64, 66–70, 78–82, 154)*
GPT	Generative Pretrained Transformer *(pp. 140–142, 145, 152)*
GSP	Gross State Product *(pp. 97, 111)*
HDOP	Horizontal Dilution of Precision *(pp. 66, 67)*
HILDA	Household, Income and Labour Dynamics in Australia *(pp. 85, 89, 94, 98, 100, 101, 103, 104, 107, 110, 111, 120, 121, 123, 126, 128, 137)*

IP Internet Protocol *(p. 39)*
IPF Injected Preference Fusion *(p. 18)*
IQR Interquartile Range *(pp. 40, 46, 47, 49)*

J-PCMCI$^+$ Joint-PCMCI+ *(pp. 123, 128, 129)*

KB Knowledge-Based *(pp. 20, 27, 30)*
KG Knowledge Graph *(pp. 18, 20, 21, 31)*

LLM Large Language Model *(pp. 17, 18, 21, 31–34, 139, 141, 142, 145)*

MCI Momentary Conditional Independence *(p. 129)*
MEPI Multidimensional Energy Poverty Index *(pp. 85, 87, 90, 91, 93, 110, 111, 113, 120, 123, 126, 127, 135)*
MF Matrix Factorization *(pp. 18, 29)*
mRMR minimum-Redundancy-Maximum-Relevance *(pp. 64, 68, 69, 73, 78, 82, 153)*

NB Naive Bayes *(p. 20)*
NLP Natural Language Processing *(p. 17)*

ODP Open Directory Project *(p. 17)*
OvR One-vs-the-Rest *(pp. 94, 113)*

PCMCI$^+$ Peter and Clark Momentary Conditional Independence *(pp. 123, 128, 129)*
POI Point of Interest *(pp. 16, 17)*

RDP Ramer–Douglas–Peucker *(pp. 144–146, 152)*
RF Random Forest *(pp. 88, 89, 111)*
ROC AUC Receiver Operating Characteristic - Area Under Curve *(pp. 70, 94–96, 113, 116, 119, 121)*
RS Recommender Systems *(pp. 9–12, 14, 17, 20, 24, 26–31, 33)*

SAH Self-Assessed Health *(pp. 127, 128)*
SHAP SHapley Additive exPlanations *(pp. 2, 4, 5, 85–87, 94, 95, 99–102, 105, 153)*
SVM Support Vector Machine *(pp. 19, 64, 69–73, 75–80, 82)*

TF-IDF Term Frequency–Inverse Document Frequency *(pp. 15, 16, 18)*
TMER Temporal Meta-path Guided Explainable Recommendation *(p. 31)*

UCPR User-Centric Path Reasoning *(p. 31)*
UEFA Union of European Football Associations *(p. 63)*

XAI eXplainable Artificial Intelligence *(p. 30)*
XGBoost Extreme Gradient Boosting *(pp. 88, 89, 110, 111, 125)*

CHAPTER 1

Introduction

Every time a user creates a new account on a digital platform (e.g., for online shopping or movie streaming), the system begins constructing a profile for this user, considering their interactions, browsing history, and previous purchases. This process is known as *user profiling* and consists of creating a structured representation of an individual's (or those of a group of individuals) preferences, behaviors, and characteristics within a system (i.e., a user profile). Based on this user profile and patterns observed in similar users, the system can model user preferences and interests, recommend products or services they might be interested in, and personalize their experience on the platform. As the user interacts more with the platform, their profile becomes more detailed. This leads to more accurate and relevant suggestions.

User profiling has two main applications. One is in recommender systems, where user profiles are used to suggest relevant items, such as movies, music, and books, using techniques like collaborative filtering, content-based filtering, and hybrid filtering [Bob+13; IFO15]. The other application is in personalized or adaptive systems. This includes customizing user interfaces [Bro90] or refining search results [HF15]. The focus of this thesis, however, is to use user profiling for predictive monitoring. This helps to identify individuals at risk of adverse outcomes, allowing targeted interventions and proactive decision support.

The process of user profiling consists of five steps. The first step is information retrieval, where data about users is collected. These data can be *explicit*, such as preferences users directly specify, and *implicit*, which includes browsing habits, interactions, and purchase histories [IFO15]. These attributes can be static (e.g., date of birth) or dynamic (e.g., browsing history or changing preferences).

After retrieving the information, the raw data undergo feature extraction and representation [Eke+19]. These data are then systematically organized into a user profile, denoted as P. This process corresponds to the second and third steps of user profiling and can be formally defined as follows [Fer+24]:

$$P = \ll \text{label}_1, \text{label}_2, ..., \text{label}_n \gg, \tag{1.1}$$

where each label_i represents a characteristic that assumes a finite set of possible values. For instance, if label_1 represents “marital status”, the possible values could be {“single”, “married”, “divorced”, “widowed”}. Alternatively, user profiles can be categorized into keyword-based profiles and knowledge-based profiles [Gao+14; TVB09]. In keyword-based profiles, the user’s interests are represented through keywords chosen based on their activities, often ranked by their importance. In knowledge-based profiles, the profiles are designed as semantic networks (e.g., ontologies), where each profile is a network of related concepts that represent the user’s interests.

Once features are established and user profile representations are defined, the next step is preference learning. In this phase, statistical or, more recently, machine learning techniques analyze user behaviors to identify interests, predict future preferences, and determine the importance of different profile attributes [Tan+10]. However, as these models become more complex, they often operate as “black boxes”, making their decisions opaque. This opacity gives rise to the need for *explainability*, which involves techniques to interpret model behavior and understand the reasoning behind their predictions.

Explainability can take several forms. Feature importance methods, such as SHapley Additive exPlanations (SHAP) values [LL17], quantify the contribution of each user attribute to a specific outcome, capturing both global and local patterns. A different approach uses *counterfactual explanations* in order to provide actionable recourse by answering “what-if” questions and identifying the minimal changes to a user’s profile that would alter a prediction [MST20]. Moreover, a deeper level of understanding can be achieved through *causal discovery*, which moves beyond correlation to map the cause-and-effect relationships between user attributes and outcomes.

At this stage, it is important to note that users’ interests, behaviors, and circumstances change over time. Thus, a *static* profile quickly becomes outdated and leads to suboptimal personalization. This corresponds to the last step of user profiling: *Profile Dynamics*. Profile dynamics enables the system to detect and incorporate short- and long-term changes in user preferences [HF15], thereby keeping the user profile up to date. There are two primary approaches to handling profile dynamics: periodic recalibration, where profiles are updated at fixed intervals, and continuous adaptation, where updates occur in response to user interactions [MIM10].

The need to manage profile dynamics constantly brings challenges of simplification and generalization. As profiles evolve, they accumulate vast amounts of temporal data, making it even more critical to distinguish signal from noise. This leads to the question: How can user profiles be simplified to their most essential elements while still enabling the model to make confident and generalizable predictions? This is a problem of optimal data efficiency that connects to the paradigm of machine teaching [Zhu+18].

Instead of learning from vast amounts of data, machine teaching focuses on identifying the minimal and most informative set of examples a teacher must provide for a learner to understand a concept. This thesis explores this principle using an analogy: A simple drawing serves as a visual proxy for a user profile. In other words, just as a user profile can be seen as a sketch of a person built from data points (features), a drawing is a simplified sketch of an object built from strokes. The motivation for this approach is to understand the limits of simplification. More specifically, by determining the "teaching size" (in this case, the minimum number of strokes required for a model to recognize a drawing), we can, by extension, understand how much a user profile can be simplified while retaining its predictive power. This framework aims, in the end, to design more efficient, explainable, and ultimately more intelligent user profiling systems that can operate with less data.

Modern user profiling uses complex predictive models. For these models, accuracy is not the only goal. It is also important that the systems are transparent, produce actionable results, and use data efficiently. This thesis explores these challenges through various approaches, including static and dynamic profiling, explainability, causal inference, and machine teaching. The following sections detail the specific problems, research questions, and contributions of this work.

1.1 Problem Statement

Machine learning problems rely on data to model the relationship between input features and output. In practical applications, determining which input features are most relevant for predicting the target variable or achieving the modeling objectives is often challenging. Consequently, it is standard practice to gather and examine a large set of features, often more than necessary for the task at hand [Sye23]. This is also the case in user profiling, where user profiles integrate attributes from many sources, such as browsing histories, explicit ratings, implicit feedback, contextual details, and potentially several demographic or personal descriptors, as it may not be straightforward to determine the most important features beforehand.

Machine learning methods are particularly effective for user profiling tasks [Eke+19] because they can find complex patterns in user data [Eke+19]. However, this approach presents several challenges. First, the large number of features in user profiles can lead to poor model performance. Second, complex models are often "black boxes," making their predictions difficult to understand (i.e., they are not explainable). Finally, simplifying profiles without losing information presents another relevant problem.

Feature selection involves identifying and selecting a subset of relevant features, denoted as $S \subseteq F$, from the original feature set F. Formally, if $P(C|F)$ represents the probability distribution of the class C given the complete feature set F, the objective is to find a subset $S \subseteq F$ such that $P(C|S)$ closely approximates $P(C|F)$ while minimizing the number of features in S [YL04]. By identifying and removing unnecessary variables, feature selection helps models focus on the most informative aspects of the data, leading to improved accuracy, reduced computational complexity, and enhanced explainability.

While the accuracy of model predictions is important, it is equally important to understand or be able to explain the reasoning behind these predictions. For instance, in recommendation systems, a user receiving a movie suggestion based on their past viewing habits might find the recommendation useful but may not understand why it was suggested. If the system provides an explanation—such as highlighting similar genres, shared user preferences, or important features that influenced the recommendation—users are more likely to trust and engage with the platform [TM07].

Various techniques can be used to improve model *explainability*. Feature importance analysis helps identify which input variables contribute most to a model's predictions. One example of a method to compute the feature importance is the SHAP values [LL17], which quantifies the influence of each feature on a given prediction (local explainability) or across all predictions (global explainability). Moreover, the SHAP values also indicate the direction and magnitude of their impact. Counterfactual explanations provide a different approach to model explainability by identifying the minimal and most feasible changes that could be made to an input instance to alter the model's prediction [MST20]. Rather than identifying which features contribute most to a prediction, counterfactual explanations can suggest actionable changes by answering "what-if" questions.

Still in the domain of explainability, it is equally important to understand the causal mechanisms that drive outcomes. *Causal discovery* techniques aim to discover these cause-and-effect relationships from observational data. This clarifies how different user attributes influence one another and the final outcome, and allows for more robust and reliable decision-making.

The final challenge lies in data efficiency and profile simplification. As profiles evolve due to their dynamic nature, they become more complex. Thus, it is important to distinguish signal from noise, which raises questions about the limits of simplification. This problem of optimal data efficiency connects to the paradigm of *machine teaching*, which studies the minimal and most informative set of examples a learner needs to understand a concept.

This work extends these ideas by applying feature selection, explainability, causal discovery, and machine teaching techniques to user profiling tasks. The objective is to enable better decision-making by identifying the most relevant user profile attributes, while ensuring that users and stakeholders can understand why certain recommendations or classifications are made. With this in mind, we apply our approach across different domains, considering both static and dynamic aspects of user profiling:

- **Marine litter:** Developing static user profiles based on questionnaire responses to identify user profiles that would benefit from marine litter literacy interventions.
- **Football injuries:** Building predictive models for non-contact injuries in professional football by analyzing how player characteristics change over time (i.e., user profile dynamics).

- **Energy poverty:**
  - Designing predictive models to identify households at risk of energy poverty by analyzing the evolution of their socioeconomic profiles over time. In this approach to user profiling, SHAP was used to explain the short- and long-term effects of individual and contextual factors on future energy poverty outcomes.
  - Using counterfactuals in user profile to identify the smallest necessary changes needed to lower the likelihood of becoming energy-poor.
  - Applying causal discovery to investigate the bidirectional relationship between health and energy poverty, and the temporal dynamics of this link.
- **Concept Complexity:** Using machine teaching to investigate whether the complexity of a concept is inherent and transcends its representation, using drawings as an analogy for user profiles.

This study begins with a narrative literature review to explore different profiling methodologies. It is then guided by the following research questions:

- **RQ1:** How can static user attributes be used to identify user profiles that require targeted interventions?
- **RQ2:** What are the most relevant features for profiling users in domains such as marine litter awareness, sports injury prediction, and energy poverty risk?
- **RQ3:** What is the effect of feature selection on the identification of at-risk user profiles?
- **RQ4:** How can feature importance analysis in dynamic user profiles indicate the direction and magnitude of their impact on model predictions and decision-making?
- **RQ5:** How can minimal profile modifications be identified in order to reduce the likelihood of a negative outcome?
- **RQ6:** How can causal discovery frameworks model the dynamic, bidirectional relationship between user attributes and outcomes, such as health and energy poverty?
- **RQ7:** Using drawings as an analogy for user profiles, does the minimum information needed to teach a concept remain consistent across different modalities?

1.2 Publications

From this PhD research, the following articles have been published:

- D. N. Freitas, K. Varela, and E. Fermé. *User profiling and its dynamics: A narrative review*. Accepted at The European Journal on Artificial Intelligence. To appear online. 2025.
  - This work corresponds to the narrative literature review, as per Chapter 2.
  - As per JCR, JIF=1.0; JIF quartile=Q4; JIF percentile=14.0

- S. Bettencourt, D. N. Freitas, S. Costa, and S. Caeiro. "Public perceptions, knowledge, responsibilities, and behavior intentions on marine litter: Identifying profiles of small oceanic islands inhabitants". In: *Ocean & Coastal Management* 231 (2023), p. 106406[1].
  - This work aims to answer RQ1 and RQ2.
  - As per JCR, JIF=4.8; JIF quartile=Q1; JIF percentile=96.2
- D. N. Freitas, S. S. Mostafa, R. Caldeira, F. Santos, E. Fermé, É. R. Gouveia, and F. Morgado-Dias. "Predicting noncontact injuries of professional football players using machine learning". In: *PloS one* 20.1 (2025), e0315481.
  - This work aims to answer RQ2 and RQ3
  - As per JCR, JIF=2.9; JIF quartile=Q1; JIF percentile=76.5.
- S. Budría, E. Fermé, and D. N. Freitas. "Unveiling energy poverty trajectories: A longitudinal analysis using machine learning". In: *Energy Strategy Reviews* 62 (2025), p. 101998[2].
  - This work aims to answer RQ2 and RQ4.
  - As per JCR, JIF=9.9; JIF quartile=Q1; JIF percentile=87.1
- D. N. Freitas, E. Fermé, and S. Budría. "A counterfactual approach to energy poverty mitigation: A case study for Australia (Preliminary report)". In: *Proceedings of the Workshop on Foundations and Future of Change in Artificial Intelligence (FCAI)*. Bologna, Italy, 2025, pp. 27–43
  - This work aims to answer RQ2 and RQ5.
- D. Freitas, B. Håvardstun, C. Ferri, D. Garigliotti, J. A. Telle, and J. Hernandez-Orallo. "Relative Drawing Identification Complexity is Invariant to Modality in Vision-Language Models". In: *Proceedings of the 28th European Conference on Artificial Intelligence (ECAI)*. Bologna, Italy, 2025, pp. 4507–4514.
  - This work aims to answer RQ7.
  - As per CORE, rank=A.

1.3 Document Structure

This PhD thesis is organized as a compilation of research articles. The first research contribution, presented in Chapter 2, is a narrative literature review on user profiling. The subsequent chapters each present a distinct research contribution from other published or submitted papers.

The first two studies apply user profiling to specific domains. Chapter 3 presents the work on marine litter, discussing how public perceptions are used to create user profiles. Chapter 4 details the research on football injuries, where machine learning models, in conjunction with dynamic profiles, are used to predict noncontact injuries.

[1] My contribution was the statistical analysis of the collected data, including profile discovery.

[2] My contribution was the application of the machine learning methods that supported the profile creation.

The subsequent three chapters focus on the domain of energy poverty, each from a different technical perspective. Chapter 5 uses explainability methods to predict and explain energy poverty. Chapter 6 then uses counterfactuals to identify minimal interventions to reduce energy poverty risk. Finally, Chapter 7 applies causal discovery to investigate the bidirectional relationship between health and energy poverty.

The final research chapter, Chapter 8, explores a more foundational question related to profile simplification, using machine teaching and a drawing analogy to investigate concept complexity.

The thesis concludes with Chapter 9, which summarizes the main findings, discusses their implications, and outlines future research directions.

CHAPTER 2

User Profiling and its Dynamics: A Narrative Review

Diogo Nuno Freitas, Katherin Varela, Eduardo Fermé

Abstract. *The need for personalized content has grown considerably with the increasing amount of online information. User profiles, as structured collections of user characteristics and interests, are essential for personalization because they help systems better understand individual preferences and deliver more relevant content. This review examines methods for user profiling and their adaptation over time. We organize existing literature into five categories: User Profile Modeling, Profile Dynamics, Recommendation Systems, Personalized Systems, and Adaptive Systems. Key findings highlight the importance of combining explicit and implicit data collection methods, differentiating between short- and long-term user preferences, and employing techniques such as evolutionary algorithms, context-awareness, and explainability. Additionally, we identify promising areas for future research, including multimodal data integration, scalability, privacy preservation, contextual adaptation, and universal user models. This review aims to help readers navigate the extensive literature and provide insights to support the development of practical applications based on user profiling techniques.*

2.1 Introduction

User profiles consist of a set of characteristics that describe how users interact with a given system. These profiles represent users' interests, needs, preferences, demographic traits, and desires [AS99]. User profiles are created and managed through a process called *User Profile Modeling*, which consists of two main steps [Lop+19, p. 9]. First, it is necessary to define what user information is to be stored and how this information will be represented. Second, it is required to establish an appropriate methodology for constructing, updating, and maintaining these profiles over time.

To build accurate user profiles through user modeling, systems must collect relevant data about users. Typically, user data can be gathered in two ways: explicitly and implicitly. Explicit data collection typically involves directly asking users to provide ratings or feedback on items. Implicit data collection, on the other hand, involves observing and analyzing users' behavior without direct input. Examples of implicit data include monitoring songs played, applications downloaded, websites visited, or books read [IFO15].

User profiles can be categorized into two main types: static and dynamic. Static profiles contain fixed information about users that does not change over time. Such static information, however, carries the risk of becoming outdated, potentially leading to ineffective personalization [HF15]. Indeed, users' interests, preferences, and needs often evolve. To address this issue, the concept of *Profile Dynamics* was introduced. Profile dynamics refers to techniques that account for the evolving nature of user profiles. These techniques allow systems to detect both short- and long-term changes in user interests and characteristics. By identifying and incorporating these changes, systems can maintain updated user profiles. Consequently, profile dynamics ensures that the personalization provided by the system remains accurate, relevant, and helpful.

User modeling is important for personalized systems, especially given the current data explosion where the amount of digital data being generated and stored is growing exponentially. In this context, user modeling enables systems to understand users' characteristics and adapt their behavior dynamically. Such adaptation typically involves tailoring the content to individual users or groups of users according to their specific behaviors, needs, and preferences. Furthermore, user modeling can also help identify and suggest new interests to users, i.e., interests they may not be aware of but that align closely with their existing preferences [Mou97].

One of the primary applications of user modeling is information filtering, which is, in turn, used, for example, in Recommender Systemss (RSs) across various domains. These domains include web personalization, computational advertising, e-learning systems [HF15; Le+09; Wu+22; Yin+15], and search engines [Eke+19]. A RS collects information about users' preferences regarding specific items, such as movies, songs, or books. Using this information, the system suggests new items that closely align with users' interests and needs [Bob+13; IFO15].

User modeling also benefits prediction systems. In these systems, user modeling can help estimate ratings for items even without explicit user feedback [Far+18a]. Furthermore, adaptive systems can also rely on user modeling to customize the user interface. Customization might include adapting font sizes or displaying context-aware information, thus improving the effectiveness of the interaction between the system and the user [Bro90].

In order to provide context for how the selected studies were identified and organized, the next section describes the methodological approach used to conduct this narrative review.

2.2 Methodology

This article adopts a narrative review approach to synthesize and organize the existing literature on user profiling and its dynamics. The literature search and retrieval process started in March 2023 and concluded in June 2025, using Web of Science, ACM Digital Library, and Google Scholar databases. The search used combinations of keywords such as “user profile,” “profile dynamics,” “recommender systems,” and “adaptive systems.” No restrictions were applied regarding publication dates. Additionally, further relevant publications were identified through references cited within the initially retrieved articles.

Publications were reviewed for relevance based on titles, abstracts, and full texts. Works were included if they contributed to understanding how user profiles are constructed, maintained, or applied in personalized and adaptive systems. In total, 117 publications were selected, and their full texts were extracted and analyzed. Among these publications, 53 were conference articles, 51 were journal articles, 3 were books, 3 were book chapters, and 7 were pre-prints.

The publications have been grouped into five categories, namely: User Profile Modeling ($n = 57$), User Profile Dynamics ($n = 44$), RSs ($n = 52$), Personalized Systems ($n = 29$), and Adaptive Systems ($n = 30$). This taxonomy, as depicted in Figure 2.1, helps the reader understand the diverse fields of user modeling and profile dynamics. This categorization also improves the clarity and organization of the review, making it easier to navigate and comprehend the extensive body of research in this area. Note that each publication can belong to one or more of these categories.

This article follows a structured approach, in which the relevant literature for each stage of user profile modeling is introduced. While the suggested taxonomy improves the understanding of the different domains associated with user modeling and profile dynamics, the systematic step-by-step approach facilitates developing user profile modeling applications. Therefore, following the introduction, we first discuss user profile modeling, focusing on relevant topics such as information retrieval and feature representation. The next section addresses profile dynamics and summarizes techniques for handling evolving user preferences and interests. This discussion is followed by a section on recommendation systems, including their definition and important sub-classifications. We then describe personalized/adaptive systems, highlighting their relationship with both user profiles and profile dynamics. Finally, we close the review with the concluding remarks and directions for future work.

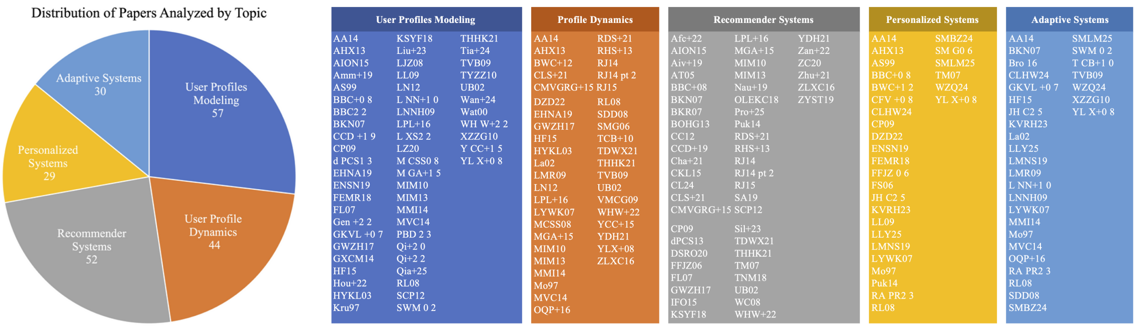


Figure 2.1: Distribution of the papers analyzed by topic. The alphanumeric labels (e.g., MIM10, DZD22) correspond to the list of selected studies provided in Appendix A.

2.3 User Profile Modeling

User profiles play a fundamental role in RSs [IFO15] and personalized or adaptive systems [BKN07]. They provide a comprehensive representation of individual users or user groups. A user profile is an instantiation of a user model within the system [Eke+19]. Typically, user profiles consist of two main components: demographic information, such as age, gender, and occupation, and a set of keywords or concepts associated with corresponding values. These keywords and concepts serve to represent and estimate user intentions, interests, and other pertinent information for profile modeling, both for long-term and short-term time frames [DZD22]. However, some profiles can be more complex, using users' behavioral data, such as clicking behavior or time spent on a web page [AX13].

While there is a consensus in the literature that a profile represents a set of characteristics for individual users or user groups, there exists a gap when it comes to a formal definition of what constitutes a profile and its inherent structure. An exception to this gap can be found in [Fer+24], where a user profile $P_{\mathbb{L}}$ and a language for managing profile dynamics are rigorously defined. In this context, a profile P is defined as a tuple $\ll \text{label}_1, \ldots, \text{label}_n \gg$, where $\text{label}_i \in \mathbb{N}_0$. Informally, each element within the profile tuple represents a characteristic that can assume a finite number of possible values. While natural numbers were used to define the content of each $label_i$, it is worth noting that this definition can be adapted to incorporate linguistic labels. For example, if label_1 represents marital status, possible values such as "single/married/separated/widowed" among others, can be utilized. To illustrate this profile structure, consider a simple example of a profile structure: $\ll$ age, weight, height $\gg$, where a possible profile might be John $= \ll 20, 80, 178 \gg$.

As mentioned earlier, user profiles are generated through a process known as user modeling, which is subsequently employed for inferring unobservable information about users. The user modeling process involves several distinct steps, namely information retrieval, feature extraction and representation, profile construction, preference learning,

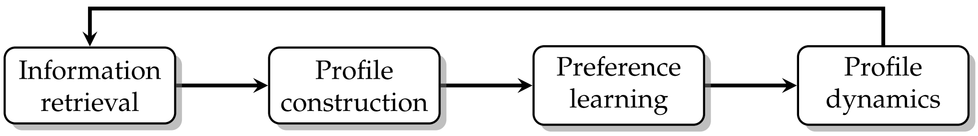


Figure 2.2: User modeling phases required to build user profiles [RL08].

and profile dynamics (also referred to as profile maintenance) [RL08; Zho+12]. Figure 2.2 provides an overview of these different stages. The subsequent sections provide a comprehensive description of each step.

User profiling can raise concerns regarding data privacy. These concerns have become increasingly important since the introduction of the General Data Protection Regulation (GDPR). One possible solution to address these privacy issues involves federated systems. In federated systems, users do not need to share their raw data directly. Instead, they only share model parameters through a decentralized approach. These parameters can then be aggregated centrally to build an overall model, thereby improving privacy protection for users. Liu et al. [Liu+23a] proposed a solution that employs hierarchical information for user profiling. The resulting user profiles are then aggregated to build an overall model. Federated RS that uses graph neural networks [LXS22; Tia+24] can also be a solution to protect user privacy. Other approaches exist where a federated collaborative filter is introduced (see, e.g., [Qi+20] or [Amm+19]).

2.3.1 Step 1—Information Retrieval

In order to accurately identify users' needs and interests, the system needs to collect as much information about the user as possible; therefore, the information retrieval phase is important to make an effective user personalization. Information retrieval can be classified as explicit, collected by getting users' feedback or manually editing their profile; implicit, collected by analyzing users' interaction and behavior within the system; or hybrid, combining explicit and implicit information [IFO15].

Information about users may include demographic data, such as age, gender, and education. These data can be utilized to classify users and generate baseline recommendations based on demographically similar individuals. This technique is known as Demographic Filtering [AT05] and can be particularly useful for addressing the data sparsity problem, which occurs when limited information about users is available.

In addition to demographic characteristics, the user's behavior within the system, their past searches, and previously expressed preferences can also be used to infer their interest in specific items or particular needs. Such information can be represented using graph structures, in which nodes represent entities such as users or items, and edges refer to interactions or relationships between these entities [PBD23]. However, user profiling systems typically concentrate only on users' positive interactions with the system, overlooking other relevant information. In fact, relying solely on positive signals might not be sufficient to fully capture the user's interests and preferences. As argued by Leung and Lee [LL09], incorporating knowledge about negative user preferences can provide a broader perspective on the user profile, allowing systems not only to understand what users like but also to explicitly represent what they dislike.

In any case, the collected data undergoes a preprocessing process to extract label values, known as user profile features. These features are then employed in constructing the user profile [Eke+19], which can apply to either an individual user or a group of users.

As already mentioned, the information used for profiling can be categorized as explicit (provided by users), implicit (automatically collected), or hybrid. In the following subsections, we will explore these techniques in detail. The highlighted papers for this subsection are presented in Figure 2.3.

2.3.1.1 Explicit Information

Explicit information retrieval relies on information explicitly provided by users. This may include demographic information, such as age, gender, and place of residence, which can be used to create demographic clusters and infer users' preferences [Kru97].

Users can also provide positive or negative feedback information using ratings or dislike buttons [Dau+20; DZD22; Xie+20]. Additionally, users can provide information about their interests [Gau+07] and manually assign weights to or order their preferences within the system.

Although simple and highly effective [Zha+16], this information retrieval strategy has areas for improvement. The process is time-consuming and requires users to have domain knowledge of the concepts used to describe preferences and the rating scale for weighting preferences [Dau+20]. Furthermore, not all users are willing to spend time providing feedback and information to the system, even when encouraged to do so [MIM10]. For these reasons, implicit information retrieval methods have arisen.

2.3.1.2 Implicit Information

Implicit information retrieval occurs when the user's information is collected automatically without user intervention in the process [BKN07]. There are many ways to implicitly collect information about users' preferences, feedback, and interests.

One method is by analyzing browsing behavior, including the websites visited by the user, their contents, timestamps, and duration of the visit on each website [DZD22]. Another approach is to analyze the user's query logs, examining their past searches and creating query-flow graphs to represent a user's search behavior. This technique is particularly useful for identifying logical sessions and recommending new queries [Bol+08].

Additionally, observing which options the user selects or clicks can provide valuable insights [Lop+19]. For instance, if a system presents options sorted by preference level and the user does not choose one of the top options, a comparison can be made between the selected option and the top ones to identify differences. This information can then be used to modify the weights of a given model, ensuring that similar options to the selected one are prioritized in the subsequent recommendation process [Zha+16].

This method requires, however, considerable computational power due to the large amount of data collected, as well as data mining techniques. Furthermore, the implicitly collected data may be noisy [Xie+20], and the confidence in the obtained results is occasionally low [MIM10], especially in the initial stages. On the other hand, this information retrieval process has the advantage of being capable of collecting data continually, and thus, receiving updated information [Eke+19].

We argue that contextual information can also be categorized as implicit information, which can be used effectively for user profile modeling [Rav+21]. Contextual information refers to any details that help describe the user's current situation. Typical examples include the user's location, current activity, and social relationships [HPV16].

2.3.1.3 Hybrid Information

Hybrid information retrieval is a methodology that makes use of the strengths of both explicit and implicit feedback to verify collected information and strengthen the effectiveness of the system, particularly in the context of user profile dynamics [RJ15]. One way to obtain hybrid feedback is by using implicit feedback to verify the veracity of explicitly collected information [IFO15]. This methodology creates unbiased user models by considering both positive and negative feedback during the training phase of the model [Xie+20], enabling the same model to better capture the user's intention [DZD22].

Logesh and colleagues [Rav+21] argue that the optimal performance of RSs requires the use of both implicit behavior and explicit feedback information from users. By incorporating both of these information retrieval approaches, the authors suggest that the resulting models will become more transparent and trustworthy. This increased transparency and trustworthiness is expected to encourage users to provide more information, thereby enabling the model to be trained more effectively.

Still within the domain of RSs, Interactive RSs aim to integrate feedback into the recommendation process. These systems suggest items and subsequently receive feedback from users, which can be utilized to generate more refined and personalized recommendations [Zho+20]. A similar application can also be found in Conversational RSs [Gao+21], where implicit interaction can be used as a means to evaluate recommendation models.

2.3.2 Step 2—Feature Extraction and Representation

After the information retrieval phase, the feature extraction is conducted. Feature extraction involves extracting, from the raw data, the features that can potentially be used to model the user profile. For example, an important aspect of this phase is to identify and extract the most relevant keywords present within a document [Lau02]. Another example is categorical features, which can be represented using encoding techniques such as one-hot encoding. In one-hot encoding, each category is converted into a binary vector;

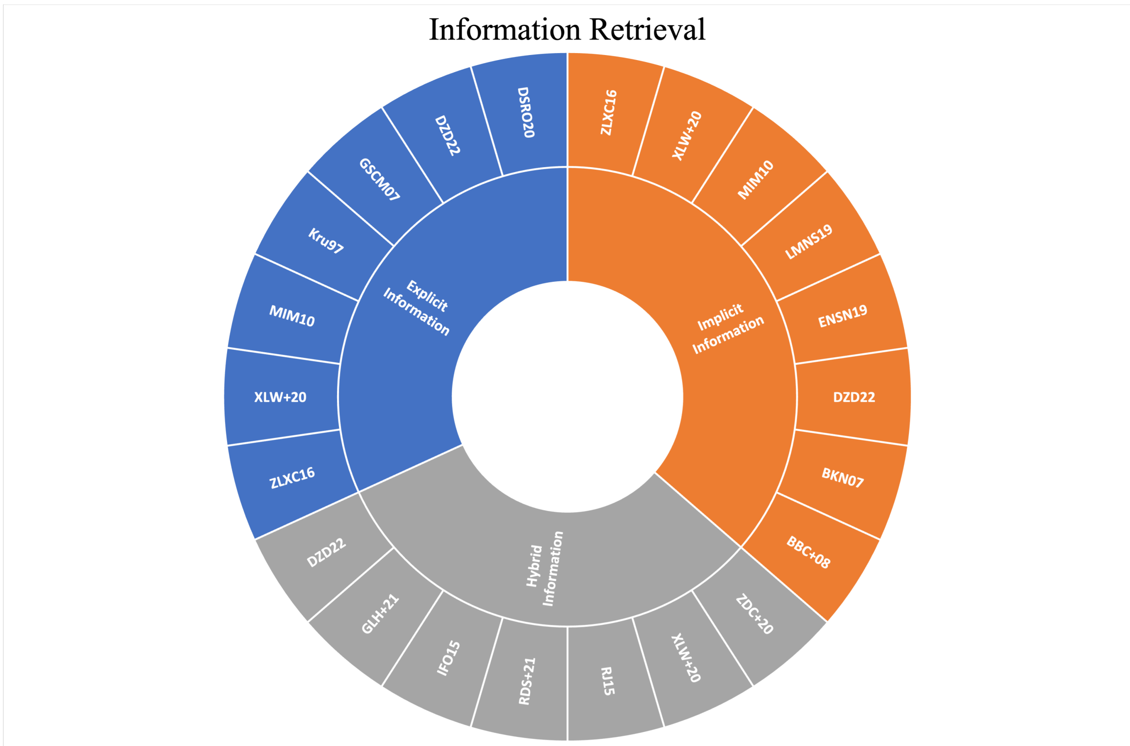


Figure 2.3: Highlighted papers on Information Retrieval. The alphanumeric labels (e.g., MIM10, DZD22) correspond to the list of selected studies provided in Appendix A.

this strategy can be applied to both users and items [Wu+22]. Eke and colleagues [Eke+19] summarized the possible features to extract depending on the application.

Following feature extraction, feature representation is the next step for user profile construction. In this step, the previously extracted information is encoded into a latent representation (embeddings), transforming features into a format suitable for further analysis, such as user profile modeling. Feature representation has received increased attention, especially due to the recent emergence of deep learning techniques [BBC22], and has been successfully applied across various systems for multiple purposes.

Li and Zhao [LZ20] suggested that feature representation can be performed using either structured datasets (e.g., matrices) or unstructured sequences (e.g., purchase histories). And it is important to note here that feature extraction, and later feature representation, can be applied to both users and items.

The highlighted papers of this subsection are displayed in Figure 2.4.

2.3.2.1 Text Representation

For text representation, one widely used method is the Term Frequency–Inverse Document Frequency (TF-IDF) classifier, a text vectorization technique that calculates the weight of a keyword in a document based on its frequency within that document [AT05]. For instance, in a study by Lauschke and Ntoutsi [LN12], the TF-IDF classifier was utilized

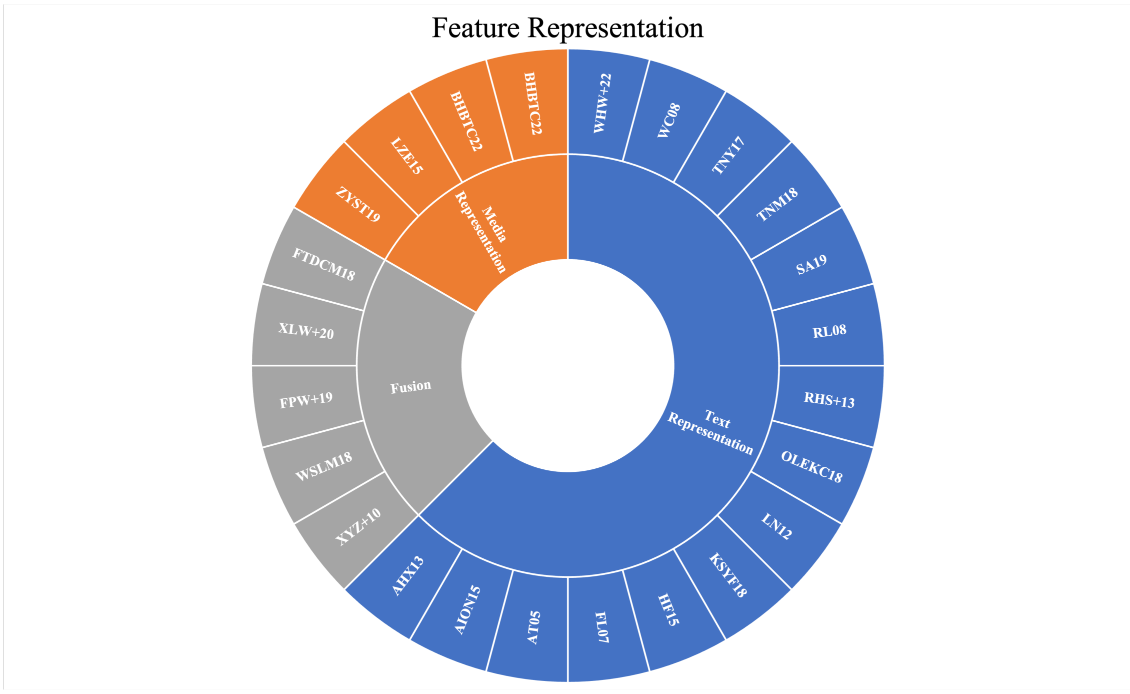


Figure 2.4: Highlighted papers on Feature Representation. The alphanumeric labels (e.g., MIM10, DZD22) correspond to the list of selected studies provided in Appendix A.

to extract features (in this case, keywords) in a system designed to monitor user profiles and their evolution on the social platform X (formerly known as Twitter).

Another approach, known as the Bag-of-Words method, is a simpler alternative to TF-IDF. It analyzes the frequency of words input by users to generate a set of keywords representing their interests. This autoencoder technique is particularly effective in systems that rely on explicit data, such as microblog text [AX13]. Moreover, n-gram models [SA19], word2vec [DZD22] and attention models [Wu+22] are also commonly employed for feature representation.

An ontology, on the other hand, is a formal knowledge representation of a specific domain in a structured manner that can be understood by both humans and machines [Obe+18]. One of the key characteristics of ontologies is their capability to establish and describe relationships between classes [FL07]. As a result, ontologies can be considered a form of feature representation, more specifically, a representation of semantic features [RL08], which can be shared and reused in different systems. Moreover, ontologies can solve the cold-start problem, i.e., when there is insufficient information about the user to make accurate user profiles [TNY17].

Ontologies represent entities, such as users, products, or services, within the system [TNM18] and concepts. Subsequently, these ontologies are compared to predict users' needs, particularly in the context of recommendation systems [Ruo+13; WC08]. For instance, She et al. [Liu+18] developed a framework where the ontology structure incorporates both the Point of Interests (POIs) and the users. Recommendations are then generated by evaluating the similarity between the POIs and the users' ontologies using the Jaccard index.

Similarly, Obeid and colleagues [Obe+18] implemented a semantic RS wherein ontologies were employed to represent higher education institutions, employment, and students. However, machine learning techniques were utilized to generate tailored recommendations for university degrees for students.

Reference ontologies, such as Open Directory Project (ODP), also serve an important role in representing semantic features since they provide an improved and standardized framework for defining, representing, and relating entities and concepts [TNM18]. As a result, they provide a richer representation of information when compared to ontologies developed from scratch.

For example, Hawalah and Fasli [HF15] mapped web pages to concepts present in a reference ontology for web personalization. Interestingly, these authors also proposed techniques that utilize the established mapping for constructing and maintaining ontological user profiles (including session-based, long-term, and short-term profiles). These profiles were subsequently used to adapt the system according to user behavior.

Amini and colleagues [Ami+15], on the other hand, integrated various reference ontologies into a unified and more comprehensive ontology for profiling scholars' knowledge.

Finally, with the advent of Large Language Models (LLMs), semantic embedding has also emerged as an effective option for text representation [Gen+22; Wan+24], enabling the implementation of a new RS paradigm called *Generative Recommendation*, which will be explored later in this paper.

2.3.2.2 Media Representation

Features from other sources, such as images, text, audio, and video, are often extracted and projected into latent spaces using Convolutional Neural Networks (CNNs) [BBC22] for user modeling. These networks are well-suited to process and model unstructured multimedia data through convolution and pooling operations [Zha+19].

For instance, in the study conducted by Hassen et al. [BBC22], a transfer learning technique and an autoencoder were employed to extract the latent features of the images representing items. This was not the first study to use pre-trained models to represent features at a high-level abstraction [LZE15].

Attention mechanisms are extensively employed in the field of Natural Language Processing (NLP), as well as in the extraction of visual features. These mechanisms are applied to the input data to allow the predictions to focus on the most relevant components of the inputs [Zha+19], thereby improving the quality of predictions generated.

From this perspective, a potential strategy could involve employing two attention mechanisms to extract pertinent features associated with users and multimedia items (see, for instance, [Liu+19] and [Che+17]). Subsequently, this extracted information can be incorporated into, e.g., recommendation models.

2.3.2.3 Fusion

Equally important are deep fusion methods, which have been suggested as a viable approach for modeling multiple data sources in user profiling, as shown by Golnoosh and colleagues [Far+18b]. This technique is capable of integrating data from multiple sources (i.e., heterogeneous information [Zha+17]), such as the integration of implicit and explicit feedback [Xie+20], into a unified representation. This integration can occur either at the decision level or at the feature level [Fu+19].

Wen and collaborators [Wen+18] proposed an effective visual background recommendation for dance performances that combined textual information with the visual content of images. In addition, Liang et al. [Xia+10] suggested a method named Injected Preference Fusion (IPF) for combining long- and short-term user profiles in recommendation calculations.

Moreover, with more recent multimodal LLM architectures, one can also develop multimodal feature extraction and representation. We argue that this represents a promising research opportunity and suggest exploring it further in future work, following proposals similar to the one presented by Qiang et al. [Qia+25].

2.3.3 Step 3—User Representation

Various methodologies can be used to represent the user profile based on the information collected in the previous step. The Matrix Factorization (MF), or more generally, Tensor Factorization, serves as an initial point of reference. It is an embedding model that encompasses the diverse interactions, such as ratings, between users and items [BBC22], and can be employed to construct the user profile, considering both the current and predicted interactions. This matrix can be further expanded by incorporating Knowledge Graphs (KGs), for instance, to identify higher-ordered relationships [YDH21]. MF gained wide popularity due to the Netflix Prize Contest [BL07].

Three other main methods are commonly used in the literature to represent user profiles [BKN07; Eke+19]. These three methods can be further categorized into keyword-based and knowledge-based methods [TVB09]. The first approach involves the creation of keyword profiles (see, e.g., [On-+16]), where user profiles are constructed as vectors consisting of pairs of keywords and their respective weights [AX13] according to the degree of interest. The keyword profiles may comprise explicit and implicit information covering long-term and short-term preferences. Furthermore, the weights associated with these profiles can be determined through various methods, such as employing the TF-IDF, and a similarity measure [Far+18a] is then utilized to identify the contents of interest to the user.

However, keyword-based profiles are susceptible to the inherent ambiguity of natural languages [Lop+19]. Specifically, relying solely on keywords may not capture the complete contextual information, polysemy, or underlying intention behind a statement or a textual segment [PCS13]. Due to these limitations, novel approaches were suggested, including semantic network profiles. Semantic network profiles represent user profiles as weighted semantic networks of nodes, where the weights of the nodes and their interconnections denote the user's level of interest [Gao+14]. In this methodology, each node within the network corresponds to a set of keywords that can be mapped to a specific higher-level concept [BKN07], either manually or by using lexical databases such as WordNet [Mil95], or ontologies [HF15; TVB09]. This mapping is user-specific and helps to overcome challenges associated with keyword profiles, particularly those associated with polysemy [Gau+07], as words related to the same concept are clustered within the same node.

Like semantic network profiles, concept-based profiles analyze the interrelationships between words to identify abstract topics or concepts that represent the user's interests [Far+18a]. Nevertheless, unlike semantic network profiles that construct user profiles through the mapping process, concept-based approaches utilize pre-existing concept mappings [BKN07]. In essence, concept-based profiles can be viewed as an expansion of keyword-based profiles, in which concepts replace keywords, and the weights continue to signify the extent of user interest in a specific concept.

It is also important to mention the work by Hou and colleagues [Hou+22], who proposed a universal user and item representation that can be transferred to other systems (e.g., new recommendation systems) without the need to retrain the models for each specific context.

2.3.4 Step 4—Preference Learning

After collecting the user's information and determining how the profile will be constructed, weights can be calculated and associated with the user's preferences. The values of these weights can be static or dynamic. In a later part of this article, the process of profile dynamics will be explored in more detail. For now, the most common [Tan+10] preference-learning techniques are listed below.

Machine learning techniques such as k-Nearest Neighbors (k-NN) have found utility in profiling users based on their shared interests [BBC22]. This method can be used to categorize users according to their interests. Specifically, this method assesses users by measuring their similarity to the k previously known profiles of the closest individuals, commonly referred to as the nearest neighbors. Another clustering technique, the k-Means method, is also a clustering technique that clusters data points into k groups or clusters based on the distances between data points and their respective cluster centroids. One significant advantage of this method is its ability to group users based on their characteristics, preferences, and behaviors without the need for predefined profiles [IFO15].

Support Vector Machines (SVMs), also a machine learning technique, are powerful models that operate by identifying the optimal hyperplane to separate data points of distinct classes while maximizing the margin between them. In the literature, SVMs have been employed to, e.g., update user profiles using positive and negative feedback [Zha+16], as well as to identify relevant documents for users from the Web [Tan+10].

Moreover, CNNs can also be effectively applied to preference learning tasks. Although CNNs were originally developed and widely adopted in computer vision for extracting visual information, their capability to detect local patterns and relationships also makes them applicable to other types of data. Recent studies have confirmed, for instance, that CNNs can capture relevant textual features for preference modeling [Qi+22]. At the same time, their visual feature extraction capabilities continue to be highly relevant for RSs, as illustrated by their effective use in visual preference learning scenarios (e.g., [BBC22]).

In the area of statistical methods, Naive Bayes (NB) employs Bayes' theorem in order to compute the probability of an event happening, given prior information about the conditions associated with the event. In the context of user profiling, NB can be applied to classify unrated Web pages [AT05], predict the web pages that are likely to capture a user's interest [IFO15], and model user preferences [BKN07, p. 390].

In a multi-agent system developed for the purpose of user profiling, diverse agents collaborate and interact during distinct personalization phases, such as information collection [Eke+19]. As a result, the system is capable of identifying hidden structures and capturing the inherent complexity of user behavior.

There are several filtering techniques used in recommendation systems for preference learning [IFO15]. These techniques include content-based filtering, collaborative-based filtering, and hybrid-based filtering [RJ14b]. Content-based filtering involves recommending items based on their attributes and users' characteristics. Collaborative-based filtering, on the other hand, recommends items based on the preferences of similar users. Hybrid-based filtering combines different filtering techniques to provide more accurate recommendations.

Finally, it is relevant to mention here the use of KGs for user profiling. Ontologies were mentioned in Step 2 as a means for textual representation, as they enable knowledge to be represented in a structured manner. Associated with the ontologies are the KGs. If, on the one hand, ontologies are semantic data models that serve as a schema layer for a KG, on the other hand, the KG represents the population of this schema, containing the actual entities and relationships defined by the ontology. As Di Noia et al. [Di +18] mentioned, KGs represent an advantage in selecting the most relevant features to constitute a user profile (c.f., Step 2), since the use of the ontological schema enables selecting the features by semantic relationships. However, they can also be used for preference learning using data summarization techniques, which find subsets of the KG that are most representative of a user's preferences.

In a recent review, Zhang et al. [Zha+24] categorized methods for Knowledge-Based (KB)-based preference learning in recommender systems into three main groups. The first is the two-stage learning method, where representations for entities and relations are learned from the KG first and then used as input for a separate recommendation model. The second is the joint learning method, which combines the objective functions of the KG embedding and the recommendation algorithm into a single objective for preference learning. The third is the alternate learning method, which alternates the training of KG embeddings and the recommendation model.

It is also important to highlight here that the use of KGs has the advantage of addressing some limitations of the traditional recommendation systems, such as data sparsity and the cold-start issue. Furthermore, they can make the recommendations explainable, which increases user trust and acceptance [Cao+19; Zha+24].

KGs also have applications in mobile user profiling. For example, the recent work by Liu et al. [Liu+23b] showed that KGs can be constructed to model mobile users' interactions with their urban environment, and with this, move beyond basic mobility features to also include semantic knowledge.

Another approach consists of combining KGs with reinforcement learning [Wan+20; ZW25] in order to enable dynamic, incremental user profiling (c.f., Step 5), where the KG provides a structured semantic representation of entities and relations while reinforcement learning continuously refines user models through interaction feedback. With the advent of LLMs with reasoning capabilities, a new research direction has emerged that explores the integration of LLMs with KGs, as it allows for evaluating the content generated by LLMs and reducing hallucinations. Cui and collaborators [Cui+25] reviewed this new research direction for healthcare applications, e.g., for personalized clinical decision support.

2.4 Step 5—User Profile Dynamics

Profile modeling can be considered complete at this stage. In this case, user profiles are regarded as static entities, meaning they remain unchanged over time. Nonetheless, it is important to acknowledge that user profiles can rapidly become obsolete, making personalized systems ineffective over time. For instance, in social networks [LN12], user profiles may undergo updates due to age, new hobbies, and qualifications. If these updates are reflected in the user profile, they can lead to an accurate representation [HF15] of their current short- or long-term interests.

To address that issue, profile dynamics techniques have been devised to accommodate the dynamic nature of users' interests and requirements, thus creating dynamic profiles. Profile dynamics enables a system to identify short-term and long-term alterations in user characteristics and incorporate them into their profiles. This ensures the system can provide up-to-date and pertinent personalized recommendations, enhancing user experience and system effectiveness.

The highlighted papers of this subsection are displayed in Figure 2.5.

2.4.1 Long-Term vs. Short-Term Interests

One way of obtaining a dynamic user profile is by creating a user profile, based on weights, for the user's long-term interests—which contains user interests that tend to be more stable over time—and another for the user's short-term interests—which represents users' current interests that change constantly. Short-term interests are usually constrained to periods, such as the previous month, week, or last session [HF15].

Several works have used this technique, making relevant conclusions about the functionality of long- and short-term behavior in identifying user preferences. For example, Lin and collaborators [Li+07] used independent models for both short-term and long-term user preferences. The long-term model integrated a taxonomic hierarchy, whereas the short-term model exploited the user's recent search history. In order to accommodate evolving user preferences, the researchers implemented dynamic user profile strategies based on click-history data.

Paul et al. [Ben+12] studied long- and short-term activity interactions. Their findings showed that long-term interests are most helpful at the beginning of a search process, while short-term interests become increasingly valuable during extended searches. Moreover, the researchers suggested that combining both short-term and long-term profiles leads to superior outcomes compared to using either profile in isolation.

Lucas and colleagues [MIM13] introduced an algorithm that incorporates dynamic user profiles by analyzing both short-term user changes (in this case, the latest user interactions) and long-term data, i.e., the user's several interactions with the system. Their study concludes that, similarly to the findings of Paul et al. [Ben+12], combining these two approaches improves both adaptation speed and robustness.

In the context of location recommendations, Logesh et al. [Rav+21] proposed a recommendation framework that used two distinct agents for long-term and short-term user interests. The suggested approach prioritizes the short-term agent as it captures the user's current preferences more effectively.

Chenlong et al. [DZD22] present a model that integrates positive and negative feedback to enhance personalized search. The model is divided into long-term and short-term modules. The long-term module analyzes users' historical search logs, while the short-term module investigates the specifics of the user's ongoing search interactions in the current session.

Context information at a given moment, such as the user's current location, can also be considered a component of short-term user profiles [Yin+15]. In other words, the system will use this information to adapt its behavior according to the user's needs in that given context and circumstance. Additionally, it has been proposed that visualization of temporal graphs [Tch+10] can be employed to construct long- and short-term profiles rather than solely focusing on preference weights at specific time instances.

2.4.2 Evolution vs. Recalculation

When user interests or preferences change, user profiles must be updated accordingly to maintain accurate personalization. Generally, there are two main approaches for updating user profiles over time: recalculation and evolution [MIM10; MMI14; RJ15].

In recalculation, the user profile is recalculated entirely at fixed periodic intervals (e.g., monthly), without explicitly using previous profiles. Suppose we have two consecutive profiles P_t and P_{t+1}. Let R be the set of user activities or interactions occurring between these profiles. In recalculation, the new profile P_{t+1} is obtained directly from R, ignoring the previously computed profile.

In contrast to recalculation, the evolution approach progressively adjusts the existing profile by incorporating recent user activities, thus explicitly considering the user's profile history. At any given moment, the new profile is calculated incrementally by integrating the previous profiles with recent information. Typically, the evolution approach can be summarized in the following steps:

1. Calculate intermediate profiles at fixed intervals by applying, e.g., machine learning techniques.

2. Integrate these intermediate results with previous profile information: $P_t + P_t' \rightarrow P_{t+1}$.

3. Gather information about user actions or interactions that occurred between profiles P_t and P_t'.

4. Convert the gathered information into training data (or training rules).

5. Repeat steps 3 and 4 until the profile stabilizes and reaches convergence.

To make the distinction between evolution and recalculation clearer, consider a movie recommendation scenario in which a user's preferences are modeled over time. Initially, the user profile may fluctuate as new interactions (such as watching action films or rating romantic comedies) alter the weight of genre preferences. However, as the system collects sufficient data, the profile begins to stabilize, i.e., to converge toward long-term user preferences, with subsequent interactions only slightly adjusting the preference vector. In contrast, with recalculation, the system periodically recalculates the entire profile from scratch using only the most recent data. In this case, stability may be disrupted; however, the resulting profile will more accurately reflect short-term or context-dependent preferences.

Another example could be the analysis of an academic CV. The evolutionary approach looks at a researcher's entire history to show their long-term expertise. In contrast, the recalculation approach focuses only on the last five years to highlight their current activity and recent trends.

2.4.3 Evolutionary/Genetic Algorithms

Evolutionary algorithms are based on the concept of evolution, such as the natural selection process, as popularized by Charles Darwin. In user profile modeling and evolution, these algorithms are used to find and update users' interests by constantly applying, e.g., genetic operators like selection, mutation, and crossover until an optimal solution is obtained.

El Houda and colleagues [ENA19] proposed a genetic algorithm that updates the user's interests by using their queries and current interests. In the proposed algorithm, the weight of the queries or interests is used to create genes, which are then used to create chromosomes. Chromosomes are iteratively transformed by applying genetic operators until an optimal solution is found or a stopping criterion is satisfied.

Rana and Jain [RJ14a], on the other hand, propose an evolutionary clustering algorithm called EVAR (Evolution VARiance clustering algorithm). This algorithm aims to group similar users into clusters and then evolve these clusters to accurately represent the users' evolving preferences over time.

2.4.4 Adaptation Algorithms/Rules

In this algorithm class, algorithms adapt their behavior based on the information fed to them before they are run. Several works [MIM13; Zha+16] propose an adaptation algorithm where the options selected by the user are constantly analyzed to make items similar to those selected rank higher and nonselected items rank lower in the future.

Marin et al. [MIM10] introduced a similar algorithm. In their algorithm, adaptation occurs in two phases, namely: 1) an online phase where the preference is decreased for undesired attributes and increased for desired attributes; 2) an offline phase where the over-ranked items collected over time are analyzed to find which characteristics appear more frequently (over a certain threshold) and then increase the preference for them.

López-Jaquero and collaborators [LMR09] developed adaptation rules and applied them to interface development, whereas Pukkhem [Puk14] worked on a set of adaptation rules for a system in the e-learning context.

2.4.5 Context Awareness

Another way to achieve dynamic user profiles is through context awareness, a technique that presents the system with information about the context in which the user is inserted (e.g., location). Several studies have shown the successful application of context-awareness systems, even in the domain of profile evolution [TVB09]. For instance, Le et al. [Wu+22] introduced a method that incorporates contextual information as a part of user data, with the objective of making content-enriched models also context-aware. Robbie and colleagues [SMG06] presented a context-aware framework designed for the customization of smart home environments with multiple devices situated in diverse contexts within the home. Similarly, SMARTMUSEUM [Ruo+13] is a context-aware ubiquitous RS designed specifically for tourists. This system offers real-time recommendations of cultural information, catering to the unique needs of each individual tourist.

Ontologies have shown their efficacy as tools for modeling user context, with the proposal of standard ontologies in the process of modeling user context [Mar+07]. Sutterer et al. [SDD08] introduced a user profile selection method that selects a profile based on the user's current environment. This method also leverages ontologies to facilitate the modeling of context.

In the domain of content selection and presentation, Xu and collaborators [Xu+10] proposed a method that handles various types of contextual information. This method goes beyond conventional context attributes such as time, device, and location, also making use of sensor data such as temperature and ambient light levels, as well as outputs from other context-reasoning systems, such as user activity recognition and mood detection.

In the domain of user modeling, Qian et al. [Gao+14] developed a context-aware personalized approach that tracks users' digital behavior and query patterns to update their preference profiles. Yin and colleagues [Yin+15], on the other hand, proposed a temporal context-aware mixture model that combines intrinsic user interests with temporal context to analyze social media behavior.

Furthermore, Colombo-Mendoza et al. [Col+15] presented the RecomMetz, a context-aware, knowledge-based mobile movie recommendation system. Interestingly, this system takes into account location, time, and crowd information for recommendations, incorporating a domain ontology as the main component of its recommendation process.

2.4.6 Belief Revision

The AGM belief revision model [AGM85; FH18] is a formal framework to represent the dynamics of the belief of a rational agent. To the best of our knowledge, there are not many works that relate belief revision to users' profiles. In particular, we can mention two papers: Fermé et al. [Fer+24] proposed a method, based on AGM, for creating and dynamizing user profiles to represent and review information about users' interests. It also focuses on defining and enforcing the principle of consistently making minimal changes to the user profile. That means ensuring that changes to users' interests do not happen abruptly and that no significant changes can happen; instead, changes to the knowledge base should be slow and steady [Lau02]. That article also proposed a way to create, represent, and update user profiles by presenting and characterizing four operators to achieve profile dynamics through a belief revision-inspired approach.

The second paper is by Raymond and Song [LS12]. In that paper, the authors developed a service recommendation agent based on belief revision logic to handle the non-monotonicity problem in web service recommendation. They applied belief revision-based reasoning to determine the most suitable context for the initial service request based on the beliefs stored in the user's profile. After service request reasoning, the set of potential web services is identified and ranked. The highest-ranked services are considered to be the most desirable ones that match the user's specific interests.

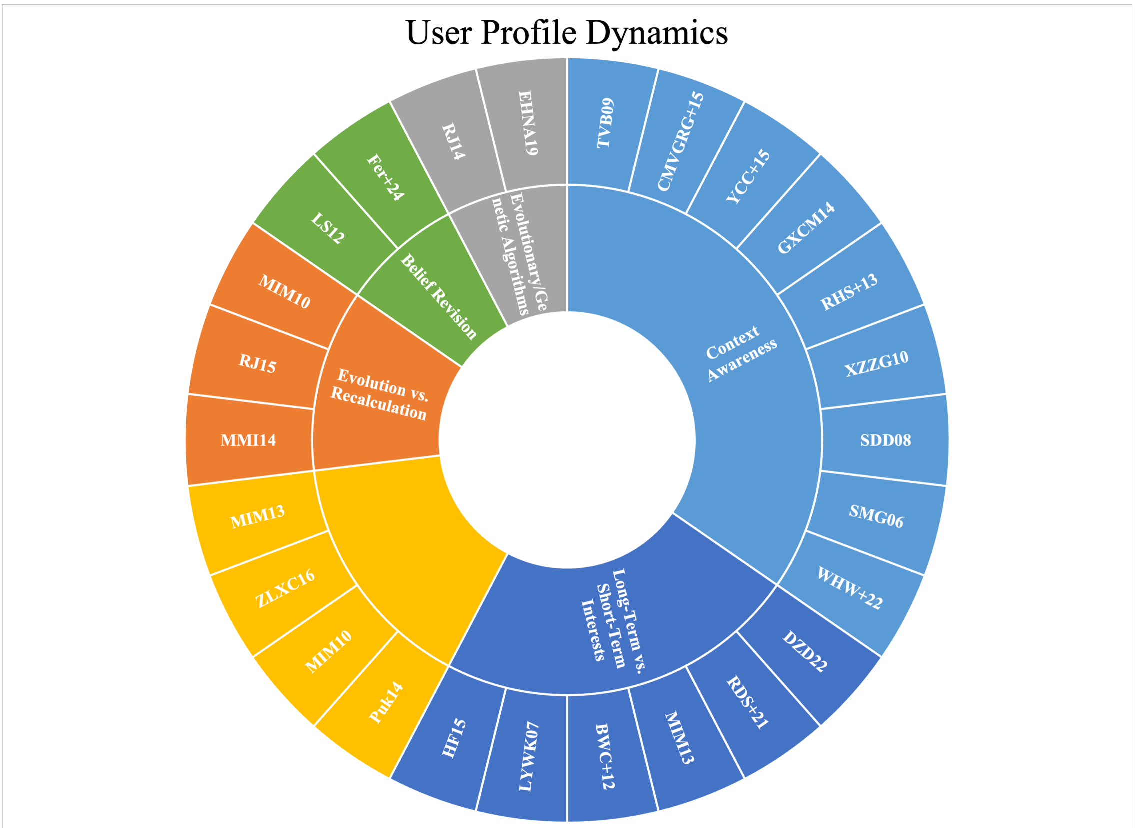


Figure 2.5: Highlighted papers on Profile Dynamics. The alphanumeric labels (e.g., MIM10, DZD22) correspond to the list of selected studies provided in Appendix A.

2.5 Applications

User profiles are primarily used in two main types of systems: RSs and personalized/adaptive systems. RSs use information from user profiles to suggest specific items (such as products, songs, or movies) that align closely with users' interests and preferences. On the other hand, personalized or adaptive systems use user profiles in a broader sense. They not only deliver personalized content but also dynamically adjust their behavior, interfaces, and functionalities according to users' characteristics, needs, and context.

2.5.1 Recommendation Systems

RSs are systems that use the information provided by the user profile about the user's interest in a particular item (e.g., movie, song, or book) to recommend other similar items to the user [Bob+13]. This study area emerged in the mid-1990s when researchers focused on predicting the ratings of items the user had not yet seen.

It is important to highlight that RSs have been successfully applied in various heterogeneous, human-centered, and socially relevant fields. Examples include applications in education [Sil+23], healthcare [Cha+21a], and social media [Aiv+19].

RS can be classified as Collaborative Filtering (CF) Recommendation, Content-Based (CB) Recommendation, and KB Recommendation. Additionally, hybrid recommendation approaches exist, combining these different recommendation techniques. Hybrid methods aim to overcome the limitations of one approach by leveraging the strengths of another, thus achieving more effective results [AT05; IFO15].

Before describing the different recommendation approaches, we highlight the Cross-Domain Recommendation (CDR) systems. CDR systems integrate user preferences and interaction data collected from multiple domains to enhance recommendation quality. Since they can integrate data from multiple domains, these systems can also help reduce common issues such as data sparsity and the cold-start problem [Zan+22].

It is worth mentioning here the work by Berkovsky et al. [BKR07], which represents, to the best of our knowledge, one of the earliest studies addressing CDR systems in the context of collaborative filtering. More recently, Chen and Lee [CL24] integrated the notion of profile evolution into a CDR system. We note that if a CDR system aims to improve recommendation accuracy across multiple domains simultaneously, it is referred to as a *multi-target* recommendation system. For an in-depth exploration of CDR systems, we refer the reader to the recent survey by Zhu and team [Zhu+21].

2.5.1.1 Collaborative Filtering Recommendation

CF compares the target user to other users and finds similarities between them. The users with greater similarity are then grouped into neighborhoods or clusters. The previous ratings of these groups are analyzed to find items that have not yet been rated by the user but might be of interest based on the positive ratings.

This is the most widely used recommendation technique and is best suited for recommending items that are difficult to describe in an objective manner, such as movies, art pieces, or music [IFO15].

CF can be classified as Model- and Memory-based. Memory-based methods consist of heuristics that use the information collected about the interests stored in the user profile to make predictions (e.g., similarity metrics to obtain the distance between users and items). On the other hand, model-based methods use the interests information to fit a model that will later be used to make recommendations by predicting their effectiveness in satisfying the user's needs [AT05; Bob+13].

Table 2.1 presents some of the most common techniques used in CF according to their classification. In this table, clustering appears as both a model- and memory-based method. This is because clustering builds abstract user or item groups that function as models, yet its construction and updating rely directly on stored interaction data to compute similarities and reassign members. Other approaches, such as regression or neural networks, cannot be considered hybrid since their learned parameters replace the need for direct access to historical data during inference, while purely memory-based methods rely exclusively on those stored interactions.

Table 2.1: A list of the most used CF techniques [AT05; Bob+13].

Category	Techniques
Model-Based	Genetic Algorithms [UB02] Fuzzy Logic [MMI14] Neural Networks [Nau+19; Wu+22] Bayesian Networks [AA14] Linear Regression [CP09] Probabilistic Models [SA19] Clustering [RJ14a] Latent Semantic Indexing (LSI) [Bob+13] Singular Value Decomposition (SVD) [Bob+13]
Memory-Based	Similarity Measures [SC12] Aggregation Approaches [TNM18] Nearest Neighbor (Cosine Correlation) [AT05] Clustering [SC12] Graph Theory [AT05]

CF strategy works well when dealing with difficult items (such as non-objective items) to describe because they depend on factors like opinions. It can also make new and effective recommendations that are not represented in the user profile by detecting interests that the target user has yet to show, but that similar users have already expressed.

However, one of the main limitations in CF is known as the cold-start problem, which happens at the beginning of the interaction when the system does not have enough information about the user to find their neighbors and make accurate recommendations [Bob+13]. Data sparsity is also a common problem because users are often unwilling to spend time writing reviews or providing feedback. As a result, a lot of information is missing from the database [IFO15]. Due to the processing of vast amounts of data, scalability is also considered a problem, especially when the number of users grows. MapReduce can, however, minimize this problem [CKL15]. Finally, synonymy problems or polysemy may occur when similar items have different names in the database but similar meanings, making it challenging to distinguish similar items [IFO15].

Nevertheless, several studies have explored the use of CF for real-world recommendation problems. For example, Wei and Seung-Taek [CP09] proposed a recommendation framework for Yahoo! News that applied CF and addressed the cold-start issue by enriching user profiles with dynamic features. Chhavi and Sanjay Kumar [RJ14a; RJ14b], on the other hand, introduced an evolutionary clustering algorithm designed to identify groups of similar users more accurately.

In the context of digital libraries, Lun-Chi et al. [CKL15] proposed an RS based on pairwise user similarity computed via an ontology model. Additionally, Aminu and collaborators [Dau+20] developed a deep learning model for extracting product features, aiming to improve CF-based recommendations.

In another study, Logesh et al. [Rav+21] presented a location recommendation system built upon a multi-agent framework, which also integrated short- and long-term user behaviors for dynamic recommendations. Finally, Zhongqi et al. [Lu+16] proposed a method called Collaborative Evolution (CE), which integrates MF and autoregressive vectors into a unified learning framework. In their method, users' latent interests at different points in time are iteratively learned through CF, and the temporal evolution of these interests is explicitly modeled to guide the learning process further.

2.5.1.2 Content-Based Recommendation

This method analyzes past user interactions with the system, identifies which items the user liked in the past, and searches for similar ones to recommend. This type of recommendation is ideal for document recommendations, such as books and news.

CB systems can be classified into case-based reasoning and attribute-based techniques. Case-based reasoning recommends items that are most related to items the user previously liked. On the other hand, the attribute-based technique focuses on the features of the items liked in the past and recommends new items with those same features [TNM18]. Some of the techniques used in this kind of system are Bayesian Classifiers, Clustering, Decision Trees, and Artificial Neural Networks.

These systems have the benefit of not depending on the information from other users, thus relying only on the information of the target user; to adapt their recommendations quickly; to be more private because users do not have to share their profile information for the system to function; to support explainable recommendation since they base their recommendations on item features which facilitate explainability [IFO15]. However, they depend on the items being described accurately and extensively to make practical and accurate recommendations. They can also suffer from overspecialization, which happens when users get similar items consistently and cannot discover new interests. The cold-start problem can also challenge the usability of these systems [AT05].

Several studies have applied CB recommendation methods. For instance, Mehrpoor and collaborators [Meh+15] presented a RS designed to facilitate knowledge accessibility in an engineering environment, benefiting from both CB and collaborative filtering techniques. Additionally, Liu et al. [Liu+18] proposed a CB RS that leverages ontological information.

2.5.1.3 Knowledge-Based Recommendation

This method performs a deep analysis of the items and uses this knowledge to find ways to satisfy specific user needs [Fel+06]. This deep knowledge is sometimes accomplished by using ontologies to represent it because they can be used to describe objects semantically without the constraints of other models [Col+15; TNM18].

These kind of system have an advantage over CF and CB systems. More specifically, they do not suffer from problems that stem from new users and new items (e.g., cold-start problems) or rating sparsity problems because they make use of the deep knowledge acquired on the items to calculate recommendations [AT05; TNM18].

Some limitations of KB systems are the need for an understanding and knowledge of engineering skills [TNM18] and the need for deep knowledge acquisition of the problem at hand. For this reason, the majority of these types of systems are developed for areas where domain knowledge is easily available [AT05].

Applications of KB recommendation methods can be found in the literature. For example, Felfernig et al. [Fel+06] developed a KB recommender called CWAdvisor, illustrating its use in two distinct scenarios: as a digital camera advisor and as a financial service advisor. Another study by Codina and Ceccaroni [CC12] presented KB RSs that use semantic relationships between items to enhance recommendation performance, describing various approaches for creating semantically enriched models. Additionally, Ruotsalo et al. [Ruo+13] proposed a KB recommender focused specifically on tourism sites, such as museums. Finally, Colombo-Mendoza et al. [Col+15] introduced a context-aware KB recommendation system designed to recommend movie showtimes.

2.5.1.4 Explainability in Recommender Systems

Over time, user profiling methods have evolved from simple rule-based systems toward more complex machine learning models [Pro+25]. With this increasing complexity, the resulting recommendation systems have also become less transparent, often behaving as black-box models lacking clear explanations about how they generate recommendations.

This lack of transparency motivates the need for explanations in RSs, leading to the emergence of EXplainable Artificial Intelligence (XAI) techniques in the area of recommendations. Explanations help users better understand why [ZC20] certain items were recommended to them. Besides improving the overall user experience [Afc+22], explanations can provide several additional benefits, according to the literature [TM07], namely:

- Transparency—Show the user how the system chose the recommended item.
- Scrutability—Give the user the ability to correct the system in the case of an ineffective recommendation (e.g., "not interested" button).
- Trust—Measure how much the user trusts the system, which can be achieved by questionnaires or by analyzing customer loyalty and sales numbers.
- Persuasiveness—The ability of the system to convince the user to choose an item they would typically not be interested in.
- Effectiveness—Help users make the correct decision according to their preferences. Additionally, the level of satisfaction with the item before and after an action can be analyzed to see if it really was a good fit for the user.
- Efficiency—The user should be able to use the system and select recommendations quickly. This can be measured by seeing how long the user took to complete a task.
- Satisfaction—Aims to increase the users' satisfaction with their choices and with the system.

Additionally, Zhang and Chen [ZC20] proposed categorizing explainable recommendation system problems according to the dimensions of what, when, who, where, and why (5W).

There are several examples of systems that incorporate explanations into their recommendation processes. One example is a General Knowledge Enhanced Framework for Explainable Sequential Recommendation (GFE) [YDH21]. This system captures users' preferences along with the evolution of their dynamic interests. Another example is the User-Centric Path Reasoning (UCPR) [Tai+21]. This framework creates explainable recommendations by guiding the recommendation search process according to user demands, thereby enhancing the diversity of explanations provided. Lastly, a system called Temporal Meta-path Guided Explainable Recommendation (TMER) [Che+21] achieves explainable recommendations by sequentially modeling evolving user-item interactions within a dynamic KG.

2.5.1.5 Generative Recommendation

Unlike traditional discriminative models that rank a fixed set of candidates, generative recommendations utilize LLMs to formulate recommendations as a language generation task. These recommendations are typically interactive, explainable, and contextually adaptive. Nevertheless, Wu et al. [Wu+24] categorize LLM-based recommendation systems into discriminative and generative LLMs. In discriminative approaches, the model serves as an encoder that extracts semantic embeddings to improve ranking or matching accuracy. This means that LLMs can also perform feature extraction tasks, producing representations that serve as input for RSs. In contrast, generative models operate as autonomous sequence generators that synthesize personalized outputs directly from text prompts containing user, item, and contextual features.

Several studies have explored how LLMs can maintain and update user profiles over time [Jia+25; Sab+25]. Jiang et al. [Jia+25], for example, evaluated how LLMs can construct, remember, and adapt evolving user personas across multiple interactions. Sabouri and team [Sab+25] suggested an explainable temporal user profiling framework in which a LLM generates concise natural-language representations of short- and long-term user histories.

Chen et al. [Che+24] argue that LLMs will enable current recommendation systems to be more than content filtering, transforming them into interactive service generators where models can plan, reason, and create customized responses through natural language. Kirk et al. [Kir+23], with the focus on model alignment, complements this view and suggests applying reinforcement learning with human feedback as a mechanism for fine-grained adaptation of model outputs to user preferences and values.

Regarding user segmentation, Li et al. [LLY25] applied a generative model to consumer segmentation, and showed that LLMs can synthesize realistic persona-driven conversations and predict segment-specific preferences with high consistency. Similarly, Salemi et al. [Sal+24] introduce a benchmark, which evaluates an LLMs' capacity to adapt its outputs across seven personalized text classification and generation tasks. These two works suggest that LLMs can simultaneously encode user intent, retrieve relevant historical data, and generate adaptive recommendations.

A natural extension is to use the multimodality capacity of some of these models to improve the task of user profiling (in other words, to make the user profile richer in information). Indeed, these models can be used for multimodal embedding fusion, in which image, audio, and text features are aligned within a shared latent space to support generative recommendation [Qia+25; She+24]. This multimodal capacity aligns closely with emerging paradigms of Artificial Intelligence (AI) agents [Zha+25], where, through collaboration within an ecosystem of multimodal agents, one can obtain enriched user profiles with contextual, multimodal information, thus enabling better generative and adaptive recommendations. For example, Zhang [Zha+25] considered users and items as intelligent agents capable of perceiving, reasoning, and interacting across modalities, while continuously exchanging information to co-evolve their understanding of preferences and contexts. In other words, this allows each agent to continuously learn and collaboratively refine user profiles and recommendations.

2.5.2 Personalized/Adaptive Systems

Personalized systems use information from user profiles to adapt their interface and behavior according to users' needs and preferences. Typically, personalization aims to display content that interests users and filter out irrelevant items. Additionally, personalization can include other more detailed adaptations, such as adjusting the font size and style based on the user profile [Bro90]. According to Fan and Poole [FP06], personalization can be described across three main dimensions: (1) *what is personalized*, referring to the type of content presented, the presentation style, the information delivery channel, or the functionality itself; (2) *the target of personalization*, referring either to specific user groups (e.g., pet owners, families with small children) or individual users; and (3) *who performs the personalization*, reflecting the degree of automation involved in the user-modeling process. When users explicitly provide information to the system, personalization is called explicit personalization. Conversely, if the system automatically performs profile adaptation, this is known as implicit personalization.

Various studies have explored personalization approaches in different domains, including personalized search and e-learning. In the area of personalized search, Moukas [Mou97] introduced a system called Amalthaea, which provides personalized search capabilities through information filtering. Achemoukh and Ahmed-Ouamer [AA14] proposed a method for modeling and representing long- and short-term user profiles based on Dynamic Bayesian Networks. Similarly, Hawalah and Fasli [HF15] presented a multi-agent personalized search system, which dynamically builds user profiles by analyzing both short-term and long-term user interests inferred from previously visited web pages.

In the context of e-learning, several personalization approaches can also be found. For example, Le et al. [Le+09] proposed a generic model capable of providing personalized learning resources and services within blended learning contexts. Razmerita and Lytras [RL08] introduced an ontology-based user modeling framework, tested within the context of a semantic learning portal, highlighting the explicit use of user modeling in adaptive systems. Another work by Le and collaborators [Le+10] discusses the application of user profiles in Adaptive Instructional Systems, aiming for improved utility in blended-learning scenarios.

It is important to note that although many authors use the terms "personalized" and "adaptive" systems interchangeably [Le+10], a key difference separates the two concepts. Adaptive systems explicitly take into consideration changes in user profiles over time; in other words, adaptive systems explicitly incorporate profile dynamics.

In this context, LLMs have started to be considered as promising systems capable of combining personalization (via, e.g., personas) and adaptation capabilities. User profiles constructed from previous interactions with these models can be used to personalize the model's future outputs and responses [Sal+24]. Besides user segmentation [LLY25], LLMs can also perform feature extraction tasks, producing representations that serve as input for RSs, or even directly generating recommendations expressed in natural language [Wu+24]. Another recent approach involves reinforcement learning based on human feedback, which helps to better capture the user profiles, thus providing higher-quality recommendations [Che+24]. Nonetheless, existing techniques still face limitations in effectively capturing users' current situations and the dynamic nature of user preferences [Jia+25]. On the positive side, recommendations and user profiles generated by LLMs tend to be easier to explain compared to traditional deep learning models [Sab+25]. However, it is important to highlight recent concerns regarding alignment issues in LLMs, emphasizing the need for careful system design and evaluation to avoid undesired system behaviors [Kir+23].

2.6 Conclusion

This review has explored the literature on user profile modeling and profile dynamics. We believe that the proposed taxonomy, which includes user profile modeling, user profile dynamics, RS, personalized systems, and adaptive systems, has helped clarify the discussion and provided a structured overview of existing research in the field.

We highlighted how accurate and frequently updated user profiles enhance the effectiveness of personalized and adaptive systems. Specifically, we discussed several methodologies for creating user profiles, techniques for keeping them up-to-date over time, and the primary application domains where these profiles are employed.

Several conclusions can be drawn from our analysis. First, user profiles play a key role in enabling personalization across different systems, such as RS, personalized interfaces, and adaptive applications. Second, both explicit and implicit data collection methods

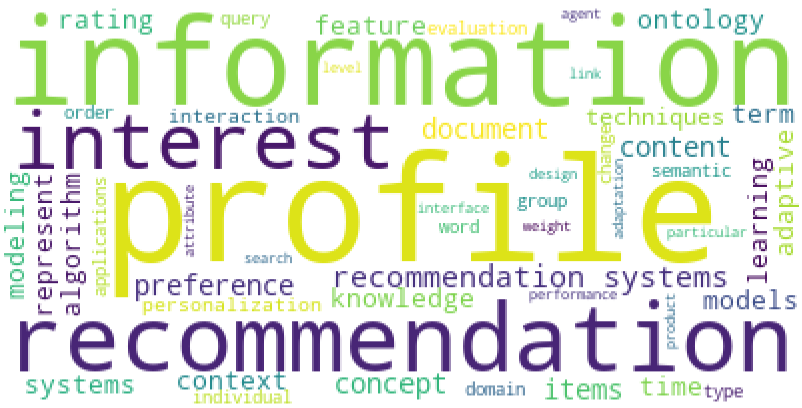


Figure 2.6: Word cloud showing the most frequently used terms in the extracted papers.

provide valuable input when creating user profiles. Hybrid approaches, in particular, help improve profile quality by combining the strengths of implicit and explicit user information. Third, differentiating between short-term and long-term user interests is important for dynamic user profiling, as capturing both timeframes leads to better personalization and recommendation accuracy. Fourth, profile dynamics techniques, such as context-awareness methods and belief revision frameworks, help address natural user profile evolution effectively. Fifth, explainability has become increasingly important for recommendation systems, as it can make recommendations easier to understand, improve user trust, and provide greater user satisfaction. Finally, recent developments involving LLMs have shown promise for achieving advanced user modeling and producing personalized natural-language recommendations. However, alignment and ethical considerations still need further attention.

Although substantial progress has been made, some important research directions remain open for exploration. First, it is worth exploring multimodal and heterogeneous data integration further. In particular, new approaches leveraging multimodal deep learning architectures may help create richer and more informative user profiles. Second, developing scalable user profile update methods that can handle dynamic changes in real-time is another significant challenge, especially as personalized systems continue to grow in size and complexity. Third, given the rising concerns about user privacy and regulations such as the GDPR, research into privacy-preserving approaches, e.g., federated learning, is expected to become more important. Fourth, more effective methods to capture contextual and situational information will enhance the adaptation of user profiles, particularly in mobile and ubiquitous computing scenarios. Finally, another promising research direction would be the development of universal and holistic user profiles capable of generalizing across multiple platforms, applications, and usage contexts.

The word cloud presented in Figure 2.6 was generated from the corpus of extracted articles and visually highlights the most frequently occurring terms related to user profiling. In the word cloud, the size of each term directly reflects its frequency in the analyzed literature.

As illustrated in the figure, the terms "recommendation" and "recommendation systems" appear frequently, suggesting that recommendation systems represent one of the primary applications of user profiling. Within the scope of recommendation systems, several studies propose approaches that dynamically adapt recommendations to the user's current contextual environment.

CHAPTER 3

Public Perceptions, Knowledge, Responsibilities, and Behavior Intentions on Marine Litter: Identifying Profiles of Small Oceanic Islands Inhabitants

Sara Bettencourt, Diogo Nuno Freitas, Sónia Costa, Sandra Caeiro

Abstract. *Marine litter is a global threat, particularly on oceanic islands where the problem is exacerbated. Perceptions, knowledge, awareness, and attitudes towards the theme are crucial in its mitigation and prevention. This study assessed these points through a questionnaire to the inhabitants of a Portuguese archipelago. Data revealed that people associate marine litter with plastic and its impacts and are well informed about its sources and pathways. Yet, the degradation rates of marine items were frequently underestimated and the problem of marine litter was attributed, among others, to littering, single-use products, and excessive packaging. Some individuals did not consider themselves responsible for reducing marine litter, attributing responsibilities to third parties. The youngest group, men, and students were the ones who reported less litter-reducing intentions and behaviors. Distinct profiles were traced using the questionnaire's answers, highlighting who needs marine litter literacy. Individuals who do not consider marine litter a current threat and live in a community that does not care about marine litter (profiles 1 and 2) were the groups that needed deeper intervention, due to their low perception and understanding of the problem. Marine litter literacy, management, and governance measures are necessary so that the public recognizes marine litter as a current threat, is worried about its impacts, avoids plastic use, and chooses reusable products (profile 4). In the studied oceanic islands, results indicated marine litter is not fully*

perceived by the public. A global and transformative shift in the way people are educated and behave towards waste and pollution is required, thereby highlighting the importance of increasing public perceptions assessment and marine litter literacy in the society.

3.1 Introduction

Marine litter is a worldwide concern that poses severe threats to the environment, society, and economy. Urgent actions to avoid and reduce marine litter are needed. At the G20 Summit in 2017 (Hamburg, Germany), a "G20 Action Plan on Marine Litter" was adopted to "Promote public information campaigns for citizens and businesses to reduce waste generation, to re-use and to facilitate their participation in waste collection systems and to avoid littering" [Bar+18, p. 9].

Marine litter, also known as marine debris, can be defined as the persistent, manufactured, or processed solid materials discarded, disposed of, or abandoned in the marine and coastal environment [UNE09]. These materials can be deliberately dumped into the sea, coasts, rivers, and beaches or brought indirectly to the ocean by rivers, sewage, stormwater, currents, tides, or wind. The accidental loss also contributes to the accumulation of litter in the sea [Vei+16]. The sources of marine litter can be classified as land- or ocean-based, depending on how debris enter the sea. The land-based sources include the recreational use of the coast, unprotected landfills, dumping of garbage (households and industries), public littering, and sewage overflows [GHM15]. It is estimated that these sources contribute to about 80% of the total litter in the ocean [All+06]. Nevertheless, this assumption is poorly documented and should be validated [WO18]. The ocean-based sources are associated with human activities and actions at sea: fishing, merchant shipping, research, and military vessels, recreational boats, cruise ships, and offshore petroleum installations [SR07].

Worldwide, plastic is the most common material found in the marine environment, followed by cigarette butts, metal, glass/ceramics, wood, paper/cardboard, and textiles/fabrics [Hah20]. The continuous input of marine litter items causes severe impacts. Entanglement, smothering, ingestion, ecosystem deterioration, and dispersal of non-native species are some of the widely known environmental consequences of marine debris [Fos+18; GT15; KRF15]. Social [PH11; SR07] and economic [MLB10; New+15] threats can equally be directly or indirectly attributed to marine litter. These environmental, social, and economic impacts are typically interconnected, being difficult to tackle separately.

Addressing the problem requires understanding and engagement of citizens with the topic [Dil+19; Ray+19]. Public perceptions can be measured as a combination of awareness, knowledge, and concern [Oos+22]. Furthermore, identifying public perceptions is crucial for developing management measures and policies [GPM20; Ols+20]. Prevention, mitigation, removal, and behavior change are the four actions proposed to reduce marine litter inputs and impacts. Prevention is one of the most favorable approaches, given that it prevents the generation of debris and their entry into the sea, as well as behavior change measures, which have a long-term impact. Education can be used to support behavior change [BCC21; Har+18b; Wic+22], as it influences and promotes the adoption of sustainable behaviors [Che15; Ran+20], particularly important when tackling marine litter and waste management problems.

There has been a growing number of studies assessing people's perceptions of the marine litter problem, its impacts, and solutions [Dil+19; Fer+21; Fil+21; Har+18a; SGC12]. Nevertheless, few authors linked what people know about marine litter, what people's efforts to prevent it are, and which demographic factors influence it. With such data, it is possible to quickly establish profiles that determine who needs marine litter literacy interventions. If this is not understood, interventions to increase marine litter literacy cannot be properly designed.

Considering marine pollution is a prevalent problem on oceanic islands [Kie+17], with some ecosystems acting as crucial marine litter accumulation zones [Por+22], we examined the perceptions, knowledge, responsibilities, and behavioral intentions of people from two oceanic islands about litter. We combined that data to trace different profiles and investigate whether people living in these locations better perceive the marine litter issue and act accordingly to prevent it. If the hypothesis that marine litter is not fully perceived by the public becomes validated, measures to manage the problem and guide future research can be adapted.

3.2 Methods

3.2.1 Study Site, Participants, and Sampling

Marine litter is a significant problem frequently exacerbated on oceanic islands. The present work was carried out on two small oceanic Portuguese islands—Madeira and Porto Santo—which are the inhabited islands of the Autonomous Region of Madeira. Madeira is an important touristic destination awarded eighth times as the "Europe's Leading Island Destination" by World Travel Awards, and Porto Santo the "Best Beach in Europe 2022" by European Best Destinations. Atmospheric circulation and the southward Portugal currents are important pathways of marine litter intercepting Madeira Island [CC21].

The 2021 census revealed that Madeira and Porto Santo have a population of 251,060 individuals. To determine the sample size of the population, the adjusted Yamane's formula was applied. The equation was adjusted to allow the determination of optimum sample sizes for continuous and categorical variables at all levels of confidence: $n = N/(1 + N\epsilon^2)$, where n is the minimum returned sample size, N the population size, and ϵ the degree of accuracy expressed as a proportion [Ada20]. A sample of 271 valid questionnaires was determined to be needed, considering a margin of error of ±5% and a confidence level of 90%.

The geographical dispersion—Madeira (741.00 km^2) and Porto Santo (43.00 km^2)—required an extra effort to survey residents of all municipalities, particularly those who live in more rural areas (e.g., by sending the questionnaire to the different municipalities and parish councils). Questionnaires were randomly distributed in printed (for the less digitally literate) and electronic forms (social media, email, and QR codes) from April to July 2021 following a convenience sampling technique, where data is collected until the required sample size is reached [SLT09]. To reach a wider and diversified sample, the questionnaire was spread through social media advertising to Internet Protocol (IP) addresses in Madeira and Porto Santo islands.

Lime Survey, which is an online survey software, was used (version 2.06) to collect the answers. Individuals were informed about the research and the confidentiality of responses and provided consent to participate in the study. A total of 350 valid questionnaires were obtained. For the questionnaire answers to be valid, the respondents had to be older than 18 years old and reside on the island of Madeira or Porto Santo and the survey had to have all the questions answered.

3.2.2 Questionnaire

Considering the research question "What do people know and how do they perceive marine litter?", a questionnaire was prepared following the recommendations to be based on existing studies and use standardized questions [Lot+18]. Questions were adapted from the MARLISCO survey "Perceptions about marine litter" [Har13], and new ones were added according to the survey's population and geographical context. The questionnaire was pre-tested with a small group as suggested by Bolarinwa [Bol15]. The test group included 25 people of different ages, gender, profession, and place of residence. With feedback from respondents, minor adjustments were made in the questions and answer options. The final version of the questionnaire comprised 20 questions (1 short answer and 19 closed answers) and space for optional comments/observations. All questions and response options were written in Portuguese and the estimated time of completing the survey was 10 min.

The short-answer question was the first query of the questionnaire. It requested the participants to write two words immediately associated with marine litter, which were subsequently clustered into eight categories. That was the first question in the questionnaire to avoid biased answers, as the subsequent words used in the questionnaire could influence the answers. The following questions were closed-answers of multiple choice or single choice Likert scale and covered distinct topics (Appendix B, Table I).

As some municipalities had small sample sizes, the following grouping was performed, as proposed by Hermida and Costa [HC20]: Funchal, Câmara de Lobos, South-West Coast (including the municipalities of Ribeira Brava, Ponta do Sol, and Calheta), North Coast (including Porto Moniz, São Vicente, and Santana), South-East Coast (Machico and Santa Cruz), and Porto Santo Island. The professional activity was codified according to the groups established in the "International Standard Classification of Occupations" [ILO12] and used by the Portuguese National Institute of Statistics [INE11]. Three additional groups were introduced to facilitate the self-filling of the questionnaire: student, don't know how to answer, and other. Some of the answers provided in the "don't know how to answer" and "other" fields were manually introduced in the respective professional groups by the authors when the data were analyzed.

3.2.3 Statistical Analysis

The data from the questionnaires were statistically analyzed using Python programming language (version 3.8), with auxiliary external libraries, such as the SciPy library. Each statistical test was conducted considering the significance level of .05 (alpha) [Coo14]. Due to a low number of respondents, the conducted statistical analyses did not include respondents with households of zero members or respondents that answered "NA" in the gender. Likewise, the only occupations considered were technicians and associate professionals, clerical support workers, and students, as it would not be possible to draw meaningful and general conclusions about other respondents.

All Likert scale questions were considered on an ordinal scale, with responses encoded as a value from 1 through 5. For this reason, instead of reporting the mean and standard deviation, the median and Interquartile Range (IQR) are used in the descriptive statistics, as they provide a more accurate measure of central tendency [Coo14]. The IQR can be interpreted as a measure of the spread of the data of the central group. In this view, if IQR = 0, then a significant number of respondents provided the same answer.

Moreover, Likert scale questions were analyzed with non-parametric tests to test if the difference between the examined groups was significant [BB12; Jos+15]. More particularly, the Kruskal–Wallis one-way analysis of variance (which is an extension of the Mann–Whitney U test for more than two groups) was used to test whether the difference in distributions of three or more groups concerning one factor is significant. Since the Kruskal–Wallis test does not specify which groups are statistically significantly different, the Dunn's test, suitable for testing the difference between two conditions, was used for the pair-wise post-hoc tests, thus detecting the groups that differ. A two-sided alternative hypothesis was tested for the latter statistical test and the p-values were adjusted according to the Bonferroni adjustment for multiple comparisons [Coo14].

The Chi-Squared (χ^2) test of independence was used for the multiple-choice questions to test whether the association between categorical variables was statistically significant. It is important to note that the Chi-Squared test of independence assumes that the observations are independent of each other and requires the expected frequencies to be at least 5 in no less than 80% of the cells in the contingency table. This last prerequisite was not satisfied for some statistical tests. In these cases, the statistical results were marketed as

N/A (not available). Finally, the Chi-Squared test of independence was also used for the pair-wise post hoc tests; Yates' correction for continuity was applied, and the p-values were adjusted according to the Bonferroni adjustment for multiple comparisons [Wue11].

In the Kruskal–Wallis, the effect size was calculated using the eta squared formula and the Cramér's ϕ was reported for the significant Chi-Squared tests of independence [Coo14].

It is also of interest, based on the data from the questionnaires, to identify the respondents' profiles. The method for selecting respondents' profiles consists of two steps: cluster analysis and Decision Trees (DTs) splitting. The k-means clustering method [Llo82] was used to group respondents into five clusters based on the average data in each cluster. However, this method does not identify the rules used to split the groups. In this manner, DTs were used. The input samples for DTs were the data from the questionnaires, and the target values were the groups identified by the cluster analysis. The algorithm starts by selecting the root node corresponding to the question that best separates the respondents. After that, the DT continues this process for the next important feature, splits the participants, and creates and adds new branches to the DT [Für10]. This process is repeated until the maximum depth of the tree (in the case of this work, 5) is attained. It is noteworthy that the Gini Impurity function was used to identify the DT's root node and subsequent question splits.

3.3 Results

3.3.1 Survey Sample–Demographic Data

A total of 350 valid questionnaires were obtained, achieving more than the minimum of the required sample. Table 3.1 clusters the respondents' sociodemographic variables. Most of the respondents were between 36 and 50 years old, were female, and lived in the municipality of Funchal. Households with three members were the most common. Concerning the respondents' occupations, the majority were professionals (e.g., science and engineering; health; teaching; business and administration professionals), followed by technicians and associate professionals (e.g., science and engineering; health; business and administration associate professionals), and students (secondary and higher education).

3.3.2 Definition

The respondents' concept of marine litter was evaluated through an open question, where people wrote the words they associate with marine litter (Figure 3.1). "Plastic", "microplastic", and "bioplastic" were referred to 189 times (27.0%). Terms linked to marine litter impacts and "pollution" were also frequently mentioned. Solutions to reduce marine litter were the less written ideas (2.9%).

Table 3.1: Sociodemographic characteristics of the respondents.

Item	Variable	No.	%
Gender	Female	195	55.7
	Male	153	43.7
	Not answered (NA)	2	0.6
Age groups	19–25	53	15.1
	26–35	81	23.1
	36–50	129	36.9
	≥51	87	24.9
Residence area	Funchal	215	61.4
	Câmara de Lobos	26	7.4
	South-West Coast	26	7.4
	North Coast	6	1.7
	South-East Coast	70	20.0
	Porto Santo	7	2.0
No. of household members	0	11	3.1
	1	38	10.9
	2	74	21.1
	3	115	32.9
	4	82	23.4
	≥5	30	8.6
Occupations	Managers	10	2.9
	Professionals	168	48.0
	Technicians and Associate Professionals	47	13.4
	Clerical Support Workers	33	9.4
	Services and Sales Workers	14	4.0
	Skilled Agricultural, Forestry, and Fishery Workers	3	0.9
	Craft and Related Trades Workers	1	0.3
	Plant and Machine Operators and Assemblers	1	0.3
	Elementary Occupations	9	2.6
	Armed Forces Occupations	4	1.1
	Student	42	12.0
	Don’t know how to answer	5	1.4
	Other	13	3.7

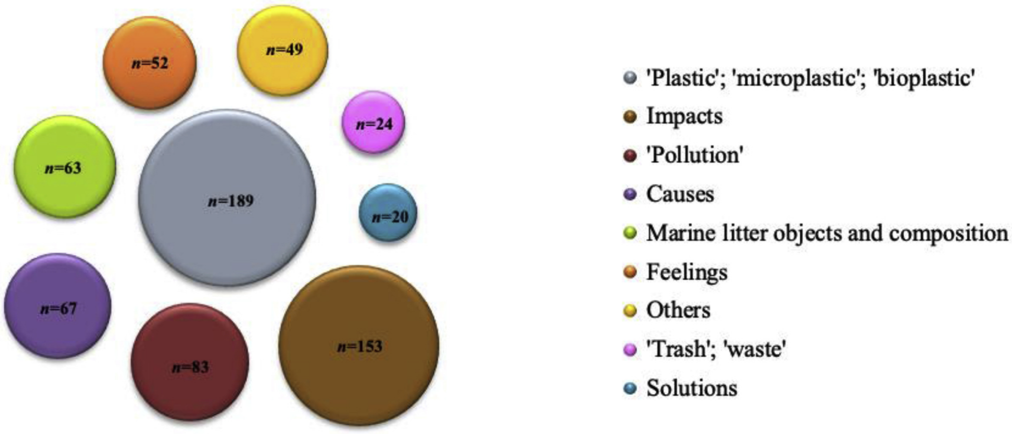


Figure 3.1: Representation of the first words participants linked with the concept of marine litter. Colors refer to different groups of words, and the circle diameter indicates the word frequency.

3.3.3 Composition, Degradation Time, and Microplastics

Plastic, cigarette butts, and sanitary waste were the three most common marine litter composition materials ranked by the participants. 98.0% of citizens admitted that plastic belongs to the top 3 marine litter material composition. Cigarette butts were indicated by 73.1% of the surveyed participants and sanitary waste by 28.6%.

Knowledge regarding the estimated degradation rates for frequent marine litter items was evaluated having as reference values the rates from National Oceanic and Atmospheric Administration and Woods Hole Sea Grant (Figure 3.2). More than half of the respondents believed that cigarette butts take less than 20 years to decompose, with 24.3% correctly estimating the degradation time (1–5 years). This was the item with the largest percentage of correct answers. Men, older participants (≥51 years old), and clerical support workers (e.g., general and keyboard; customer service; numerical and material recording clerks) were the groups of participants who most correctly estimated the time. The degradation times of plastic bags and aluminum cans were correctly estimated by 12.3% and 14.6% of the respondents, respectively. Plastic bottles were recognized by almost 40% as having a degradation time between 400 and 500 years. Disposable diapers were the marine litter object with the highest discrepancy of values: it is estimated that disposable diapers take 450 years to degrade, but most people believe that it takes 10–20 years to disappear, approximately 30 times less time than in reality. A χ^2 analysis of the frequencies difference between gender, age group, and occupation across the number of correct answers was not significant ($p > .240$ for all cases). In other words, the statistical analysis did not show evidence that gender, age group, and occupation (positively or negatively) influence the knowledge regarding the estimated degradation rates.

For the question “What is the name given to the plastic particles with less than 5 mm in size often found in the ocean?”, microplastics were pointed as the correct answer by 95.1% of the surveyed people, followed by mesoplastics (3.4%) and macroplastics (1.4%).

	Cigarette butt		Plastic bag		Aluminum can		Plastic bottle		Disposable diaper	
Time	No.	%	No.	%	No.	%	No.	%	No.	%
1-5 years	85	24.3	7	2.0	5	1.4	0	0.0	21	6.00
10-20 years	107	30.6	43	12.3	23	6.6	22	6.3	77	22.0
50 years	58	16.6	45	12.9	54	15.4	41	11.7	63	18.0
100 years	46	13.1	79	22.6	63	18.0	68	19.4	60	17.1
200 years	17	4.9	53	15.1	51	14.6	58	16.6	36	10.3
300 years	10	2.9	20	5.7	40	11.4	23	6.6	26	7.4
400 years	8	2.3	33	9.4	34	9.7	46	13.1	28	8.0
500 years	19	5.4	70	20.0	80	22.9	92	26.3	39	11.1

Figure 3.2: The estimation of the degradation time of different marine litter objects. The color gradient indicates the percentage of responses (darker: more responses; lighter: fewer responses). The predicted correct degradation time for each item is marked by a green box (cigarette butt: 1–5 years, plastic bag: 10–20 years, aluminum can: 200 years, plastic bottle: 450 years, and disposable diaper: 450 years).

3.3.4 Perceived Sources/Pathways, Factors, and Impacts

When asked about the sources and pathways that contribute to litter reaching the coast and sea, the median of answers attributed the maximum score to three paths: (i) direct release of litter into the ocean, (ii) litter resulting from the discharge of sewage to the coast, and (iii) litter transported through streams, irrigation channels, and rivers. Direct release of litter on the coast was the source less mentioned (Table 3.2). Statistical results (Appendix B, Tables II and III) indicated that participants similarly perceived the direct release of litter on the ocean and coast, and sewage discharge to the coast, regardless of their age and occupation. This indicates that respondents recognize both land- and ocean-based activities as marine litter sources.

Statistically significant differences were detected in the perception of the litter transported through streams, irrigation channels, and rivers. The age group 19–25 assigned a lower contribution of these pathways when compared to the age groups of 36–50 and ≥51 years old. In the same direction, students showed a lower perception of the contribution of streams, irrigation channels, and rivers as pathways for litter accumulation, especially when compared to professionals.

Various factors were rated by the participants as contributing to litter reaching the marine environment. The excess of single-use products and packaging; fishers and boats that dump garbage into the sea; people littering in the streets, irrigation channels, and streams; over-packaged products that are difficult to recycle; and extensive use of plastic in everyday products and packaging were perceived as the most important causes of marine litter accumulation (Table 3.2). Stores along the beaches and coasts, people littering the beaches, and lack of bins in public areas were perceived by the population as less relevant causes of marine debris accumulation. Statistical tests (Appendix B, Tables II and III) indicated there is no evidence that age or occupation significantly alters each contributing factor's ratings. However, this excludes the lack of bins in public areas, which the age group 26–35 considers a more important factor than the 36–50 years old group.

The degree of marine litter impacts on the quality of the marine environment, tourism, human health, shipping and fishing industry, and the appearance/aesthetic of the beaches and coast were equally evaluated. Participants reported marine litter to pose different levels of threat, with the quality of the marine environment and human health being the most identified consequences (Table 3.2). The interquartile range for these impacts was zero and the median five, indicating that respondents were mostly convinced that these are the topmost impacts. It was possible to identify a trend in the students and clerical support workers regarding the perception of the impacts of marine litter on human health, shipping and fishing, and tourism. These two groups were the ones that most often had opposite opinions, with students always attributing a lower impact of marine litter on those items (Appendix B, Table III). The appearance/aesthetic of beaches and coasts also had a median of five, but the interquartile range was one, indicating that some respondents considered it a medium-level impact. The shipping and fishing industries and tourism had a "medium" impact median. It is interesting, however, that the younger-age group attributes a lower impact of tourism in the marine debris (Appendix B, Table II), especially when compared to older groups (i.e., 36–50 and $\geq$51).

3.3.5 Risk Perception and Social Norms

The assessment of risk perception revealed that most of the participants perceived lasting damages caused by marine litter and recognized it as a topic that deserves more attention at the national level (Table 3.3). Similarly, most respondents agreed that the amount of debris is increasing and reported being worried about the consequent impacts. Interestingly, only the item related to the concern about the impacts of marine litter showed to be statistically different depending on age and occupation. Results showed that people aged 36–50 and more than 51 years old were more worried than the inhabitants between 19 and 25 years old, and students were less concerned than professionals (Appendix B, Tables IV and V). When asked about marine litter being a problem in the region where they live and not just in other countries, the inhabitants agreed that marine pollution is a real concern in the place where they live. The statement "marine litter is a future environmental threat rather than a current threat" gave rise to a greater dispersion of opinions, with some agreeing and others disagreeing, mainly the youngsters (19–25 years old), who considered marine litter a more distant threat. Nevertheless, the Kruskal–Wallis tests revealed that the risk perception did not vary significantly with age, only with occupation, with professionals recognizing more clearly marine litter as a present threat than clerical support workers (Appendix B, Tables IV and V).

The perceptions of the respondents' about others (i.e., family, friends, and local community) were equally assessed. The median showed that most of the participants partially

Table 3.2: Perceived sources, pathways, contributing factors, and impacts of marine litter. Mdn = Median; IQR = Interquartile range.

Item	Mdn	IQR
Sources and pathways (1–none to 5–a lot)		
Direct release on the ocean	5.0	1.0
Sewage discharge to the coast	5.0	1.0
Litter transported through streams, irrigation channels, and/or rivers	5.0	1.0
Direct release on the coast	4.0	1.0
Contributing factors (1–not at all important to 5–very important)		
Single-nature use of several products and packaging	5.0	1.0
Fishers and boats that release garbage into the sea	5.0	1.0
Littering in the streets, irrigation channels, and streams	5.0	1.0
Over-packaged products that are difficult to recycle	5.0	1.0
Extensive use of plastic in everyday products and packaging	5.0	1.0
Businesses along the beaches and coast	4.0	1.0
Littering on the beach	4.0	2.0
Lack of bins in public areas	4.0	2.0
Impacts (1–none to 5–high)		
Quality of the marine environment	5.0	0.0
Human health	5.0	0.0
Appearance/aesthetic of the beaches and coast	5.0	1.0
Shipping and fishing	4.0	1.0
Tourism	4.0	1.0

agree that their family and friends think it is important to reduce marine litter and support them in taking actions to reduce marine litter (Table 3.3). Inhabitants were unsure of whether their families and friends knew what marine litter was and talked about it. Concerning the local community's care about marine litter, respondents were undecided whether people care or not, with more than 20% partially or totally disagreeing that their local community cares about marine litter questions.

The recognition that most respondents' family and friends know what marine litter is, talk about it, and think it is important to reduce it were the social norms that varied statistically significantly with age. The older ones (51 or more years old) reported that their relatives and friends know and talk more about marine litter than the contacts of the 19–25 and 26–35 years old. Similarly, family and friends of older respondents (51 or more years old) think it is more important to reduce marine litter, especially when compared with the younger ages' groups (i.e., 19–25 and 26–35 years old groups). The social norms perception about marine litter by professional occupation was equally analyzed and did not show statistically significant differences. The statistical tests that support these conclusions are depicted in Appendix B, Tables VI and VII.

Table 3.3: Risk perception and social norms about marine litter. Mdn = Median; IQR = Interquartile range.

Item	Mdn	IQR
Risk perception (1 – totally disagree to 5 – totally agree)		
Marine litter is an important topic that deserves more attention at national level	5.0	0.0
The amount of litter on the coast and in the sea is increasing	5.0	1.0
The marine litter impacts are a cause of concern for me	5.0	1.0
The ocean is so large that it is unlikely that marine litter cause lasting damage	1.0	0.0
Marine litter is a problem in other regions, but not on Madeira Island	1.0	1.0
Marine litter is a future environmental threat rather than a current threat	1.0	3.0
Social (1 – totally disagree to 5 – totally agree)		
Most of my family and friends think it is important to reduce marine litter	4.0	1.0
Most of those close to me support me in taking actions to reduce marine litter	4.0	1.0
Most of my family and friends know what marine litter is and talk about it	4.0	2.0
Most people in my local community care about marine litter	3.0	1.0

3.3.6 Responsibility

As important as verifying if respondents recognize the necessity of reducing marine litter, is identifying whom they hold responsible for it. A set of individual and collective entities were listed. Results indicate that all groups are responsible. Nevertheless, it is interesting that respondents first consider the general public responsible for reducing marine litter and only then recognize their own role (Figure 3.3). Governments and town councils were also regarded as key actors in tackling marine litter. In contrast, companies that collect and recycle garbage, scientists, and researchers were perceived as the least responsible. Results were also compared by age and occupation to explore the influence of these variables on perceived responsibility. The χ^2 analysis revealed that almost all age groups equally perceive responsibility. The exception were the participants aged 51 and over, who were generally more likely to attribute more responsibility to teaching professionals, bathers, and commercial users of the coast and sea when compared to younger groups (Appendix B, Table VIII).

Regarding the relationship between perceived responsibility and occupation, the results that stand out are the ones that refer to managers, clerical support workers, and services and sales workers. The last two groups are highlighted because the role of these individuals in reducing marine litter in their archipelago just appears in the fourth place, translating their low individual perception of responsibility. The managers' group (e.g., chief executives, senior officials, and legislators; administrative and commercial managers) did not mention themselves in the top five of those who must reduce marine

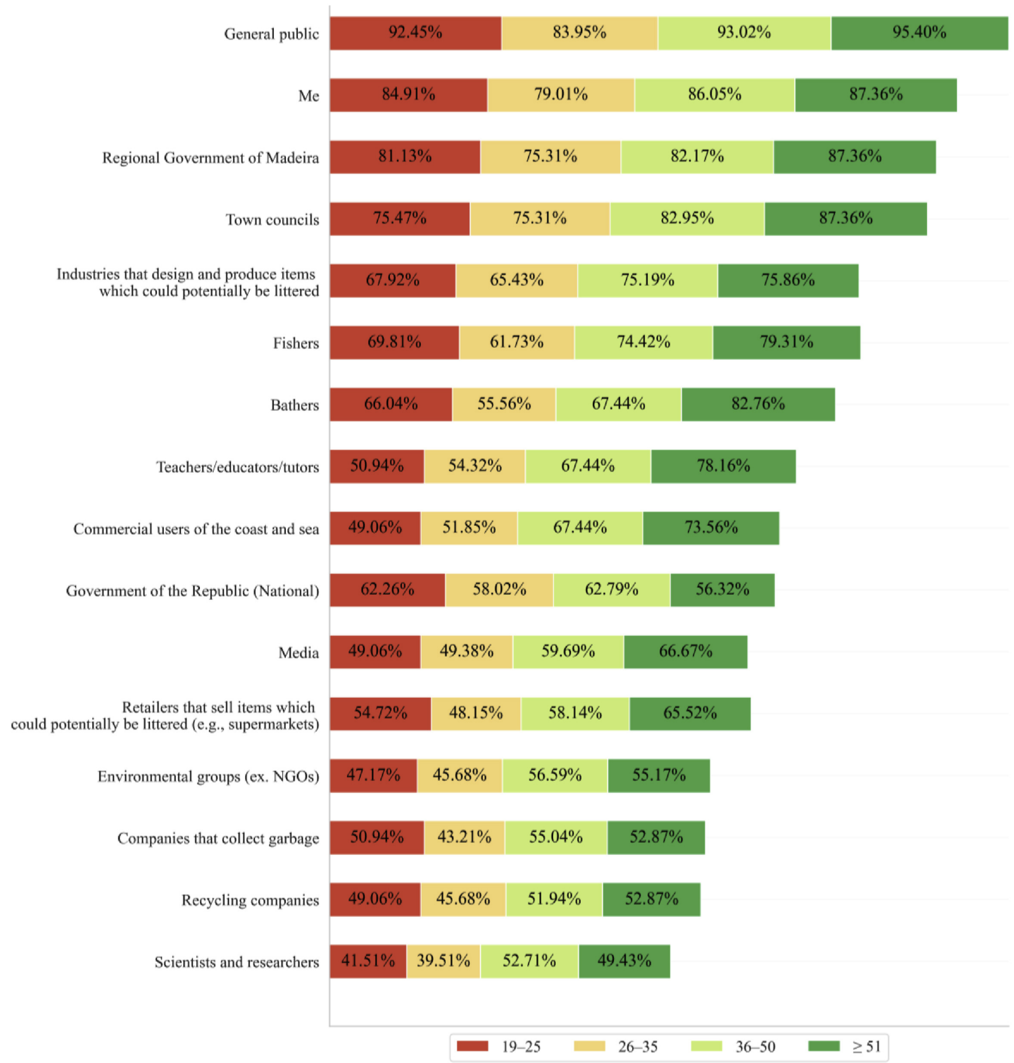


Figure 3.3: Perceptions of the responsibility for reducing marine litter, according to respondents' age.

litter, attributing the responsibility to town halls, regional government, general population, fishers, and beachgoers. It is noteworthy that the χ^2 pair-wise tests (Appendix B, Table IX) showed that the professionals, compared to the students, attribute more responsibility for marine litter reduction to the occupations of education and research. Moreover, professionals also perceive a higher responsibility to bathers and fishers than the clerical support workers.

3.3.7 Self-Perceived Responsibility and Solutions

Different results for self-perceived responsibility and competence were obtained: all people believe that they can contribute to reduce marine litter, however, not everybody considers themselves responsible for that (Table 3.4). The Kruskal–Wallis test results showed that age significantly affects how respondents consider themselves contributors

Table 3.4: Self-perceived role, solutions, and litter-reducing intentions to diminish marine litter. Mdn = Median; IQR = Interquartile range.

Item	Mdn	IQR
Self-perceived role (1 – totally disagree to 5 – totally agree)		
I can contribute to reducing marine litter	5.0	0.0
I am responsible for reducing marine litter	5.0	1.0
Solutions (1 – totally disagree to 5 – totally agree)		
People changing their behavior is part of the solution to the marine litter problem	5.0	0.0
If there are no radical changes in people's behavior, the marine litter problem will not be solved	5.0	0.0
Recycling is part of the solution to the marine litter problem	5.0	1.0
Litter-reducing intentions (1 – extremely unlikely to 5 – extremely likely)		
Use own bags when shopping	5.0	0.0
Don't litter the irrigation channels and streams	5.0	0.0
Don't litter	5.0	0.0
Don't leave trash on the beach	5.0	0.0
Don't leave trash in the mountains	5.0	0.0
Sort the garbage and recycle	5.0	0.0
Advise my family and friends to adopt sustainable practices	5.0	1.0
Avoid the use of plastic bags	4.5	1.0
Buy products with environmentally friendly packaging	4.0	1.0
Use re-useable products, rather than single-use and disposable ones	4.0	1.0
Avoid plastic use	4.0	2.0
Ask people to pick up their litter if littering in the street	4.0	2.0

and responsible for marine litter reduction: the youngest group (19–25 years old) is the one that seems to be more indecisive about their self-perceived responsibility, being, statistically significantly different from the older groups. Concerning the self-perceived competence, the only two age groups that showed statistically significant differences were the age groups of 26–35 and 36–50, being the group 26–35 the one with a higher self-perceived competence (Appendix B, Table X). Considering the occupations, professionals consider themselves more responsible than students (Appendix B, Table XI). No other pair-wise comparison revealed statistically significant differences among the groups.

When inquired about what is necessary to tackle marine litter, people changing their behaviors, and doing so in a radical way, was pointed out by all the respondents as the most necessary aspects, without significant statistical differences among the different age groups and occupations. Recycling was a solution that did not gather unanimity among the participants. Respondents between 26 and 35 years old were the ones that considered this measure less necessary when compared to the remaining age groups (Appendix B, Table X). Comparisons between the different occupations were not significant (Appendix B, Table X).

Participants were then asked to rate the likelihood of performing certain actions to diminish marine litter. Using their own bags when shopping; not disposing of litter on irrigation channels, streams, floors, beaches, and mountains; and sorting and recycling the garbage were the intentions with a median of five and an interquartile range of zero. This indicates it is extremely probable that respondents practice the referred actions. Avoiding plastic use and asking people to pick up litter that is left on the floor were the intentions that participants reported being less willing to accomplish. When the results are analyzed by age, some differences can be observed mainly between the groups 19–25 and 36–50 years old (Appendix B, Table XII) in the following intentions: avoid plastic bags; avoid plastic use; buy products with environmentally friendly packaging. Mainly, the oldest (36–50 years old) indicate that they are more likely to take these actions than younger people (19–35 years old). When the intention of sorting the garbage and recycling was ascertained, the 36 or more years-old respondents said to be more enthusiastic about doing that in comparison to younger inhabitants (35 years old or less). Two trends were identified regarding occupations (Appendix B, Table XIII). On one side, students are less keen to use plastic bags and general plastic items when compared to the other occupations. On the other side, technicians and associate professionals have a higher predisposition for leaving trash on the beach, floor, and littering the irrigation channels and streams compared to other occupations. As gender affects behavior, results were also compared between females and males. Women reported being more willing to use their own bags when shopping, avoid the use of plastic items and bags, advise family and friends to adopt sustainable practices, sort the garbage and recycle, and not to litter (Appendix B, Table XIV).

When the municipality area was crossed with the inhabitants' intentions, some differences were observed. Sorting the trash and recycling was an intention reported by most of the respondents in Funchal, South-East Coast, Porto Santo, and North Coast. Porto Santo was the area that collected the highest number of behavioral intentions with the maximum punctuation (extremely probable of practicing the action). As per the statistical tests (Appendix B, Table XV), Funchal and South-East coast inhabitants exhibited the same intention of sorting the garbage and recycling. This intention was higher than the one shown by the inhabitants of Câmara de Lobos and the South-West coast.

Together with the intentions, the behaviors were registered through multiple choice selection. As depicted in Figure 3.4, almost 30% of the respondents reported not littering the floor/irrigation channels/streams and using their own bags when going the supermarket. Drawing attention to someone's incorrect waste disposal behavior was the least-practiced action by the inhabitants in the week prior to the collection of data.

Age, gender, and household were factors that significantly influenced actions (Appendix B, Tables XVI, XVII and XVIII). Older participants (≥51 years old), women, and, as outlined in Figure 1 in Appendix B, people living in smaller families reported having practiced more sustainable actions in the week before the application of the questionnaire.

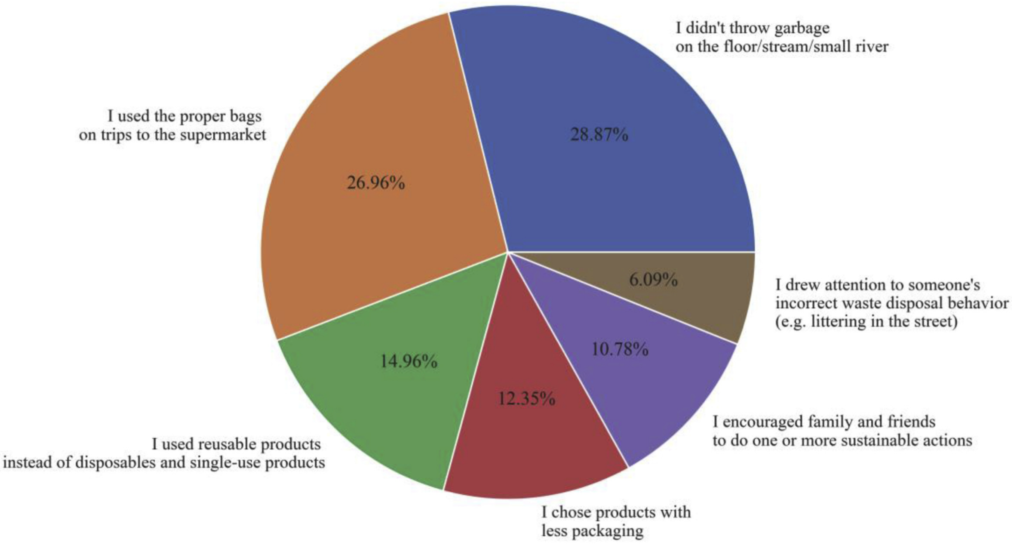


Figure 3.4: Sustainable actions taken by the participants.

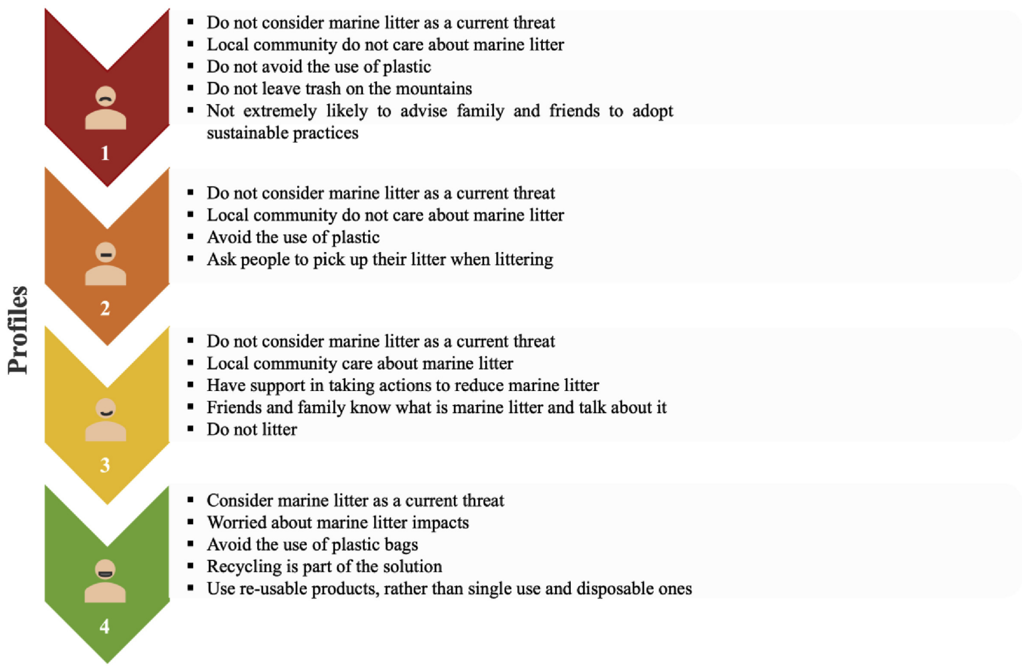


Figure 3.5: Four main respondents' profiles.

3.3.8 Respondents' Profiles

Four main profiles (Figure 3.5) were identified at the end of the analysis of the questionnaire, following the cluster analysis using k-means clustering and DT splitting (Figure 2 in Appendix B). For this work, the DT splitting obtained a mean accuracy of 89.3%.

Each profile revealed important characteristics that respondents tend to have, directing distinct educative and raise-awareness interventions to such targets. Although the DT depicted in Figure 2 in Appendix B used all the questions from the questionnaire, some questions were considered unimportant according to the maximum defined depth (i.e., 5) and information gain metrics (i.e., the Gini Impurity function). As a result, those questions were not used by the DT splitting to draw the four profiles.

Profile 1 (n = 22) relates to people who do not consider marine litter a current threat and live in a community that does not care about marine litter. Therefore, these people do not avoid the use of plastic nor advise family and friends to adopt sustainable practices. The positive point this group has is not littering in the mountains. Profile 2 (n = 46) shares commonalities with people in profile 1: neither do they consider marine litter a current threat, nor do they have a caring local community. However, individuals with this profile avoid using plastic and ask people to pick up their litter when littering. Individuals fitting profile 3 (n = 66) do not litter, have support in taking actions to reduce marine litter, and live in a good environment, with a caring community, and with family and friends knowing and talking about marine litter. Nevertheless, these individuals do not consider marine litter a current threat. On the contrary, people who consider marine litter a current threat (profile 4, n = 77) are worried about its impacts, so they adopt sustainable practices: avoid using plastic bags, recycle, and use reusable products.

3.4 Discussion

This study examined the perceptions about marine litter that inhabitants of oceanic islands have. A questionnaire was used to survey the public and to establish distinct profiles, which can be used to propose measures to tackle the problem.

3.4.1 Survey Sample–Demographic Data

A diverse audience responded to the questionnaire. The characteristics of the respondents are in line with those registered in the Autonomous Region of Madeira, which has 53.1% women and 46.9% men according to the last census [Est21]. These censuses were made in 2021, so the available results are preliminary. The previous censuses were collected ten years ago, and the statistics are no longer up to date for comparison. Hence, data on population by age groups, households, and occupation was not available and no comparisons were carried out between the studied sample and the inhabitants of the Autonomous Region of Madeira inhabitants. Regarding the residence, compared to the 2021 census, the proportion of respondents per area approximately corresponds to the percentage of the population living in the municipalities.

3.4.2 Marine Litter Knowledge and Perceptions

Different questions were used to evaluate what people know about marine litter, namely the definition, material composition, and degradation time. Overall, a moderate baseline knowledge was revealed by inhabitants.

Cigarette butts and plastic items belong to the list of common debris items on land [Har+21]. According to the Awi-Litterbase database, plastic-affiliated debris makes up approximately 75% of the total marine litter [Tek+21]. Littered cigarette butts are heterogeneously distributed [GPN14], being the dominant items in several beach litter studies [AC19; Nov+09]. These data are very much in accordance with the perception of the surveyed participants, as 98.0% classified plastic items and 73.1% the cigarette butts as the most common marine litter items. The presence of sanitary waste on some of Madeira's beaches (unpublished data) may lead to respondents to extrapolate the reality of their region to the global reality, pointing to sanitary waste as the third most common marine litter category.

Regarding the degradation time, people estimated approximately the reported degradation rates, with a tendency to undervalue the times of degradation. Underestimates of decomposition time are in line with other studies [Den+20], showing that despite people recognizing that marine litter takes several years to degrade, the order of magnitude is sometimes unknown. A reinforcement of people's knowledge about marine litter's long decomposition rates (through images on the products or even on the label, for example), could encourage correct waste disposal practices curbing the entry of tons of items into the sea every year. Visualization, where images are linked with emotions, can indeed lead to changes in consumption and disposable behaviors [PWT17]. The high number of anti-littering cigarette butts campaigns undertaken in the last years could have positively influenced the percentage of correct answers compared to other items, thus justifying the registered differences.

A similar explanation can be attributed to the high number of correct responses to the question "What is the name given to the plastic particles with less than 5 mm in size often found in the ocean?". Microplastics were pointed to as the correct answer by 95.1% of the surveyed inhabitants. The growing number of studies on microplastics in the marine environment [BG15] probably lead to the wide-spread knowledge of the term among the population, and even though people are not sure about the theoretical definition, microplastic is a word they associate with marine litter. Also, the public frequently hears in the media sensational headlines when studies related to microplastic discoveries [Ges+19; ICF17; Rag+21] come out.

The inhabitants were surveyed about the sources and pathways that contribute to litter reaching the coast and sea, to verify if people perceive that marine litter is not just "produced" on the marine environment, as one might be misled by the term. The authors also wanted to verify the degree of perception of residents regarding the streams as transport routes for marine litter, as 47.4% pointed out in the questionnaire that Funchal's streams are very dirty. Results showed that most of the inhabitants identify streams and irrigation channels as garbage routes, an output extremely important considering the region's characteristics (Madeira Island does not have rivers, instead, it has small streams). Like rivers [Rec+14], streams act as a source of marine litter because they have a typical radial drainage pattern, characteristic of oceanic islands, that transport high amounts of solid materials to the sea in the rainy season [PSC05; RRG20]. Studies in other islands equally showed the influence of waterways [Car+13].

Together with the identification of marine litter sources and path-ways, the participants perceived different factors contributing to litter reaching the marine environment. The nature of the products (single-use and over-packaged) and human behaviors (fishers that release garbage into the sea and people littering) were pointed out as significative causes of litter accumulation. Such results are in accordance with previous surveys, where extensive use of plastic and it single use nature, together with human behaviors were highlighted as very important factor contributing to marine litter [Har13].

Regarding the negative impacts of marine debris, inhabitants recognized that litter poses several consequences. The quality of the marine environment was perceived as the biggest consequence, a result observed in other studies [Har13]. This result was expected, as people associate charismatic animals (e.g., whales and turtles) with ingestion or entanglement on litter [FN20]. Furthermore, several anti-littering campaigns that were developed in Madeira Island use marine species to demonstrate the impact of marine debris, reinforcing the perception that the main consequences of litter are on the fauna. Human health was also recognized as an important challenge created by marine litter. Campbell et al. [Cam+16] survey showed that people recognize beach litter injuries as a major concern, mainly wounds. Additionally, sewage-related debris linked to water quality concerns has been gaining relevance as a public health hazard [PH11; WGT05]. The appearance/aesthetic of the beaches and coast was ranked as the third most relevant impact of marine litter. As the study was conducted in an archipelago (and so the respondents have more opportunities to visit beaches), it was predictable that respondents perceived this impact as being higher. In fact, most respondents only visit Funchal's beaches in the summer months/bathing season (40.3%) and consider them a little dirty (41.4%). Similar expectations were observed for the shipping and fishing industries and tourism, which had a "medium" impact median. Being a touristic destination awarded as a "Europe's Leading Island Destination", it was expected that respondents from Madeira Island attributed a higher impact of marine litter on tourism. According to Mouat et al.[MLB10], reduced recreational opportunities, loss of visual amenities, negative publicity and reputation, and ultimately a decline in the number of tourists and lower revenues, are some of the potential consequences of marine litter on coastal tourism. Estimates indicate that dirty beaches can reduce tourism revenue by up to 52% [BRT00], considering that litter and sewage discharge evidences are factors considered in coastal scenic evaluations [Mes+20; RWA18]. A survey conducted in the Netherlands confirms that: tourists prioritize clean destinations when choosing a coastal holiday destination [Wer+16]. Therefore, reinforcement in the awareness of the population about the impacts of marine litter on tourism is necessary, especially in regions and oceanic islands whose economic activity is heavily dependent on tourism. Similarly, increasing people's literacy on the impacts marine litter has on shipping and fishing is necessary, as marine litter costs each Portuguese vessel an average of €2930 per year, usually due to fouled propellers [MLB10].

3.4.3 Risk Perception and Social Norms

Six items surveyed the risk perception of oceanic islands' inhabitants. Globally, respondents considered marine litter an important topic and were concerned about it, stating it is a local and actual problem that requires more attention. Inhabitants agreed that marine pollution is a real issue in the place where they live, an expected result considering that people living closer to the coast or shoreline tend to show a higher sense of marine citizenship and to be aware of the impacts [Hec+18; MF10]. Furthermore, it was recognized by almost all respondents that marine litter deserves more attention at the national level, a result that reinforces that marine litter management and governance measures are necessary (Section 3.4.6.3).

What family, friends, and the local community know and think about marine litter and how they behave regarding it were also assessed. Part of the respondents' family and friends recognized it is important to reduce marine litter and supported them in doing so. Nevertheless, few family members and friends know and talk about it. A reinforcement of campaigns raising awareness about the theme (Section 3.4.6.3) appears to be necessary. Respondents were undecided about how their local community cares or not about marine litter, a result equally observed by Hartley [Har13]. These authors predicted the respondents do not know the community well enough to answer the question [Har13].

3.4.4 Responsibility

Surveyed inhabitants believed all groups indicated in the questionnaire were somewhat responsible for reducing marine litter. Other people, the respondents themselves, governments, town councils, and industries that design and produce items that can potentially be littered were regarded as key actors in tackling marine litter. Companies that collect and recycle garbage, scientists, and researchers were perceived as the least responsible. Similar results were obtained in McKinley and Fletcher's interviews, with the responsibility of marine environment management being mainly attributed to the government [MF10]. In Dilkes-Hoffman et al. [Dil+19] survey, companies and industries were perceived as the ones holding the highest level of responsibility for altering the way plastic is used, whereas Campbell et al. [Cam+16] results shown that beach users are the main responsible for beach litter avoidance. Oppositely, Gkargkavouzi et al. [GPM20] survey showed that respondents consider scientists as being far more competent in managing the marine environment in comparison to the central government and local authorities. These differences indicate respondents' perceptions about who is responsible for reducing marine litter varies among the regions.

3.4.5 Self-Perceived Responsibility and Solutions

Responsibility and competence were surveyed: although people believe they can contribute to reduce marine litter, they do not consider it as their responsibility. Most inhabitants believed people must change their behavior and do it radically to tackle the problem.

Using their own bags, not littering, and sorting and recycling the garbage were the litter-reducing behavioral intentions respondents admitted being more prone to perform. Reducing the use of plastic and asking people to pick up litter they threw away were the intentions that registered a lower propensity to be practiced by the respondents. Circumscribe the use of plastic is something that is not entirely on the side of the respondents, as there are still not many alternatives to packaged products provided by brands and industries. Additionally, there are very few bulk stores on the studied islands and insularity forces many of the daily consumption products to come packaged to guarantee their quality and hygiene. About not asking people to pick their litter, Hartley [Har13] observed the same low intention and believed it may be due to the little ease of the task.

Together with intentions, behaviors were recorded. Not littering and using their own bags when shopping were the actions that the participants admitted the most to practice. Results are in line with those of Martinho et al. [MBP17], which showed a reduction in the use of plastic bags by about 74% in Portugal a few months after the implementation of a plastic bag tax. Nevertheless, people report that they cannot completely give up on buying plastic bags due to convenience, namely their price and availability [Den+20; Sun+17]. Drawing attention to someone's improper waste disposal behavior was the action respondents practiced the least. This result was expected, as people often feel constrained to reprimand others and fear counter-punishment or retaliations [BNR16].

3.4.6 Respondents' Variables

3.4.6.1 Respondents' Demographic Characteristics

Older participants, women, and people living in smaller families reported having practiced more sustainable actions. Such results are in accordance with Deng et al. [Den+20] observations that reported senior people buying plastic bags less frequently than youngsters and females being more willing to reduce pollution than men. The gender hypothesis that females are more environmentally responsible than males, more likely to consider environmental concerns in their purchases, and more prone to recycle has been presented by Mainieri et al. [Mai+97] and confirmed by Zambrano-Monserrate and Ruano [ZA20]. The latter observed that houses in which women are heads of the household are more prone to choose non-plastic bags for shopping. Hartley group [Har+18a] also verified that women and older participants report more behavioral intentions to reduce marine litter. Our results are consistent with the literature and reinforce that marine litter education interventions must target the youngest, as several authors have proven education efficiency among the younger ones [HTP15; LMA19; Tor+19]. About occupations, the same trend was verified and reinforces the above: students are the ones who less perceive the problem, thus interventions must rely on this youngest group.

When the municipality area was crossed with inhabitants' intentions, some differences were observed, mainly at litter deposition intentions. A possible explanation for this might be due to the differential waste management collecting system existent between the areas.

3.4.6.2 Respondents' Profiles

We used a decision tree to establish respondents' profiles. Defining groups of individuals who share common variables facilitate the establishment of approaches to reduce marine pollution and manage the waste. Several educative actions and raise awareness initiatives already developed [Kus+20; TKL19; Tor+19] can be adapted according to the target audience.

People who fall in profile 1 are the ones that need more educational interventions, as well as their community. The intervention should be broad and encompass different topics on marine litter. Individuals fitting profile 2 do not consider marine litter a current threat nor live in a caring community but avoid plastic use and ask people to pick up their litter when littering. Educational and awareness-raising actions are also needed for this group but focusing above all on raising awareness of the community about the actuality and dimension of the theme. Profile 3 people do not litter, have support in taking actions to reduce marine litter, family and friends know and talk about marine litter, and the community cares about marine litter. Yet, these individuals do not consider marine debris a current threat. Interventions for this group must focus on raising awareness, explaining marine litter is an up-to-date and worrying threat. On the contrary, people in profile 4 consider marine litter a current threat and are worried about its impacts, adopting sustainable practices. These people are the group who better perceive and know marine litter. For instance, initiatives and educative interventions that focus on marine litter, its impacts, pro-environmental practices, attitudes, and social norms must be created for all who do not fall into profile 4. Additionally, literate communities was found to be important, with family members and friends claimed as important sources of information [MRČ20].

It is important to note here that in the case of the Madeira archipelago, only 27.6% of the respondents fall into profile 4, which shows that marine litter perception is not something to take for granted for all. Some of the inhabitants do not recognize marine litter as a current threat, a recognition that is essential to engage people with the issue and that appears as an essential step in adopting good practices and behaviors. Yet, more than half of the surveyed public is concerned with the issue and has positive behavioral intentions. Since this study was conducted on small oceanic islands, and in these regions marine litter and waste management problems are exacerbated, it was expected a higher percentage of inhabitants in profile 4. Yet, our findings are in line with the perceptions of the public from other oceanic islands. In Cape Verde, participants of public participatory sessions showed concern about the causes and consequences of marine litter [Fer+21] and in Azores participants identified marine litter as the most severe threat to the marine environment and recognized the value the marine environment has for the island's economy [Abe+13]. On Rapa Nui Island few litter was found on beaches and the inhabitants were environmentally conscious, adopting coastal protection actions (e.g., picking up litter) and waste-reduction measures (e.g., recycling, volunteering for beach clean-ups) [Kie+17].

Together with evaluating the perceptions Madeira Autonomous Region inhabitants have, the questionnaire here used serves as a quick and cheap tool to assess and identify in other geographical contexts which targets need to be intervened and to what degree, allowing the development of different educational and awareness-raising strategies according to the profiles.

3.4.6.3 Marine Litter Management and Governance

Evaluation of the perceptions, knowledge, awareness, and attitudes of the population towards marine pollution is essential for the definition of prevention and behavior change actions. Questionnaire results showed that increasing literacy and raising awareness activities among individuals are necessary, particularly for the youngest group, men, and students. These actions must be, however, complemented with management and governance measures, both from global to local levels, as an integrated approach is necessary to tackle the issue [Che15]. The same conclusions were obtained by Wichmann et al. [Wic+22], who found that just participating in environmental citizen science projects is not enough, auxiliary educational activities are required. The authors believe inhabitants need to realize the importance of managing and preserving the ocean and coast so that the marine litter problem can be better perceived. Yet, this ocean-oriented cultural identity must be supported by other actions.

Regional governments and town councils were identified by respondents of all ages as very responsible entities for reducing marine litter. That is important, as their identification can drive action: "the ban (. . .) was facilitated by a broad concern among the general public, which led to a bottom-up movement culminating in the national government taking stakes in the issue" [Cri+20, p. 105079]. Respondents referred in the comments/observations section of the questionnaire that effective management practices and governance approaches were needed: "it is the superior entities that have to take the necessary measures", "there should be measures on the part of the municipalities, stricter for those who produce the most garbage, who do not make the proper separation (. . .) penalties, for example in the bill for water, electricity", "the town council seems to have some faults, with regard to pumping stations", "it is necessary that the responsible entities, namely the Regional Government, carry out awareness campaigns", "a measure to be implemented in the region could be the use of bioplastics", "there is a lack of containers for waste separation in many locations", "lobbies/economic interests are delaying the implementation of measures that should have been underway for a long time", and "it is a matter of the will of people and governments" [direct citations of respondents' comments]. These statements are in line with Williams and Rangel-Buitrago's [WR19] observations that demonstrated that weak coastal governance, insufficient financial support, poor political practice, lack of commitment, and nature of public participation can hinder the formation of integrated coastal zone management regulations.

It becomes necessary to diversify the policy tools and move from a passive to an active governance to manage seas and marine pollution effectively. The central governments, local governments, enterprises, and the public are the main stakeholders of marine environmental governance, with their participation being necessary to ensure a successful governance [YB19]. In this view, over the years there has been a shift into policies that control waste discharge from multiple sources, that strengthen the monitoring and treatment of existing wastes, and that stimulate recycling [YB19].

However, irrespective of the region, a principle that cannot be neglected when implementing an approach to litter management is the necessity of always thinking of adaptive management [WR19]. Additionally, due to the transboundary nature of the marine litter issue, a global response that turns to holistic solutions is necessary: "No single solution exists to cope with the litter issue. However, legally binding global governance that will effectively limit and control the magnitude of litter pollution is greatly needed" [Ran+22, p. 113546].

Some recommendations to better manage litter and promote the conservation of ocean and coasts can be found in the literature [Gjy+20; Pra+19; SJC20]. In the Madeira archipelago, questionnaire data suggest that measures that stimulate the choice of reusable products (e.g., through benefits or taxes), regulation on production and consumption of over-packaged products, themes related to ocean conservation on political agendas, and raise awareness of the 10R's-principles (refuse, rethink, reduce, re-use, repair, refurbish, remanufacture, repurpose, recycle, and recover) are necessary [Pot+17]. Although these measures were derived from the questionnaire presented to inhabitants of a Madeira archipelago, they are generic and can be applied to any region.

3.4.7 Study Limitations

The present study focused on self-reported concerns, perceptions, intentions, and behaviors, which can be affected by social desirability bias. This means that participants can answer in a way they consider socially accepted by others to gain their approval [KB00]. To avoid this effect, the questionnaires were completed without the presence of the researchers, thus circumscribing the putative inflated answers [Oka+02].

A random sampling of individuals was undertaken during this study. Nevertheless, as some questionnaires were obtained through uncontrolled methods, it is possible that they have increased the participation of individuals who have a closer link to the sea theme. Additionally, some of the municipalities registered a low number of answers, a limitation difficult to surpass considering the high prevalence of older people in those zones. The distribution of paper questionnaires aimed to surpass this constraint, as electronic distribution directs responses to younger audiences and with internet access. Still, the number of answers obtained in this questionnaire is in line with similar studies: 127 responses in a study with 16 European countries [Fil+21], 374 answers in a questionnaire spread over 11 months in Greece [Cha+21b], and 107 participants in focus groups about marine biodiversity [TL17]. Furthermore, to avoid people withdrawing from the questionnaire, questions about the income or educational level were not included, even though the questionnaires guaranteed the anonymity of respondents. Other authors have reported this limitation [Den+20], and questions exclusion is in line with guidance for writing effective questionnaires [PJC15].

3.5 Conclusion

The public survey focused on perceptions, awareness, knowledge, responsibilities, and behavioral intentions regarding marine litter, a problem of growing global concern. Results show that people are aware of the problem, linking the theme mainly to plastics and their impacts. Yet, the population is not completely aware of the order of magnitude of the degradation rates of some items. Less than 20% of the respondents accurately estimated the decomposition time of plastic bags, an everyday item. We believe raising awareness about the extensive degradation times of several items is hence a fundamental point to avoid their presence in the marine environment. Most of the respondents identified the sources and pathways of marine litter and perceived it as a actual problem. Excess of single-use products and packaging, fishers and boats releasing garbage into the sea, and people littering were pointed out as some important causes of marine pollution. Nevertheless, when intentions and behaviors were analyzed, not all participants reported, among others, avoiding plastic use or choosing environmentally friendly packaging. It was pointed out by respondents that one of the waste management measures may involve greater inspection and government action, through the application of fines and the existence of more sustainable alternatives (e.g., bioplastics). This reinforces that end users sometimes have to choose poor environmentally friendly products because the producers and industries impose them to consumers, not assuming their role in marine litter reduction. Similarly, individuals perceived others to be responsible for curbing increases in marine debris. If marine litter is not seen as a universal problem that requires everyone's action, including the producers, it is unlikely to be resolved. Additionally, the youngest, men, and students were the ones who reported less litter-reducing intentions and behaviors. Overall, the questionnaire results allowed us to identify distinct key profiles (from profile 1—the less aware of the theme, to profile 4—the more aware and informed on the subject of marine litter), validating the hypothesis that marine litter is not fully perceived by the public of the studied oceanic islands. Besides, knowing the profile of the respondents is important when planning educational initiatives, highlighting to whom marine litter literacy is needed. These insights can be used to motivate and empower individuals who reported being less conscious about the severity and dimension of the marine litter problem, thus becoming exemplar models for the marine litter fight. At the same time, results show that it is necessary to ensure effective marine litter, ocean, and coastal management within different frameworks so that this major environmental problem is tackled in a coordinated manner across councils and government departments. Results equally show that in the Madeira archipelago, marine litter perception is not currently something to take for granted.

CHAPTER 4

Predicting Noncontact Injuries of Professional Football Players Using Machine Learning

Diogo Nuno Freitas, Sheikh Shanawaz Mostafa, Romualdo Caldeira, Francisco Santos, Eduardo Fermé, Élvio R. Gouveia, Fernando Morgado-Dias

Abstract. *Noncontact injuries are prevalent among professional football players. Yet, most research on this topic is retrospective, focusing solely on statistical correlations between Global Positioning System (GPS) metrics and injury occurrence, overlooking the multifactorial nature of injuries. This study introduces an automated injury identification and prediction approach using machine learning, leveraging GPS data and player-specific parameters. A sample of 34 male professional players from a Portuguese first-division team was analyzed, combining GPS data from Catapult receivers with descriptive variables for machine learning models—Support Vector Machines (SVMs), Feedforward Neural Networks (FNNs), and Adaptive Boosting (AdaBoost)—to predict injuries. These models, particularly the SVMs with cost-sensitive learning, showed high accuracy in detecting injury events, achieving a sensitivity of* 71.43 %*, specificity of* 74.19 %*, and overall accuracy of* 74.22 %*. Key predictive factors included the player's position, session type, player load, velocity and acceleration. The developed models are notable for their balanced sensitivity and specificity, efficiency without extensive manual data collection, and capability to predict injuries for short time frames. These advancements will aid coaching staff in identifying high-risk players, optimizing team performance, and reducing rehabilitation costs.*

4.1 Introduction

Professional football players are confronted with high physical and mental demand levels. This condition inevitably leads to a substantial likelihood of (noncontact) injury events [EHW11a]. In this study, a noncontact injury is defined in alignment with the Union of European Football Associations (UEFA) guidelines for epidemiological research. Specifically, it refers to an acute physical pain experienced by a player during a training session or match, which occurs without any physical contact with other players. Such an injury may or may not necessitate medical intervention but results in the player being unable to participate in the subsequent training session or match [Ful+06; Häg+05].

Noncontact injuries (henceforth, injuries) are known to significantly impact the player's physical performance and psychological aspects, with repercussions for the entire team [Häg+13] and club [LS08; McC+14]. These injuries account for more than a third of all injuries that require players to stop, and more than a quarter of all injuries reported in a season [EHW11b; EWH16]. Nevertheless, the scientific evidence of the last two decades around the modifiable risk factors associated with preventing and reducing the risk of injury in professional football players is increasing [Gab20].

One way of assessing the risk of injury is by employing screening battery processes [Lóp+18; McC+14; Oli+20] or training load monitoring [Gab+17; Hal14]. In training load monitoring, Global Positioning Systems (GPS) receivers occupy a prominent place since they can be used to measure distances covered by players, and in conjunction with other sensors, quantify the number of accelerations and decelerations, as well as provide estimations of measurements of the global and metabolic external training load.

Previous studies have investigated the relationship between training planning based on GPS metrics and the risk of injury [Ehr+16; Jas+18; Mal+17a; Mal+18]. The majority of these studies are retrospective in nature and primarily focus on identifying significant statistical correlations between individual GPS-derived metrics and the incidence of injuries. Nevertheless, the causal relationship of injuries is complex and influenced by multiple factors. This means that while certain metrics tracked by GPS devices might show a correlation with the occurrence of injuries under specific conditions, there are also other important metrics that might not seem significant on their own. However, when these metrics are analyzed together with additional data, they could become important indicators of injury risk [Aya+19; Bah16]. In response to that multifaceted task, the literature suggests using machine learning techniques as an imminent solution for injury prediction [Van+21]. Nevertheless, certain studies concentrate on predicting specific types of injuries. Ayala et al. [Aya+19], for example, explored the potential of these techniques by using pre-season evaluation data, including personal, psychological, and neuromuscular measurements, to predict hamstring strain injuries. Similarly, Ruiz-Pérez et al. [Rui+21] leveraged machine learning to predict lower extremity soft tissue injuries in elite futsal players. In both scenarios, the predictive models generate an estimation of the likelihood that each player will sustain an injury over the course of the entire season.

In a different view, several studies have investigated the capabilities of machine learning methods in predicting injuries throughout a football season. A noteworthy example is the work of Rossi [Ros+18], which not only proposed a machine learning-based injury predictor but also introduced an interpretable framework for understanding the underlying causes of injuries. On the other hand, several studies [Nag+18; Ros+22; Val+20] emphasize the importance of integrating diverse data sources, such as manually collected questionnaire responses and GPS tracking data, to improve the accuracy of machine learning predictions. In particular, Vallance et al. [Val+20] successfully developed an injury predictor capable of making short-term (1-week) and medium-term (1-month) predictions. The authors utilized GPS tracking data alongside data from well-being questionnaires as inputs for various machine learning models.

The existing literature provides only a preliminary understanding of the factors that influence the risk of injury [Rud+18b], and the statistical prediction models used are still basic (e.g., cut-off values [Rud+18b]) or not accurate enough [Car+18; Oli+20] (i.e., not able to correctly identify both injury and noninjury events). Machine learning strategies, on the other hand, are associated with a high incidence of false positives. This could result in coaches erroneously sidelining players who are incorrectly flagged as having a high risk of injury or misallocation of medical resources. Additionally, existing methods focus exclusively on a singular injury type, do not provide precise temporal information about injury risk, and necessitate continuous manual data collection to enhance the accuracy of predictive models.

The primary goal of this study is to address these challenges. More specifically, this research is aimed to develop an automated system that uses machine learning techniques to predict injury risks. Unlike previous machine learning models designed for injury prediction, the proposed approach calculates the likelihood of injury for each player daily over the course of a football season. To achieve this, data from GPS devices are utilized in conjunction with variables related to the players and the match sessions. This study employed the Minimum-Redundancy-Maximum-Relevance (mRMR) [PLD05] in combination with a wrapper method for feature selection, which helps in identifying the most relevant and nonredundant features for predicting injury events.

It is important to note here that the methodology employed does not utilize data from questionnaires or any other continuous manual data collection.

More specifically, the objectives of this paper are to:

- identify the most relevant metrics with minimal redundancy for predicting injury events. Improved understanding and the elimination of overlapping features (i.e., redundant) can enable coaches to concentrate solely on the most critical features to better prevent injuries while improving player performance.
- develop predictive models based on Adaptive Boosting (AdaBoost), Feedforward Neural Network (FNN), and SVM classifiers, considering a trade-off between accurately detecting true injury events and decreasing the likelihood of incorrectly identifying noninjury events as injuries (i.e., false positives).

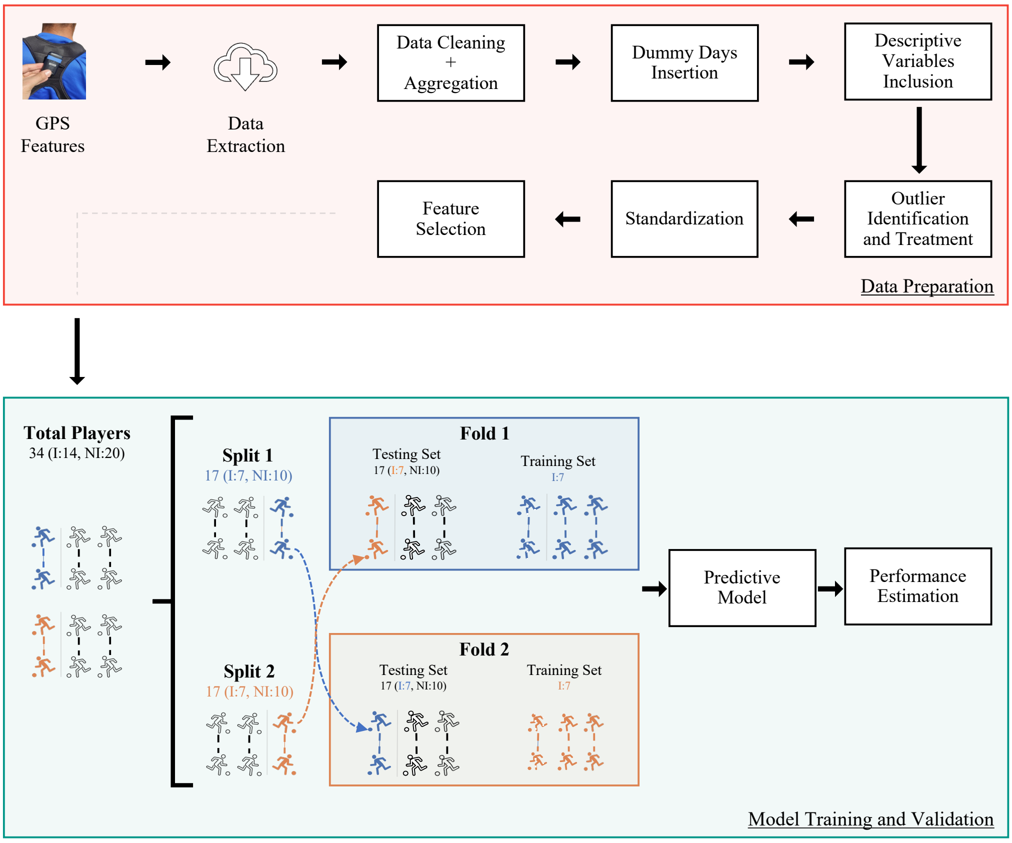


Figure 4.1: Machine learning pipeline followed in the current study for injury detection. The steps include data collection, data preparation, and development and validation of the predictive models.

- propose a strategy for incorporating the effects of sudden changes in player load into the predictive models. This approach aims to provide a more accurate assessment of injury risk by accounting for fluctuations in physical demands on athletes.

4.2 Materials and Methods

Figure 4.1 depicts the machine learning pipeline followed in the current study for injury detection. All the steps of this pipeline will be explored in the following sections.

4.2.1 Design and Data Collection

A longitudinal study was conducted among a convenience sample of 34 male professional football players from a Portuguese first-division team. The team's physical trainers collected data over 36 weeks in the 2020–2021 season, covering 217 training sessions and 38 official games. The data collection period was from September 7, 2020, to May 19, 2021.

Participants had a mean age of (26.27 ± 3.28) years, a mass of (77.54 ± 7.63) kg (collected with the InBody 770), and a height of (180.65 ± 6.60) cm measured with a stadiometer (SECA 213). These professional players were distributed among five positions. That is, 11

(32.35 %) players were defenders, 9 (26.47 %) attacking midfielders, 7 (20.59 %) forwards, 4 (11.76 %) defensive midfielders, and 3 (8.82 %) midfielders. Besides being a goalkeeper, no other exclusion criterion was applied to recruit the players for this study.

The data was collected in the context of the football club's professional contract with each player. These contracts between the players and the club included clauses for gathering data related to their performance, thus ensuring that consent for this activity was formally obtained. As already mentioned, the responsibility for collecting these data was assigned to the club's physical trainers, meaning the researchers were not involved in the data collection phase. Nevertheless, the authors submitted a request for an ethical assessment to obtain access to this data. This request was approved by the Faculty of Human Kinetics ethics committee at the University of Lisbon (statement no. 34/2021). The committee's approval verified that the study complied with national and international ethical standards, as outlined in the Convention on Human Rights and Biomedicine (Oviedo Convention) and the Declaration of Helsinki. Furthermore, all participants provided written informed consent, ensuring they were willingly participating and fully aware of the study's goals and methods. After receiving ethical approval and informed consent, the researchers accessed the data on June 15, 2021, which marked the start of the data analysis phase.

The club's medical staff reported 18 traumatic and overload injuries throughout the 2020–2021 season. Of these 18 injuries, 10 (55.56 %) injury events were due to match circumstances, and 8 (44.44 %) were during training sessions. Moreover, most of the injuries were located in the muscles and tendons (11 injuries, 61.11 %), the remaining being in the ligaments (6 injuries, 33.33 %) or contusions (1 injury, 5.56 %).

Players used GPS receivers from Catapult (GPSports EVO) placed in a skin-tight vest in the thoracic region between the scapulae, capturing the players' position data with a sampling frequency of 10 Hz. Such devices were already in use by the team when the study started. Despite this prior use, they have been shown to be valid, reliable, and accurate for measuring acceleration [Tho+19; VFA12] and the mechanical work applied to the player [Hal14]. Moreover, since these devices are equipped with accelerometers, magnetometers, and other sensors, it was possible to collect raw data, which was then processed by the Catapult system to derive a total of 1 379 parameters for this study. The average Horizontal Dilution of Precision (HDOP) recorded was 1.12, which serves as a metric to assess the geometric accuracy of the GPS satellite-based positioning. HDOP values span from 1 to 2, with values closer to 1 reflecting higher positional accuracy.

4.2.2 Data Preparation

All the players' metrics data per exercise and session were collected by the GPS receivers and extracted using the Catapult's GPSports Cloud. The information was then grouped (using the proper aggregation function, such as maximum and average) into a single record according to the player's unique identifier and session date.

Moreover, metrics from GPS that contained GPS information (e.g., maximum satellite count and HDOP), as well as invalid or missing data were removed. The latter includes, e.g., metrics with invalid data due to sensors' absence (such as heart rate and players' weight).

The unique identifier for each player was also removed from the dataset to maintain anonymity. This step also ensures that the prediction models were trained without access to player-specific information, allowing the models to remain player-independent.

Those processes resulted in a set of 424 GPS metrics that were retained for further data preprocessing steps.

4.2.2.1 Dummy Days Insertion

In order for the model to capture sudden changes in the players' load, time information called "dummy days" was added after grouping the sessions by player and date. A dummy day is attributed to a specific athlete and is denoted by a series of zeroes across all measured variables, thereby emulating a hiatus in the athlete's training regimen, which may arise from intervals of rest or recovery.

In other words, a dummy day is a record that indicates a specific day with no physical activity records from a given player. This information will help the models capture sudden changes in the player's load time series, a potential factor that increases the injury risk [Mal+17b]. Also, this strategy enables the predictive models to learn short-term dependencies between past and current period information.

Furthermore, the inclusion of dummy days ensures a continuous time scale in the dataset for all players, allowing the model to consistently analyze the temporal sequence of training and rest periods.

4.2.2.2 Descriptive Variables Inclusion

Besides all the parameters collected by the GPS receivers, it was possible to identify other descriptive variables that could increase the probability of a given player having an injury. Namely, (a) the players' position and corridor [Bar+14]; (b) the players' age (in months) at the time of the session or match [Oli+20]; (c) the day of the week; (d) the type of session (i.e., training or match session); (e) the results of the games (win, draw or loss); (f) the game location (own stadium or opponent's stadium); (g) the competition of the match session (in this case, Liga NOS [first division] or Taça de Portugal [national championship]); (h) the number of exercises that the players' did in a session and (i) the duration of the session.

It is also important to note that before feeding the information (GPS parameters and descriptive variables) to the models, categorical parameters were converted to numerical values utilizing the one-hot encoding method [Bro20].

The study utilized parameters obtained from GPS receivers as well as other descriptive variables as features. Additionally, the variable "injury" was incorporated into the dataset as a binary target variable. Nevertheless, not every feature was retained for the final model; feature selection using the mRMR [PLD05] method was conducted to exclude variables that were either irrelevant or redundant.

4.2.2.3 Outlier Identification and Treatment

Outliers were identified based on the upper and lower bounds calculated for each GPS parameter values $\mathbf{x_i}$ taking into account the data from injured players.

The upper ${x_i}^{\mathrm{U}}$ and lower ${x_i}^{\mathrm{L}}$ bounds for the i-th parameter were computed as

$$ {x_i}^{\mathrm{L}} = \tilde{x}_i - \left(\left\lfloor \frac{\mathrm{min}_i}{\sigma_i} \right\rfloor + 1 \right) \times \sigma_i, \tag{4.1} $$

and

$$ {x_i}^{\mathrm{U}} = \tilde{x}_i + \left(\left\lfloor \frac{\mathrm{max}_i}{\sigma_i} \right\rfloor + 1 \right) \times \sigma_i \tag{4.2} $$

where $\tilde{x}_i$ denotes the median, σ_i the standard deviation, and min_i and max_i the minimum and maximum value of the i-th parameter, respectively. The lower and upper bounds were derived from the Standard Deviation Method for identifying outliers [Bro20]. In contrast to the original method, which uses a predetermined number of standard deviations, the approach followed here calculates the number of standard deviations based on the maximum and minimum values of each parameter, ensuring that at least one standard deviation is always maintained.

As a result, the j-th data point of the i-th parameter was calculated as:

$$ x_{i,j} = \begin{cases} {x_i}^{\mathrm{L}}, & \text{if } x_{i,j} < {x_i}^{\mathrm{L}} \\ {x_i}^{\mathrm{U}}, & \text{if } x_{i,j} > {x_i}^{\mathrm{U}} \\ x_{i,j}, & \text{otherwise.} \end{cases} \tag{4.3} $$

It is also noteworthy that ${x_i}^{\mathrm{U}}$ and ${x_i}^{\mathrm{L}}$ were calculated based on data from injured players only. This is because injury events can be caused by abnormal parameter values (compared to noninjury records), and these values should not be considered outliers. The outlier treatment was, however, applied to all records, including noninjured records.

The decision to replace outlier values with the calculated upper or lower bounds was made to maintain the data distribution's integrity and minimize the distortion of the features' players' natural range. In other words, this method maintains the important differences between the data points of a player, which is essential for the machine learning models to correctly identify patterns related to the risk of injury.

4.2.2.4 Standardization

Parameters were separately standardized such that the overall parameters' values have a mean of zero and a standard deviation of one, that is

$$ \mathbf{x_i}' = \frac{\mathbf{x_i} - \bar{x}_i}{\sigma_i}, \tag{4.4} $$

where $\bar{x}_i$ denotes the arithmetic mean.

This process is essential since the parameters are measured in different measurement units. Besides, having all the parameters on the same scale improves the stability of the models during the learning phase [KJ19].

4.2.2.5 Feature Selection

In order to reduce the number of input variables of the predictive injury model, and thus reduce the computation time of training the model, a dimensionality reduction was performed by eliminating all features that demonstrated zero variance (i.e., constant variables). In total, 449 features were processed in this step of the machine learning pipeline (424 GPS features and 25 descriptive variables), resulting in 189 features being removed. In other words, 260 features (237 GPS features and 23 descriptive variables) were kept for further investigation.

After removing all features exhibiting zero variance, an additional feature selection analysis was conducted to obtain the most important features for the injury detection model. In this case, the analysis consisted of using the mRMR method [PLD05] to calculate the importance of each feature and rank the features based on their importance. Subsequently, the most important features were identified, and the top p-features were selected for inclusion in the injury detection model.

mRMR [PLD05] is a model-agnostic feature selection mechanism that finds an optimal set of features that minimizes the redundancy between the independent variables, and at the same time, maximizes the relevance with respect to the dependent variable (in this case, with the injury). The mRMR algorithm also ranks the features according to their importance and redundancy. Therefore, in this study, the method was used only to sort the GPS and the descriptive parameters, and only the top features were used in the models.

In order to determine the ideal number of features for each injury detection model, and thus perform feature selection, the models were tested using various feature sets derived from the standardized dataset with outliers removed. Each set of features was composed of the first p (where $p = 10, 20, \cdots, 260$) most important and less redundant features according to the mRMR method. The combination of this wrapper feature selection mechanism with the mRMR method is a novel approach in the field of automatic injury prediction and is significantly faster than a sequential forward search [MMR18].

Appendix C Table 1 provides a comprehensive list of the GPS and the descriptive parameters collected in this study. The supplementary table also includes their ranking according to the mRMR method.

4.2.3 Predictive Models

This work used SVMs, FNNs, and AdaBoost classifiers to map the GPS parameters and the descriptive variables to the recorded injury events. In other words, the predictive models were fed the most important GPS parameters (with dummy days, followed by outlier treatment and standardization) and combined with the most important descriptive variables (such as player position and corridor). It is highlighted again that, in order to simplify the data representation, the players' data per exercise and session were grouped into a single record per day, i.e., each player has one sample per day.

Each data point in the dataset has been assigned a discrete label based on the occurrence of an injury event for a given player and day (i.e., 1=injured, 0=noninjured). The predictive models utilize these labels to make accurate injury predictions. Although the injury labels presented to the models are binary, the models' output is a continuous prediction in the range of $(0, 1)$. Nevertheless, applying a threshold can convert these continuous predictions to a discrete label.

That procedure opens the opportunity to adapt this threshold according to the strategy defined for the team. For instance, decreasing the threshold at the start of the season can make the model more sensitive to potential injury events, leading coaches to reduce training sessions' intensity in order to have more players free from injury during this period. On the other hand, at the end of the season or in preparation for big games, the football coaches, coaching staff, and sports science staff could opt to increase this threshold, balancing better preparation (with higher training loads) with an increased risk of injury. It is important to note here that to report the results of our approach, the threshold was selected using the Receiver Operating Characteristic - Area Under Curve (ROC AUC). The ROC AUC is a graphical representation that illustrates the trade-off between sensitivity and the False Positive Rate (FPR) at various threshold values. The selected threshold was the point on the ROC AUC nearest to the top-left corner. This point corresponds to the highest sensitivity and minimum FPR, and is often referred to as the "elbow" of the ROC AUC, signifying the location where the sensitivity and specificity (1 - FPR) balance is optimal. It is also important to highlight that thresholds were calculated on the training sets and subsequently applied to the testing sets.

The task of predicting injury events presents an imbalanced machine learning problem since only 0.20 % of the observations correspond to injury events within the total dataset. For this reason, each model was trained using both cost-sensitive and traditional learning (i.e., cost-insensitive learning) in a supervised manner. In cost-sensitive learning, misclassification errors have different costs depending on the class, which leads to minimizing the total cost during model training. This enables the models to learn how to classify minority classes correctly (in this study, the injury class) [MH13]. On the other hand, traditional learning does not explicitly address class imbalance, theoretically making the models less prone to classify the minority class correctly. In the cost-sensitive models, the cost of each class was inversely proportional to class frequencies.

For the case of the SVMs, the calculated class weights were used to adjust the margin proportionally. For AdaBoost, the weights were applied to increase the influence of misclassified instances, ensuring that the model focuses more on injury records. Similarly, for FNNs, the class weights were applied to adjust the weighted squared error loss function, making the model more attentive to injury records during training and enhancing its sensitivity to these instances.

The stratified cross-validation [JS11] method was used to validate the predictive models. This was necessary due to the class imbalance between injured and noninjured records. Additionally, the scarcity of injury records required an effective validation method, as creating a validation set with the required number of samples was not possible.

The cross-validation strategy splits the data into two subsets (i.e., two-fold), as depicted in Figure 4.1 Each subset holds a random training dataset and a test data set. In both folds, the training datasets contain only players who were injured at least once during the season (50 % of the total players injured in each fold). The subsets are subject-independent, meaning that each player's data is exclusively in only one of the dataset splits. In other words, the cross-validation strategy was implemented on a player-by-player basis, meaning that the division of data was conducted with respect to individual players rather than individual samples.

As a result, the stratified cross-validation yielded split sets of 1 501 samples for the training set and 3 090 for the testing set in fold 1, with 11 and 7 injury records (i.e., where injury=1) included in the training and testing sets, respectively. On the other hand, for fold 2, the training set comprised 1 410 records, and the testing set comprised 3 082. As expected, the training and testing sets for fold 2 contain, respectively, 7 and 11 injury records.

It is important to highlight that the train and test datasets were created from the complete dataset after executing outlier removal, standardization, feature selection, and feature importance sorting. To help address the class imbalance, however, an undersampling technique was applied to both training sets using the k-means clustering algorithm before being fed to the models. Essentially, the k-means clustering method undersamples noninjured records (i.e., samples where injury=0) by replacing them with a cluster centroid calculated on noninjured records. The ratio of noninjured records to injured records was set at 40 % and $k = 8$.

All the predictive models were built using Python (v. 3.8) with the TensorFlow [Mar+16] (v. 2.3.0) and scikit-learn [Ped+11] (v. 0.24.2) libraries, in conjunction with other supplementary libraries, such as pandas [McK10]. Models were trained in a machine equipped with an Intel Core i7-9700F CPU, running at 3.00 GHz, with 32 GB of RAM.

4.2.3.1 Support-Vector Machine

A SVMs is a supervised learning algorithm that constructs an optimal hyperplane through an optimization strategy. This hyperplane is designed to maximize the margin between the data points of distinct classes; specifically, in the context of this work, it differentiates between noninjured and injured records. SVMs are known for their good generalization capabilities, robustness, and effectiveness, even in high-dimensional spaces.

Additionally, they require low computational requirements [Zha17]. Besides that, they are commonly used for injury forecasting [RPC21].

In order to develop an SVM-based injury model, the selected regularization parameter was the squared L2 penalty, and the radial basis function kernel was employed. Also, the classification output of the SVM was transformed into probability by using the Platt scaling [Pla+99], making thus possible to use a custom threshold for the final binary classification.

4.2.3.2 Feedforward Neural Network

A FNNs is a directed graph that processes input data through weighted connections, optimized during training to approximate the desired output. FNNs were chosen for this work mainly due to their ability to learn complex and nonlinear relationships between the inputs (i.e., the multiple predictor variables) and the desired output [Rud+18a]. Besides that, although FNNs have shown to be successful in various sport science works (see, e.g., [Mun+19] and [CL22]), their application to injury prediction is still to be further explored [Lóp+18; Van+21]. In fact, to the authors' best knowledge, only the work developed by Ruddy et al. [Rud+18a] used FNNs for injury prediction (in this case, injuries from professional Australian footballers).

For this work, FNNs with three layers were used. The number of units on the input layer was tuned by testing the different number of features. The hidden layer consisted of ten units with the hyperbolic tangent activation function. A single neuron unit with the sigmoid function was used for the output layer with the range of $(0, 1)$. A dropout regularization technique was also applied between the hidden and output layers [Sri+14]. The dropout probability was set to 60 % to prevent overfitting due to the limited data available for feeding the model. Lastly, and similarly to the SVM models, the chosen regularization parameter is the squared L2 penalty.

The FNNs were trained over 20 epochs using the Adam [KB15] optimization algorithm with a learning rate of 0.001, being the loss function the binary cross-entropy.

4.2.3.3 AdaBoost

AdaBoost is a well-known meta-learning algorithm in the boosting family due to its capabilities. The algorithm combines multiple weak classifiers into a strong one by iteratively adjusting the weights of training samples based on the performance of previous classifiers. AdaBoost has already been successfully applied for injury prediction; however, only two works were reported to the authors' best knowledge [Aya+19; Lóp+18]. For this reason, exploring the AdaBoost algorithm in other contexts was considered necessary.

AdaBoost is often used with decision trees (which was also the case for this work), where the learning models are stumps—one-level decision trees [FSA99]. The stumps are added to the ensemble at each iteration to minimize the errors from the previous weak learners [Bro10]. A maximum of 50 stumps were added to the final model.

4.3 Results

Sensitivity is the ability of the model to identify injuries correctly, and specificity is the ability of the model to identify noninjuries correctly. The average Geometric Mean (GMEAN) was calculated between the two folds to find a balance between sensitivity and specificity. This function is defined as

$$\text{GMEAN} = \left(\text{Sensitivity} \times \text{Specificity}\right)^{\frac{1}{2}}, \tag{4.5}$$

with higher values representing better models.

The features sorted by the mRMR method were tested for the best combination by sequentially adding features to the models. In order to achieve statistical significance, each combination of model, type of learning, and number of features was tested 500 times, resulting in 78 000 simulations.

The results of the current work are divided into three sections. The first section investigates the effect of each type of learning on the GMEAN value. Then, the best models found among the simulations are presented. Finally, the last section assesses the quality and stability of those models.

4.3.1 Cost-Sensitive Learning and Traditional Learning

Figure 4.2 compares, for each model, the cost-sensitive learning and traditional learning approaches regarding the average GMEAN value of the 500 runs for each number of features. It also offers the possibility of studying the number of features each model requires to reach a certain GMEAN level.

From that figure, it is possible to infer that cost-sensitive learning in the AdaBoost models does not represent a substantial advantage in terms of average GMEAN value. Regarding the training splits, if the number of features is lower than 230, both types of learning behave similarly; nevertheless, if a higher number of features are fed to the models, the traditional learning models are, on average, significantly better than the cost-sensitive learning models. In the testing splits, both types of learning follow the same trend of average GMEAN values. Nevertheless, traditional learning is better in some number of features and cost-sensitive in others (especially in the range of 180 to 210 features).

For the FNN and SVM models, however, cost-sensitive learning was always the best methodology to train the models. For these models, the average GMEAN value was consistently superior to traditional learning. This situation can be stated in both training and testing splits.

It is also important to note that for the AdaBoost and SVM models, both types of learning require approximately the same number of features to reach a certain level of average GMEAN value. On the contrary, FNNs trained with a cost-sensitive learning approach needed fewer features when compared to traditional learning. As an illustration, in order to reach 60.00 % of the average GMEAN value on the training splits, cost-sensitive learning requires 40 features, while traditional learning requires 180 (127.27 % more).

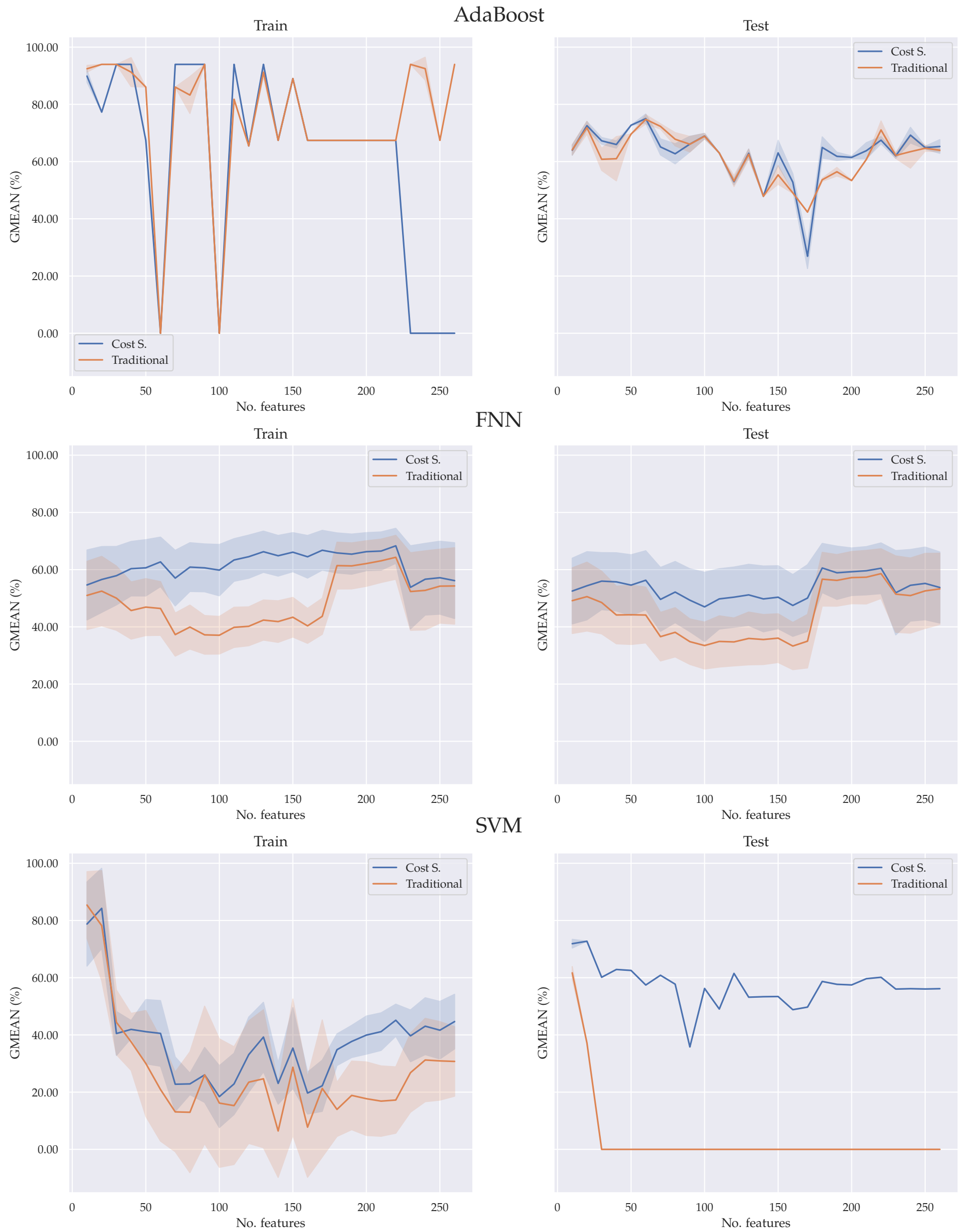


Figure 4.2: Comparison of cost-sensitive and traditional learning approaches across 500 simulations. This figure depicts the comparison of models' average GMEAN values from 500 simulations, contrasting cost-sensitive learning with traditional learning methods for both train and test data splits. It also illustrates the impact of adding more features on the GMEAN scores of classifiers. The shaded areas represent the standard deviation, highlighting the variability within each model configuration.

Table 4.1: Evaluation metrics of the best classifiers found. These classifiers were chosen considering the highest geometric mean value, calculated between sensitivity and specificity on the training splits, for each type of learning and classifier.

	AdaBoost		**FNN**		**SVM**	
Cost-sensitive	Yes	No	Yes	No	Yes	No
No. features	90	240	120	220	20	150
Train						
Sensitivity (%)	88.31	88.31	83.77	76.62	88.31	88.31
Specificity (%)	100.00	100.00	85.81	86.26	94.44	98.15
Accuracy (%)	96.60	96.60	85.80	86.21	92.65	95.29
AUC	1.00	1.00	0.88	0.87	0.98	0.99
GMEAN (%)	93.97	93.97	84.78	81.30	91.33	93.10
Test						
Sensitivity (%)	78.57	86.36	35.06	57.79	71.43	0.00
Specificity (%)	65.02	58.31	81.81	82.19	74.19	100.00
Accuracy (%)	65.08	58.38	81.67	82.13	74.22	99.71
AUC	0.72	0.73	0.46	0.75	0.85	0.73
GMEAN (%)	71.47	70.96	53.56	68.92	72.80	0.00

4.3.2 Best Predictive Models

The best predictive models were selected in a three-step process. First, the best classifiers were selected by considering the highest value of GMEAN in the training splits for each model and type of learning (independently of the number of features). Then, the same models were tested on data from players that were not used to train the models (i.e., unseen data). The final best models were then selected considering the highest GMEAN in the testing splits. It is highlighted again that, due to data limitations, it was not possible to create a validation set that would further improve the selection of the best models.

Those models are depicted in Table 4.1, along with their respective evaluation metrics and the best number of features. From this table, it can be concluded that the best two models used the AdaBoost learning method. Regardless of the type of learning, these classifiers obtained in the training splits a mean sensitivity of 88.31 %, a specificity of 100.00 %, and an accuracy of 96.60 %, resulting in a GMEAN of 93.97 %.

Followed by the AdaBoost classifiers, SVMs were the second-best models, obtaining a GMEAN ranging from 91.33 % to 93.10 % in the training splits. Although the AdaBoost and SVMs models presented similar sensitivity values (88.31 %), i.e., an equivalent capacity of identifying injuries on truly injury events, the mean specificity metrics dropped 5.56 % in cost-sensitive learning and 1.85 % in traditional learning.

In its turn, FNNs were the models with the lowest performances, with a GMEAN ranging from 81.30 % to 84.78 %. Consequently, these classifiers also obtained the lowest mean sensitivity and specificity metric values in the training splits (76.62 % and 85.81 %, respectively).

After choosing the best classifiers in the training datasets, the models that best predicted the injury and noninjury events on the testing splits were the AdaBoost and SVM, both combined with cost-sensitive learning. If, on the one hand, the AdaBoost classifier obtained a GMEAN value of 71.47 %, on the other hand, the SVM obtained the more balanced result (balancing sensitivity and specificity), reflected in the highest GMEAN value obtained (72.80 %). At the same time, the SVM model only required 20 features (vs. 90 features required by the AdaBoost classifier) to predict more than 70 % of injury and noninjury events. Nevertheless, these two models were selected as the best two models of this work since they achieved similar GMEAN values.

It is noteworthy, however, that the AdaBoost classifier combined with traditional learning was the model with the highest sensitivity (86.36 %) at the cost of having a low specificity and accuracy. This is a common scenario in imbalanced classification problems. At the other end of the spectrum, the model incorporating SVM with traditional learning could not detect any injury, invalidating its use in real-world scenarios.

Similarly to the training splits, FNN got one of the lowest performances of the set of models and types of learning. Indeed, the ability of the FNN trained using cost-sensitive learning to detect injuries is even inferior to a pure random classifier. The combination of FNN with traditional learning is, moreover, at the limit of its use in a real-world application since it can only detect 57.79 % of the actual injury events. Nevertheless, both models presented a high percentage of correct detections of noninjury events (81.81 % and 82.19 %, respectively).

The effect of cost-sensitive learning on the evaluation metrics compared to traditional learning is visible here. In this work, all models trained with cost-sensitive learning resulted in sensitivities in the training splits equal to or higher than traditional learning. As expected, the models' accuracies were neglected in the training splits; in the testing splits, it is interesting to note that traditional learning seems more favorable in correctly predicting injury events. However, the more balanced results (i.e., the higher GMEAN values) are usually achieved with cost-sensitive learning coupled with fewer features.

4.3.3 Quality and Stability of the Models

Figure 4.3 depicts the radar plots containing the mean and standard deviation information obtained from the 500 runs of the models' combinations, type of learning, and number of features identified in Table 4.1. This figure thus enables assessing the quality and stability (based on the standard deviation) of the injury forecasting models in the training and testing phases. At the same time, it also compares cost-sensitive learning with traditional learning, in terms of GMEAN, accuracy, specificity, and sensitivity, for both train and test splits.

AdaBoost models trained with cost-sensitive learning consistently learned to predict 88.31 % of the injury events in the training splits. Although the 500 trials exhibit, during training, the same and constant evaluation metric values from Table 4.1, the same was not

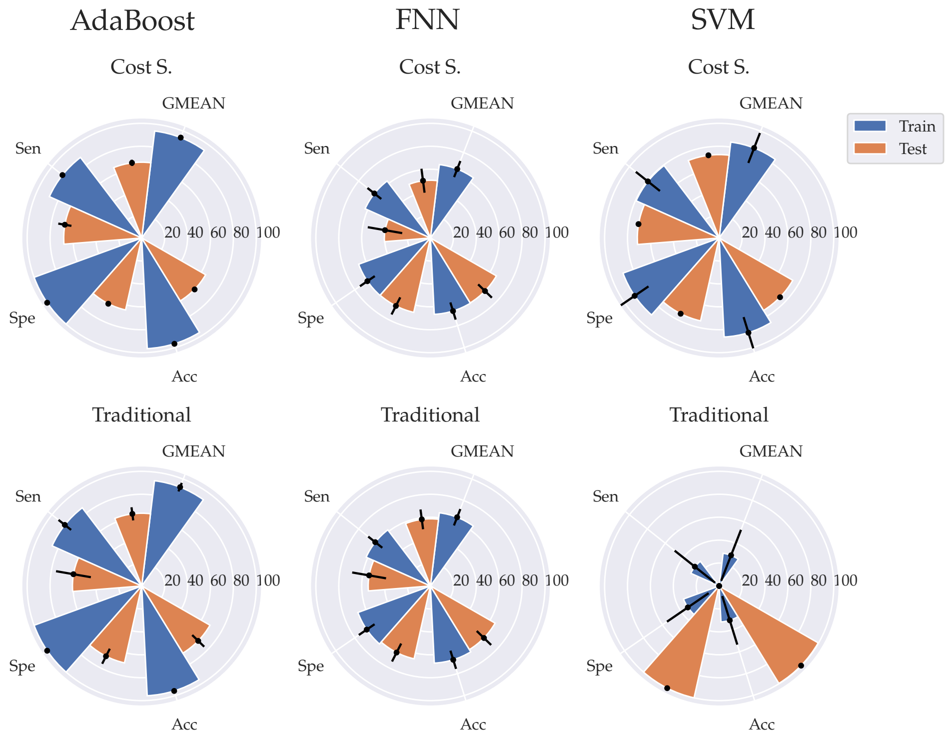


Figure 4.3: Radar plots of model performance metrics across 500 simulations. The radar plots presented here compare the quality and stability of various models and learning types over 500 simulations, based on the best combinations of model, learning type, and feature count as identified in Table 4.1. Stability is assessed through standard deviations depicted as lines on top of the bars, while quality is measured in terms of GMEAN values, sensitivity (Sen), specificity (Spe), and accuracy (Acc) for both training and testing data splits.

true for the testing phase, especially for the sensitivity metric. In testing, these models correctly predicted, on average, 68.09 % (SD=5.82 %) of the injury events and 64.33 % (SD=0.51 %) of the noninjury events, resulting in an average GMEAN value of 66.11 % (SD=2.72 %).

Cost-sensitive learning SVMs models, on the other hand, were shown to predict, on average, 79.43 % (SD=13.43 %) and 71.43 % (SD=0.00 %) of the injury events in the training and testing phases, respectively. Unlike the previous AdaBoost models, SVMs trained with cost-sensitive learning were more stable in the testing phase than in the training phase in all evaluation metrics. For instance, the specificity obtained in the training splits was 89.36 % with a standard deviation of 14.62 %; conversely, in the testing phase, a specificity of 74.10 % with a standard deviation of 0.03 % was reached. It is also important to note the average GMEAN values, which were 84.25 % (SD=14.01 %) for training and 72.75 % (SD=0.01 %) for testing.

AdaBoost models trained using traditional learning gave more emphasis to specificity and accuracy at the expense of a lower sensitivity value in the testing splits. Still, these models were able to detect correctly, on average, 60.47 % (SD=15.42 %) of the injury events and 68.76 % (SD=7.35 %) of noninjury events. On the other hand, FNN models were again dubious concerning their applicability in real scenarios, only detecting 40.25 % (SD=15.24 %) and 54.32 % (SD=15.00 %) of the actual injury events when using cost-sensitive learning and traditional learning, respectively. It is also noteworthy that, from the 500 trials, no SVM model trained using traditional learning was able to detect true injury events.

Overall, models trained with cost-sensitive learning showed to predict, as expected due to this learning's sensitivity to the minority class, more injury events correctly when compared to traditional learning. Equally important is the fact that cost-sensitive learning also produced the most stable models during the training and testing phases. However, FNN models are an exception to these two findings. In particular, although using cost-sensitive learning on these models provided, in the testing datasets, higher specificity and average accuracy values, sensitively, and GMEAN metrics were lower than traditional learning. By the same token, the differences in the standard deviation between cost-sensitive learning and traditional learning were, despite minimal, sometimes higher.

4.4 Discussion

The current investigation aimed to develop an automatic technique for forecasting injury events. This technique is based on the information from the GPS devices collected throughout games and training sessions combined with other descriptive variables (such as the player's corridor). In this view, the conducted study allowed us to derive two models. The first model consisted of an AdaBoost classifier, and the second of an SVM. Both models generalized the results with acceptable sensitivity, specificity, and accuracy and were revealed to be stable, thus suggesting their applicability in real-world scenarios.

To the authors' best knowledge, this is the first study that included the use of the mRMR algorithm to rank the features according to their relevance and redundancy for forecasting injuries. In this view, the AdaBoost required 90 features to detect 78.57 % of the actual injury events. On the other hand, the SVM model required only 20 features but obtained a lower sensitivity (71.43 %). Nevertheless, the selected variables in both cases enabled the two predictive models with a high explanatory capability for injury events.

The player's position and type of session (i.e., training or match session) were essential descriptive variables for the two injury prediction models (cf. Appendix C Table 1). The player position has not been broadly reported to be associated with an increased injury risk factor [VML18]; however, injury rates are documented to be higher for matches than training sessions [Høy+92]. Possible justifications for the increased injury rate in match sessions may arise from the variations in the intensities of training and competition [Pfi+16] and, to a certain level, from the type of training before the next match (e.g., tactical practice) [WHE05]. Thus, the two developed models seem to capture this information and use it to predict injuries in conjunction with the other predictors.

The day of the week is another descriptive variable the models require for predicting injuries. To the authors' best knowledge, no other study on injury prediction in professional football reported a relationship between the day of the week and injury risk. All things considered, the added descriptive variables were essential to leverage the model's performance in predicting injury and noninjury events.

It is also interesting to note that AdaBoost used the player load and all the velocity and acceleration bands in order to predict injuries. The information about the player load across all axes of accelerometer movement, distance, duration, and effort count performed in all bands makes the model detect when players are in undertraining or overtraining situations. Both situations are boosters of injury events, with some studies reporting a U-shaped curve between these parameters and injury risk [Gab16; Mal+17a].

Interestingly, those models do not use the players' age for forecasting injuries. Indeed, the literature that studied the relationship between age and risk of injury is inconsistent. Some works suggest that the risk of injury increases with age, and others report insufficient evidence to infer a significant effect of age on injury risk [Hug+17].

The use of cost-sensitive models is an approach followed by several studies for injury detection (see, e.g., [Aya+19; Rui+21]) due to the problem's imbalanced nature, i.e., due to a significant difference between the number of injury and noninjury events. However, to the authors' best knowledge, this technique was only employed in studies made based on the data collected on screening battery processes, leaving its use with longitudinal GPS data to be investigated in this work.

In the screening studies [Aya+19; Rui+21], cost-sensitive models performed better than traditional learning classifiers. Despite having a different data collection process, this study demonstrates to be also in line with these results; however, AdaBoost coupled with a cost-sensitivity train did not show to be superior to traditional learning in terms of average GMEAN value. A plausible explanation is that, unlike the FNN and SVM models, the AdaBoost algorithm is an accuracy-oriented classification algorithm. Thus, even with a cost-sensitive learning approach and oversampling, the specified cost for the minority class could have been insufficient to incline the boosting strategy to the minority class [Sun+07].

The most recent studies investigated, on the other hand, the use of tree-style classifiers for injury detection since these models provide the classification rules and the most critical features for injury prediction (see, e.g., [Ros+18; Ros+22; Val+20]). However, a trade-off between performance and interpretability must be made. As a result, tree-style classifiers have a performance that is usually inferior to other methods. The current work, in its turn, is focused on the performance metrics, leaving the interpretability aspect for future works.

In that view, after conducting the 500 runs for each combination, the AdaBoost and the SVM were the best models identified. These models use cost-sensitive learning and can

Table 4.2: Comparative analysis between results reported by the state-of-the-art works and the results attained in this work.

Work	Model	Sen. (%)	Spe. (%)	Acc. (%)	AUC (%)
Rossi et al. [Ros+18]	Decision tree	80.00	†	†	0.76
Naglah et al. [Nag+18] ‡	SVM	82.22	†	†	0.76
Vallance et al. [Val+20] ‡,§	Gradient Boosting	96.81	†	96.75	0.97
Vallance et al. [Val+20] ‡,¶	Gradient Boosting	96.56	†	96.63	0.97
Rossi et al. [Ros+22] §	Gradient Boosting (XGB)	65.00	†	†	†
This work	AdaBoost	78.57	65.02	65.08	0.72
	SVM	71.43	74.19	74.22	0.85

† Information not reported.
‡ Data was inferred from the plots.
§ One-week prediction.
¶ One-month prediction.

detect, respectively, 78.57 % and 71.43 % of the injuries in the testing datasets while keeping an acceptable (> 65 %) true negative rate. Although there is a known trade-off between correctly predicting more injuries and incorrectly flagging noninjury events [Car+18], the SVM model is the most balanced of the two models, obtaining an Area Under the receiver operating characteristic Curve (AUC) of 0.85.

A comparison between the results reported by the previous state-of-the-art works and the results attained in this work is presented in Table 4.2. It is important to note that the best model for the studies that reported multiple classifiers was selected based on the highest sensitivity since not all studies reported the metrics used in the current study to compute the GMEAN.

Besides being the first work that combined machine learning with GPS data to predict injuries, the work of Rossi et al. [Ros+18] can be considered the most influential in this area. It was established in their work an injury forecaster capable of predicting 80.00 % of true injury events. It also provided an interpretable framework between injury risk and training performance. However, the AUC metric suggests that the model might create a significant number of false alarms and thus unnecessarily bench players before the next game or training session. Although this situation is also visible in the proposed AdaBoost model, the SVM model remedies this situation by obtaining a more balanced result between sensitivity and specificity at the expense of lower sensitivity.

The models proposed in this study do not require constant manual data collection to forecast injury events accurately, thus being cost- and time-effective. This, however, is not the case for some studies in the literature that combine GPS data with other pieces of information. For example, Naglah et al. [Nag+18] obtained one of the highest sensitivity values. Although meritorious, their proposal uses GPS data combined with players' questionnaire data. Requiring players to fill out questionnaires frequently is a strategy that can be time-consuming and challenging to incorporate into players' routines. Furthermore, the increase in sensitivity (about 2 %) is not significant to the point of requiring questionnaires. In the same view, Rossi et al. [Ros+22] combined GPS data with blood parameters to assess individual psychophysiological responses to training and create an injury forecasting model that predicts injury events in the subsequent seven days. Although only three blood samples were collected, on average, from each player, the post-blood collection procedure is complex, costly, time-consuming, and requires specialized personnel to be conducted. Besides that, unfortunately, they were unable to predict more than 65.00 % of the injury events.

Vallance et al. [Val+20] presented the best results in the literature, detecting almost every injury event while balancing sensitivity and specificity. Contrary to the approach presented in the current proposal, Vallance et al. [Val+20] generate injury predictions for the forthcoming week or month. The superior performance of Vallance et al.'s [Val+20] method, when compared to the present approach and other methodologies in the scientific literature, could be attributed to the difference in prediction time frames. Generally, forecasts covering a more extended period tend to be more accurate because they allow for a wider margin of error, leading to less precise predictions. For example, predicting an injury for the next week suggests a potential occurrence at any time during those seven days, which is inherently less precise than a prediction pinpointing a specific day for the potential injury.

The main findings of this study will help the coaching staff to identify football players in high-risk situations for injury and improve their decision-making. This will inevitably leverage the team's performance, and simultaneously, reduce the club's economic cost due to injury events.

Besides, it will enable constant monitoring of multiple parameters without manual intervention and the analysis from the coaching staff, which is limited due to the large number of parameters collected by the GPS receivers. Indeed, knowing training and competition effects on the injury risk will also improve the training design and ensure that players receive an adequate training session before and after matches by keeping the correct balance between high and low intensities.

Ultimately, having the possibility to adjust the threshold used to convert continuous injury predictions to discrete labels will enable the coaching staff to draw more informed football tactics. This will enable to control the risk of injury events according to, for example, the team's position in the championship table.

This study is to be seen in the light of some limitations, which can be, at the same time, possible directions for future works. The number of injuries needed to be larger to test more complex models (eventually, recurrent models such as long short-term memory networks) or to assess the models' prediction capabilities fully. Additional instances from, e.g., another season or different cohorts (for example, U23 and U19 teams) would remedy this situation. The authors thus highlight that the models were only validated for the analyzed football team.

This study did not specifically measure muscle and body fatigue or include certain types of exercise, such as cardio workouts, that players might have engaged in on their rest days (for example, during the dummy days). These factors might also be connected to the risk of injury and could improve the injury prediction models. Nonetheless, this study aims to develop an automated system to predict injury risks daily throughout a football season without requiring continuous manual data collection, such as that needed for assessing muscle and body fatigue (for example, through self-reports). And, although direct measurements of muscle fatigue or specific activities on off days were not part of the data collected, it is reasonable to assume that the information obtained from the GPS devices provide a reliable indication of the player's physical condition.

In future studies, it would be beneficial to include additional physiologic parameters of the players (e.g., history of prior injuries and rating of perceived exertion for each session) to enhance the models further. Unfortunately, this information was not available at the time of this study. Moreover, future studies should focus on enhancing model interpretability without significantly sacrificing performance.

4.5 Conclusion

This work used three machine learning methods (SVM, FNN, and AdaBoost) to predict injuries from professional football players. Besides using the information from the GPS receivers, the models incorporated the effect of sudden changes in player load by including dummy days (i.e., records with zeros for all parameters). Descriptive variables, such as player position and day of the week, were also included and showed to leverage the ability of the models for injury prediction.

Before feeding information to the models, features were sorted and selected according to their redundancy and relevance to injury risk using the mRMR method. This procedure revealed the player's position, type of session, velocity bands, and acceleration bands as essential features for injury prediction.

In turn, the predictive models were shown to be able to accurately detect injuries and noninjuries events, especially the AdaBoost and the SVM, trained with cost-sensitive learning. These models were able to predict more than 70.00 % of new injury and noninjury events and be stable in terms of performance metrics. Comparing these results with the ones available in the literature, the models developed in this work stand out for (a) being the most balanced ones (between sensitivity and specificity), (b) not requiring lengthy and manual data collection processes, and (c) the ones that predict injury for short time frames (in this case, one day).

Although the number of injuries was not large enough to fully assess the models' prediction capabilities, the current models can be used in real-world scenarios. Models will help the coaching staff to identify football players in high-risk situations and, thus, leverage the team's performance while minimizing rehabilitation costs.

CHAPTER 5

Unveiling Energy Poverty Trajectories: A Longitudinal Analysis Using Machine Learning

Santiago Budría, Eduardo Fermé, Diogo Nuno Freitas

Abstract. *Identifying at-risk populations is essential for designing effective energy poverty interventions. Using data from the HILDA Survey, a longitudinal dataset representative of the Australian population, and a multidimensional index of energy poverty, we develop a machine learning model combined with SHAP (SHapley Additive exPlanations) values to document the short- and long-term effects of individual and contextual factors—such as income, energy prices, and regional conditions—on future energy poverty outcomes. The findings emphasize the importance of policies focused on income stability and may be used to shift the policy focus from reactive measures, which address existing poverty, to preventive strategies that target households showing early signs of vulnerability.*

5.1 Introduction

Energy poverty, a condition where households are unable to access or afford adequate, reliable, and clean energy services, has emerged as a critical global issue. Recent estimates suggest that 750 million people still lack access to electricity worldwide, and more than 2 billion people lack access to clean cooking fuels [Int24]. In economic literature, energy poverty has garnered independent attention and is now studied as a distinct subject. This is because energy poverty is only moderately correlated with income poverty and yet negatively related to relevant economic outcomes, including human capital formation and employment opportunities [Sha+24a], well-being [NL24], and health [Pon+24].

In this context, the identification of populations at risk is essential for formulating effective policy interventions. While there has been significant research on the socioeconomic determinants of energy poverty [ABR22; FFT22; KA22], most studies look only at contemporaneous relationships, assuming that control variables fully represent the information influencing the observed outcome. This approach may overlook the potential role of long-memory processes and the enduring influence of contextual factors. If energy poverty is indeed a chronic state shaped by an individual's history, such perspectives could provide an incomplete understanding. Evidence on the long-term effects of specific characteristics on energy poverty remains scarce, highlighting the need for further research in this area.

This paper takes a step forward by analyzing the short- and a long-term effects of individual variables and contextual factors—such as income, energy prices, regional conditions, and other socioeconomic variables—on future energy poverty outcomes. We use the 2007–2021 waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey[1], a micropanel survey representative of the Australian population, which allows us to track people over up to 15 consecutive years. Since energy poverty is a multifaceted construct, we utilize a Multidimensional Energy Poverty Index (MEPI) incorporating five items that capture both objective (expenditure-based) and subjective (self-assessed) dimensions. We then use machine learning models to forecast the MEPI and employ SHAP [LL17] to interpret these predictions. SHAP quantifies the contribution of each variable, highlighting how specific factors at different points in time influence future energy poverty outcomes. This approach provides an understanding of the importance of each variable and the temporal dynamics shaping energy poverty trajectories.

This paper makes three significant contributions to the literature. First, the paper aims to advance the methodological toolkit for studying energy poverty. SHAP has been applied successfully to the study of financial time series data [MHB19], short-term load forecasting [Lee+23], and aviation's predictive maintenance [AA24]. To the best of our knowledge, this is the first analysis to combine SHAP techniques with high-quality micropanel data to identify household-level drivers of energy poverty. By using SHAP, the paper highlights how its application in energy poverty research extends beyond the capabilities of traditional analytical methods.

Second, the paper contributes to the growing body of literature employing machine learning techniques to analyze the determinants of energy poverty. This approach is still limited but increasingly recognized for its potential to guide alleviation strategies[2] [DSZ21;

[1] https://melbourneinstitute.unimelb.edu.au/hilda.

[2] Unlike traditional regression methods, which require prior assumptions about potential correlations and their functional forms, machine learning techniques allow these relationships to naturally emerge during model training and excel at capturing complex, non-linear dependencies.

GJA24; HDZ22; SRL23]. However, much of the existing machine learning-based evidence relies on contemporaneous relationships between explanatory variables and energy poverty, largely due to the prevalence of cross-sectional or short-duration datasets. Our study, on the other hand, utilizes a 15-year longitudinal dataset to explore predictive dynamics and, by employing SHAP, we are able to quantify and disentangle the contribution of each feature to predictions, both overall and at specific points in time, offering insights into how past conditions influence future energy poverty. We note that, although our context does not fully align with the high-dimensional settings typically emphasized in the broader machine learning literature, applying machine learning to a focused set of policy-relevant variables remains valuable for prediction, risk identification, and effective communication of results.

Third, historically, public initiatives that address energy poverty—particularly in developed nations—have primarily focused on providing financial assistance and energy subsidies to individuals currently classified as (energy) poor. This approach operates on the assumption that immediate interventions can effectively alleviate energy poverty in the short term. However, this focus overlooks households that are at risk of becoming energy-poor in the future, leaving a significant portion of the potentially vulnerable population unaddressed. Our approach challenges this perspective by emphasizing the importance of both the timing and the magnitude of key variables, such as income stability and energy prices, in shaping energy poverty trajectories. By identifying these "to-be energy-poor" households, our paper paves the way for more proactive policies that tackle the historical, individual-level gradient of energy poverty, moving beyond temporary relief measures.

We consider Australia to be a compelling subject for our research. Escalating energy costs have been a major concern over the last decade in Australia as electricity prices have almost tripled [Pro22]. Forward electricity prices for 2023 delivery in Australia's National Electric Market surged from approximately $48 in 2021 to $156/MWh in 2022 (the 52-week average), peaking around $247/MWh in October 2022 [Sim23]. This substantial surge, relative to household income, has placed a heavier burden on household budgets and exacerbated issues related to energy access and affordability [OEC23]. Furthermore, despite its fragmented system of energy assistance—varied across states and territories—existing programs are largely focused on mitigating costs through price subsidies and welfare payments for energy bills [Wil22]. These measures heavily rely on means-testing and target low-income groups [SM23], often overlooking individuals who are not currently energy-poor but are at risk of becoming so. By identifying to-be energy-poor individuals before they fall into vulnerability, our paper shifts the focus from reactive policy interventions to preventive, forward-looking interventions.

The paper shows that historical household income levels are pivotal in forecasting energy poverty outcomes, particularly over longer time horizons and in more severe cases. It also highlights the significant impact of income variations, independent of static income levels. This effect intensifies when transitioning from short-term to long-term poverty, suggesting that income volatility is particularly harmful in the long run. Additionally, energy prices have a moderate, non-linear effect in the short term but become less relevant for longer horizons. These findings emphasize the importance of policies focused on income stability and may be used to shift the policy focus from reactive measures, which address existing poverty, to preventive strategies that target households showing early signs of vulnerability.

The paper is structured as follows: Section 5.2 reviews the relevant literature on energy poverty, its determinants, and the application of machine learning methods in this context. Section 5.3 describes the data, key variables, and the construction of the MEPI. Section 5.4 outlines the methodological approach, including model development and the use of SHAP for interpretability. Section 5.5 presents the results, highlighting the predictive performance of the models and the temporal dynamics of key variables. Section 5.6 discusses sensitivity analyses and robustness checks. Finally, Section 5.7 concludes with key findings, policy implications, and limitations of the study. The paper includes three appendices with technical details.

5.2 Review of the Literature

Energy poverty can be defined as a household's inability to afford or access energy services needed to support adequate living conditions and human development. While translating into practice conceptual definitions of energy poverty is typically a challenge and has been the object of extensive discussion in the literature (for an overview see Sy and Mokaddem [SM22]), the focus has generally been put on the inability of households to afford and have access to adequate energy services.

The global interest in energy poverty stems from its far-reaching consequences, which are multifaceted. Research based on international macroeconomic data shows that the prevalence of energy poverty negatively affects development, health outcomes, and average schooling levels [BMM21]. Moreover, energy access and affordability are crucial dimensions of multidimensional poverty, and, as such, they can be negatively related to economic growth [BL24]. Studies based on microeconomic panel data are consistent with this notion, showing that energy poverty significantly affects a number of personal-level outcomes, including subjective well-being [LO21] and health [Pon+24; Zha+21]. Energy poverty is also negatively related to children's academic performance [ZAK21] and human capital formation [Sha+24a].

Using international comparable data, research shows that country-level factors such as education, governance quality, technology advancements, economic development, and health expenditures are relevant determinants of household-level energy poverty depending on the country's GDP [Boţ+24]. Moreover, income inequality and, to a lesser extent, climate conditions also play a role [IM22]. Furthermore, the sources of electricity production also contribute to shaping energy poverty outcomes, reflecting the importance of a country's energy mix [KUO23]. Additionally, high energy costs, accessibility, and the types of energy sources further shape these outcomes [KR25]. Inefficient building structures, dwelling size, age, thermal insulation, floor area, and heating system can be significantly correlated with various forms of energy deprivation [KŚ20]. Spatial disparities—particularly between rural and urban areas—also exacerbate energy poverty [Uma+24].

At the household level, income constraints, coupled with high energy prices, can culminate in the difficulty of paying bills, energy debt, and even the disconnection of energy supplies [ABR22; MM24]. Educational attainment is inversely correlated with energy poverty, primarily due to energy-saving practices and an improved economic situation. Education enhances knowledge and the capacity to make choices that benefit household welfare, leading to better living conditions through improved decision-making and the adoption of more efficient energy sources [CAF19]. Place of residence, gender, and household size also exhibit a statistically significant relationship with multidimensional poverty due to increased energy consumption needs [Abb+20]. Additionally, age effects may arise from life cycle patterns, household arrangements, and risk-taking behavior, while poor health conditions may hinder access to energy services and goods by altering spending priorities and consumption patterns [FFT22]. Labor market status, as well as marital status, are frequently found to be significantly associated with energy deprivation, with the effect being particularly pronounced in developing economies [Abb+20; ABR22; MM24]. Cultural characteristics and parental behavior [PAS22], and energy subsidies also contribute to shaping energy deprivation outcomes [Hos+23].

Despite these advances, a significant gap in the literature persists: understanding how current circumstances shape energy poverty outcomes later in life. The studies discussed above primarily emphasize contemporaneous relationships between explanatory factors and energy poverty, regardless of whether the findings stem from cross-sectional or panel data analyses. Studies on energy poverty dynamics are scarce, with only a few papers addressing this issue through dynamic panel models in which energy poverty is allowed to depend on past energy poverty [AD20; DJ21; HK23].

5.2.1 Machine Learning Models in Energy Poverty Research

A recent body of literature has introduced machine learning techniques to predict energy poverty outcomes. Evidence based on an Extreme Gradient Boosting (XGBoost) framework to predict the risk of experiencing energy poverty in the Netherlands identifies income, house value, and house ownership as the main drivers of energy poverty [DSZ21]. In a similar setting, and based on 11 European countries, income, household size, and floor area were consistent predictors [HDZ22]. Evidence based on an Random Forest (RF) classifier across the European Union uncovers household- and country-level predictors like dwelling conditions, energy efficiency, and gas supplier switching rates [SRL23].

While the previous studies are based on a single energy poverty indicator, other studies define a multidimensional energy poverty index similar to ours. These studies showed that in Asian and African countries, wealth, marital status, and residence attributes are significant predictors of poverty [Abb+20]. Recent research has further advanced these methodologies by employing ensemble models, such as XGBoost, combined with RF and Artificial Neural Network (ANN), revealing the critical importance of education and food security indicators in determining energy poverty [GJA24].

5.2.2 Measurement

The literature typically distinguishes between objective (expenditure-based) and subjective (self-assessed) approaches. Because poorer households often spend higher proportions of their budget on energy-related expenses relative to higher-income households [SM22], expenditure-based measures label a household as energy-poor when the income that households spend on energy is above a specific threshold. For instance, a household may be classified as energy poor if i) its share of income spent on energy is greater than twice the national median (the 2M indicator); ii) its share of income spent on energy exceeds 10% (the Ten Percent Rule, TPR); or iii) its actual energy expenditures are above the national median and, at the same time, their income net of energy costs is below the official national income poverty line (the Low Income High Costs indicator, LIHC). These measures have been used extensively in the literature [ABR22; FFT22; MM24].

However, while expenditure-based measures are objective and transparent, they may overlook intentional reduction in energy consumption by low-income households. If vulnerable households limit their energy consumption to prioritize other services and goods, measures based on the actual energy costs may underestimate the true prevalence of energy poverty. Moreover, low-income families can resort to energy credits and repayments to smooth their monthly energy costs over time. To overcome these limitations, applied research has relied on individuals' self-evaluations of their ability to afford and access specific energy services [PAS22; SRL23]. Following this criterion, several multidimensional energy poverty indexes have been proposed, gathering information related to basic energy services, including cooking, lighting, and household appliances in developing countries [Abb+20; GJA24].

5.3 Data and Key Variables

We use the HILDA Survey, a comprehensive, nationally representative longitudinal study that examines the economic, social, and demographic dynamics of Australian households. Initiated in 2001 and conducted annually, it tracks individuals and households over time, providing important information about income, labor market activities, health, education, and family relationships, among other factors. The original 2001 sample included approximately 7,600 households and 13,000 individuals, with periodic updates to account for attrition. While panel data is subject to selection and attrition bias, potentially limiting the generalizability of findings, HILDA has a high average retention rate of over

90% across waves. Nonetheless, to address concerns about attrition bias, several sensitivity checks are presented in Section 5.6.

We utilize a balanced panel, allowing for varying durations. Our benchmark analysis relies on data spanning up to $T = 8$ consecutive years, enabling us to conceptualize energy poverty at time T as a function of characteristics from the previous $T - 1$ periods. This approach yields 57,160 observations from a cohort of 7,145 individuals with complete records. To enhance the robustness of our findings, in Section 5.6 we present additional results for panels spanning $T = 2$, $T = 4$, $T = 12$, and $T = 14$ years.

We model energy poverty as a function of socioeconomic factors that are standard in the literature. These include household income, employment status, schooling, age, marital status, parenthood, health status, and household size. We also include controls for remoteness, region of residence (the six states and two territories of Australia, reference: New South Wales), and wave-specific effects. Due to their potential impact on energy poverty, we use annual electricity and gas prices at the state level drawn from the Australian Bureau of Statistics [Aus24]. All income and price variables used in the paper are transformed using the OECD equivalence scale and normalized into real terms using the yearly consumer price index. We also include variables to control for macroeconomic conditions at the regional level. The economic cycle affects the chance to find and keep jobs, and it also impacts the likelihood of having a stable income source. We include controls for the regional unemployment rate, per capita GDP, and GDP growth. We also include the regional participation rate to capture competition effects in the labor market and the labor force share of part-time workers to account for the fact that areas with a higher proportion of temporary and/or part-time contracts typically experience greater uncertainty in work hours and income stability. In Appendix D.1 we provide a detailed summary of the variables used in the analysis.

Energy Poverty

Energy poverty is a multifaceted construct; therefore, we rely on five items that capture both expenditure-based and subjective dimensions. The expenditure-based measures include the 2M, TPR, and LIHC indicators, which are widely recognized in the energy poverty literature and detailed in Section 5.2.2. We also consider two self-assessed indicators based on the household's inability to pay to heat their home because of a shortage of money (Heat) and pay electricity, gas, or telephone bills on time (Arrears).

The MEPI index is calculated as follows: Let $J = 5$ represent the set of poverty indicators, with element j, $j \in J$ and $m = \text{card}(J)$. Let I be a set of individuals, with element i, $i \in I$, and T be a set of time periods, $t \in T$, representing a specific moment when the survey was conducted. Let EP_{ijt} denote the status of the ith individual in the j-th indicator during period t. If an individual i is poor under indicator j in the period t, then EP_{ijt} takes the value of one, and zero otherwise. Following the family of indexes typically described in the literature on material deprivation [DPX19], individual i's weighted poverty score is given by:

$$\text{MEPI}_{it} = \left(\sum_{j \in J} w_j \text{EP}_{ijt} \right), \quad \forall i \in I,\ t \in T_i;\ T_i \subseteq T, \tag{5.1}$$

where w_j denotes the weight assigned to the poverty indicator j, with $\sum_{j \in J} w_j = 1$. Hence, the MEPI_{it} ranges from 0 to 1 and captures the percentage of dimensions in which the individual is deprived. An individual i is regarded as energy poor if $\text{MEPI}_{it} > \bar{m}$, where $\bar{m}$ is a cut-off point. Thus, our dependent variable is a binary variable that takes value one if the individual is energy-poor, and zero otherwise. For the baseline parametrization, we set $\bar{m} = 0$. In Section 5.6, we provide robustness checks with alternative cut-off points, namely $\bar{m} = 0.2$ and $\bar{m} = 0.4$.

While it is common to assign equal weights to the indicators, we emphasize the indicators where deprivation is less common, the so-called frequency-based weighting approach [DL13]. The weight given to an indicator is proportional to the percentage of individuals *not* classified as poor under that specific indicator within a particular state. In other words,

$$w_j = \frac{(1 - n_j)}{\sum_{j \in J}(1 - n_j)}, \tag{5.2}$$

where n_j is the proportion of poor individuals in dimension j. This choice is motivated by the idea that not having access to common items should be a more relevant determinant of deprivation than less common items. Additionally, the weights are based on the distribution of achievements in society without considering any value judgment about what the trade-offs between items should be. For greater granularity and accuracy, the weights are calculated separately for each wave. There are two advantages to using that approach. Firstly, it allows the poverty of a given individual to increase if their conditions do not change and the conditions of all others improve. Secondly, it adapts automatically over time, considering economic conditions and social and cultural preferences when accessing items.

The MEPI shows two desirable characteristics, as it can be used to measure the prevalence and average intensity of energy poverty in a population. Prevalence is given by:

$$p = \frac{q}{\text{card}(I)}, \tag{5.3}$$

where q is the number of deprived individuals, $q = \sum_{i \in I, t \in T_i} \mathbb{I}(\text{MEPI}_{it} > \bar{m})$, where the indicator function $\mathbb{I}(\cdot)$ equals one if its argument holds, and zero otherwise. The intensity of energy poverty, i.e., the average poverty score of individuals identified as energy poor, is:

$$a = \frac{\sum_{i \in I, t \in T_i} \text{MEPI}_{it} \times \mathbb{I}(\text{MEPI}_{it} > \bar{m})}{q}. \tag{5.4}$$

The average population MEPI is then:

$$\text{MEPI} = a \times p. \tag{5.5}$$

The advantages of these axiomatic properties have been highlighted in previous work of Crentsil, Asuman, and Fenny [CAF19].

5.4 Methodological Approach

We model energy poverty at time T as a function of individual characteristics observed in periods $T-1$, $T-2$, $T-3$, and so on. This structure is not intended to imply that contemporaneous characteristics are irrelevant or lack explanatory power. Rather, the exclusion of period T covariates is a deliberate modeling choice aimed at avoiding potential endogeneity that could arise if current individual characteristics were themselves affected by the experience of energy poverty in the same period [Pon+24]. By lagging the covariates, we reduce the risk that the direction of causality is reversed—i.e., that energy poverty influences the covariates rather than being influenced by them—thus strengthening the causal interpretation of the estimated relationships. Notwithstanding, in Appendix D.4, we provide estimates when contemporaneous predictors are included in the model and compare their explanatory power with that of lagged variables.

We then integrate the machine learning techniques with an interpretability framework. This integration allows us not only to predict energy poverty outcomes but also to understand the contribution of each historical factor to these predictions. This involves a systematic process of data preparation, model development, and the application of feature importance and explainability techniques.

5.4.1 Data Preparation

To capture the temporal dynamics of the variables, we created lagged features, which serve as the input to the predictive models. Generically, for each original feature, we obtained new features representing their values from each of the previous years. This transformation ensures that the model has access to the full temporal history of each variable, enabling it to learn patterns and relationships that may influence the energy poverty indicator in the T-th year. We split the dataset into training, validation, and test subsets to facilitate model development and evaluation. For our baseline estimates ($T = 8$), out of the 7,977 participants in our dataset, 6,382 (80%) were randomly selected for training and validating the predictive models, while the remaining 1,595 participants (20%) were included in the test set. The test set was held out and used exclusively to evaluate the final performance of the models, providing an unbiased estimate of their forecasting accuracy.

To avoid data leakage across the splits, each individual was assigned exclusively to one subset, ensuring that no participant's data appeared in more than one split. Additionally, we removed any user identifiers from the data to prevent the models from learning user-specific patterns, which could limit their generalizability. The year variable (“wave" variable) was also excluded from the input features to ensure that the models focus on patterns within the socioeconomic and demographic variables rather than relying on specific temporal markers.

Before training the models, we standardized the data to ensure consistency and reliability in our modeling process. This involved removing the median and scaling the data according to the interquartile range, a method particularly effective at managing outliers and recommended as a best practice in machine learning [SWW21]. Such standardization is crucial in predictive modeling; it normalizes all input features to a similar scale, thereby enhancing the model's generalization capabilities and preventing variables with larger magnitudes from disproportionately influencing the learning process [Mah+24]. To prevent data leakage, the scaling parameters were calculated using only the training set and then applied to the test sets.

Feature engineering was explored in this study to improve the forecasting power of our models. Specifically, we expanded the set of socioeconomic, geographical, and contextual factors by including a range of interaction terms and decomposing variables into levels and yearly variations. While this approach increased the model's ability to identify energy-poor households to 78.04% compared to the final model (cf. Table 5.1), it reduced the overall accuracy, with the ability to correctly classify non-energy-poor households dropping to 59.31%. Moreover, the increased complexity introduced by additional variables would pose practical challenges for policymakers, making the results harder to interpret and apply. Consequently, we retained the model configuration that provided a better balance between performance and practical usability for policy design. The results of the models with the expanded feature set can be provided by the authors upon request.

5.4.2 Model Development

We treat the energy poverty forecasting task as a classification problem. Specifically, households are classified as energy-poor depending on whether their MEPI is greater than $\bar{m}$ (cut-off point), where $\bar{m} = 0$ for the baseline model.

The dataset used in this study exhibited a significant class imbalance, with most participants (73.55%) being classified as not energy-poor and a smaller proportion (26.45%) classified as energy-poor. This imbalance poses challenges for predictive modeling, as standard machine learning methods tend to favor the majority class, potentially leading to poor performance in identifying the minority class [Pro00]. To address this issue, we tested three ensemble classifiers, namely random under-sampling boost classifiers [Sei+10], balanced bagging classifier (for a review on bagging classifiers, see, e.g., Galar, Fernandez, Barrenechea, Bustince, and Herrera [Gal+12]), and easy ensemble classifier [LWZ09]. Due to space constraints, we describe here only the balanced bagging classifier. The descriptions of the other classifiers are provided in Appendix D.2.

A balanced bagging classifier is an ensemble technique that combines the predictions of multiple base models, e.g., decision trees, in order to improve the robustness and accuracy of the outcomes. This method specifically addresses class imbalance by ensuring that each decision tree in the ensemble is trained on a balanced subset of the dataset. These subsets are created by resampling the original training data, wherein each subset contains a representative distribution of both minority (energy-poor) and majority (not energy-poor) classes. In order to further refine the modeling approach, we implemented the classifiers in an One-vs-the-Rest (OvR) binary classification framework [Mur12]. OvR decomposes the problem into multiple binary classification tasks, where each class is treated as a separate binary problem against all other classes. Although OvR is commonly used for multi-class classification tasks, this methodology fits one classifier per class, which enables the models to focus on the distinctions between the two groups.

We optimized the hyperparameters of our classifiers using a grid search, which tested various configurations to identify settings that maximize model performance. For details on the specific hyperparameters and grid configurations, see Appendix D.2. We employed 5-fold cross-validation on the training dataset to ensure the robustness of the hyperparameters across different data splits, selecting the best set based on the highest training ROC AUC score. This metric is crucial for datasets with class imbalances, like the HILDA Survey, as it fairly assesses the model's discriminatory power between energy-poor and non-energy-poor households.

The final model was trained on the complete training set using the identified optimal hyperparameters and subsequently evaluated on a completely separate held-out test set of 1,595 participants who were never observed during training, predicting outcomes for entirely unseen households in future time periods. This approach ensured an unbiased assessment of the model's forecasting accuracy. In addition to this, all experiments were conducted with a fixed seed to guarantee reproducibility.

5.4.3 Feature Importance and Explainability

To interpret the forecast of our model and understand the contributions of individual features, we employ the SHAP. SHAP is a well-known method for explainability in the literature due to its theoretical consistency and ability to provide both local and global explanations of model behavior [LL17]. It is rooted in cooperative game theory and assigns each feature a contribution value toward the model's prediction, being thus model agnostic. The SHAP value (ϕ) for a given feature k is given by,

$$\phi_k = \sum_{S \subseteq N \setminus \{k\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{k\}) - f(S) \right], \tag{5.6}$$

where N is the set of all features, S is a subset of features excluding feature k, and $f(S)$ is the model's prediction based only on the features in the subset S. This equation ensures that each feature's contribution is fairly allocated by accounting for all possible combinations of features.

We use SHAP values to evaluate the importance of each input variable in the model and to identify when a particular variable had the most significant influence on the predictions.

We did not perform feature selection before training the models, despite its potential to improve the overall model performance. This decision ensured that no variables were excluded prematurely, allowing the model to consider all socioeconomic, geographical, and contextual factors and their interactions. With this approach, we can identify not only which variables to target for interventions but also the optimal timing for these interventions.

It is important to note that, in this work, we use the traditional SHAP method, which calculates each predictor's contribution by averaging its effect across all possible combinations of features. This approach captures both individual and combined effects, so it does not require predictors to be independent. By considering all feature subsets, SHAP naturally accounts for both independent and correlated effects among predictors [AJL21].

5.5 Results

5.5.1 Model Evaluation

Among the models evaluated, the balanced bagging classifier achieved the highest average ROC AUC of 73.22% ± 4.48%, outperforming the random under-sampling boost classifier (69.50% ± 6.22%) and the easy ensemble classifier (70.59% ± 4.82%). Based on these findings, the balanced bagging classifier was selected for detailed analysis. Performance metrics for the other models' parametrization are provided in Appendix D.2 Table D.3. A grid search was conducted to optimize the balanced bagging classifier's configuration for the baseline window. The best setup included 100 estimators with bootstrapping of features but not samples. Each estimator sampled 50% of the data, and the sampling strategy ensured an equal representation of energy-poor and non-energy-poor instances. Replacement was used in the resampling process.

Table 5.1 presents the performance metrics. Sensitivity reflects the model's ability to identify energy-poor households, specificity measures its ability to identify non-energy-poor households, and ROC AUC assesses overall discrimination performance. The baseline window ($T = 8$) achieved a ROC AUC of 70.01%, with a sensitivity of 73.25% and a specificity of 66.77%.

The results in Table 5.1 reveal how window size influences performance. Shorter windows ($T = 2$) yield balanced sensitivity (70.39%) and specificity (69.23%), with a ROC AUC of 69.81%. At $T = 4$, specificity improves significantly to 73.65%, leading to a higher ROC AUC of 72.24%. The baseline window ($T = 8$) prioritizes sensitivity, achieving the highest value 73.25%, but with a slightly lower specificity. For longer windows ($T = 12$ and $T = 14$), performance varies: $T = 14$ achieves the highest ROC AUC of 72.52% by increasing sensitivity to 78.74%, though specificity stabilizes at 66.31%. On the other hand, $T = 12$ obtains the lowest ROC AUC.

Overall, shorter windows ($T = 2$) favor specificity, while longer windows enhance sensitivity. This can be related to the nature of shorter windows capturing more immediate

Table 5.1: Predictive performance of the balanced bagging classifier model across varying time windows.

Window size (T)	Sensitivity (%)	Specificity (%)	ROC AUC (%)
2	70.39	69.23	69.81
4	70.82	73.65	72.24
8 (baseline)	**73.25**	**66.77**	**70.01**
12	71.60	65.79	68.69
14	78.74	66.31	72.52

Notes: This table highlights the trade-offs between sensitivity, specificity, and ROC AUC. Sensitivity reflects the model's ability to correctly identify energy-poor households, while specificity measures its ability to correctly identify non-energy-poor households. ROC AUC evaluates the model's overall capacity to discriminate between energy-poor and non-energy-poor households across varying decision thresholds. These results were obtained from the evaluation of unseen data (i.e., unseen participants).

and recent information, which tends to reduce false positives and improve specificity. In contrast, longer windows incorporate cumulative historical data, allowing the model to better detect patterns associated with energy poverty over time, which enhances sensitivity by reducing false negatives. The choice of window size thus depends on the specific policy objective, whether it prioritizes minimizing false positives or false negatives.

To assess the robustness of our findings, we also estimated models that included contemporaneous variables. The results, reported in Appendix D.4, confirm that historical predictors retain substantial predictive power even when current conditions are observed. While contemporaneous household income at time T emerged as the single most important feature—accounting for 18.79% of the total normalized importance—lagged income variables at $T-1$, $T-2$, and $T-3$ together contributed over 14%. Moreover, income change variables from previous periods (e.g., $T-1$ through $T-7$) accounted for an additional 8% of model importance.

Table D.2 (Appendix D) presents the predictive performance of the model, as used in the main analysis, across time windows of $T = 2, 4, 8, 12$, and 14 years. As expected, including contemporaneous information improves classification performance across all time windows, as reflected by consistently higher ROC AUC scores. In most cases, sensitivity and specificity increase with the addition of contemporaneous variables. The only exception is the baseline time window ($T = 8$), where sensitivity decreases slightly from 73.25% to 70.86%, while specificity improves substantially from 66.77% to 78.54%. This trade-off still results in a net gain in model performance. For the baseline window, ROC AUC increases from 70.01% to 74.70%, a relative gain of 4.69 percentage points. These results confirm that contemporaneous factors add useful short-term predictive information but do not eliminate the value of historical data.

Figure D.1 (Appendix D) in compares the relative importance of predictors in the combined model (historical and contemporaneous) for the baseline time window ($T = 8$). In the left panel, which aggregates predictors regardless of time, household income accounts for 42.34% of total normalized importance—indicating a dominant role in prediction. However, when temporal structure is preserved (right panel), we see a more nuanced picture: contemporaneous income and income change together account for 27.33% of the total importance, but lagged variables—particularly household income from $T-1$ through $T-5$—also make substantial contributions, exceeding 14% when combined. Historical income change variables across several previous periods (e.g., $T-1$ through $T-7$) add another 8% or more.

Taken together, these results demonstrate that while current conditions provide strong signals, long-term socioeconomic trajectories remain highly relevant and predictive. Most variables contributing at least 1% to total importance are historical, underscoring that energy poverty is shaped not only by present hardship but also by the cumulative effects of past disadvantage. This reinforces the central argument of this paper: leveraging historical information is essential for early identification and proactive intervention.

5.5.2 Main Explanatory Factors and Initial Policy Recommendations

In this section, we take our benchmark model as a reference to examine the relative contribution of historical factors, while the discussion of their directional effects is addressed in the next section. Figure 5.1 shows those factors that contribute at least 1 % to the observed outcome over the entire 8-year time window, ranked by order of importance. Household income emerges as the most critical determinant, contributing 38.84% to the total predictive importance. Notably, changes in household income rank as the second most influential predictor, accounting for 11.29%. Variables with medium explanatory power include the part-time employment rate (7.31%), which underscores the role of labor market dynamics in shaping energy poverty, and household size (6.82%), likely due to the balance between higher energy consumption needs and economies of scale. Energy prices (5.80%) emerge as the fifth predictive factor, and years of education (5.67%) emphasize the interplay between human capital and energy poverty. Lower-contribution factors include poor health (3.09%), employment status (2.78%), macroeconomic indicators such as the unemployment rate (2.53%), Gross State Product (GSP) per capita (2.38%), GSP per capita growth (2.26%), and the total labor force participation rate (2.11%). Finally, demographic and family characteristics such as the number of children at home, age groups, and marital status round out the analysis.

Figure 5.2 breaks down the results from Figure 5.1 across the different time lags. Household income consistently stands out as the most critical predictor, with its impact peaking at $T-1$ (10.85%) and gradually diminishing over longer lag periods ($T-2:$ 8.55%, $T-3:$ 4.50%, $T-4:$ 4.41%, $T-5:$ 4.17%, $T-7:$ 3.99%). Household income changes are also among the top predictors, particularly at $T-1$ (2.55%) and $T-2$ (2.21%). Additional

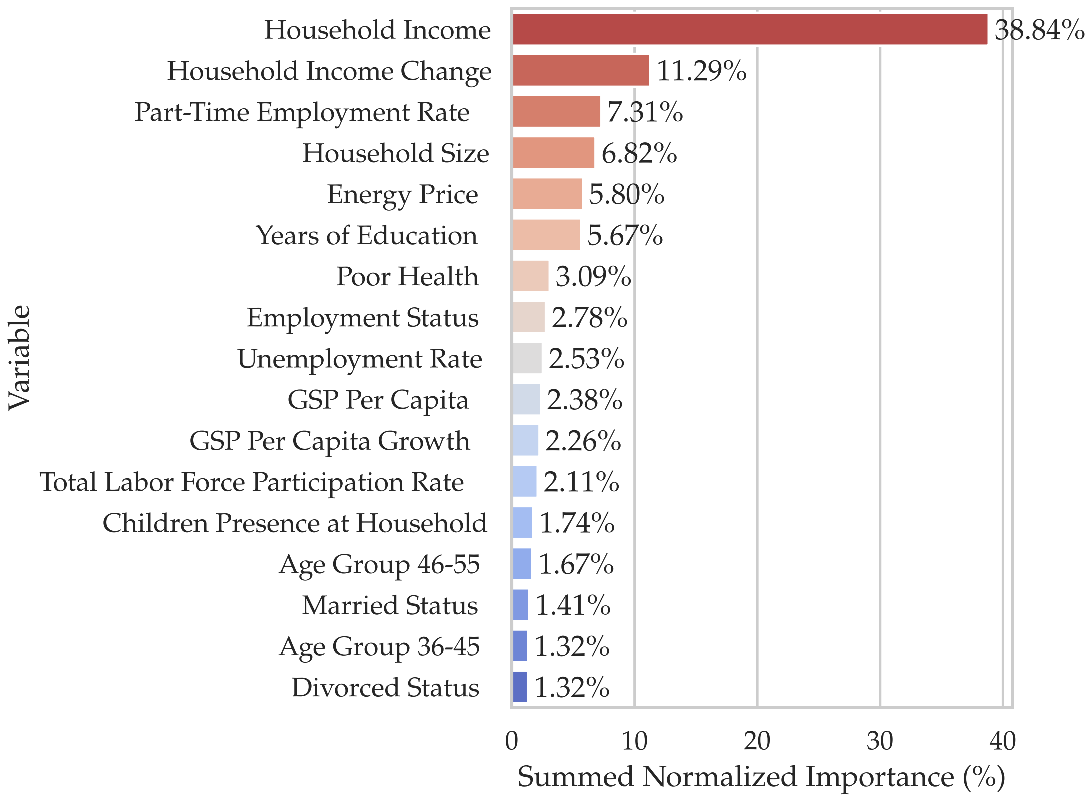


Figure 5.1: Relative contribution (%) of predictive variables for energy poverty outcomes across a 8-year time window.
Notes: i) The figure presents the top predictors with a summed normalized importance of at least 1% for energy poverty outcomes; ii) Source: HILDA 2007–2021 waves.

contributions come from household income changes at $T-3$ and $T-6$, suggesting that historical fluctuations in income continue to influence household energy vulnerability years later. Energy prices operate mainly through a one-year lag, highlighting the effects of short-term fluctuations. Household size at $T-1$ (1.80%) and $T-2$ (1.22%) reflect the immediate impact of family composition on energy poverty. The part-time employment rate also emerges as an important variable, particularly at $T-2$ (1.55%) and $T-6$ (1.50%), pointing to the relevance of the regional employment structure for household energy vulnerability.

Overall, from a policy perspective, the results offer a set of initial insights. First, the strong association between income across all lags and current energy poverty suggests that income can serve as an indicator to identify individuals at risk of energy poverty, even in the long-term. Second, Figure 5.3 focuses on the top 5 contributing variables and their relative contribution over time. The growing importance of household income and income changes toward $T-1$ suggests that policies aimed at stabilizing income in the short term can have a great impact on mitigating immediate energy poverty risks. According to our results, such policies may benefit not only those with low incomes but also individuals with moderate incomes who experience above-average income volatility.

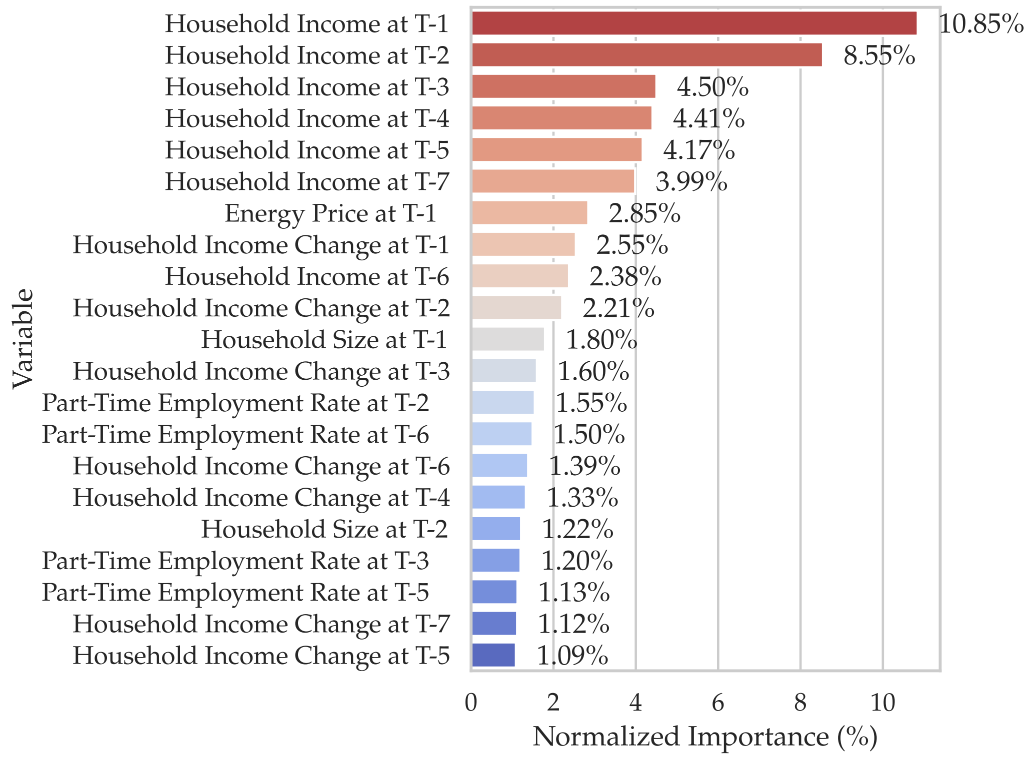


Figure 5.2: Relative contribution (%) of predictive variables for energy poverty outcomes across an 8-year time window—discriminated by period.
Notes: (i) The figure presents the normalized importance of predictive variables for energy poverty outcomes across individual time lags. Only predictors with a normalized importance of at least 1% at any lag are shown; (ii) The suffix "$T - j$" indicates the time lag of the feature relative to the prediction for year T; (iii) Source: HILDA 2007–2021 waves.

Third, the contribution of energy prices to energy poverty rises from $T - 3$ onwards, reflecting the fact that the energy burden over the last 3 years is partly responsible for current energy poverty outcomes. Therefore, price stabilization strategies that extend beyond just one year or rely on occasional interventions could be beneficial for policy. Finally, the contribution of household size also grows steadily over the time window, suggesting that energy poverty is critically influenced by recent adjustments in household arrangements and the changes in energy needs and economies of scale associated with them. In the next section, we identify key household sizes.

5.5.3 How Key Predictive Variables Shape Energy Poverty Outcomes

This section explores how key predictive variables influence their SHAP contributions. A positive SHAP value indicates a higher probability of energy poverty, while a negative value reflects a reduced risk. The results, shown in Figure 5.4, are suggestive of some non-linear relationships. To facilitate interpretation, a fourth-degree least squares polynomial fit was applied to highlight the main trends. As household income increases, the SHAP value decreases sharply. However, this effect is more intense at low and moderate income

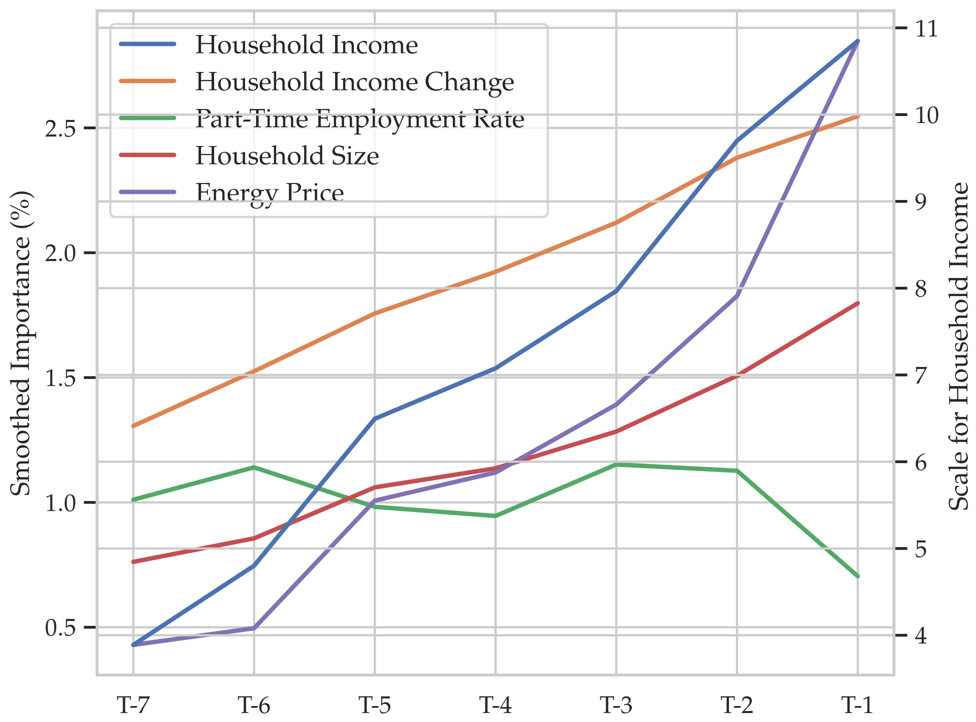


Figure 5.3: Evolution of the relative contribution (%) of the top five predictive variables for energy poverty outcomes across a 8-year time window.
Notes: i) the figure highlights the temporal trends, persistence, and shifts in the influence of different variables across the individual time lags; ii) Source: HILDA 2007–2021 waves.

levels than at high incomes. Similarly, the scatter plot for yearly income variations is suggestive of a somewhat asymmetric effect, with income losses being relatively more relevant for energy poverty than income gains. This pattern reinforces earlier insights that interventions like income insurance, unemployment benefits, or programs aimed at shielding households from income shocks are essential for mitigating these risks.

The part-time employment rate contributes to the energy poverty risk, particularly in areas where the part-time employment rate exceeds 30%. One possible explanation is that part-time jobs reflect labor market and income instability. These positions often lack critical benefits, such as health insurance or retirement plans, which heightens financial vulnerability. Additionally, fluctuating hours and earnings further amplify economic uncertainty. At lower part-time employment rates (below 25%), SHAP values remain relatively stable, indicating a minimal influence. These findings indicate that policies promoting income stability, benefits for part-time workers, and access to full-time employment opportunities are crucial for tackling energy poverty in regions with high part-time employment rates. Additionally, the results in Figure 5.3 reveal that regional labor market dynamics can have delayed impacts on energy poverty, suggesting that such policies could produce lasting effects.

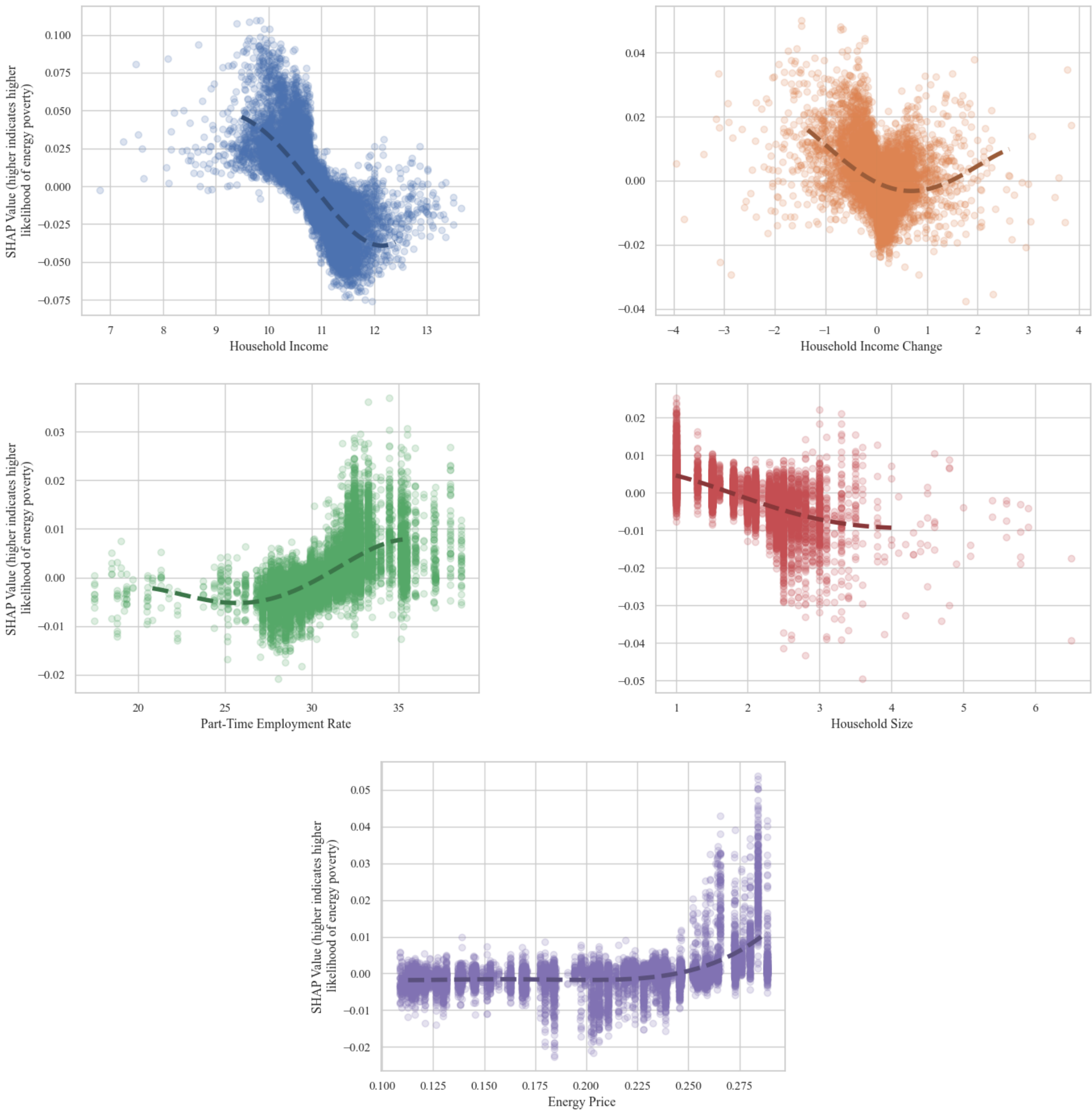


Figure 5.4: Relationship between key predictive variables and SHAP values for energy poverty outcomes.
Notes: i) Each point represents a household, with the x-axis indicating the feature value and the y-axis showing the SHAP value, which reflects the feature's contribution to the forecasting. Positive SHAP values indicate a higher likelihood of energy poverty, while negative values suggest a reduced risk. The plots highlight how changes in the variables influence the model's forecasts; ii) The dashed lines summarize the underlying trends and were calculated using a least squares polynomial fit of degree 4; iii) Source: HILDA 2007–2021 waves.

The relationship between household size and SHAP values highlights a clear risk group: people living alone or in two-person households. This is likely because fixed energy costs pose a disproportionately heavy burden on them. As household size increases to 3–4 members, the likelihood of energy poverty decreases, likely reflecting economies of scale in energy consumption, which reduce the per-capita cost burden.

Lastly, energy prices display a notable pattern, suggesting that below a certain threshold, they are not relevant for energy poverty. At low price levels, SHAP values remain relatively stable, but they increase steadily above \$0.228 and rise significantly beyond \$0.266. This is a relevant finding, as most representations in the literature describe the effect of energy prices on energy poverty in a linear, average manner. However, the results indicate that the relationship between energy prices and energy poverty is non-linear and nearly flat within certain ranges. In this context, energy subsidies and price controls may be ineffective within these ranges, whereas informational campaigns and targeted support for individuals exposed to high prices could play a crucial role.

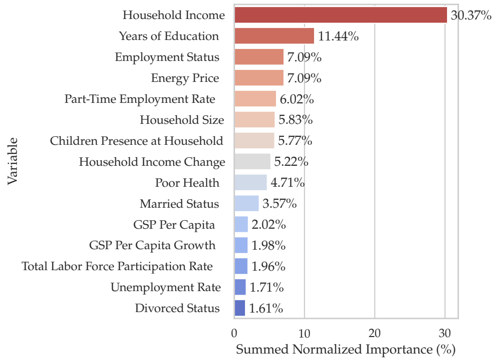
5.6 Sensitivity Checks

In this section, we conduct a series of supplementary analyses. Specifically, we explore the sensitivity of the results to variations in i) the time length considered for the analysis and ii) the chosen cut-off point for defining the energy poverty line. We also examine to what extent our findings might be affected by selection and attrition bias.

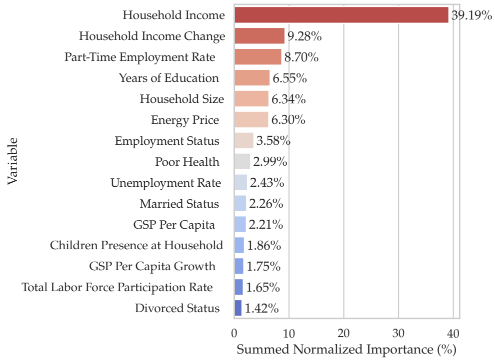
In Figure 5.5 we depict the relative contribution of the predictive variables for alternative time spans. The results show robust consistency across scenarios, with household income emerging as the most significant determinant of energy poverty, irrespective of the time span. Notably, the predictive power of income changes significantly, increasing more than threefold from about 5% when $T = 2$ to over 17% when $T = 12$ or $T = 14$. This suggests that income volatility and the uncertainty it creates are crucial factors influencing long-term energy poverty outcomes.

Additionally, energy prices are relatively important in the short term (6–7%), but their relevance decreases over the long-term (< 4%). Similarly, improvements in education levels are more strongly associated with short-term energy poverty outcomes than with long-term ones. Household size maintains a consistent level of importance across both short- and long-term periods, reinforcing its stable role as a determinant of energy poverty.

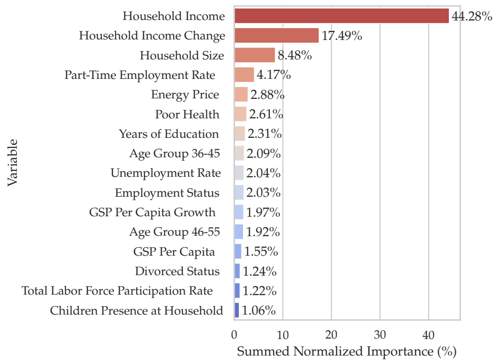
In Figure 5.6 we discriminate across the different time lags. Perhaps the most relevant finding is that household income in the immediate past ($T-1, T-2, T-3$) holds less accumulated relevance for long-term energy poverty outcomes compared to short-term outcomes. This underscores the notion that energy poverty is influenced by a “long memory" process, where the individual’s entire history—albeit with diminishing weight—plays a critical role. Lagged energy prices are in the list of top contributors for energy poverty outcomes at $T = 2$ and $T = 4$. However, they disappear for $T = 14$, suggesting that in



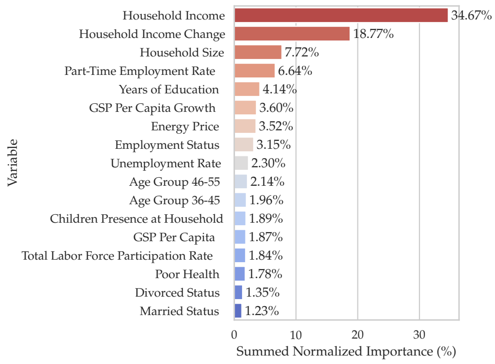
(a) ($T = 2$)



(b) ($T = 4$)



(c) ($T = 12$)



(d) ($T = 14$)

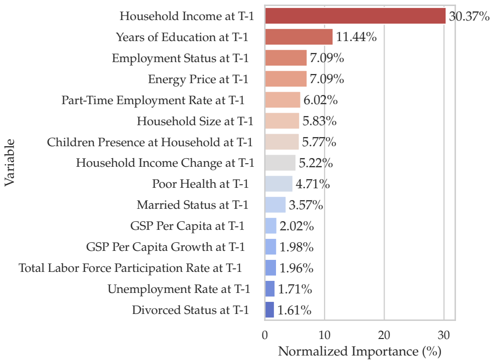
Figure 5.5: Relative contribution (%) of predictive variables for energy poverty outcomes across alternative time windows.
Notes: i) This figure presents the top predictors with a summed normalized importance of at least 1% for energy poverty outcomes. ii) Source: HILDA 2007–2021 waves.

the long-term, the structural aspects of the individual are relatively more relevant than energy prices.

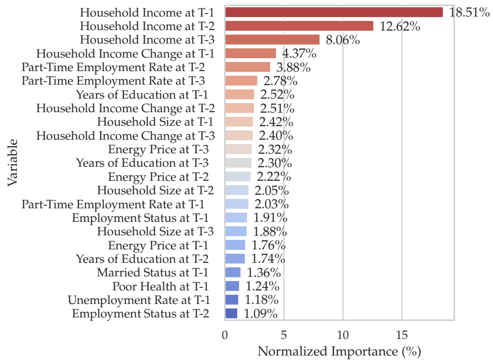
In Figure 5.7 we conduct additional sensitivity checks and present results using more stringent criteria for energy poverty ($\bar{m} = 0.2$ and $\bar{m} = 0.4$). The estimates are based on $T = 8$, as in the baseline estimates. The contribution of income rises from approximately 39% in the baseline estimates ($\bar{m} = 0$) to 50.3% when $\bar{m} = 0.4$. Reversely, the contribution of energy prices falls from 5.8% in the baseline model to 1.5% when $\bar{m} = 0.4$, suggesting that energy prices are not a primary driver of severe energy poverty. Additionally, marriage emerges as a protective factor against stricter definitions of poverty (>2.5%), highlighting its buffering effect in more vulnerable contexts. Finally, Figure 5.8 documents the timing effects, with income in the previous year gaining importance when accounting for the most stringent definition of energy poverty.

Is Attrition Endogenous?

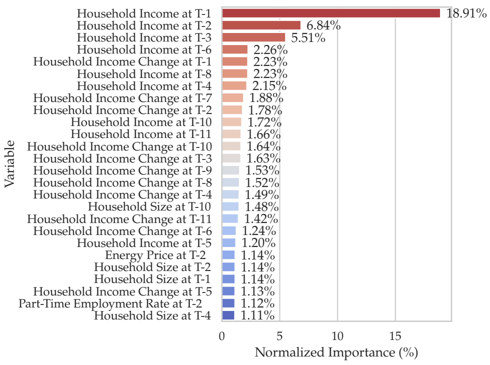
Although the average entry rate (individuals not in the sample in the previous period who are in the current period) and exit rate (individuals who leave the sample) are very



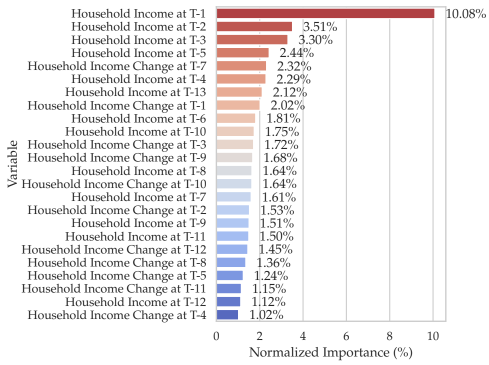
(a) $(T = 2)$



(b) $(T = 4)$



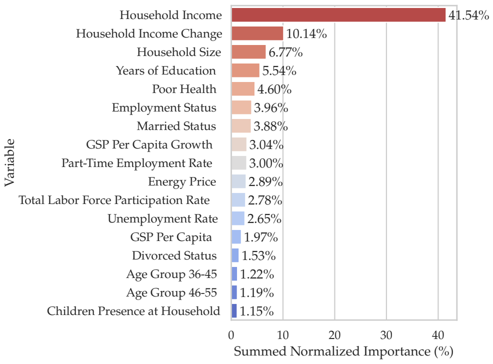
(c) $(T = 12)$



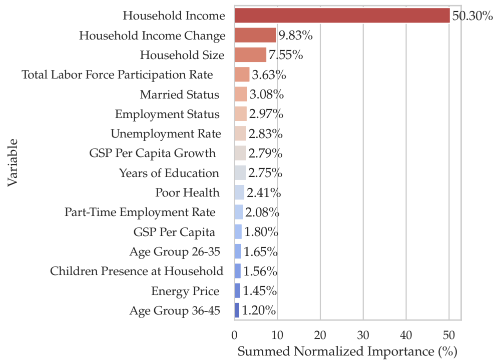
(d) $(T = 14)$

Figure 5.6: Relative contribution (%) of predictive variables for energy poverty outcomes across a T-year time window- discriminated by period.

Notes: i) This figure presents the normalized importance of predictive variables for energy poverty outcomes across individual time lags. Only predictors with a normalized importance of at least 1% at any lag are shown; ii) The suffix "$T - j$" indicates the time lag of the feature relative to the prediction for year T; iii) Source: HILDA 2007–2021 waves.



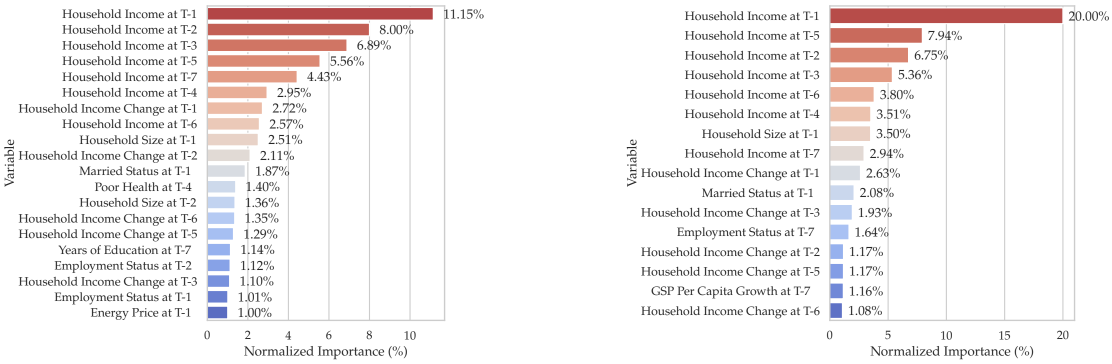
(a) $(\bar{m} = 0.2)$



(b) $(\bar{m} = 0.4)$

Figure 5.7: Relative contribution (%) of predictive variables for energy poverty outcomes across a 8-year time window for different cut-off values.

Notes: This figure presents the normalized contribution of predictive variables for energy poverty outcomes across individual time lags. Only predictors with a normalized importance of at least 1% at any lag are shown. The results are for a $T = 8$ year time window.



(a) ($\bar{m}$ = 0.2) (b) ($\bar{m}$ = 0.4)

Figure 5.8: Relative contribution (%) of predictive variables for energy poverty outcomes across a 8-year time window for different cut-off values - discriminated by period.
Notes: Notes: This figure presents the normalized contribution of predictive variables for energy poverty outcomes across individual time lags. Only predictors with a normalized importance of at least 1% at any lag are shown. The results are for a $T = 8$ year time window.

moderate in our sample (8.9% and 7.4%, respectively), the non-random exit and entry of individuals for reasons related to energy poverty is a potential concern. To address this issue, we conducted a regression using a dummy variable that equals one if the individual exits the sample in the following year and zero otherwise, against energy poverty and all controls and obtained a coefficient equal to -0.003 (p-value = 0.454). In other words, leaving the sample is not significantly related to energy poverty. We proceeded likewise with individuals entering the sample, and energy poverty showed a barely significant negative effect -0.005 (p-value = 0.082). This suggests that the incorporation of new panelists in the sample over the years is not completely random, with a slight tendency to incorporate people who are less likely to suffer energy poverty. These individuals may be either less difficult to contact or more ready to join the panel, although once they decide to participate, their attrition is mostly random.

5.7 Discussion and Conclusions

This study highlights the potential of AI-based methodologies, particularly SHAP, for analyzing the dynamics of energy poverty. It examines the short- and long-term effects of key variables and contextual factors—such as income, energy prices, and regional conditions—on future energy poverty outcomes. By capturing both the timing and magnitude of past events, the study offers a perspective on how these factors shape energy poverty over time. This approach sets our research apart from previous studies, which predominantly rely on static models or contemporaneous relationships between energy poverty and explanatory variables.

The paper shows that current energy poverty is the outcome of historical trajectories. The results are robust to a battery of sensitivity checks, including alternative definitions of multidimensional energy poverty and varying time spans. Income levels emerge as the most critical factor, particularly for long-term outcomes and under strict definitions of poverty. While the contemporaneous relation between income and energy poverty has been highlighted in previous work [DSZ21; HDZ22], our results uncover the association between income across all lags and current energy poverty. From a policy design perspective, we provide evidence that income can serve as an effective screening tool for identifying “future" energy-poor individuals—those at risk of becoming energy-poor in the years ahead. Moreover, our results emphasize the role of income *changes*. Historical income fluctuations have lasting effects on household energy vulnerability, persisting over time. This insight introduces a new dimension to combating energy poverty, showing that beyond income levels, individuals experiencing income volatility and uncertainty constitute a high-risk group.

Consistent with numerous studies, we find a positive association between energy prices and energy poverty [SRL23]. However, our study adds that energy prices have a significant impact in the short term and under less stringent definitions of poverty. In contrast, their influence diminishes when addressing long-term energy poverty or more severe cases. Furthermore, our findings suggest the existence of a price threshold beyond which energy prices become particularly detrimental. In this context, measures such as energy subsidies, price controls, and informational campaigns specifically aimed at individuals facing high energy prices could play a critical role.

The AI-approach used in the paper provides insights that may be used to shift the policy focus from reactive measures, which address existing poverty, to preventive strategies that target households showing early signs of vulnerability. Specifically, our findings suggest that policymakers can enhance resilience and reduce long-term socioeconomic disparities by balancing immediate relief measures—such as energy price support, energy benefits, and income transfers—with structural policies addressing systemic vulnerabilities identified in our study, particularly income volatility, labour market conditions and small households.

This exploratory study has several limitations that warrant further investigation. A key shortcoming is the failure to account for the endogeneity of life events, such as income shocks, which may be driven by unobserved behaviors, situational factors, or omitted variables. Addressing these issues in future research is essential and could involve incorporating more advanced econometric and AI techniques to ensure a better understanding of the mechanisms at work. Another limitation is the uniform treatment of households, which overlooks heterogeneity in responses to energy poverty predictors. In other words, this approach implicitly assumes that all households above the threshold experience energy poverty in a similar way. However, in practice, households may face very different forms of deprivation. Factors such as income, age, education, and personal traits likely influence how individuals experience and respond to energy challenges [Con+22]. Future research could improve the granularity and relevance of conclusions by conducting separate analyses based on these dimensions or within socio-economic groups or household clusters. Finally, this study is focused on a single dataset, the HILDA Survey. While this dataset provides rich, longitudinal information about Australian households, testing the methodology on additional datasets from other regions and contexts with varying degrees of energy poverty would help assess the generalizability of the findings and derive equally meaningful insights for other jurisdictions.

CHAPTER 6

A Counterfactual Approach to Energy Poverty Mitigation: A Case Study For Australia (Preliminary Report)

Diogo Nuno Freitas, Santiago Budría, Eduardo Fermé

Abstract. *Energy poverty is a persistent global issue where households lack access to adequate energy services. Research often focuses on contemporaneous factors, overlooking the predictive power of long-term socioeconomic trajectories. This study addresses two questions: How do past socioeconomic conditions affect future energy poverty? And what are the minimal interventions that could prevent a household from becoming energy-poor? We aim to shift the focus from reactive mitigation to proactive prevention by developing a framework that forecasts risk and identifies actionable pathways to avoid it. Using seven years of Australian longitudinal data, we train a machine learning classifier to predict energy poverty in the following year. We then apply counterfactual analysis to identify minimal, interpretable changes that alter the predicted outcome. The model successfully predicts future energy poverty with a ROC AUC of 70.01%. The counterfactual analysis consistently reveals that modest increases in household income, often less than 5%, are the most effective single intervention. Other factors, such as decreases in energy prices and reductions in unemployment, also contribute to preventing energy poverty, often in combination with income gains. "What-if" scenarios suggest that external shocks, like a sudden rise in energy prices or job loss, can be offset by small, timely income adjustments. The effectiveness of these changes, whether recent or past, highlights the importance of long-term financial stability. Also, it enables proactive policy by identifying at-risk households a year in advance and suggesting targeted interventions as a more efficient alternative to broad measures.*

6.1 Introduction

Energy poverty remains a pressing global challenge, affecting the quality of life for millions of people. It refers to a situation where households cannot access or afford adequate, reliable, and clean energy services for their daily needs. Despite international efforts to address this issue, approximately 750 million people still do not have access to electricity, and over 2 billion lack access to clean cooking fuels [Int24].

The consequences of energy poverty extend far beyond basic comfort. Limited access to energy reduces educational opportunities, restricts access to jobs [Sha+24a], and negatively affects well-being [NL24] and health [Pon+24]. For instance, Nawaz [Naw21] showed that households facing energy poverty are 9 to 13% more likely to experience poor health compared to those who are not energy-poor. Poor indoor conditions, such as a lack of adequate heating, have a considerable impact on both physical and mental health. Long periods of low indoor temperatures have been linked to higher rates of illness and death [Bal+22; DM11], as well as worsening mental health issues [Ben+23]. Vulnerable groups, including infants, are especially at risk for respiratory diseases [Moh21; SFL24]. There is also evidence that energy poverty can increase the risk of developing metabolic disorders like diabetes [ST23].

Many factors contribute to energy poverty. These include rising energy prices, economic instability, and the presence of inefficient or poor-quality housing [PD15]. Research has extensively examined the socioeconomic aspects of energy poverty, with the aim of identifying households at risk and designing targeted policy solutions [ABR22; FFT22; KA22]. However, most studies focus mainly on contemporaneous socioeconomic conditions and assume that only current variables are needed to explain energy poverty. As a result, there is limited understanding of the long-term impact that past socioeconomic factors have on future energy poverty.

Besides identifying households at risk, it is equally important to consider what actions can help prevent energy poverty. In other words, if a household is predicted to become energy-poor, what actions could be taken, or could have been taken, to prevent this outcome? Ideally, these actions should require minimal intervention but provide the greatest benefit in terms of prevention.

Our work addresses these two gaps in the existing research: First (Contribution 1), we explore how past factors, such as income, energy prices, regional circumstances, and other socioeconomic variables, affect *future* energy poverty outcomes. We apply a machine learning algorithm that uses household-level data from seven previous years to predict energy poverty status in the following year. Second (Contribution 2), we use counterfactual analysis on the model developed in Contribution 1 to identify specific, minimal, and human-interpretable changes that could help a household shift from being energy-poor to non-energy-poor. These changes might result from personal actions or targeted policies. We refer to these as *chains of minimal changes*. We also present "what-if" scenarios that identify shocks which increase poverty risk and the minimal counterfactual response required to mitigate that risk.

Contribution 1 and Contribution 2 are complementary, and allow us to move from reacting to energy poverty after it occurs to taking proactive measures to prevent it. Our approach identifies households at risk one year in advance and provides practical suggestions for actions that require the least effort but have the greatest impact on preventing energy poverty. In this way, we bridge the gap between prediction and real-life intervention, providing practical guidance for policymakers, community organizations, and individuals seeking to reduce the long-term effects of energy poverty.

To conduct our analysis, we use the HILDA Survey[1], which is a nationally representative, long-term study that tracks the same individuals and households in Australia from 2007 to 2021. The survey collects data on income, jobs, education, health, and family life, providing a strong foundation to study how energy poverty changes over time. For measuring "energy-poverty", we employ the MEPI [NBM12], which combines both objective (expenditure-based) and subjective (self-assessed) indicators of energy poverty into a single, comprehensive metric.

6.2 Literature Review

Energy poverty can be defined as a household's inability to afford or access energy services needed to support adequate living conditions and human development. While conceptual definitions of energy poverty have been the subject of extensive discussion in the literature (for an overview see Sy and Mokaddem [SM22]), the focus has generally been on the inability of households to afford and have access to adequate energy services.

Based on the literature, energy poverty is a complex phenomenon stemming from a wide range of factors. These include macroeconomic conditions like GDP, governance, and a country's energy mix [BMM21; BL24; Boţ+24; IM22; KUO23; KR25], as well as household-level characteristics such as income, dwelling type, and size [ABR22; KŚ20; MM24]. Individual attributes like educational attainment, health status, age, and employment also play a significant role [Abb+20; CAF19; FFT22; Pon+24; Sha+24a; ZAK21; Zha+21]. Furthermore, factors like spatial disparities, cultural behaviors, and energy subsidies add to the complexity [Hos+23; PAS22; Uma+24].

Given the multi-layered nature of these determinants, a recent body of literature has introduced machine learning techniques to predict energy poverty outcomes. Evidence based on a XGBoost framework to predict the risk of experiencing energy poverty in the Netherlands identifies income, house value, and house ownership as the main drivers of energy poverty [DSZ21]. In a similar setting, and based on 11 European countries, income, household size, and floor area were consistent predictors [HDZ22]. Evidence based on an

[1] https://melbourneinstitute.unimelb.edu.au/hilda

RF classifier across the European Union uncovers household- and country-level predictors like dwelling conditions, energy efficiency, and gas supplier switching rates [SRL23].

While the previous studies are based on a single energy poverty indicator, other studies define a multidimensional energy poverty index similar to ours [NBM12]. These studies have shown that in Asian and African countries, wealth, marital status, and residence attributes are significant predictors of poverty [Abb+20]. Recent research has further advanced these methodologies by employing ensemble models, such as XGBoost, combined with RF and ANN, revealing the critical importance of education and food security indicators in determining energy poverty [GJA24].

6.3 Materials and Methods

We use the HILDA Survey, a comprehensive, nationally representative longitudinal study that examines the economic, social, and demographic dynamics of Australian households. Initiated in 2001 and conducted annually, it tracks individuals and households over time, providing important information on income, labor market activities, health, education, and family relationships, among other factors. The original 2001 sample included approximately 7,600 households and 13,000 individuals, with periodic updates to account for attrition.

The variables used to model energy poverty are described in Appendix D.1, and include labor market indicators (such as part-time employment rate, unemployment rate, and labor force participation), economic measures (GSP per capita and energy prices), and household characteristics (income, household size, and region). We also include individual information such as age, years of education, marital status, employment, health, and the presence of children in the household.

In this study, we approach energy poverty as a classification problem. Let $\mathcal{X} \subseteq \mathbb{R}^d$ be the feature space, and denote an element (feature vector) by $\mathbf{x} \in \mathcal{X}$. We employ a machine learning model $f : \mathcal{X} \to [0, 1]$, where $f(\mathbf{x})$ predicts the likelihood that a household, characterized by the features $\mathbf{x}$ from the past seven years, will experience energy poverty in the upcoming eighth year. This prediction relies on the MEPI indicator, which serves as our dependent variable.

Outcome—The MEPI Indicator: The MEPI variable captures both expenditure-based and subjective dimensions. The expenditure-based measures include the 2M, TPR, and LIHC indicators. The subjective dimension is reflected by two self-assessed indicators: the household's inability to pay for heating due to a lack of money (Heat), and the inability to pay electricity, gas, or telephone bills on time (Arrears).

The MEPI index is calculated as follows: Let $J = \{1, \ldots, m\}$ be the set of $m = 5$ poverty indicators. Let $\mathcal{I}$ be a set of individuals, with element i, $i \in \mathcal{I}$, and $\mathcal{T}$ be a set of time periods, $t \in \mathcal{T}$, representing a specific moment when the survey was conducted. Let EP_{ijt} denote the status of the i-th individual in the j-th indicator during period t. If an individual i is poor under indicator j in the period t, then EP_{ijt} takes the value of one, and zero otherwise. Following the family of indexes typically described in the literature on material deprivation [DPX19], individual i's weighted poverty score is given by:

$$\mathrm{MEPI}_{it} = \sum_{j \in J} w_j \mathrm{EP}_{ijt}, \quad \forall\, i \in \mathcal{I},\ t \in T_i,\ T_i \subseteq \mathcal{T}, \tag{6.1}$$

where w_j denotes the weight assigned to the poverty indicator j, with $\sum_{j \in J} w_j = 1$. Hence, the MEPI_{it} indicator ranges from 0 to 1 and captures the percentage of dimensions in which the individual is deprived.

Although it is common to give the same importance (w) to each indicator, we focus more on the indicators where deprivation is rare. This method is known as the Frequency-Based Weighting Approach [DL13]. The weight given to an indicator is proportional to the percentage of individuals *not* classified as poor under that specific indicator within a particular state. In other words,

$$w_j = \frac{(1 - n_j)}{\sum_{j \in J}(1 - n_j)}, \tag{6.2}$$

where n_j is the proportion of poor individuals in dimension j. This choice is based on the belief that lacking access to everyday items should be considered a more significant indicator of deprivation than lacking access to less common items. The weights are calculated separately for each wave.

In the context of this work, an individual i is regarded as energy poor if $\mathrm{MEPI}_{it} > 0$ [NBM12]. This means that a household is considered "energy-poor" if it is deprived in any one of the five dimensions. Therefore, the variable of interest in this study, which indicates whether a person is energy-poor, is binary: it is 1 if the person is energy-poor, and 0 if not.

Variables and Data Preparation: We model energy poverty at time $T = 8$ as a function of socioeconomic and demographic characteristics observed in periods $T - 1, \ldots, T - 7$. We exclude the period T to ensure that our forecasting relies entirely on historical data. Our focus is on the role of past factors. Including contemporaneous variables could potentially mask the effects of lagged factors, especially if there is autocorrelation in the data. More importantly, the inclusion of contemporaneous variables may introduce reverse causality between energy poverty and socio-demographic characteristics such as health and education [PK19; Pon+24]. By considering only past variables, we eliminate the risk that current energy poverty influences these characteristics.

To capture the temporal dynamics of the variables, we created lagged features, which serve as the input to the predictive models. Generically, for each original feature, we obtained new features representing its values from each of the previous years.

We then split the dataset into training, validation, and test subsets to facilitate model development and evaluation. Out of the 7,977 participants in our dataset, 6,382 (80%) were randomly selected for training and validating the predictive models, while the remaining 1,595 participants (20%) were included in the test set. The test set was held out and used exclusively to evaluate the final performance of the models, providing a fair estimate of their forecasting accuracy. Moreover, the dataset is split by individual to prevent data leakage.

Before training the model, we standardized the data to ensure consistency and reliability in our modeling process. The standardization parameters were estimated solely from the training set to avoid information leakage. Specifically, we removed the median and scaled the data using the interquartile range, as described in [SWW21]. These parameters were then applied to transform the training, validation, and test data.

Model Development: We treat the energy poverty forecasting task as a classification problem. Specifically, households are classified as energy-poor depending on whether their MEPI is greater than 0 (cut-off point). To model the relationship between the socioeconomic and demographic factors and the MEPI indicator, we used a balanced bagging classifier.

A balanced bagging classifier [Gal+12] is an ensemble technique that combines the predictions of multiple base models, in our case, decision trees, to improve the robustness and accuracy of the outcomes. In order to further refine the modeling approach, we implemented the classifier in an OvR binary classification framework [Mur12].

We optimized the hyperparameters of our classifier using a grid search. For details on the specific hyperparameters and grid configurations, see Appendix D.2. We employed 5-fold cross-validation on the training dataset to ensure the robustness of the hyperparameters across different data splits, selecting the best set based on the highest ROC AUC score.

The final model was trained on the complete training set using the identified optimal hyperparameters and subsequently evaluated on a held-out test set of 1,595 participants.

In addition to the balanced bagging classifier, we tested two other class-imbalance ensembles. Namely, we benchmarked the random under-sampling boosting [Sei+10] and the easy ensemble [LWZ09]. All models used the same feature set, preprocessing pipeline, and 5-fold cross-validation on the training partition, with hyperparameters tuned via grid search to maximize ROC AUC. However, the balanced bagging classifier was the model that achieved the highest ROC AUC. Therefore, we chose the balanced bagging classifier as our final model.

Counterfactual Explanations: Counterfactuals are a post-hoc means to understand and explain the model predictions [Ver+24]. In the context of this work, counterfactuals are used to generate alternative household profiles, and in combination with the machine model f, help determine if one change (e.g., an increase in household income) or a set of changes in the user profile can increase or decrease the likelihood of energy poverty.

Additionally, they are also used to identifying the minimum response needed to mitigate the risk of becoming energy poor after a shock that increases poverty.

Our focus is on *minimal* changes in order to ensure that the counterfactual recommendations remain practical and actionable, allowing individuals to make small adjustments that could have a meaningful impact on their future energy poverty risk. The minimal changes align with the concept of *proximity*, where the suggested profile modifications are as close as possible to the individual's original state. Besides proximity, another relevant aspect is *diversity*, meaning that multiple plausible pathways out of energy poverty profiles should be considered.

In this work, counterfactual explanations are formulated as a constrained optimization problem, as proposed by Mothilal, Sharma, and Tan [MST20]. Given an individual feature vector $\mathbf{x}$ with $f(\mathbf{x}) = 1$ (classified as energy poor at $T = 8$), we search for a new feature vector $\mathbf{x}'$ such that $f(\mathbf{x}') = 0$. The goal is to find $\mathbf{x}'$ that is as close as possible to $\mathbf{x}$, so only minimal and realistic changes are required.

Formally, if k is the total number of counterfactual examples to be generated, then the set of counterfactuals $\{\mathbf{x}'_i\}_{i=1}^k$ is obtained by solving the following optimization problem:

$$\begin{aligned}\{\mathbf{x}'_i\}_{i=1}^k = \arg\min_{\mathbf{x}'_i} \Bigg\{ & \frac{\lambda}{k} \sum_{i=1}^{k} \sum_{j \in \mathcal{F}} m_j w_j \cdot d_j(x_j, x'_{i,j}) \\ & + \frac{1}{k} \sum_{i=1}^{k} \ell(f(\mathbf{x}'_i), 0) - \gamma D(\{\mathbf{x}'_i\}_{i=1}^k) \Bigg\}\end{aligned}, \tag{6.3}$$

where the first term measures the weighted distance between the original instance $\mathbf{x}$ and each counterfactual $\mathbf{x}'_i$. The per-feature distance $d_j(x_{i,j}, x'_{i,j})$ measures the change for feature $j \in \mathcal{F}$. The term m_j is a binary mask, where $m_j = 1$ if feature j is mutable and $m_j = 0$ otherwise. The weight w_j reflects the difficulty of changing feature j and is defined as the inverse of the median absolute deviation (MAD).

The second term is a hinge-style loss on the logit of the predicted probability:

$$\ell\big(f(\mathbf{x}'_i), 0\big) = \max\!\big(0,\ 1 + \operatorname{logit}\big(f(\mathbf{x}'_i)\big)\big), \quad \text{where } \operatorname{logit}(p) = \ln \frac{p}{1-p}. \tag{6.4}$$

This penalizes any counterfactual $\mathbf{x}'_i$ whose predicted probability $f(\mathbf{x}'_i)$ is not confidently below the decision boundary for class 0.

The third term promotes diversity among the generated counterfactuals by maximizing their pairwise distances. To avoid overloading notation, we distinguish between a per-feature distance $d_j(\cdot,\cdot)$ and a vector-level distance $\Delta(\cdot,\cdot)$:

$$D\big(\{\mathbf{x}'_i\}_{i=1}^k\big) = \sum_{i=1}^{k} \sum_{a=i+1}^{k} \Delta(\mathbf{x}'_i, \mathbf{x}'_a), \tag{6.5}$$

with

$$\Delta(\mathbf{x}, \mathbf{x}') = \sum_{j \in \mathcal{F}} d_j(x_j, x'_j), \tag{6.6}$$

and, following Mothilal, Sharma, and Tan [MST20], the per-feature distance is

$$d_j(u,v) = \begin{cases} |u-v|, & j \in \mathcal{F}_{\text{cont.}}, \\ \mathbb{1}[u \neq v], & j \in \mathcal{F}_{\text{cat.}} \end{cases}. \tag{6.7}$$

That is, d_j corresponds to the L^1 change for continuous features, while categorical feature modifications are penalized uniformly to discourage unnecessary changes [MST20].

The parameters λ and γ control the trade-off between proximity, classification change, and diversity. In our analysis, we set $\lambda = 5$ and $\gamma = 2.5$. We used a genetic algorithm to solve Equation (6.3) and generated $k = 50$ counterfactuals for each of the prototype profiles described below.

In our analysis, we set $\lambda = 5$ and $\gamma = 2.5$, based on preliminary tests with alternative parameter values. These tests indicated that a lower λ resulted in counterfactuals needing larger changes in features, making them unrealistic for households. On the other hand, higher λ values limited proximity too much. Similarly, varying γ showed that values that were too small limited diversity, and values that were too large reduced interpretability. The selected values offered the best trade-off between minimal, realistic changes and sufficient diversity.

To encode immutability explicitly, let $\mathcal{M} \subseteq \mathcal{F}$ denote the set of mutable features and $\mathcal{I} = \mathcal{F} \setminus \mathcal{M}$ the set of immutable features. We impose the following constraints, for all $i = 1, \ldots, k$:

$$x'_{i,j} = x_j, \qquad \forall\, j \in \mathcal{I}. \tag{6.8}$$

To ensure realism and actionability, we also restrict changes in continuous mutable features to lie within a ±5% band around their original values. Let $\rho = 0.05$. Then, for all $i = 1, \ldots, k$ and all $j \in \mathcal{F}_{\text{cont.}} \cap \mathcal{M}$,

$$(1-\rho)\, x_j \;\leq\; x'_{i,j} \;\leq\; (1+\rho)\, x_j. \tag{6.9}$$

We note here that counterfactual explanations are closely related to, and often considered a form of, contrastive explanations [Ste+21]. The literature, however, draws a subtle distinction between the two. Contrastive explanations primarily identify the features responsible for a classification (e.g., “The household is energy-poor *because of* its low income”) [Dhu+18], whereas counterfactuals identify the minimal changes that would alter the outcome (e.g., “The household would *not be* energy-poor *if* its income were increased by X”) [Ver+24]. As our work focuses on providing actionable recourse by showing *what must change* to mitigate energy poverty, we have adopted the term “counterfactual explanation” as it most precisely describes our objective.

Prototype Profiles: To obtain a small set of representative borderline cases, we first define borderline profiles as those for which the model predicts class 1 with low confidence. Specifically, when the predicted probability $f(\mathbf{x})$ lies in the range $[0.5, 0.5+\varepsilon]$, with $\varepsilon = 0.05$. That is, the borderline sample set is defined as

$$B = \left\{ \mathbf{x}_i \,\middle|\, f(\mathbf{x}_i) \in [0.5,\, 0.55] \text{ and } y_i, \hat{y}_i = 1 \right\}, \tag{6.10}$$

where $\hat{y}_i = \arg\max_y f(\mathbf{x}_i)$ denotes the predicted class label.

We then compute the empirical mean of the set $B = \{\mathbf{x}_i\}_{i=1}^n$ as

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i. \tag{6.11}$$

The first prototype p_1 is chosen as the observation in B closest to this mean, i.e.,

$$p_1 = \arg\min_{\mathbf{x}\in B} \|\mathbf{x} - \overline{\mathbf{x}}\|_2 . \tag{6.12}$$

Subsequent prototypes $\{p_j\}_{j=2}^m$ are selected using the farthest-first traversal rule:

$$p_j = \arg\max_{\mathbf{x}\in B}\left[\min_{1\le L<j}\left\|\mathbf{x} - p_L\right\|_2\right], \quad j = 2, \ldots, m, \tag{6.13}$$

ensuring that each new prototype maximizes its minimum Euclidean distance to the already-selected set. This procedure yields m diverse, actual observations from the borderline region, with p_1 capturing the central tendency and the remaining $\{p_j\}_{j\ge 2}$ spanning the range of variability in feature space.

In our application, we selected $m = 3$ prototypes.

6.4 Results

Predictive Model (Contribution 1): A grid search was conducted to optimize the balanced bagging classifier's configuration for $T = 8$. The best setup included 100 estimators with bootstrapping of features but not samples. Each estimator sampled 50% of the data, and the sampling strategy ensured an equal representation of energy-poor and non-energy-poor instances. Replacement was used in the resampling process.

Our approach predicts energy poverty status in the eighth year for each household, using information from the previous seven years as input features. The model achieved a ROC AUC of 70.01%, which indicates that the model can discriminate between energy-poor and non-energy-poor households across varying decision thresholds. For class-specific metrics, sensitivity was 73.25%, meaning that the model correctly identified most energy-poor cases. The specificity of 66.77% indicates an acceptable rate of correctly classifying non-energy-poor households.

These results show that the model favors sensitivity. This is useful for policy design, where identifying the most energy-poor households is more important than avoiding the misclassification of some non-energy-poor households. High sensitivity supports early interventions from a prevention perspective, while reasonable specificity keeps misclassifications within an acceptable range.

Overall, our results support the use of balanced ensemble methods with appropriate sampling and parameter selection to model energy poverty using historical data.

Counterfactuals (Contribution 2): To interpret the predictive model and develop actionable insights, we conducted a counterfactual analysis on three representative prototype profiles, as described in Section 6.3. Each prototype represents a borderline energy-poor household, where the model predicts energy poverty with relatively low confidence. It is important to note that, for interpretability, all features were transformed back to their original scale by applying the inverse of the preprocessing transformation. Moreover, the immutable features include demographic attributes such as age and age group classifications, marital status (including being married, divorced, or widowed), the presence of a disability, the presence of children in the household, and the household size.

For each prototype, we generated 50 diverse counterfactual examples that change the predicted class from energy-poor to non-energy-poor (i.e., from class 1 to class 0), while requiring only minimal and realistic changes to the input features. All generated counterfactual profiles satisfy the classification constraint $f(\mathbf{x}') = 0$.

We examined the full set of generated counterfactuals and analyzed the types of changes required to shift a household's classification. Among these, we identified a subset of counterfactuals that involved only a single actionable modification. In all such one-action counterfactuals, the feature that was changed was household income. That is, all one-action classification changes across the three prototypes were achieved solely through an increase in income. No other individual feature was sufficient to produce a change in predicted energy poverty status when modified in isolation.

The timing of these interventions varied across different counterfactuals. Successful income changes tended to occur in more recent years, most frequently in $T = 6$ and $T = 7$, but also appeared in earlier years, including $T = 1$ and $T = 2$. This indicates that while recent changes in income have strong predictive power, improvements in earlier years can also contribute meaningfully to reducing energy poverty risk. The scale of required income changes was generally modest, but varied slightly depending on the year. In year $T = 1$, the required increase ranged from 3.18% to 3.81%; in $T = 2$, from 2.76% to 4.03%; in $T = 6$, from 1.83% to 4.93%; and in $T = 7$, from 1.30% to 4.82%. These ranges show that effective interventions can occur across multiple time points and typically demand income increases of less than 5%. This suggests that relatively small financial improvements, whether recent or distributed over a longer time horizon, can be sufficient to prevent energy poverty for households at the margin.

Figure 6.1 shows several connected counterfactual (with more than one action) pathways that lead from an energy-poor to a non-energy-poor household profile at time $T = 8$. Each path represents a sequence of minimal changes to selected features in earlier years. The figure illustrates the range of time steps and features through which this transition can be achieved.

Across all paths, an increase in household income appears as a consistent component. Income changes are present at multiple time points, including $T = 1$, $T = 2$, $T = 6$, and $T = 7$. The required increases range from 2.26% to 4.36%, confirming earlier results that show small changes in income are effective in changing the classification of households that are near the energy poverty threshold.

In addition to income, other features involved in the transitions include energy price, unemployment rate, and part-time employment rate. Decreases in energy prices at $T = 1$ and $T = 3$, and reductions in unemployment at $T = 4$ and $T = 7$, contribute to multiple successful pathways. Several paths combine income changes with labor market improvements, such as a decrease in the part-time employment rate at $T = 6$ or $T = 7$.

The timing of the changes is also significant. While most changes occur in the final time step ($T = 7$), effective changes in earlier years, especially $T \leq 4$, also appear. This suggests, similar to the one-action counterfactuals, that both recent and earlier interventions can influence future energy poverty status.

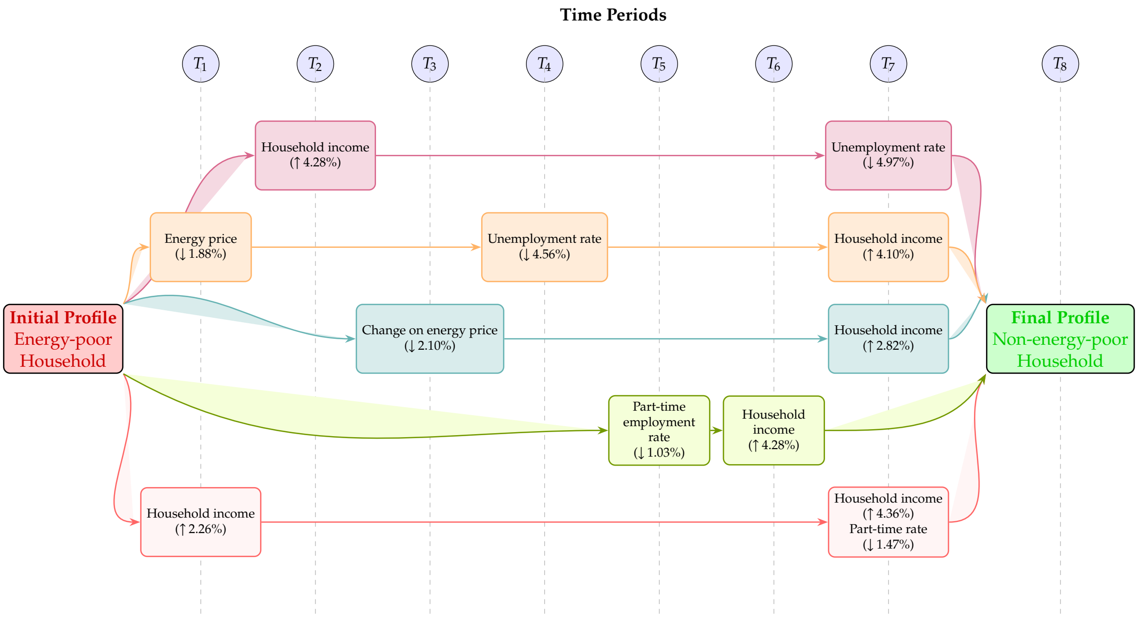


Figure 6.1: Connected pathways showing updated cases for transitioning from energy poor to non-energy-poor profile, with specific feature changes at each time step.

Shocks (Contribution 2): To further explore how households may avoid energy poverty under different conditions, we present in Table 6.1 a set of scenarios based on selected multi-feature counterfactuals. These scenarios describe situations in which a change in a macroeconomic or household-level factor (in other words, a shock) increases the risk of energy poverty. For each situation, we identify the minimal change that the model associates with successfully avoiding this outcome.

In one scenario, energy prices rise by 3.03% in year $T = 4$. The model finds that a compensating income increase of 2.92% in year $T = 7$ is required to prevent the household from entering energy poverty. Similarly, when energy prices increase by 3.16% in year $T = 6$, a larger income adjustment of 4.82% in $T = 7$ is needed to offset the rising cost burden.

Table 6.1: Counterfactual Scenarios for Avoiding Energy Poverty. The table shows the initial shock that increases poverty risk and the minimal counterfactual response required to mitigate that risk, as identified by the model.

	Shock (Risk Increase)				Required Response (Mitigation)		
Scenario	**Factor**	**Change**	**Time**		**Factor**	**Change**	**Time**
1	Energy Price Increase	3.03%	$T = 4$	→	Income Increase	2.92%	$T = 7$
2		3.16%	$T = 6$			4.82%	$T = 7$
3	Labor Force Drop	–2.15%	$T = 5$	→	Income Increase	3.23%	$T = 7$
4	Energy Price Volatility	3.46%	$T = 2$	→	Income Increase	3.73%	$T = 7$
5	Becomes Unemployed	—	$T = 2$	→	Income Increase	1.83%	$T = 6$

Changes in labor force participation also play a role. When the total labor force participation rate drops by 2.15% in year $T = 5$, the model identifies a 3.23% income increase in $T = 7$ as sufficient to prevent energy poverty. Shifts in energy cost volatility show a similar pattern. When the energy price change rate increases by 3.46% in year $T = 2$, future income must increase by 3.73% in $T = 7$ to avoid a negative classification.

A final scenario involves a complete loss of employment, with the individual becoming fully unemployed in year $T = 2$. In this case, a modest income increase of 1.83% in year $T = 6$ is sufficient to counterbalance the risk introduced by the unemployment shock.

These scenarios show that while the causes may differ, income remains the most frequent feature requiring adjustment. Small income increases, aligned with observed external conditions, are often enough to reduce the likelihood of energy poverty according to the learned model.

6.5 Discussion

Our study suggests a novel approach to energy poverty prevention by combining a machine learning algorithm with counterfactual analysis. We used a machine learning algorithm to predict energy poverty one year in advance and identify minimal-effort interventions. With this, our work shifts the focus from reactive mitigation to proactive prevention.

The primary contribution of this work is a framework that not only identifies households at risk of future energy poverty but also provides actionable, human-interpretable pathways to avoid it. Our model achieved a ROC AUC of 70.01%, and can therefore distinguish between future energy-poor and non-energy-poor households using only historical data. The model's sensitivity (73.25%) is significant from a policy perspective as it focuses on accurately identifying households at risk, which is essential for successful early intervention. This provides the Contribution 1 of this work.

Our counterfactual analysis highlights the central role of household income. In all the prototype profiles, a modest rise in household income was always the most effective way to change a household's predicted condition from being energy-poor to not being energy-poor. This was true for both single-action and multi-feature counterfactuals. The required income increases were often less than 5%, suggesting that for households on the borderline, relatively small financial improvements can make a significant difference.

Our results also highlight the long-term nature of energy poverty. The use of data from the preceding seven years confirms that energy poverty is not merely a consequence of immediate circumstances but is related to a household's long-term socioeconomic trajectory. This challenges the conventional focus on contemporaneous factors and suggests that policies must consider the cumulative impact of past conditions. Interestingly, our analysis revealed that interventions could be effective at various time points. For example, both recent income gains (in years $T = 6$ and $T = 7$) and earlier improvements (in years $T = 1$ and $T = 2$) were shown to reduce future risk. This implies that both sustained financial health and timely support can be critical for prevention. We highlight here that the contemporaneous relation between income and energy poverty has been highlighted in previous work [DSZ21; HDZ22]; however, our results reveal the effect of lagged income and future energy poverty, indicating that a household's long-term financial trajectory, not just its present condition, is an important determinant of its future risk.

Other indicators of energy poverty include the unemployment rate, the rate of part-time employment, and the cost of energy. Concerning energy prices, their connection to energy poverty aligns with recent research [SRL23].

In addition to the counterfactual analysis, we examined whether small year-to-year income increases are associated with improvements in MEPI values in the observed data. The results indicated that households experiencing a modest income increase of less than 5% exhibit a lower average MEPI (0.051) compared to households without such an increase (0.062). Regression results in Table D.4 of Appendix D confirm that these differences are statistically significant. A small income increase is associated with an average reduction in MEPI of 0.011 ($p < 0.001$), even after controlling for year fixed effects and clustering standard errors at the household level. These findings support the plausibility of our counterfactual scenarios, showing that the small changes in income identified through the framework align with realistic transitions found in the HILDA dataset.

All things considered, the findings of this study offer concrete guidance for policymakers and social support organizations. The predictive model can serve as a tool to identify households that are likely to face energy poverty. This allows for early, targeted support to prevent the problem before it fully develops. The counterfactual analysis complements the model's predictions by generating interpretable scenarios for intervention. It also provides evidence-based strategies with minimal interventions to prevent energy poverty.

For policymakers, this framework supports the design of more efficient and cost-effective programs. Instead of implementing broad, untargeted subsidies, governments could deploy precise interventions, such as small, targeted income supplements, financial counseling, or employment support for households flagged by the model. The scenarios presented in Table 6.1 illustrate how such support could be tailored to offset specific external shocks, like rising energy prices or a drop in labor force participation. This counterfactual analysis constitutes the Contribution 2 of our work.

For non-governmental and community organizations, our findings reinforce the importance of programs aimed at improving income stability. The evidence that even small income increases can be highly effective provides a strong rationale for prioritizing such initiatives.

However, we acknowledge several limitations in this explanatory study. First, the predictive power of our model (ROC AUC of 70.01%) is good but not perfect. Misclassifications are inevitable, and it is important to consider their real-world consequences. A false positive (wrongly flagging a household as at-risk) may lead to inefficient allocation of resources, while a false negative (failing to identify a household that will become energy-poor) means a missed opportunity for prevention. Second, our counterfactual analysis is based on the patterns learned by the model, but these should not be interpreted as definitive causal links. For instance, while an income increase is associated with avoiding energy poverty, our study does not model the underlying cause of that increase (e.g., a new job, a promotion, or a government benefit), which could have its own complex effects. The prototype approach does not distinguish cases where profiles are close in feature space but differ in which features contribute to that similarity. Another limitation is the uniform treatment of households, which overlooks heterogeneity in responses to energy poverty predictors. In other words, this approach implicitly assumes that all households above the threshold experience energy poverty in a similar way. However, in practice, households may face very different forms of deprivation. Factors such as income, age, education, and personal traits likely influence how individuals experience and respond to energy challenges [Con+22]. Finally, our findings are based on data from the HILDA survey. The specific drivers of energy poverty and the effectiveness of certain interventions may differ in other countries with different climates, energy markets, economic conditions, and social safety nets. Therefore, the generalizability of our results to other contexts should be approached with caution.

CHAPTER 7

Health and Energy Poverty: A Causal Discovery Analysis of a Bidirectional Relationship

Diogo Nuno Freitas, Santiago Budría, Eduardo Fermé

Abstract. *Energy poverty, the inability of a household to afford adequate energy services, is a policy concern given its negative effects on health and development. While the literature has established the effect of energy poverty on health, the reverse causal link remains underexplored. This study addresses this gap by investigating the bidirectional relationship between health and energy poverty. We use panel data for 4,770 individuals from the HILDA survey (2012-2019) and apply a causal discovery algorithm (J-PCMCI+) to a multidimensional energy poverty index (MEPI) and a latent health stock variable. Motivated by the hypothesis of a bidirectional causal relationship, we find that the link is temporally asymmetric. A negative health shock acts as an immediate cause of energy poverty. The reverse effect, from energy poverty to poor health, emerges more gradually over a one- to three-year period. We also find strong state dependence in energy poverty. The temporal asymmetry of the results suggests that policy requires a dual approach: short-term support to shield households from health shocks and long-term structural interventions to address the persistence of energy poverty and its cumulative health impacts.*

7.1 Introduction

Energy poverty is a policy concern affecting a large share of the global population. Recent estimates suggest that 750 million people still lack access to electricity worldwide, and more than 2 billion people lack access to clean cooking fuels [Int24]. The European

Commission [Eur23] defines energy poverty as the inability of a household to afford or access the energy services needed to support adequate living conditions and human development. These services include lighting, cooking, access to technology and communication, as well as heating and cooling. Standard definitions of energy poverty consistently highlight the potential negative consequences of insufficient access or affordability of energy on individual well-being. Factors such as climate change and increasing energy prices have likely intensified the impacts of energy poverty. For this reason, researchers have examined its effects on various dimensions of health and well-being in multiple countries [AS21; PBL21; Pon+24; ZAK21].

Given that energy poverty intersects with various socioeconomic vulnerabilities, it is important to identify the causal drivers in order to support policy interventions. We use a panel dataset from HILDA[1], a micro-survey representative of the Australian population, and apply a causal discovery framework to MEPI, a multidimensional index of energy poverty. The analysis characterizes a system of socioeconomic causal relationships, with a focus on the drivers of energy poverty and health status, and their bidirectional link. In doing so, the study makes four contributions to the literature.

First, this paper applies the Joint-PCMCI+ (J-PCMCI$^+$) [GNR23] causal discovery algorithm to examine the relationship between poor health and energy poverty. This framework is an extension of the original Peter and Clark Momentary Conditional Independence (PCMCI$^+$) algorithm [Run+19], and is able to learn causal relationships from multiple datasets (in our case, one for each individual in the panel). This approach allows us to identify time-lagged causal relationships while accounting for unobserved heterogeneity across individuals and time. To our knowledge, this is the first analysis to apply a formal causal inference framework like J-PCMCI$^+$ to micropanel data for this purpose. It is important to note here that our causal method differs from the feature importance techniques used in machine learning models [LL17]. A variable with high feature importance may only be correlated with energy poverty and not necessarily be a direct cause. Our analysis, in contrast, identifies the specific variables that have a direct causal influence on future energy poverty outcomes.

Second, by using a 15-year longitudinal dataset, our study moves beyond contemporaneous associations. While traditional econometric and machine learning models have identified the determinants of energy poverty in static contexts [DSZ21; GJA24; SRL23], they often do not capture the timing of the causal effects. Our method complements this traditional analysis with a more dynamic view. We map how past and present conditions causally affect future vulnerability. Our approach, therefore, allows us to identify not only *if* a causal relationship exists, but also *when* it manifests. As an example, in the context of this study, we aim to identify a bidirectional relationship between health and energy poverty. Interestingly, however, the causal pathway from health to energy poverty operates on a different timescale than the reverse pathway.

[1] https://melbourneinstitute.unimelb.edu.au/hilda

Third, we extend the literature on the determinants of energy poverty by addressing a largely overlooked driver: health status. Much of the existing work focuses on income, energy prices, and housing conditions as primary explanatory factors. The role of health has received less attention. Specifically, the possibility that poor health leads to energy poverty, rather than resulting from it, is not well understood. A growing body of evidence shows that energy and fuel poverty negatively affect physical and mental health [AS21; BMM21; Naw21; Oli+21; Oum19; PBL21; Pon+24; Zha+21]. However, the reverse causal pathway from health shocks to energy poverty remains underexplored. Recent research has started to address this gap. Studies such as Brown and Vera-Toscano [BV21] and Budría, Li Donni, and Zucchelli [BLZ25] find that the impact of ill health on energy poverty is statistically significant and quantitatively stronger than the reverse effect. This suggests health deterioration may act as a trigger for falling into energy poverty, particularly among vulnerable populations.

Fourth, in addition to the main causal connection between health and energy poverty, our analysis also describes a wider network of cause-and-effect relationships. We identify the contemporaneous and lagged drivers of both energy poverty and health status, confirming the roles of factors like employment and education while also highlighting strong persistence effects. Furthermore, our study extends the causal dynamics at two levels: at the individual and household level, revealing pathways linking health shocks to economic outcomes through the labor market; and at the macroeconomic level, discovering a feedback loop between state-level labor markets and energy prices.

This study is motivated by the hypothesis of a bidirectional causal relationship between health and energy poverty. We hypothesize that poor health can increase a household's risk of falling into energy poverty, and that the experience of energy poverty, in turn, can lead to a deterioration in health.

The remainder of this paper is structured as follows. Section 7.2 reviews the relevant literature on the determinants of energy poverty. Section 7.3 details the data and methodology, including the construction of the energy poverty and health indicators and the causal discovery framework. Section 7.4 presents the results of the causal analysis. Finally, Section 7.5 discusses the findings, considers their policy implications, and concludes the paper.

7.2 Literature Review

The global interest in energy poverty arises from its far-reaching consequences, which are multifaceted. Research based on international macroeconomic data indicates that the prevalence of energy poverty negatively affects development, health outcomes, and human capital formation [BMM21; PK19]. Moreover, energy access and affordability are a crucial dimension of multidimensional poverty and, as such, they can be negatively related to economic growth [BL24]. Studies based on microeconomic panel data are consistent with this notion, showing that household-level energy poverty significantly affects a number of personal-level outcomes, including subjective well-being, health, and quality of life [Pon+24; Zha+21].

Cross-national studies underscore that energy poverty is influenced by a range of macroeconomic and institutional factors, including education, governance quality, technological progress, health expenditures, and the level of economic development [SKR23]. The relationship between these factors and household energy deprivation often varies by GDP levels, as shown by Boţa-Avram, Apostu, Ivan, and Achim [Boţ+24]. Income inequality, internal conflict, and climatic conditions also contribute, though to varying degrees, depending on regional contexts [IM22; KSC24]. Moreover, the energy mix of a country—reflecting the sources used for electricity production—plays a pivotal role in shaping energy outcomes [KUO23].

At both the regional and national levels, energy poverty arises from a complex interplay of structural and household-specific factors. Key determinants include energy prices, availability, and inefficiencies in building infrastructure, such as insulation quality, heating systems, and floor space, all of which significantly influence vulnerability to energy deprivation [CZN23; KŚ20]. Additionally, the urban-rural divide plays a crucial role in shaping energy access and affordability [WZZ+24]. At the household level, financial constraints, such as low income and high energy costs, intensify risks, often resulting in unpaid bills, energy debt, or supply disconnections [BB19b]. Moreover, individual characteristics play a pivotal role. Higher educational attainment is negatively associated with energy poverty, as it fosters energy-efficient behaviors and economic resilience [CAF19; LYR24]. Factors such as household size, marital status, and geographic location further interact with energy needs and costs, while age influences energy deprivation through life-cycle patterns and risk preferences [Abb+20; DJ21]. Lastly, poor health can reshape household spending priorities, limiting access to adequate energy services, while income and employment status remain consistent predictors of energy poverty [Abb+20].

The application of AI-based methods to predict energy poverty has gained attention recently, reflecting its potential to enhance both understanding and mitigation strategies. Although still in its infancy, this approach provides new insights into the drivers of energy poverty. For instance, studies using advanced machine learning frameworks such as XGBoost have identified critical predictors in diverse contexts. In the Netherlands, income, house value, and homeownership emerged as the most significant determinants of energy poverty [DSZ21]. Complementarily, research utilizing random forest classifiers highlights the importance of both household-level factors, such as dwelling conditions and energy efficiency, and country-specific elements, including gas supplier switching rates [SRL23]. While the previous studies are based on a single energy poverty indicator, other studies define a multidimensional energy poverty index similar to ours. These studies have examined Asian and African contexts, showing that wealth, marital status, and place of residence play pivotal roles [Abb+20]. Methodological advancements have also emerged, with ensemble models like XGBoost, random forest, and artificial neural networks uncovering novel predictors, such as education and food security, which significantly influence energy poverty risks [GJA24]. Similarly, high-performing machine learning models have proven effective at predicting and understanding the mechanisms driving access to cooking with clean energy [MWB23].

7.3 Materials and Methods

7.3.1 Data Source: The HILDA Survey

In this study, we use the data from the HILDA survey. HILDA is a nationally representative longitudinal study initiated in 2001 with a sample of approximately 7,600 households and 13,000 individuals. The survey annually tracks these individuals and their households, collecting information on income, labor market status, housing conditions, family structure, and health. We note here that HILDA has an average retention rate of over 90% between waves, reducing concerns related to panel attrition.

The causal discovery analysis requires a balanced panel structure. For this, we use the final eight available waves of the survey, spanning from 2012 to 2019, which creates a panel of $T = 8$ consecutive years for each household in our sample. This structure is necessary to examine how characteristics from previous periods causally influence energy poverty at time $T = 8$. The final sample consists of 38,160 observations from a cohort of 4,770 individuals for whom there is complete information on all relevant variables for the entire eight-year period.

Our analysis includes a set of variables to account for external and internal factors. We account for regional conditions using state-level gross product per capita and its growth rate, labor market indicators, geographic remoteness, and energy prices. Household and individual characteristics include age, sex, marital status, the presence of children, and years of education. We also control for individual economic circumstances through measures of labor force status, personal income, and social capital. The two central variables in our analysis are the household's energy poverty status (the *MEPI* indicator) and the individual's health status (the *health stock* indicator). The construction of these two measures is detailed in the following sections. A complete description of all variables used in the analysis is provided in Appendix D.7.

7.3.2 Measuring Energy Poverty: The MEPI Indicator

To measure energy poverty, we construct a MEPI. This index is multifaceted and captures both expenditure-based and subjective dimensions of deprivation. The expenditure-based measures are the 2M, TPR, and LIHC indicators. The subjective dimension is captured by two self-reported indicators: a household's inability to pay for heating due to a shortage of money and its inability to pay utility bills (electricity, gas, or telephone) on time.

The MEPI for an individual i at time t is calculated as a weighted sum of these five indicators. Let J be the set of five indicators. Let EP_{ijt} be a variable that takes the value one if individual i is deprived in indicator j at time t, and zero otherwise. The score for individual i is given by:

$$\mathrm{MEPI}_{it} = \sum_{j \in J} w_j \mathrm{EP}_{ijt}, \tag{7.1}$$

where w_j is the weight assigned to indicator j, with $\sum_{j \in J} w_j = 1$. An individual is classified as energy poor if their score exceeds a predefined cutoff, $\mathrm{MEPI}_{it} > \bar{m}$. For our main analysis, we set the cutoff $\bar{m} = 0$, meaning a household is considered energy poor if it experiences deprivation in at least one of the five dimensions.

While equal weighting is a common approach, we adopt a frequency-based weighting scheme [DL13]. This method assigns a higher weight to indicators where deprivation is less common. The weight for indicator j is calculated as:

$$w_j = \frac{(1 - n_j)}{\sum_{k \in J}(1 - n_k)}, \tag{7.2}$$

where n_j is the proportion of individuals in the sample deprived in dimension j. This approach is motivated by the idea that deprivation in a dimension where most of the population is not deprived (e.g., the ability to heat one's home) is a stronger signal of hardship. These weights are calculated for each survey wave, allowing the relative importance of each indicator to adapt to changing economic conditions and social norms over time.

7.3.3 Measuring Health Status: The Health Stock Indicator

The other central variable in our analysis is health status. It is important to note that standard self-reported health indicators are subject to measurement error and reporting biases that can compromise causal inference. We therefore construct a latent health stock variable using an instrumental-variable-type strategy from the empirical health economics literature. This involves a Generalized Ordered Probit (GOP) estimation where Self-Assessed Health (SAH) is specified as a function of detailed physical health indicators, such as functional limitations and pain intensity. The method produces a hidden health stock variable that compensates for personal reporting biases and acts as our main health indicator.

This approach is frequently applied in studies of health and labor outcomes (e.g., Brown, Roberts, and Taylor [BRT10], Harris, Zhao, and Zucchelli [HZZ21], and Jones, Rice, and Roberts [JRR10]). It helps reduce biases caused by subjective reporting, although it does not completely eliminate the problem of reverse causality. By considering different health-related aspects, like limitations in daily activities, physical restrictions, and bodily pain, the resulting latent index provides a more indicative measure of actual health status than just using a single self-reported item.

The SAH variable in HILDA is an ordinal indicator derived from the question: "In general, would you say your health is excellent, very good, good, fair, or poor?." The health indicators used as explanatory variables in the GOP estimation cover various physical abilities and health-related limitations. These include challenges with medium and intense activities, lifting weights, climbing stairs, walking long distances, or managing personal care tasks. This estimation also accounts for work limitations caused by health issues and the degree of pain and its interference with daily functioning.

To allow for heterogeneous reporting behavior, we perform separate GOP estimations for men and women and let the reporting thresholds vary by demographic and socioeconomic factors, including age, education, ethnicity, employment status, and income. Following Harris, Zhao, and Zucchelli [HZZ21], we assume that any residual link between SAH and socioeconomic characteristics, once specific health conditions are controlled for, reflects reporting bias rather than actual health differences. Although this assumption has its limitations, it helps us create a more accurate and unbiased indicator of overall health.

7.3.4 Causal Discovery Analysis

To identify the causal pathways between health, energy poverty, and other socioeconomic factors, we employ the J-PCMCI$^+$ causal discovery algorithm [GNR23], which allows for both contemporaneous and lagged links between variables. J-PCMCI$^+$ is an extension of the PCMCI$^+$ framework [Run+19] designed for causal discovery from multiple datasets, such as the panel data from the HILDA survey. It allows for the presence of both observed and unobserved contexts that differ between datasets (i.e., individuals) and over time. This characteristic is important for our work as it allows us to control for unobserved individual-specific, time-invariant characteristics (spatial contexts) and common shocks affecting all individuals in a given year (temporal contexts), thus mitigating the risk of false positives.

The algorithm operates within the graphical causal model framework [SGS01]. Let $\mathbf{X}_t = \{X_t^i\}_{i \in \mathcal{I}}$ be the set of system variables at time t, $\mathbf{C}_t$ the set of observed context variables, and $\mathbf{D} = \{D_{\text{space}}, D_{\text{time}}\}$ the set of dummy variables. The core idea of J-PCMCI$^+$ is to explicitly model contexts to de-confound relationships by incorporating the dummy variables directly into the analysis. This allows the algorithm to distinguish genuine causal links from spurious correlations induced by unobserved heterogeneity.

The causal interpretation of the resulting graphs relies on a set of standard assumptions [Run+19]. These include *Causal Sufficiency* (i.e., all common drivers are included in the dataset), the *Causal Markov Condition*, and the *Faithfulness* assumption. While Causal Sufficiency is a strong assumption, the comprehensive nature of the HILDA survey, which includes a wide range of socioeconomic and demographic variables, helps to mitigate concerns about major omitted variable bias. We also assume stationarity in the underlying processes and explore time lags up to $\tau_{\max} = 3$ years.

The J-PCMCI$^+$ algorithm estimates the causal graph through a multi-stage procedure based on conditional independence tests. For this analysis, we use the linear partial correlation test, which assumes approximately linear relationships, and explore time lags up to $\tau_{\max} = 3$ years. The algorithm proceeds in four main steps.

In the first step, the PC_1 algorithm is run on all system variables and observed temporal contexts $\{\mathbf{X}, \mathbf{C}_{\mathrm{time}}\}$ to discover a superset of lagged parents for each variable, denoted $\hat{\mathcal{B}}_t^-(\cdot)$. Then, the Momentary Conditional Independence (MCI) test is used to identify links between observed context variables ($\mathbf{C}$) and system variables ($\mathbf{X}$). For each pair (C^k, X^j), it tests for conditional independence, for instance:

$$C_{t-\tau}^k \perp\!\!\!\perp X_t^j | \mathbf{S}, \hat{\mathcal{B}}_t^-(X_t^j) \setminus \{C_{t-\tau}^k\}, \hat{\mathcal{B}}_{t-\tau}^-(C_{t-\tau}^k), \tag{7.3}$$

where $\mathbf{S}$ is a subset of contemporaneous adjacencies. This step identifies the set of observed contextual parents for each system variable.

In the third step, the MCI is run to identify links between the dummy variables ($\mathbf{D}$) and system variables ($\mathbf{X}$). It tests for conditional independence:

$$D \perp\!\!\!\perp X_t^j | \mathbf{S}, \hat{\mathcal{B}}_t^C(X_t^j), \tag{7.4}$$

where $\hat{\mathcal{B}}_t^C(X_t^j)$ is the set of all lagged and observed-context parents of X_t^j found in the previous steps. This step identifies the influence of latent confounders.

Finally, the MCI test is run on all system-variable pairs (X^i, X^j) to identify the core causal structure. It tests for conditional independence:

$$X_{t-\tau}^i \perp\!\!\!\perp X_t^j | \mathbf{S}, \hat{\mathcal{B}}_t^{CD}(X_t^j) \setminus \{X_{t-\tau}^i\}, \hat{\mathcal{B}}_{t-\tau}^{CD}(X_{t-\tau}^i) \tag{7.5}$$

where $\hat{\mathcal{B}}^{CD}$ is the full set of parents (lagged, context, and dummy) discovered so far.

The remaining adjacencies are oriented using the standard rules of the PCMCI$^+$ algorithm. For our analysis, we evaluate the method considering a significance level of $\alpha_{\mathrm{PC}} = 0.05$.

We treat each individual's 8-year history as an independent realization of the same structural dynamic system. Moreover, to mitigate unobserved time-invariant heterogeneity and common time shocks, we apply two-way demeaning to each numeric variable X:

$$\tilde{X}_{it} = X_{it} - \bar{X}_{i\cdot} - \bar{X}_{\cdot t} + \bar{X}, \tag{7.6}$$

where $\bar{X}_{i\cdot}$ is the individual mean over t, $\bar{X}_{\cdot t}$ is the cross-sectional mean in year t, and $\bar{X}$ is the grand mean. We then standardize continuous variables to a zero mean and unit variance.

Before running the causal discovery algorithm, all categorical variables were converted using one-hot encoding. However, the partial correlation test used in our analysis assumes continuous inputs. Thus, to handle discrete variables while preserving rank information, we add minimal Gaussian jitter $\epsilon \sim N(0, \sigma^2)$ with $\sigma = 10^{-6}$ to all variables.

Table 7.1: Causal drivers of energy poverty for the lagged specification ($\tau_{\min} = 1$).

Causal Driver	**Lag (Years)**	**Partial Corr.**	***p*-value**
Energy Poverty	1	0.284	<0.001
Energy Poverty	2	0.149	<0.001
Energy Poverty	3	0.038	0.001
Employment Status	1	-0.074	<0.001

7.4 Results

We now present the causal discovery results under two specifications for the time lags. We analyze only lagged effects in the first specification by setting the minimum and maximum lags to $\tau_{\min} = 1$ and $\tau_{\max} = 3$ years, respectively. This excludes contemporaneous effects. In the second specification, we include contemporaneous relationships by setting $\tau_{\min} = 0$ while maintaining $\tau_{\max} = 3$. For both specifications, the significance level for the parent selection stage is $\alpha_{\mathrm{PC}} = 0.05$. The analyses were performed on non-normalized data.

7.4.1 Causal Discovery Excluding Contemporaneous Effects

7.4.1.1 Causal Drivers of Energy Poverty

Table 7.1 presents the statistically significant drivers of energy poverty identified in the lagged-only analysis. The results indicate persistence in energy poverty status and show that employment has a protective (negative) effect on future energy poverty.

More specifically, the causal discovery suggests a strong auto-causal effect of energy poverty. A household's status in the three preceding years is a direct cause of its current energy poverty status. This finding shows a strong state dependence; in other words, households in energy poverty have difficulty escaping this situation. We also note here that the partial correlation is strongest for the one-year lag and decreases as the number of lagged years increases, suggesting that more recent experiences of hardship have a greater causal impact.

The analysis also reveals a lagged protective effect of employment. Being employed in the previous year has a significant, negative causal effect on the likelihood of experiencing energy poverty. This could indicate that the financial stability gained from employment acts as a buffer, providing resources that help households withstand economic shocks in the subsequent year.

7.4.1.2 Causal Drivers of Health Status

The results for the drivers of health status, presented in Table 7.2, also shows strong persistence. More importantly, this analysis reveals a significant lagged causal effect from energy poverty to health.

Table 7.2: Causal drivers of health status for the lagged specification ($\tau_{\text{min}} = 1$).

Causal Driver	**Lag (Years)**	**Partial Corr.**	**p-value**
Health Status	1	0.357	<0.001
Health Status	2	0.178	<0.001
Health Status	3	0.127	<0.001
Energy Poverty	1	-0.032	0.010
Energy Poverty	3	-0.037	0.002

Indeed, an individual's health is strongly determined by their health in prior years. The key finding here is the emergence of statistically significant causal links from energy poverty to health status at both one-year ($p = 0.010$) and three-year ($p = 0.002$) lags. This suggests that the experience of energy poverty has a detrimental causal effect on an individual's health, and this effect begins to manifest within one year and persists for several years.

7.4.2 The Role of Contemporaneous Effects

To test for more immediate causal mechanisms, we adjusted the causal analysis to include contemporaneous effects ($\tau_{\text{min}} = 0$), keeping all other parameters the same. This modification complements the base case findings by revealing causal pathways that act within the same year. The results of this analysis are presented in Table 7.3.

This analysis identified two significant contemporaneous drivers of energy poverty: health status ($p = 0.002$) and years of education ($p < 0.001$). In contrast, no contemporaneous causal drivers were identified for the health status variable. Regarding the lagged structure, no new causal pathways were discovered. The lagged drivers of energy poverty remained consistent with the previous specification. For health status, however, the one-year lagged causal effect from energy poverty became non-significant, while the three-year lagged effect persisted.

The contemporaneous link from health status suggests that health issues can lead to immediate energy poverty. This could happen due to a health shock, which could immediately decrease work capacity and income, or increase health-related expenses, thus causing a household to experience energy poverty within that year.

The second contemporaneous link is with years of education. Years of education can be a proxy for human capital and long-term earning potential. Thus, its connection to energy poverty may suggest that individuals with lower levels of education may be in occupations that are more vulnerable to income shocks or have fewer financial resources.

7.4.3 Temporal Asymmetry in the Health-Energy Poverty Relationship

When comparing the results from the lagged-only ($\tau_{\text{min}} = 1$) and contemporaneous ($\tau_{\text{min}} = 0$) specifications, we identify a bidirectional relationship between health and energy poverty. This relationship, however, is characterized by a distinct temporal asymmetry.

Table 7.3: Causal drivers for the specification including contemporaneous relationships ($\tau_{\text{min}} = 0$).

Factor	**Causal Driver**	**Lag (Years)**	**Partial Corr.**	***p*-value**
Energy Poverty	Years of Education	Cont.	-0.038	<0.001
	Health Status	Cont.	-0.032	0.002
	Energy Poverty	1	0.281	<0.001
	Energy Poverty	2	0.147	<0.001
	Energy Poverty	3	0.036	0.004
	Employment Status	1	-0.045	<0.001
Health Status	Health Status	1	0.354	<0.001
	Health Status	2	0.177	<0.001
	Health Status	3	0.126	<0.001
	Energy Poverty	3	-0.034	0.006

The causal pathway from health to energy poverty is an immediate phenomenon. A negative health shock contemporaneously increases the risk of energy poverty (partial corr. = -0.032), as this link is only identified when same-year effects are included in the analysis. The likely mechanisms are a sudden loss of income due to reduced work capacity or an increase in health-related expenditures, which immediately constrain a household's budget.

Conversely, the causal pathway from energy poverty to health operates over a longer time horizon. In the lagged-only analysis, the detrimental effect of energy poverty on health manifests at both one- and three-year lags. While the one-year effect becomes non-significant in the contemporaneous specification, the three-year lagged effect remains robust across both models (partial corr. = -0.034 in the contemporaneous specification), suggesting a persistent long-term impact. This observation indicates that the persistent stress and insufficient resources associated with energy poverty gradually impairs an individual's health over an extended period.

7.4.4 Extended Causal Discovery Findings

7.4.4.1 Individual and Household Causal Dynamics

It is also important to explore beyond the direct link between health and energy poverty. This analysis employs a causal model incorporating current factors to reveal the causal pathways linking various individual and household characteristics, as illustrated in Figure 7.1.

An individual's educational level is identified as an important socioeconomic factor with immediate causal consequences. We find a contemporaneous causal effect from years of education to household income, social capital, and energy poverty. This suggests that education directly shapes a household's immediate economic outcomes and social resources, which in turn influence its vulnerability to energy poverty.

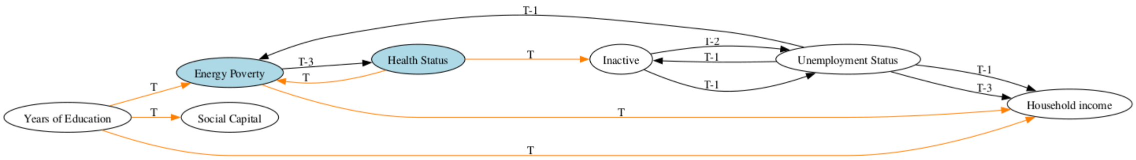


Figure 7.1: Causal graph of individual and household-level variables. Orange arrows indicate contemporaneous effects (T), while black arrows indicate lagged effects.

The analysis also reveals a causal chain linking health to economic outcomes through the labor market. Health status has a contemporaneous causal effect on being inactive, indicating that a negative health shock can immediately remove an individual from the labor force. Labor market status itself is characterized by persistence and transitions; we observe a bidirectional, one-year lagged causal link between being inactive and unemployment status. This dynamic affects household finances, as unemployment status is a lagged cause of lower household income at both one- and three-year horizons. This pathway illustrates how health shocks can initiate a series of adverse economic effects that develop over multiple years.

7.4.4.2 Macroeconomic Causal Dynamics

The analysis also identified a system of causal relationships among macroeconomic indicators at the state level, as shown in Figure 7.2. In this system, GSP per capita functions as a key exogenous driver. Regional economic output has a contemporaneous causal effect on the total labor force participation rate and lagged effects on both energy price (three-year lag) and labor participation (one- and three-year lags). This positions regional economic prosperity as an important determinant of both labor market conditions and energy market dynamics.

The labor market indicators exhibit a high degree of interconnectedness. A causal relationship exists between the part-time employment rate and the unemployment rate across multiple time lags (one, two, and three years). These indicators, in turn, have lagged causal effects on the total labor force participation rate.

Interestingly, energy price is not an independent factor but is part of a feedback loop with the labor market. The unemployment rate has a one-year lagged causal effect on energy price, which in turn causally influences both the unemployment rate (one-year lag) and the part-time employment rate (two-year lag). This reveals that labor market conditions can affect energy prices, likely through changes in aggregate demand, and these price changes then feed back into the labor market, impacting employment outcomes in the following years.

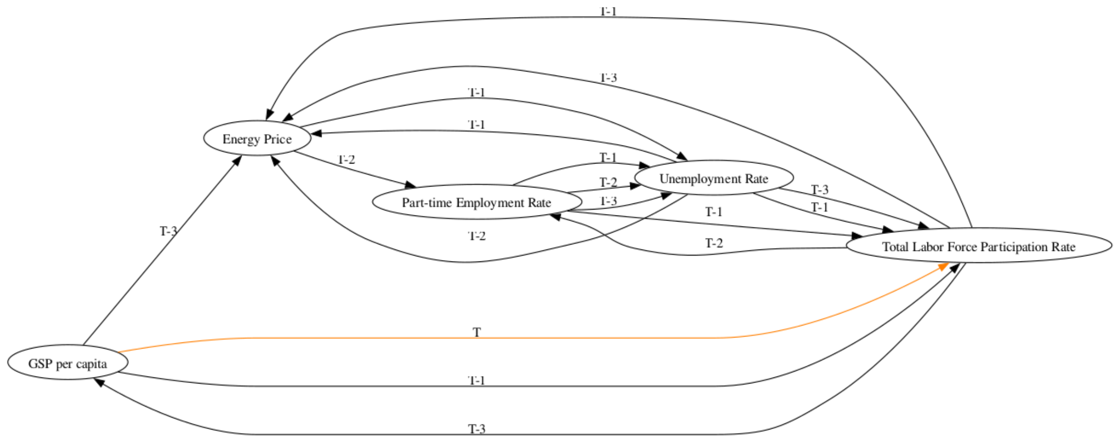


Figure 7.2: Causal graph of macroeconomic variables. Orange arrows indicate contemporaneous effects (T), while black arrows indicate lagged effects.

7.5 Discussion

In this study, we do *not* focus on the task of predicting energy poverty. Instead, we use a causal discovery framework to map the dynamic, time-lagged pathways leading to energy poverty. We identify causal triggers and their timelines by considering two specifications: one that includes contemporaneous factors and one that does not. In both cases, however, we consider a maximum lag of three years. Although we concentrate on how health and energy poverty affect each other, we also explore how they causally relate to other factors.

The asymmetric bidirectional relationship between health status and energy poverty Our main finding reveals a bidirectional but temporally asymmetric causal relationship between poor health and energy poverty. This dynamic suggests that although the two factors are closely linked and influence each other in a continuous cycle, the causes that connect them work on different lags.

First, the causal pathway from health to energy poverty manifests as an immediate, contemporaneous phenomenon. As shown in our analysis that includes same-year effects, a negative health shock is a direct and instantaneous cause of a household falling into energy poverty. This direct connection is likely driven by two primary mechanisms: a sudden loss of income due to an inability to work and a simultaneous increase in personal spending on health-related costs. These twin financial shocks can instantly strain a household's budget, pushing them below the energy poverty threshold. These two financial shocks can quickly put pressure on a family's budget, making them fall into energy poverty.

On the other hand, the causal pathway from energy poverty to poor health operates over a much longer and more gradual time horizon. Our findings indicate that the negative impact of energy poverty on health becomes apparent gradually, within a period of one to three years. This delayed impact reflects the cumulative toll of sustained hardship. The literature supports this interpretation: households facing chronic difficulties in affording adequate heating, cooling, or clean fuels experience slow deterioration of physical health (e.g. respiratory and cardiovascular disease from cold, damp homes and indoor air pollution [Abb+20; LC15; Naw21; Zha+21]) and increasing psychological stress that negatively affects their mental health [BV21]. This lagged effect shows that continuous exposure to a lack of energy resources causes long-term harm, leading to more illness and shorter life expectancy over time [PBL21].

We highlight here the work developed by Brown and Vera-Toscano [BV21], who used the same dataset as we do in this study. The findings of that study align with ours, highlighting a bidirectional connection: poor health can cause energy poverty, and energy poverty can worsen health. This cycle forms what are known as "poverty traps." Our contribution is to extend these findings by identifying *when* these causal links operate. We show that the pathway from health to energy poverty is immediate, whereas the reverse pathway materializes only after sustained exposure.

Of course, the asymmetry in timing has important policy implications. Interventions addressing health shocks may prevent households from falling into energy poverty in the short term, whereas long-term reductions in energy poverty are critical to improving population health outcomes. For example, short-term measures could include targeted energy bill subsidies for households facing sudden health shocks, medical cost relief programs, or temporary income support schemes. Long-term measures may require structural investments such as housing renovation and insulation programs, subsidies for clean cooking fuels, and expansion of affordable renewable energy access.

The dynamics of energy poverty: The most significant driver of energy poverty is its own persistence. Our analysis reveals a strong auto-causal relationship, where a household's energy poverty status in previous years is the most significant predictor of its current status. This suggests that energy poverty is a structural state from which households struggle to exit. This finding aligns with previous research in the field. For instance, using German panel data, Drescher and Janzen [DJ21] identified a similar state dependence, noting that households experiencing energy poverty were 7.5 more likely to remain so in the subsequent year. Furthermore, the work of Phimister, Vera-Toscano, and Roberts [PVR15] corroborates this persistence across different measurement types, which is particularly relevant to our composite MEPI index. They found that 44.1% of individuals in expenditure-based energy poverty remained so one year later, with this figure rising to 52% for those identified by subjective measures. Similarly, Karpinska and Śmiech [KŚ21] found evidence of persistence and identified factors that can lock households into energy poverty, such as tenant status, old age, low education, and unemployment.

The persistence of energy poverty has important policy implications. While short-term measures like energy vouchers can offer temporary relief during crises, they do not address the structural nature of the problem [KŚ20], meaning households will eventually fall again into the "energy-poverty trap". Long-term solutions are needed to lift households permanently out of energy poverty. For example, such policies include investing in energy-efficient housing, strengthening regulations for rental properties, and creating renovation programs targeted at vulnerable households. Furthermore, a single national policy may be insufficient [Mar+19]. Research suggests that local strategies are also necessary to address specific regional conditions.

Our findings also indicate additional protective factors against energy poverty, specifically employment and education. Some earlier studies (for example, [DJ21; SRL23]) also recognized these factors; however, our research examines how they function across various timescales. Employment provides a lagged protective effect, where the financial stability gained in one year acts as a buffer against experiencing energy poverty in the next. In contrast, education functions as a contemporaneous shield. This immediate protection likely stems from the higher income, financial literacy, and economic resilience associated with higher educational attainment.

Extended causal dynamics: The causal model also reveals socioeconomic and macroeconomic dynamics that have implications beyond energy and health.

At the individual level, the analysis shows that educational attainment is a contemporaneous driver of household income and social capital. This is consistent with the literature that links education to economic resilience [CAF19; LYR24]. As suggested by Lu, Yu, and Ren [LYR24], this could be because households with higher education levels usually have better access to and use more energy-efficient products.

The causal analysis also identifies a bidirectional lagged relationship between unemployment and inactivity (and as expected, the latter with health status), indicating that these labor market conditions are persistent. This relationship is also examined in labor market models [BRT10; ZHZ12]. These studies find that poor health substantially raises the rate of economic inactivity. Additionally, ill health leads to an increase in the transition from employment to inactivity. There is also strong evidence of state dependence, meaning that previous periods of inactivity make it more likely for an individual to be inactive in the future.

At the macroeconomic level, the findings show that regional economic output (GSP per capita) is an exogenous driver of state-level labor market conditions and energy prices. This confirms that regional economic output is a determinant of these conditions. Indeed, there is evidence that the economic output of a country strongly conditions the impact of energy poverty on health and education outcomes [BMM21]. Some studies even report that regional economic/geographical conditions are stronger predictors of fuel poverty than household sociodemographics [BB19a].

Finally, the analysis also identifies a feedback loop between the labor market and energy prices, where unemployment rates have a lagged causal effect on energy prices, which in turn affect employment outcomes. An important implication here is that energy and labor market policies are interconnected. Actions taken in one domain can affect the other. This suggests that integrated policy approaches could be more effective.

Policy implications: The findings of this study have policy implications. The identified temporal asymmetry in the health–energy poverty relationship suggests that interventions need to be twofold. On the one hand, the immediate causal pathway from poor health to energy poverty calls for proactive, short-term support mechanisms. For example, social safety programs could be set up to automatically offer financial help for energy bills after a significant health incident, thus preventing families from experiencing energy poverty. On the other hand, the lagged causal effect of energy poverty on health highlights the importance of energy policy as a tool for long-term public health. This implies that structural investments in energy efficiency and affordable clean energy are not just economic measures but also preventative health strategies.

Furthermore, our finding that energy poverty is a persistent state, or a "trap", corroborates earlier research suggesting that short-term policies are insufficient [KŚ20]. To be effective, policy must focus on structural solutions that enable households to permanently exit energy poverty. Such policies could include large-scale housing insulation programs or support for adopting low-cost renewable technologies. Finally, the protective effects of employment and education suggest that broader socioeconomic policies that promote job stability and educational attainment can also serve as a buffer, enhancing household resilience against energy poverty.

To conclude, this study was motivated by the hypothesis of a bidirectional causal relationship between health and energy poverty. The results of our causal discovery analysis confirm this hypothesis. We find that the causal pathways are temporally asymmetric: a negative health shock is an immediate cause of energy poverty, whereas energy poverty has a detrimental effect on health that emerges over a longer time horizon of one to three years.

Limitations: The main limitation of our analysis is the assumption that all households are the same, i.e., that all households above the threshold experience energy poverty in a similar way. This overlooks the diverse ways households experience energy poverty based on their income, age, education, and personal traits [Con+22]. Additionally, our conclusions are based on data from the HILDA survey. Because the drivers of energy poverty and the success of policies depend heavily on local context, our findings may not be applicable elsewhere. Thus, any generalization of these results requires careful consideration. Another limitation of our analysis is the use of a linear partial correlation test within the PCMCI algorithm. This choice implies that the relationships between variables are linear, an assumption that may not hold for all the complex dynamics of health and socioeconomic factors. As a result, our analysis may have failed to detect non-linear causal links.

CHAPTER 8

Relative Drawing Identification Complexity is Invariant to Modality in Vision-Language Models

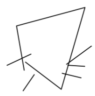
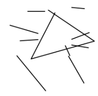
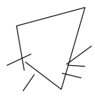
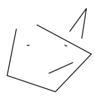
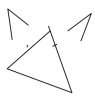
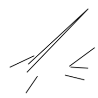
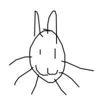
Diogo Nuno Freitas, Brigt Håvardstun, Cèsar Ferri, Dario Garigliotti, Jan Arne Telle, Jose Hernandez-Orallo

Abstract. *Large language models have become multimodal, and many of them are said to integrate their modalities using common representations. If this were true, a drawing of a car as an image, for instance, should map to a similar area in the latent space as a textual description of the strokes that form the drawing. To explore this in a black-box access regime to these models, we propose the use of machine teaching, a theory that studies the minimal set of examples a teacher needs to choose so that the learner captures the concept. In this paper, we evaluate the complexity of teaching vision-language models a subset of objects in the Quick, Draw! dataset using two presentations: raw images as bitmaps and trace coordinates in TikZ format. The results indicate that image-based representations generally require fewer segments and achieve higher accuracy than coordinate-based representations. But, surprisingly, the teaching size usually ranks concepts similarly across both modalities, even when controlling for (a human proxy of) concept priors, suggesting that the simplicity of concepts may be an inherent property that transcends modality representations.*

8.1 Introduction

As children, when we transform images of the world into drawings and other simplified sketches, we have the intuition that some objects are simpler than others [CC84; LFF18]. For instance, six segments are enough to represent a `house` that everybody can recognize, while a bit more is necessary to represent a `cat`. This intuition is epitomized by some guessing games where one person picks a concept from a card deck and has to draw something quickly for their team to identify the concept. We can easily describe and recognize some very simple visual concepts, such as letters, with verbalized descriptions. For instance, the letter `T` is a horizontal segment on top of a vertical segment. However, humans struggle to describe more complex shapes with verbal descriptions [SF22] or objects, such as a `cat`, using a series of segments.

Table 8.1: The simplest drawings (applying RDP algorithm on an original drawing) identified for the concept `cat`.

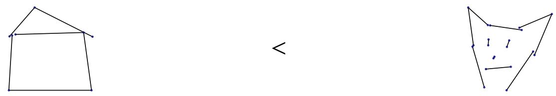
Model	Original (images)	Simplified (images)	Original (coordinates)	Simplified (coordinates)
Claude				
Gemini				
GPT-4 Turbo				
GPT-4o				
Llama				
Pixtral				

LLMs can identify objects from a textual representation of their coordinates [Bub+23]. Thus, we aim to discover whether this understanding maps to similar capabilities for the multimodal versions of these models. Also, we do not know whether this is independent of the modality. We ask two research questions:

- Q1 (*Absolute Invariance*): If we randomly sample a concept from a concept class, $c \in C$, would it take the same number of segments to identify it if represented as a bitmap drawing as if represented as a set of coordinates?
- Q2 (*Relative Invariance*): If we randomly sample two concepts from a concept class, $c_1, c_2 \in C$, and c_1 requires fewer segments than c_2 when represented as a bitmap, will this order prevail when expressed as coordinates?

Question Q1 refers to whether a concept represented as a bitmap drawing is easier or harder to recognize than the same concept as coordinates in text, while question Q2

Bitmap:



Coordinates:

```
\draw (10, 169) - (0, 0) - (250, 0)
- (226, 178) - (20, 172);
\draw (2, 166) - (78, 255) - (253,
165);
```

<

```
\draw (49, 9) - (13, 134);
\draw (67, 199) - (163, 185);
...
```

`House` `Cat`

Figure 8.1: In this paper, we address two research questions. First, Q1 (absolute invariance): When using a vision-language model, are bitmaps (top) equally efficient representations for drawings as coordinates (bottom)? The second question is Q2 (relative invariance): Are the orders (left vs. right) of simplicity preserved across modalities?

is about the relative ranking. For instance, consider that c_1 is a `house` and c_2 is a `cat`. In Figure 8.1, if a `house` is easier than a `cat` when using the bitmap of the drawing (top of the figure), is it also easier when represented as segment coordinates (bottom of the figure)? This is the *relative invariance*. Note that we are not comparing with photographic images of the object since other features would come into play, such as a striped texture to distinguish a `tiger` from other felines. Such distinctions are particularly evident in machine vision systems [Gei+23].

However, how can one determine the notion of simplicity of a concept from its drawings? The idea we pursue in this paper is based on the field of machine teaching [Zhu+18], and in particular, the notion of teaching minimality. A concept is as simple as a teacher can communicate the concept to a learner with as little information as possible. This captures our intuition that a `house` needs six segments while a `cat` needs more segments. Given a concept, the teacher has to find the simplest drawing in terms of the number of straight-line segments—the teaching size—that enables the learner to consistently recognize the concept. We use two different types of language representations (bitmaps of the drawing and coordinates in TikZ code) to present the concepts to the learner. Multiple models, including Generative Pretrained Transformer (GPT)-4 [Ach+23], Llama [Gra+24], Gemini [Tea+24], Pixtral, and Claude, are employed as the "learners". The resulting collection of the simplest images identified, across all concepts, all modalities, and all models, is intriguingly diverse. As a preview of our findings, see Table 8.1, showing the simplest identified images for the concept `cat`.

It is also important to note that priors play a role in machine teaching. When in doubt, the learner will more likely associate the evidence with the most common concept (e.g., a `house` is more common than an `envelope`). Accordingly, a Bayesian prior will be used to disentangle this effect when looking at the concept simplicity rankings.

The contributions of this paper are:

- A novel machine teaching framework for evaluating the complexity of concepts, which can be applied to drawings in coordinate- and image-based modalities.

- Use of the teaching size specifically to evaluate how simply and effectively the concept can be taught across both modalities.
- A comparison of both modalities across multiple models, including GPT-4, Llama, Gemini, Pixtral, and Claude, according to the number of concepts identified, accuracy, frequency of errors, and teaching size.
- A way to disentangle the effect of the learner's prior knowledge in the concept identification task.

These contributions are generic and can be applied to other problems and modalities. In our particular case, we show that bitmaps are more efficient than coordinates, but surprisingly, the order of complexity between the concepts is preserved to some extent. This suggests that either the representations of both modalities are tightly connected in the latent space of the model, or the simplicity of concepts is an inherent property that transcends modalities.

8.2 Related Work

Drawing (or Sketches) Recognition: Eitz, Hays, and Alexa [EHA12] provided a dataset of human drawings, including 250.00 concepts and 20,000.00 drawings. They introduced a support vector machine model to recognize these drawings and observed that humans outperformed its performance. Since then, AI models have been closer or even achieved higher accuracy than that of human classification for drawing recognition (e.g., Schneider and Tuytelaars [ST14], Yang, Ismail, Pang, Kebande, Al-Dhaqm, et al. [Yan+24], Yu, Yang, Song, Xiang, and Hospedales [Yu+15], and Zhang, Qi Zou, Pei, Zhang, et al. [Zha+20]). Using the *Quick, Draw!* dataset, Ha and Eck [HE17] proposed `sketch-rnn`, a model designed to create drawings of common objects that resemble those drawn by humans. A similar version of this model has also shown capabilities in drawing recognition [Baj17]. Other neural approaches studied for this task include convolutional neural networks [Kab20], and graph neural networks applied over drawings represented as graphs [XJB22].

Drawing Capacities of LLMs: Sharma, Shaham, Baradad, Fu, Rodriguez-Munoz, et al. [Sha+24b] assess the visual abilities of different language models. They conduct experiments that prompt the models to create code that draws images based on text descriptions and improve image generation code iteratively through text feedback. They show that: (a) LLMs possess limited ability to recognize concepts represented in code, and (b) these models sometimes fail to recognize concepts that they can accurately draw. Note that the authors addressed the problem as a multi-class classification problem. Moreover, the online interface for collecting human drawings limits components to basic shapes like ellipses, possibly restricting participants' ability to create complex drawings. In their initial experiments with GPT-4, Bubeck, Chandrasekaran, Eldan, Gehrke, Horvitz, et al. [Bub+23] present an example of drawing generation, showcasing text-to-image capabilities using TikZ. They show tasks such as GPT-4 drawing a unicorn and constructing TikZ code through a multi-step prompt process. In another study, Pourreza, Bhattacharyya, Panchal, Lee, Madan, et al. [Pou+23] introduce the *Painter*, a modified LLM that creates drawings using virtual brush strokes based on user-provided text descriptions. Additionally, Cai, Huang, Li, Ojha, Wang, et al. [Cai+23] evaluated GPT-4's ability to understand visual data in SVG format across various visual tasks, including image classification, visual reasoning, and image generation. Vinker, Shaham, Zheng, Zhao, Fan, et al. [Vin+24] propose *SketchAgent*, showing that while LLMs iteratively generate sketches, they struggle with spatial reasoning.

Machine Teaching: Machine teaching is a research area that focuses on identifying the optimal set of examples that allow a learner (e.g., a human or a machine) to identify a given concept [Zhu+18]. To illustrate the underlying idea of machine teaching, assume the teacher wants the learner to identify the concept of prime numbers. To achieve this, the teacher uses the set $S_1 = \{2, 3, 5, 7, 11, 13\}$ and succeeds. However, would it not be enough for the learner just to see the smaller set $S_2 = \{19, 23\}$? Of course, that depends on the learner. In general, optimal teaching will depend on the model the teacher has of the learner. Machine teaching presents an alternative framework to machine learning (where examples are not chosen but sampled from a distribution) to answer the question of whether some concepts are inherently more complex than others. The connections between machine teaching and computational learning theory are strong; see, e.g., the works by Doliwa, Fan, Simon, and Zilles [Dol+14] or Moran and Yehudayoff [MY16], with machine teaching putting the emphasis on the minimal evidence that distinguishes the concept from all the rest. To determine how easy it is to teach a concept, the teaching dimension [Zhu+18]—the minimum number of examples the learner needs to identify a concept—was traditionally used. Telle, Hernández-Orallo, and Ferri [THF19] introduced a new metric named teaching size. This metric puts the focus on the sum of the sizes of the examples needed to identify a concept, rather than only the number of examples.

8.3 Methods

The drawings used in this work come from the *Quick, Draw!* dataset [HE17; Jon+16], which includes over 50.00 million drawings of 345.00 concepts. Collected by Google Creative Lab via an interactive game, participants had 20.00 seconds to draw a concept while a neural network attempted real-time recognition. The dataset is the largest collection of doodles in the world, with contributions from more than 15.00 million participants.

Each drawing in the Simplified Drawing files that we use is stored as vectors of distinct pen strokes, i.e., distinct continuous movements of the pen without lifting. Each stroke s_i is represented by a sequence of (x, y) coordinates $\{(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \ldots, (x_{in}, y_{in})\}$. Note that each pair of consecutive points in a stroke creates a segment. Additionally, for each drawing, a binary flag r indicates whether the game's neural network correctly recognized the concept.

The following sections cover concept selection, corresponding drawings, learners, the machine teaching setting, and the drawing selection conducted before testing the framework.

8.3.1 Teaching Size

Let D denote an infinite space of possible drawings (and their simplifications, as will be explained later), and let C be a set of concepts. We use D_c to denote all the drawings of a concept $c \in C$. For any given concept $c \in C$, the objective is to identify the simplest drawing $S \in D_c$ (represented as S^m with modality m being either bitmap or coordinates) such that a learner L successfully learns c with a probability of at least ρ over N independent trials (i.e., recognition consistency). The *teaching size (TS)* of c for the modality m can then be defined as follows:

$$\mathrm{TS}_{\rho,N,m}(c) = \min_{S \in D_c} |S^m| \text{ s.t. } \sum_{1}^{N} \mathbb{1}\left[L(S^m) = c\right] \geq \rho \cdot N, \tag{8.1}$$

where $\mathbb{1}[\cdot]$ is the indicator function, which equals 1 if the learner L correctly identifies concept c from the drawing S^m, and 0 otherwise.

We argue that a good metric for assessing the simplicity of a given drawing d can be based on the number of segments it contains. This is represented by $|S^m|$ in the above equation. This metric is intrinsic to the drawing itself, thereby avoiding dependencies on the length or verbosity of the instructions used to generate it, such as in a descriptive language like TikZ.

We also note here that while our implementation of teaching size is grounded in segment count for drawings, the framework itself is more general. Teaching size, as a proxy for descriptive complexity, can be adapted to other domains using modality-appropriate metrics.

8.3.2 Concepts

In our work, if the expected concept is `car` and the identified concept is `police car`, the identification is still considered correct because `police car` is a specific type of `car`, i.e., it is a semantically related prediction. This approach is similar to the one followed by Lamb et al. (2020). This means that if a specific sub-concept, or *hyponym*, is identified, it should still be seen as a correct identification as long as it falls under the more general expected concept. For a concept c, such as `car`, we consider a set of hyponyms $h(c)$ that corresponds to a set of concepts with a more specific meaning than c, e.g., `police car` belongs to $h($`car`$)$. For this study, we want a set of concepts that ensures that in the set of their hyponyms, there is no overlap, i.e., for any two concepts c_i, c_j, we have $h(c_i) \cap h(c_j) = \emptyset$. This rules out certain pairs of concepts available in the *Quick, Draw!*, like `van` and `car`, and it enhances the clarity and robustness of the study. We thus select the following subset of 20.00 concepts from the 345.00 concepts available in *Quick, Draw!*, with no overlap among their hyponyms:

`apple`, `banana`, `car`, `cat`, `computer`, `cup`, `door`, `envelope`, `fish`, `grass`, `hockey puck`, `house`, `key`, `radio`, `string bean`, `sun`, `sword`, `television`, `The Great Wall of China` and `tree`.

In Table 1 in the Appendix E, we list each concept from the dataset and the accepted hyponyms that are considered correct. This correspondence is established by human inspection and after the execution of the drawing selection phase (cf. Sect. 8.3.6) and the machine teaching framework experiments, with the results then analyzed based on these mappings.

8.3.3 Drawings

After choosing the concepts to study, we only include drawings that the game's neural network correctly identified (i.e., $r = 1$) in our research. For every concept, approximately 50.00 drawings are selected by a proportional random stratified sampling method [Tah16], which groups drawings into bins based on their number of segments. (This number is approximate, as there may be rounding errors when calculating the number of samples for each bin according to its proportion.) The bin width was obtained using the minimum bin width between the Sturges's rule and the Freedman–Diaconis Estimator, ensuring that drawings of any concept are represented in a way that reflects the distribution of stroke counts for all correctly identified drawings of that concept in the dataset.

To simplify the drawings in our study, we employ the Ramer–Douglas–Peucker (RDP) algorithm [DP73; Ram72] on each stroke s of a given drawing d. RDP reduces the number of segments in each stroke while preserving its overall shape. Specifically, given a stroke s with a sequence of points $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the RDP algorithm iteratively selects the most distant point (x_d, y_d) from the line segment connecting the first and last points of the stroke. If this distance is below a predefined threshold ϵ, then this stroke is simplified to a single segment $\{(x_1, y_1), (x_n, y_n)\}$ on the first and last points. However, if the distance to (x_d, y_d) exceeds ϵ, the algorithm keeps this point and recursively processes the two sequences of points formed by $\{(x_1, y_1), \ldots, (x_d, y_d)\}$ and $\{(x_d, y_d), \ldots, (x_n, y_n)\}$. This ensures that the essential characteristics of the stroke, up to distance ϵ, are preserved. This process continues until all points in the stroke fall within the threshold, resulting in a simplified representation of the stroke with fewer segments. By incrementing the threshold parameter, from an initial value of $\epsilon = 2$ [1], until each stroke is reduced to one segment, we generate simplified versions of each original drawing associated with a given concept c, resulting in new drawings $\{d\}_\epsilon \subseteq D_c$. Figure 8.2 illustrates a drawing simplification.

We note here that image and coordinate representations are generated differently, but both encode the same visual information. While not equivalent in all respects, the coordinates in TikZ are a form of structured data that, by reflecting a sequence of drawing actions, yield the same shape as the image once rendered.

[1] The strokes stored in the Simplified Drawing files of *Quick, Draw!* have already been simplified by the RDP algorithm using $\epsilon = 2$, so this initial value did not simplify any drawing further.

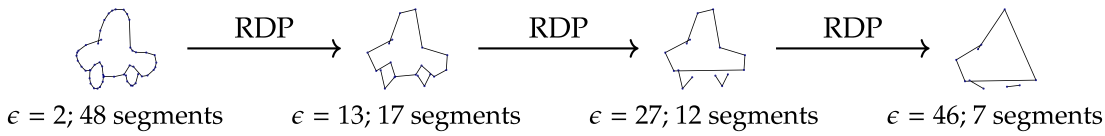


Figure 8.2: Example of a drawing simplification for the concept car using the RDP algorithm. As the value of ϵ increases, the drawings become progressively simpler.

8.3.4 Learners (L)

We utilize multiple LLMs, including two GPT-4 models (`gpt-4-turbo` and `gpt-4o`) from OpenAI, Llama (`Llama-3.2-90b-vision-instruct`) from Meta, Gemini (`gemini pro-1.5`) from Google DeepMind, Pixtral (`pixtral-large-latest`) from Mistral, and Claude (`claude-3-5-sonnet`) from Anthropic. These models are capable of processing visual and language inputs to produce text outputs. To conduct the experiments of this work, all models were accessed via their respective Application Programming Interfaces (APIs). Additionally, we set the temperature parameter T to 1 for the experiments carried out within the machine teaching framework, and we set $T = 0$ for the drawing selection phase. $T \in [0..2]$ controls the behavior of the models' outputs: the lower T is, the more deterministic (predictable) results it leads to [Ope24]. Thus, by setting $T = 0$ in the drawing selection phase, our goal is to obtain deterministic and predictable results, which are essential for creating a consistent baseline of drawings where the concepts were correctly identified. On the other hand, setting $T = 1$ in the experiments of the machine teaching framework is intended to introduce a controlled level of variability.

We consider two different representations for each concept: a visual representation and a text-based representation. Accordingly, we develop and test two prompt templates, one for each modality. For the vision-based modality, the drawings are presented as images generated from the sequence of coordinates (cf. Prompt 1 in the Appendix E). For the text-based modality, the pen stroke vectors are coded using the TikZ language (cf. Prompt 2 in the Appendix E). Both prompts ask for an open-ended answer (not multiple choice), allowing the learners to consider a wide range of possible concepts when identifying a given concept, including any that is not in our 20.00-concept set.

Data contamination occurs when language models are tested and evaluated using information from their training data, such as drawings already seen during training [Rav+24]. However, in this study, the drawings are consistently simplified using the RDP algorithm. This algorithm alters the coordinate information, thereby modifying the TikZ code and the visual representation. Consequently, we argue that these modified drawings are not part of the training set used to train the learners. Therefore, contamination tests are not required for this experiment.

It is important to note that although the models are not trained during our experiments, we refer to them as "learners", since this is aligned with the standardized terminology of machine teaching.

8.3.5 Concept Priors

As we argue in the introduction, some concepts, such as a `house`, are more common than others, such as an `envelope`. This sets a strong prior bias, especially in cases of doubt. For each of the 20.00 concepts, we use the 2022 English corpus of Google Books Ngram [Goo10], providing the prior of a given concept as a normalized number between 0.00 and 1.00, representing the relative frequency of the concept. The rationale for using word frequency from Google Books Ngram as a proxy for human priors lies in the historical and cultural representativeness of a corpus. The assumption underlying our approach is that the frequency of specific words and phrases in written text correlates with their prominence in human thoughts, discussions, and collective knowledge at particular times [TT11]. Given that LLMs are trained on large text corpora that include books, articles, and other written materials, it is reasonable to assume that the Google Books Ngram priors closely align with the priors embedded in LLMs.

The priors were obtained in a case-insensitive manner. Each concept is treated exclusively as a noun to prevent confusion with its verb form (i.e., `fish` is interpreted as the animal and not the fishing activity).

8.3.6 Drawing Selection Phase

Before applying the machine teaching framework, we first conduct a drawing selection phase. This process identifies which drawings are reliably recognized by each model across modalities. These filtered examples form the basis for estimating teaching size. Hence, our minimization of Eq. 8.1 is sufficiently accurate.

As already mentioned, the drawings are simplified using the RDP algorithm, starting with a threshold of $\epsilon = 2$ on the raw drawings and continuing until each stroke in the drawing consists of a single segment. For each ϵ, the learner is prompted using Prompt 1 for visual-based identification and Prompt 2 for text-based identification (cf. Appendix E). Then, based on the completions from the learner, we obtain, by human inspection, the correspondence (between concepts and their respective accepted hyponyms) described in Table 1 in the Appendix E, and we analyze the results based on those mappings. The accuracy and frequency of mistakes for each concept are obtained from the drawing selection phase.

In total, for the drawing selection phase, we run tests on each learner separately, generating a total of $21,896$ prompts—half $(10,948)$ for coordinates and half for images. These prompts were checked by human visual inspection, producing Table 1 (Appendix E). We then use the drawings that are correctly identified to test and evaluate the machine teaching framework proposed in Eq. 8.1, and thus obtain, for each concept, the teaching size.

8.4 Results

8.4.1 Concepts Identified

Out of the 20.00 concepts evaluated, all were identified in the image-based modality by at least one model. However, for the coordinates representation, `television`, `sword`, `radio`, `car`, `door`, `hockey puck`, `string bean`, and `The Great Wall of China` were never recognized by any model. We hypothesize that not only the complexity but also the prior of each of these latter concepts is behind their failed identification.

The image-based modality is thus more effective than the coordinate-based modality in identifying a broader range of concepts. This observation aligns with the typical human learning patterns, where visual information is often easier to process and understand than abstract textual-numerical data.

8.4.2 Accuracy

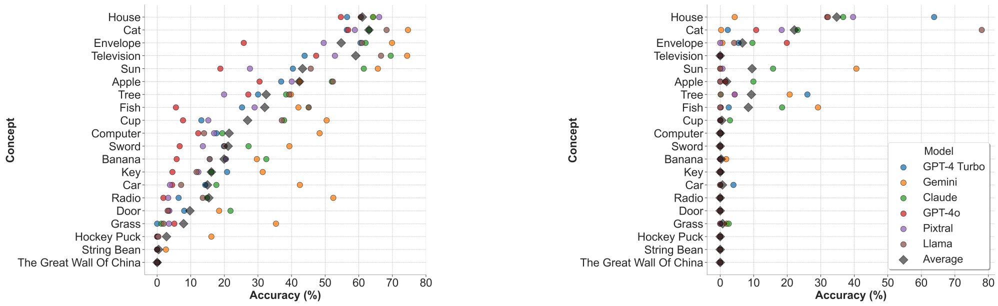


Figure 8.3: Accuracy for each concept in the vision-based (images; left) and text-based (coordinates; right) modality representations. ♦ represents the average accuracy value for the concept.

We begin by evaluating the accuracy on each concept c, Accuracy(c), defined here as

$$\text{Accuracy}(c) = \frac{1}{N_c} \sum_{i=1}^{N_c} \mathbb{1}\left[L(S_i) = c\right], \tag{8.2}$$

where N_c corresponds to the total number of tests (in this case, prompts) conducted on L for the concept c on the drawing selection phase, with $\{S_i\}_{i=1}^{N_c} \subseteq D_c$.

Figure 8.3 depicts each concept's accuracy across the two modality representations. The image modality shows a wider range of recognition accuracy, with average performance metrics spanning up to 65%. In contrast, the coordinate modality exhibits a much narrower range, largely confined to 0–25% average accuracy. This discrepancy likely reflects the models' ability to leverage richer visual features in image-based representations compared to the sparse and abstract nature of coordinate-based inputs. The richer detail in images provides more cues for concept identification, while the textual coordinates impose a more constrained and abstract recognition task.

Nevertheless, the results suggest that some concepts are fundamentally challenging to recognize, regardless of the modality. `House` and `cat` achieve relatively high accuracy across both modalities, indicating their simplicity or recognizability regardless of representation. In contrast, more complex or less visually distinct concepts, such as `hockey puck` and `The Great Wall of China`, show zero accuracy in the coordinate modality and only marginal performance in the image modality.

Among the models evaluated, Gemini emerges as the best-performing model in the image modality, consistently achieving better results across a broader range of concepts. Notably, it stands apart from the other models, which appear to form a distinct cluster in terms of performance. This suggests that while certain concepts are uniformly challenging across all models, Gemini is better equipped to handle a wider range of visual representations. In the coordinate modality, no single model shows clear superiority, likely due to the shared constraints of the textual representation.

The precise accuracy of each concept, categorized by model and modality, can be found in Table 2 in the Appendix E.

We also study the relationship between the number of segments (i.e., complexity) and the accuracy of concept identification for both image- and coordinate-based representations, as shown in Figure 8.4. For image-based representations, there is a clear positive relationship between accuracy and the number of segments. Starting from an accuracy of around 0.30 % in the $(0, 4]$ interval, the accuracy increases steadily, reaching approximately 50.00 % in the $(29, 69]$ interval.

Conversely, for coordinate-based representations, the average accuracy remains significantly lower and follows a more modest increasing trend. Beginning at roughly 1.00 %, it gradually rises to around 8.00 % in the $(16, 19]$ interval before stabilizing and fluctuating slightly in the higher segment intervals. This indicates that increasing the number of segments in coordinate-based representations provides only minimal benefits in accuracy.

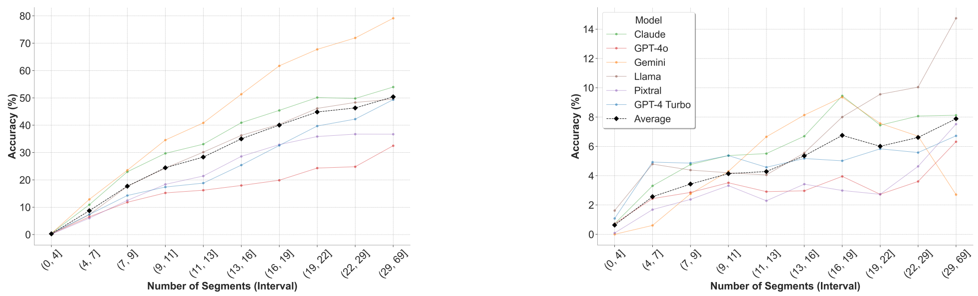


Figure 8.4: Relationship between the number of segments and accuracy for both modalities (images; left) (coordinates; right).

8.4.3 Frequency of Mistakes

Accuracy measures how well the learner has identified the correct concepts. However, the model can also respond with "I don't know" answers (or something that is not a concept) or by identifying a different concept that is incorrect. We focus on the latter case and refer to this performance metric as the *frequency of mistakes* for a given concept c in model m, $\mathrm{FOM}_m(c)$.

Formally,

$$\mathrm{FOM}_m(c) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\left[S_i \notin D_c \wedge L_m(S_i) = c\right], \tag{8.3}$$

where $N = 10,948$ is the total number of tests (prompts) conducted on L during the drawing selection phase on each modality.

To simplify the interpretation of the frequency of mistakes for a given concept, we average the $\mathrm{FOM}(c)$ across all models. We also explore whether there is a relationship between the frequency of mistakes and the prior probability of each concept. We have included in Tables 6 to 17 of the Appendix E the confusion matrices for each model and modality. These tables show how well the model performs across various concepts by detailing the true positives and the frequency of errors for each concept. Figure 8.5 shows that the vision modality exhibits a lower percentage of observed mistakes than the coordinate-based modality.

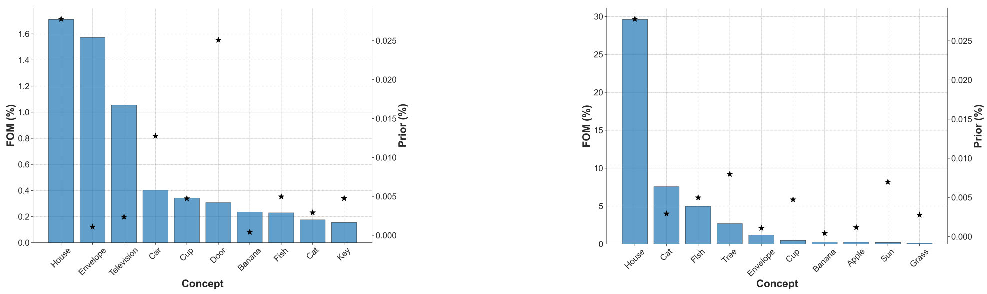


Figure 8.5: Top-10 concepts with the highest frequency of mistakes (averaged across the models) in the visual-based modality (images) (left) and text-based modality (coordinates) (right). The little star represents the prior probability for each concept.

Interestingly, as shown in Figure 8.3, the concept `house` in both modality representations, `television` only in the visual-based modality, and `cat` only in the text-based modality, shows the highest accuracy. However, these concepts also have the highest frequency of mistakes, indicating that while they are often correctly identified, they are also frequently guessed when wrong. This indicates that although these concepts are generally easily recognizable, variations in attributes like size and shape may introduce ambiguities that complicate the identification of these concepts. In other words, the models often guess these concepts, whether they are correct or not.

When calculating the Pearson correlation between the frequency of mistakes and the prior probability, we obtain a correlation of 0.91 for the coordinate-based modality and 0.43 for the vision-based modality for all concepts. This suggests that in the textual modality, the learner is more susceptible to responding based on their pre-existing biases when confronted with unfamiliar concepts. In contrast, this tendency is reduced in visual representation.

8.4.4 Teaching Size

To calculate the teaching size for each concept, we set T to 1, ρ to 0.50, and N to 50.00, meaning that a correct identification needs to happen at least 25 times out of 50 trials even with some stochasticity in the model. The aim is to determine the simplest drawing for each modality representation that the learner can identify consistently in at least 25.00 out of 50.00 trials. We highlight that this procedure is different from the one conducted in the previous sections, where the results came from the drawing selection phase.

We present the results for teaching size of images and coordinates in Tables 3 and 4 in the Appendix E. Table 5 of the Appendix E shows the respective simplest drawings identified for each concept, modality and model. The data suggest that, on average, the teaching size values for coordinates (11.46, SD=8.60) with successful identification (12.00) are higher than those for images (6.73, SD=2.25) with successful identification (20.00), regardless of the model. Even when considering only the 12 concepts that are well identified using coordinates, the mean teaching size remains lower for images. This indicates that there is no absolute invariance, answering our question Q1 in the negative. In other words, the number of strokes required for a concept to be identified by the learners is generally higher when using textual coordinates compared to bitmap images.

Table 8.2: Concept teaching size comparison for images and for coordinates, showing Kendall Rank correlation coefficient for the subset of concepts that are identified (*), and Pearson correlation between the accuracy for all concepts.

Model	Order for Images	Order for Coordinates	Rank*	Pears
Claude	`cup` < `house` = `fish` < `envelope` < `apple` = `sun` < `cat` < `grass`	`house` < `envelope` < `apple` < `fish` < `cat` < `sun` < `cup` < `grass`	0.36	0.65
Gemini	`envelope` = `house` = `sun` = `grass` = `fish` < `banana` < `tree` = `cat`	`envelope` < `house` < `sun` = `tree` < `fish` < `banana` < `grass` < `cat`	0.57	0.21
GPT-4o	`tree` < `house` < `envelope` < `apple` < `cat`	`envelope` < `house` < `cat` < `tree` < `apple`	0.00	0.67
GPT-4T	`envelope` < `house` < `fish` < `cat` < `tree` < `car`	`envelope` = `house` < `fish` = `tree` < `cat` < `car`	0.87	0.45
Llama	`envelope` = `house` < `cat`	`envelope` = `house` < `cat`	1.00	0.50
Pixtral	`house` < `cat`	`house` < `cat`	1.00	0.63

Furthermore, it is important to highlight a weak, though similar, negative correlation between the teaching size and the prior of each concept across both modalities. The correlation coefficients are -0.02 for coordinates and -0.34 for images, over all concepts and models. This suggests that, in the image modality, the more common a concept is, the simpler its drawings need to be for the learner to consistently identify it.

Interestingly, looking at Table 8.2, the teaching size still ranks concepts in a relatively similar order between images and coordinates, but the strength of this relationship varies across models. The strongest agreement is observed in Llama and Pixtral, both of which exhibit a perfect Kendall rank correlation of 1.0, meaning their rankings are identical across the two modalities. GPT-4 Turbo (shortened as GPT-4T) also exhibits a high correlation (0.87), suggesting strong alignment in concept difficulty ordering between images and coordinates.

However, other models show lower correlations, with Claude at 0.36, Gemini at 0.57, and GPT-4o displaying no correlation (0.0) between the two rankings. To control for the influence of concept priors on teaching size, we performed an ordinary least squares regression of the teaching sizes (for both modalities) on the corresponding concept priors derived from Google Books Ngram frequencies. This yielded residuals representing the portion of teaching size not explained by prior familiarity. We then calculated the Kendall rank correlation between these residuals from different modalities and found similar correlation values.

These results indicate that while some models maintain an invariant notion of teaching size across modalities, others exhibit some discrepancies, although the number of concepts is small. The accuracy correlation between all concepts is a more robust metric, and it also calculates how well concept-wise accuracies align between the two modalities. Claude and GPT-4o exhibit relatively high accuracy correlations (0.65 and 0.67, respectively), suggesting that despite their lower Kendall rank correlations, the overall accuracy patterns remain similar. Meanwhile, Gemini and GPT-4T have lower accuracy correlations (0.21 and 0.45), compensated by the better values for ranking.

Overall, the correlations are never negative, but less or more positive depending on the model. While Llama, Pixtral, and GPT-4T exhibit strong invariance in teaching size ranking across modalities, others do not. In general, however, the answer to question Q2 tends to be positive.

8.5 Discussion

In this study, we examined how multimodal models identify the same concepts in two different modalities: image- and coordinate-based drawings. Our findings show that images are generally more effective than coordinates for identifying concepts. In particular, using images led to the recognition of more concepts than when using coordinates, indicating that images are better suited for teaching concepts to a given learner. This is supported by the higher accuracy and lower frequency of mistakes seen with image-based representations. We also use the number of segments as the teaching size to measure the complexity of a concept. Our analysis indicates that the teaching size is again more beneficial for images than coordinates (clearly answering question Q1 negatively), but ranks concepts in similar ways, regardless of the type of drawing used, even when we account for the learner's priors. While there are differences depending on the model, we tend to see a positive answer to question Q2 more often. This suggests that some concepts are naturally easier or more difficult to teach, no matter how they are represented.

We believe that our study provides a step towards the investigation of a core question in the field of multimodal AI: Whether language models can interpret structured data (like coordinates) as effectively as images. We saw that models perform better with image-based representations, even for simple concepts. This suggests a limitation in current multimodal models that is important for scientific and practical development. The observed invariance in ranking teaching size across modalities suggests that some concept properties are robust regardless of representation. This may help improve cross-modal transfer learning, where models must generalize concepts between formats.

Our machine teaching framework has several practical implications. First, it improves the design and evaluation of multimodal systems by providing a quantitative, model-agnostic way to measure how "costly" it is for any vision-language model to learn a new visual concept in different modalities. Second, our work connects cognitive and computational notions of simplicity by providing empirical evidence that segment count, a classic cognitive cue, remains predictive even in state-of-the-art large language models. This contributes to ongoing discussions in AI about whether these models learn conceptual structures or simply memorize patterns. Finally, the framework can support adaptive teaching tools by identifying the simplest representations for individual learning needs. Thus, it could be used to develop educational software that teaches geometric concepts or visual reasoning using minimal and optimally chosen examples.

Our analysis has to be seen in the light of some limitations. (a) The study concentrates on a specific set of concepts, which might affect how well the findings apply to other (potentially more complex) concepts. (b) Our use of the RDP algorithm for drawing simplification streamlines each stroke but does not totally remove any single stroke from the drawing. This should not be much of a limitation as we focus on the simplest concepts. (c) A factor that can influence the teaching size of a concept is the curvature of its drawings, i.e., the amount by which it deviates from a straight line. In this work, we have chosen not to focus on this aspect, but this could be of interest for future works.

We show that the simplest concepts usually correspond to those that humans intuitively think of as less complex, and this confirms that the simplest concepts are so across modalities. This supports the hypothesis that the representation of concepts in both modalities is tightly connected in the latent space. However, since we operate under a black-box setting with models like GPT-4 and others that do not expose their internal representations, we cannot directly inspect or confirm such latent alignments. Some other methods, especially white-box approaches that have access to weights or gradients, could give a definitive answer to this hypothesis. Still, in cases such as GPT-4 or humans, a black-box approach such as the one presented in this paper is the practical course of action. Thus, our results should be viewed as a hypothesis to explain the invariance across modalities and not as a definitive claim.

The code to reproduce our results is available [Fre+25b].

CHAPTER 9

General Discussion and Future Steps

This thesis explores user profiling methodologies, moving from static descriptions to dynamic, explainable, and causal models. We apply these approaches across three domains—marine litter, sports injuries, and energy poverty—and investigate their potential to generate actionable insights. Despite the distinct contexts, our common goal is to identify shared strategies for profiling users, selecting relevant features, and delivering explainable outputs (e.g., feature importance, SHAP values, or counterfactuals) to support informed decision-making.

Indeed, as illustrated in Figure 9.1, this thesis can be seen as a progression from simple prediction to a deeper understanding of user profiles, achieved by integrating methods for feature selection, explainability (SHAP and counterfactuals), and causal discovery. Furthermore, this work also addresses the challenge of profile simplification by using a machine teaching framework to explore the nature of concept complexity, using drawings as an analogy for user profiles.

In this view, across the different domains, a consistent methodological pipeline emerges. First, user profiling gathers a rich set of attributes to form a static (Chapter 3) or dynamic (Chapter 4, Chapter 5) representation of the user. Next, feature selection methods (e.g., mRMR) isolate the most relevant attributes. The framework then moves from prediction to understanding. Explainability techniques, such as SHAP values (Chapter 5) and counterfactuals (Chapter 6), clarify *why* a prediction was made and *how* to change it. This is complemented by causal discovery (Chapter 7), which identifies the underlying cause-and-effect relationships. The final step in the pipeline explores the limits of profile simplification itself, using machine teaching (Chapter 8) to understand how user profiles can be simplified.

In Chapter 1, we identify seven research questions (RQ1–RQ7). This discussion returns now to those questions, synthesizing the main contributions while considering new research possibilities.

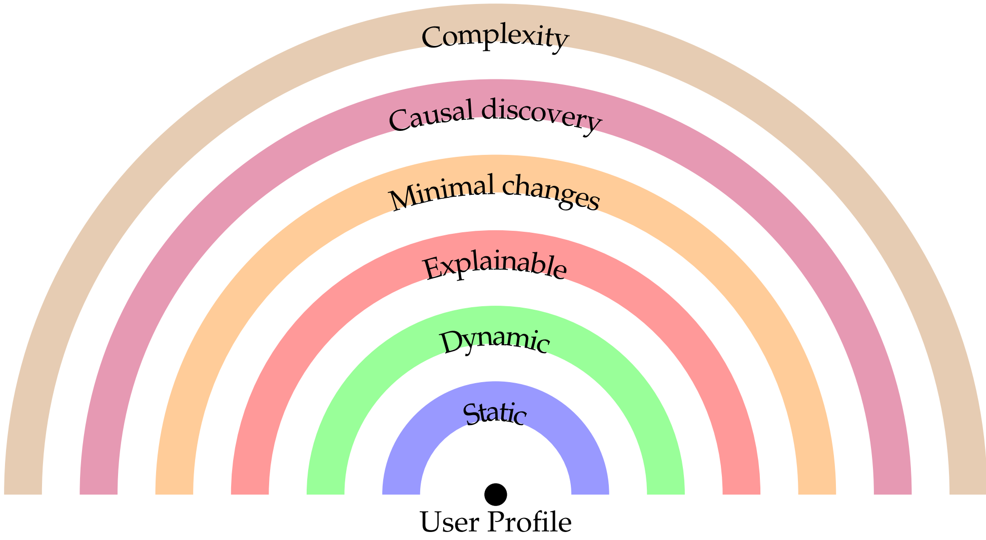


Figure 9.1: Overview of the progressive user profiling framework developed in this work. Each layer builds upon the previous one, starting with static user profiles (Chapter 3) and progressing to dynamic user profiles (Chapter 4). The framework is then enriched with explainability methods, including feature importance (Chapter 5) and counterfactuals for minimal changes (Chapter 6). It is further extended by incorporating causal discovery (Chapter 7) and finally explores concept complexity (Chapter 8).

Marine Litter Literacy (Chapter 3). Here, the focus was primarily on *static* profiling and segmenting individuals who might benefit most from awareness campaigns. The findings revealed how demographic factors (e.g., age, education level) and attitudinal measures (e.g., whether individuals perceive marine litter as a "distant" threat) can predict a lower engagement or misunderstanding of marine pollution topics. This directly addressed RQ1 (using static user attributes to identify those requiring interventions) and RQ2 (features most relevant for awareness and knowledge). The main finding was the identification of four distinct population profiles, revealing that many individuals, particularly the young, do not perceive marine litter as an immediate threat.

Football Injury Prediction (Chapter 4). This work focused, in contrast with the previous work, on *dynamic* profiles using longitudinal sensor data (GPS records) to forecast non-contact injuries. While examining RQ2 (critical features), it became evident that the player's position, session type, player load, velocity, and acceleration were the key factors associated with injury occurrences. Moreover, when answering RQ3 (the effect of feature selection), the results showed that incorporating additional features did not necessarily enhance the models' ability to identify at-risk user profiles accurately. It is also important to note that this work suggested incorporating dummy days, allowing the models to capture sudden changes in the players' profiles, possibly increasing injury risks. The key finding was that machine learning models using GPS data could predict more than 70% of non-contact injuries. The most important predictive features were player load, velocity bands, and session type, and the study showed that effective feature selection was an important step for model performance.

Energy Poverty: Explainability, Causality, and Interventions (Chapters 5 to 7). The energy poverty domain serves as a case study for a multi-faceted analysis. First, to answer RQ2–RQ4, the work in Chapter 5 found that historical income levels and income volatility are the most critical long-term predictors of energy poverty. Second, to provide actionable solutions (RQ5), the analysis in Chapter 6 identified that modest income increases (often $< 5\%$) are the most effective minimal intervention. Finally, addressing RQ6, the study in Chapter 7 revealed a temporally asymmetric causal link: poor health immediately leads to energy poverty, whereas energy poverty has delayed effects on health.

Concept Complexity and Profile Simplification (Chapter 8). Finally, the thesis explores a different challenge related to the limitations of simplifying profiles. The work in Chapter 8 answers RQ7 by using a machine teaching framework and a drawing analogy to investigate if concept complexity is an inherent property that is stable across different data representations. The main finding from this work is that while image-based representations are more efficient, the relative difficulty of concepts remains largely consistent across different data formats (e.g., images and text). This suggests that concept complexity may be an inherent property, relevant to profile simplification.

The findings of this study have important implications for policy-oriented practitioners, sports science professionals, and governmental entities. On the environmental side, governmental or non-profit campaigns aiming to reduce marine litter can use the segmentation model to pinpoint specific user profiles (e.g., communities or demographic groups) with pronounced knowledge gaps, thus ensuring more focused and efficient literacy interventions. Within sports science, integrating the dynamic user profiling tools proposed in this work can help coaches and medical staff implement proactive injury-prevention strategies. By understanding the most relevant attributes that prevent non-contact injuries, stakeholders (e.g., the coaching staff) can focus on this subset of attributes (instead of considering all parameters), and thus, adapt training loads and manage athlete recovery more precisely, mitigating the overall risk of physical harm. Finally, the case of energy poverty also has positive implications for the public policy sphere. Identifying at-risk households and suggesting minimal yet feasible profile modifications helps shift the paradigm from reactive measures toward forward-looking and preventive efforts. The study highlights how predictive modeling and explainable methods can suggest actionable recommendations, which encourage more sustainable actions and, thus, reduce energy poverty. The causal analysis complements this by suggesting the need for safety nets against health shocks that trigger poverty, and framing long-term energy policy as a public health tool. The work on concept complexity has practical implications for designing better cross-modal learning systems. Moreover, it provides a framework to measure the efficiency of different user representations and contributes to understanding whether different user representations influence the model's performance.

We highlight, however, that the empirical findings are based on specific datasets. This includes a regional survey for marine litter, sensor data from a single football club, and a national panel for Australia. The generalizability of the specific models and feature importance values to other contexts would require further validation. Methodologically, the causal discovery analysis relied on assumptions of linearity and causal sufficiency, which may not fully hold in complex real-world systems. Furthermore, the exploration of concept complexity was based on an analogy, and the direct transfer of its findings to the simplification of real-world user profiles remains a topic for future research.

Future Work

For marine litter literacy (Chapter 3), future work should focus on applying the marine litter profiling tool to other geographical contexts to assess its generalizability and adapt educational strategies accordingly. One natural step could also be to validate these profiles and create specific intervention strategies for field testing. A pre–post evaluation could be then conducted to assess changes in knowledge, attitudes, and behaviors, as well as shifts in individual user profiles.

Regarding football injury prediction (Chapter 4), future research should aim to test the models with larger and more diverse datasets, including data from additional seasons or different cohorts (e.g., U23 and U19 teams). Exploring more complex models, such as recurrent neural networks (e.g., LSTMs), and incorporating additional physiological parameters (e.g., history of prior injuries, rating of perceived exertion) would also be important. Furthermore, expanding data collection to include additional sensors and modalities (e.g., infrared thermography cameras and rate of perceived exertion records) would enrich the feature space and may improve predictive performance.

In the domain of energy poverty analysis (Chapters 5 to 7), future work should address the endogeneity of life events, potentially using more advanced econometric or AI techniques. It is also important to explore the heterogeneity in household responses by conducting separate analyses for different socio-economic groups or clusters. Testing the methodology on additional datasets from other regions and contexts is necessary to assess the generalizability of the findings. For causal discovery, investigating non-linear causal models would provide a better understanding of complex relationships.

The work on concept complexity and profile simplification (Chapter 8) can be extended by testing with a wider range of concepts and exploring different simplification algorithms (e.g., stroke removal). A potential direction is to move from the drawing analogy to experiments simplifying real-world user profile data.

In geneal terms, future work on user profiling should focus on (i) defining standardized data pattern structures to integrate data from different regions and sources (e.g., surveys, sensors, administrative records). This can include the use of ontologies and interoperable pipelines. However, if data cannot be centralized, use privacy-preserving methods like federated learning to allow collaborative modeling without sharing raw data; (ii) studying procedures for updating dynamic profiles, specifically detailing when these profiles should be updated/recalculated; (iii) integrating multi-modal data (e.g., text, images and sensors) for profiling tasks; and (iv) investigating how large language models could to support multilingual profiling, and how to mitigate cold-start conditions by using, e.g., few-shot priors.

Bibliography

[AJL21] K. Aas, M. Jullum, and A. Løland. “Explaining individual predictions when features are dependent: More accurate approximations to Shapley values”. In: *Artificial Intelligence* 298 (2021), p. 103502. DOI: doi.org/10.1016/j.artint.2021.103502 (cit. on p. 95).

[Abb+20] K. Abbas et al. “Do socioeconomic factors determine household multidimensional energy poverty? Empirical evidence from South Asia”. In: *Energy Policy* 146 (2020), p. 111754. DOI: 10.1016/j.enpol.2020.111754 (cit. on pp. 88, 89, 110, 111, 125, 135).

[AX13] A. Abdel-Hafez and Y. Xu. “A survey of user modelling in social media websites”. In: *Computer and Information Science* 6.4 (2013), pp. 59–71. DOI: 10.5539/cis.v6n4p59 (cit. on pp. 11, 16, 18).

[Abe+13] R. Abecasis et al. “Implications of community and stakeholder perceptions of the marine environment and its conservation for MPA management in a small Azorean island”. In: *Ocean Coast Manag.* 84 (2013), pp. 208–219. DOI: 10.1016/j.ocecoaman.2013.08.009 (cit. on p. 58).

[AA14] F. Achemoukh and R. Ahmed-Ouamer. “Representation and evolution of user profile in information retrieval based on Bayesian approach”. In: *Proceedings of the 21st International Symposium on Methodologies for Intelligent Systems (ISMIS)*. Roskilde, Denmark, 2014, pp. 486–492. DOI: 10.1109/isko-maghreb.2013.6728112 (cit. on pp. 28, 32).

[Ach+23] J. Achiam et al. *GPT-4 Technical Report*. 2023. DOI: 10.48550/arXiv.2303.08774. arXiv: 12303.08774 [cs.CL] (cit. on p. 140).

[Ada20] A. Adam. “Sample size determination in survey research”. In: *J. Sci. Res. Reports* 26 (2020), pp. 90–97. DOI: 10.9734/jsrr/2020/v26i530263 (cit. on p. 38).

[AT05] G. Adomavicius and A. Tuzhilin. “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions”. In: *IEEE Transactions on Knowledge and Data Engineering* 17.6 (2005), pp. 734–749. DOI: 10.1109/TKDE.2005.99 (cit. on pp. 12, 15, 20, 27–30).

[Afc+22] D. Afchar et al. "Explainability in Music Recommender Systems". In: *AI Magazine* 43.2 (2022), pp. 190–208. DOI: 10.1002/aaai.12056 (cit. on p. 30).

[Aiv+19] M. Aivazoglou et al. "A fine-grained social network recommender system". In: *Social Network Analysis and Mining* 10.1 (2019), 1–18, article no.: 8. DOI: 10.1007/s13278-019-0621-7 (cit. on p. 26).

[AGM85] C. Alchourrón, P. Gärdenfors, and D. Makinson. "On the logic of theory change: Partial meet contraction and revision functions". In: *Journal of Symbolic Logic* 50.2 (1985), pp. 510–530. DOI: 10.1007/978-3-319-20451-2_13 (cit. on p. 25).

[AD20] Y. Alem and E. Demeke. "The persistence of energy poverty: A dynamic probit analysis". In: *Energy Economics* 90 (2020), p. 104789. DOI: 10.1016/j.eneco.2020.104789 (cit. on p. 88).

[All+06] M. Allsopp et al. *Plastic Debris in the World's Oceans*. 2006 (cit. on p. 37).

[AA24] Y. Alomari and M. Andó. "SHAP-based insights for aerospace PHM: Temporal feature importance, dependencies, robustness, and interaction analysis". In: *Results in Engineering* 21 (2024), p. 101834. DOI: 10.1016/j.rineng.2024.101834 (cit. on p. 85).

[AS99] G. Amato and U. Straccia. "User profile modeling and applications to digital libraries". In: *Proceedings of the 3rd International Conference on Theory and Practice of Digital Libraries (TPDL)*. Paris, France, 1999, pp. 184–197. DOI: 10.1007/3-540-48155-9_13 (cit. on p. 8).

[Ami+15] B. Amini et al. "A reference ontology for profiling scholar's background knowledge in recommender systems". In: *Expert Systems with Applications* 42.2 (2015), pp. 913–928. DOI: 10.1016/j.eswa.2014.08.031 (cit. on p. 17).

[Amm+19] M. Ammad-ud-din et al. *Federated Collaborative Filtering for Privacy-Preserving Personalized Recommendation System*. 2019. arXiv: 1901.09888 [cs.IR] (cit. on p. 12).

[AC19] M. Araújo and M. Costa. "A critical review of the issue of cigarette butt pollution in coastal environments". In: *Environ. Res.* 172 (2019), pp. 137–149. DOI: 10.1016/j.envres.2019.02.005 (cit. on p. 53).

[On-+16] S. On-At et al. "Taking into account the evolution of users social profile: Experiments on Twitter and some learned lessons". In: *Proceedings of the 10th IEEE International Conference on Research Challenges in Information Science (RCIS)*. Grenoble, France, 2016. DOI: 10.1109/rcis.2016.7549325 (cit. on p. 18).

[Aus24] Australian Bureau of Statistics. *Consumer Price Index, Australia*. (Accessed on 29 March 2024). 2024. URL: https://www.abs.gov.au/statistics/economy/price-indexes-and-inflation/consumer-price-index-australia/mar-quarter-2024#data-downloads (cit. on p. 90).

[ABR22] A. Awan, F. Bilgili, and D. B. Rahut. "Energy poverty trends and determinants in Pakistan: Empirical evidence from eight waves of HIES 1998–2019". In: *Renewable and Sustainable Energy Reviews* 158 (2022), p. 112157. DOI: 10.1016/j.rser.2022.112157 (cit. on pp. 85, 88, 89, 109, 110).

[AS21] S. Awaworyi Churchill and R. Smyth. "Energy poverty and health: Panel data evidence from Australia". In: *Energy Economics* 97 (2021), p. 105219. DOI: 10.1016/j.eneco.2021.105219 (cit. on pp. 123, 124).

[Aya+19] F. Ayala et al. "A preventive model for hamstring injuries in professional soccer: Learning algorithms". In: *Int. J. Sports Med.* 40.5 (2019-05), pp. 344–353. DOI: 10.1055/a-0826-1955 (cit. on pp. 63, 72, 79).

[Bah16] R. Bahr. "Why screening tests to predict injury do not work—and probably never will...: A critical review". In: *Br. J. Sports Med.* 50.13 (2016-06), pp. 776–780. DOI: 10.1136/bjsports-2016-096256 (cit. on p. 63).

[Baj17] Bajaj, Payal. *The Quick, Draw! - A.I.* https://github.com/payalbajaj/sketch_rnn_classification. Accessed: 2024-08-02. 2017 (cit. on p. 141).

[BNR16] L. Balafoutas, N. Nikiforakis, and B. Rockenbach. "Altruistic punishment does not increase with the severity of norm violations in the field". In: *Nat. Commun.* 7 (2016), p. 13327. DOI: 10.1038/ncomms13327 (cit. on p. 56).

[BRT00] A. Ballance, P. Ryan, and J. Turpie. "How much is a clean beach worth? The impact of litter on beach users in the Cape Peninsula, South Africa". In: *South Afr. J. Sci.* 96 (2000), pp. 210–213 (cit. on p. 54).

[Bal+22] V. Ballesteros-Arjona et al. "What are the effects of energy poverty and interventions to ameliorate it on people's health and well-being?: A scoping review with an equity lens". In: *Energy Research & Social Science* 87 (2022), p. 102456. DOI: 10.1016/j.erss.2021.102456 (cit. on p. 109).

[BMM21] R. Banerjee, V. Mishra, and A. A. Maruta. "Energy poverty, health and education outcomes: Evidence from the developing world". In: *Energy Economics* 101 (2021), p. 105447. DOI: 10.1016/j.eneco.2021.105447 (cit. on pp. 87, 110, 124, 136).

[BL24] Y. Bao and T. Liao. "Multidimensional poverty and growth: Evidence from India 1998–2021". In: *Economic Modelling* 130 (2024), p. 106586. DOI: 10.1016/j.econmod.2023.106586 (cit. on pp. 87, 110, 124).

[BG15] L. Barboza and B. Gimenez. "Microplastics in the marine environment: current trends and future perspectives". In: *Mar. Pollut. Bull.* 97 (2015), pp. 5–12. DOI: 10.1016/j.marpolbul.2015.06.008 (cit. on p. 53).

[Bar+18] S. Barnett et al. *2017 G20 Hamburg Summit Final Compliance Report*. Tech. rep. Toronto: Center for International Institutions Research, 2018. URL: http://www.g20.utoronto.ca/compliance/2017hamburg-final/2017-g20-compliance-final.pdf (cit. on p. 37).

[Bar+14] D. J. Barron et al. "Accelerometer derived load according to playing position in competitive youth soccer". In: *Int. J. Perform. Anal. Sport* 14.3 (2014-04), pp. 734–743. DOI: 10.1080/24748668.2014.11868754 (cit. on p. 67).

[BBC22] A. Ben Hassen, S. Ben Ticha, and A. H. Chaibi. "Deep learning for visual-features extraction based personalized user modeling". In: *SN Computer Science* 3.4 (2022), 1–13, article no.: 261. DOI: 10.1007/s42979-022-01131-y (cit. on pp. 15, 17–20).

[BL07] J. Bennett and S. Lanning. "The Netflix prize". In: *Proceedings of the 13rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. San Jose, CA, USA, 2007, pp. 3–6 (cit. on p. 18).

[Ben+12] P. N. Bennett et al. "Modeling the impact of short-and long-term behavior on search personalization". In: *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (ICTIR)*. Portland, OR, USA, 2012, pp. 185–194. DOI: 10.1145/2348283.2348312 (cit. on p. 22).

[Ben+23] R. Bentley et al. "The effect of energy poverty on mental health, cardiovascular disease and respiratory health: a longitudinal analysis". In: *Lancet Reg Health West Pac* 35 (2023), p. 100734 (cit. on p. 109).

[BKR07] S. Berkovsky, T. Kuflik, and F. Ricci. "Cross-Domain Mediation in Collaborative Filtering". In: *Proceedings of the 11th International Conference on User Modeling (UM)*. Corfu, Greece: Springer Berlin Heidelberg, 2007, pp. 355–359. DOI: 10.1007/978-3-540-73078-1_44 (cit. on p. 27).

[BB19a] G. Besagni and M. Borgarello. "The socio-demographic and geographical dimensions of fuel poverty in Italy". In: *Energy Research & Social Science* 49 (2019), pp. 192–203. DOI: 10.1016/j.erss.2018.11.007 (cit. on p. 136).

[BB19b] R. Best and P. J. Burke. "Factors Contributing to Energy-related Financial Stress in Australia". In: *Economic Record* 95.311 (2019), pp. 462–479. DOI: 10.1111/1475-4932.12504 (cit. on p. 125).

[BCC21] S. Bettencourt, S. Costa, and S. Caeiro. "Marine litter: a review of educative interventions". In: *Mar. Pollut. Bull.* 168 (2021), p. 112446. DOI: 10.1016/j.marpolbul.2021.112446 (cit. on p. 38).

[Bet+23] S. Bettencourt et al. "Public perceptions, knowledge, responsibilities, and behavior intentions on marine litter: Identifying profiles of small oceanic islands inhabitants". In: *Ocean & Coastal Management* 231 (2023), p. 106406 (cit. on p. 6).

[Bob+13] J. Bobadilla et al. "Recommender systems survey". In: *Knowledge-Based Systems* 46 (2013), pp. 109–132. DOI: 10.1016/j.knosys.2013.03.012 (cit. on pp. 1, 9, 26–28).

[Bol15] O. Bolarinwa. "Principles and methods of validity and reliability testing of questionnaires used in social and health science researches". In: *Niger. Postgrad. Med. J.* 22 (2015), pp. 195–201. DOI: 10.4103/1117-1936.173959 (cit. on p. 39).

[Bol+08] P. Boldi et al. "The query-flow graph: Model and applications". In: *Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM)*. Singapore, 2008, pp. 609–618. DOI: 10.1145/1458082.1458163 (cit. on p. 13).

[BB12] H. Boone and D. Boone. "Analyzing Likert data". In: *J. Ext.* 50 (2012), pp. 1–5 (cit. on p. 40).

[Boţ+24] C. Boţa-Avram et al. "Exploring the impact of macro-determinant factors on energy resource depletion: Evidence from a worldwide cross-country panel data analysis". In: *Energy Economics* 130 (2024), p. 107341. DOI: 10.1016/j.eneco.2024.107341 (cit. on pp. 87, 110, 125).

[Bro10] G. Brown. "Ensemble Learning". In: *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010, pp. 312–320. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_252 (cit. on p. 72).

[BV21] H. Brown and E. Vera-Toscano. "Energy poverty and its relationship with health: Empirical evidence on the dynamics of energy poverty and poor health in Australia". In: *SN Business & Economics* 1.10 (2021), p. 139. DOI: 10.1007/s43546-021-00149-3 (cit. on pp. 124, 135).

[BRT10] S. Brown, J. Roberts, and K. Taylor. "Reservation wages, labour market participation and health". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173.3 (2010), pp. 501–529. DOI: 10.1111/j.1467-985X.2009.00638.x (cit. on pp. 127, 136).

[Bro90] D. Browne. *Adaptive User Interfaces*. London, United Kingdom: Academic Press, 1990 (cit. on pp. 1, 9, 32).

[Bro20] J. Brownlee. *Data Preparation for Machine Learning: Data Cleaning, Feature Selection, and Data Transforms in Python*. Machine Learning Mastery, 2020-06 (cit. on pp. 67, 68).

[BKN07] P. Brusilovski, A. Kobsa, and W. Nejdl. "User models for adaptive hypermedia and adaptive educational systems". In: *The Adaptive Web: Methods and Strategies of Web Personalization*. Vol. 4321. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2007, pp. 3–53. DOI: 10.1007/978-3-540-72079-9_1 (cit. on pp. 11, 13, 18–20).

[Bub+23] S. Bubeck et al. *Sparks of Artificial General Intelligence: Early Experiments with GPT-4*. 2023. DOI: 10.48550/arXiv.2303.12712. arXiv: 2303.12712 [cs.CL] (cit. on pp. 139, 141).

[BFF25] S. Budría, E. Fermé, and D. N. Freitas. "Unveiling energy poverty trajectories: A longitudinal analysis using machine learning". In: *Energy Strategy Reviews* 62 (2025), p. 101998 (cit. on p. 6).

[BLZ25] S. Budría, P. Li Donni, and E. Zucchelli. *Sick and Cold? Evidence on the Dynamic Interplay between Energy Poverty and Health*. Preprint, available at SSRN. 2025. URL: https://ssrn.com/abstract=5153002 (cit. on p. 124).

[Cai+23] M. Cai et al. *Leveraging Large Language Models for Scalable Vector Graphics-Driven Image Understanding*. 2023. DOI: 10.48550/arXiv.2306.06094. arXiv: 2306.06094 [cs.CV] (cit. on p. 142).

[Cam+16] M. Campbell et al. "Human health impacts from litter on beaches and associated perceptions: a case study of "clean" Tasmanian beaches". In: *Ocean Coast Manag.* 126 (2016), pp. 22–30. DOI: 10.1016/j.ocecoaman.2016.04.002 (cit. on pp. 54, 55).

[Cao+19] Y. Cao et al. "Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences". In: *Proceedings of 2019 The World Wide Web Conference (WWW)*. San Francisco, CA, USA, 2019, pp. 151–161. DOI: 10.1145/3308558.3313705 (cit. on p. 21).

[CC21] C. Cardoso and R. Caldeira. "Modeling the exposure of the Macaronesia islands (NE Atlantic) to marine plastic pollution". In: *Front. Mar. Sci.* 8 (2021), p. 653502. DOI: 10.3389/fmars.2021.653502 (cit. on p. 38).

[Car+18] D. L. Carey et al. "Predictive modelling of training loads and injury in Australian football". In: *Int. J. Comput. Sci. Sport* 17.1 (2018-04), pp. 49–66. DOI: 10.2478/ijcss-2018-0002 (cit. on pp. 64, 80).

[Car+13] H. Carson et al. "Tracking the sources and sinks of local marine debris in Hawai'i". In: *Mar. Environ. Res.* 84 (2013), pp. 76–83. DOI: 10.1016/j.marenvres.2012.12.002 (cit. on p. 53).

[Cha+21a] W. Chang et al. "A cross-domain recommender system through information transfer for medical diagnosis". In: *Decision Support Systems* 143 (2021), 1–16, article no.: 113489. DOI: 10.1016/j.dss.2020.113489 (cit. on p. 26).

[Cha+21b] A. Charitou et al. "Investigating the knowledge and attitude of the Greek public towards marine plastic pollution and the EU Single-Use Plastics Directive". In: *Mar. Pollut. Bull.* 166 (2021), p. 112182. DOI: 10.1016/j.marpolbul.2021.112182 (cit. on p. 60).

[CZN23] N. B. Cheikh, Y. B. Zaied, and D. K. Nguyen. "Understanding energy poverty drivers in Europe". In: *Energy Policy* 183 (2023), p. 113818. DOI: 10.1016/j.enpol.2023.113818 (cit. on p. 125).

[Che15] C. Chen. "Regulation and management of marine litter". In: *Marine Anthropogenic Litter*. Ed. by M. Bergmann, L. Gutow, and M. Klages. Springer, 2015, pp. 395–428 (cit. on pp. 38, 58).

[CL24] Y.-C. Chen and W.-C. Lee. "A Novel Cross-Domain Recommendation with Evolution Learning". In: *ACM Trans. Internet Technol.* 24.1 (2024). DOI: 10.1145/3639567 (cit. on p. 27).

[Che+21] H. Chen et al. "Temporal meta-path guided explainable recommendation". In: *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM)*. Virtual Event, Israel, 2021, pp. 1056–1064. DOI: 10.1145/3437963.3441762 (cit. on p. 31).

[CL22] J. Chen and Y. Liu. "Fatigue modeling using neural networks: A comprehensive review". In: *Fatigue Fract. Eng. Mater. Struct.* 45.4 (2022-01), pp. 945–979. DOI: 10.1111/ffe.13640 (cit. on p. 72).

[Che+24] J. Chen et al. "When large language models meet personalization: Perspectives of challenges and opportunities". In: *World Wide Web* 27.4 (2024), 1–45, article no.: 42. DOI: 10.1007/s11280-024-01276-1 (cit. on pp. 31, 33).

[Che+17] J. Chen et al. "Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (ICTIR)*. Tokyo, Japan, 2017, pp. 335–344. DOI: 10.1145/3077136.3080797 (cit. on p. 17).

[CKL15] L.-C. Chen, P.-J. Kuo, and I.-E. Liao. "Ontology-based library recommender system using MapReduce". In: *Cluster Computing* 18.1 (2015), pp. 113–121. DOI: 10.1007/s10586-013-0342-z (cit. on p. 28).

[CC84] M. J. Chen and M. Cook. "Representational Drawings of Solid Objects by Young Children". In: *Perception* 13.4 (1984), pp. 377–385. DOI: 10.1068/p130377 (cit. on p. 138).

[CP09] W. Chu and S.-T. Park. "Personalized recommendation on dynamic content using predictive bilinear models". In: *Proceedings of the 18th International Conference on World Wide Web (WWW)*. Madrid, Spain, 2009, pp. 691–700. DOI: 10.1145/1526709.1526802 (cit. on p. 28).

[CC12] V. Codina and L. Ceccaroni. "Semantically-enhanced recommenders". In: *Artificial Intelligence Research and Development*. Vol. 248. Frontiers in Artificial Intelligence and Applications. Amsterdam, Netherlands: IOS Press, 2012, pp. 69–78. DOI: 10.3233/978-1-61499-139-7-69 (cit. on p. 30).

[Col+15] L. O. Colombo-Mendoza et al. "RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes". In: *Expert Systems with Applications* 42.3 (2015), pp. 1202–1222. DOI: 10.1016/j.eswa.2014.09.016 (cit. on pp. 25, 29, 30).

[Con+22] S. Cong et al. "Unveiling hidden energy poverty using the energy equity gap". In: *Nature communications* 13.1 (2022), p. 2456 (cit. on pp. 106, 121, 137).

[Coo14] H. Coolican. "Multi-level analysis – differences between more than two conditions (ANOVA)". In: *Research Methods and Statistics in Psychology*. Psychology Press, 2014, pp. 570–598 (cit. on pp. 40, 41).

[CAF19] A. O. Crentsil, D. Asuman, and A. P. Fenny. "Assessing the determinants and drivers of multidimensional energy poverty in Ghana". In: *Energy Policy* 133 (2019), p. 110884. DOI: 10.1016/j.enpol.2019.110884 (cit. on pp. 88, 92, 110, 125, 136).

[Cri+20] M. Cristi et al. "The rise and demise of plastic shopping bags in Chile – broad and informal coalition supporting ban as a first step to reduce single-use plastics". In: *Ocean Coast Manag.* 187 (2020), p. 105079. DOI: 10.1016/j.ocecoaman.2019.105079 (cit. on p. 58).

[Cui+25] H. Cui et al. "A review on knowledge graphs for healthcare: Resources, applications, and promises". In: *Journal of Biomedical Informatics* 169 (2025), p. 104861. DOI: 10.1016/j.jbi.2025.104861 (cit. on p. 21).

[Dau+20] A. Da'u et al. "Weighted aspect-based opinion mining using deep learning for recommender system". In: *Expert Systems with Applications* 140 (2020), 1–12, article no.: 112871. DOI: 10.1016/j.eswa.2019.112871 (cit. on pp. 13, 28).

[DSZ21] F. Dalla Longa, B. Sweerts, and B. van der Zwaan. "Exploring the complex origins of energy poverty in The Netherlands with machine learning". In: *Energy Policy* 156 (2021), p. 112373. DOI: 10.1016/j.enpol.2021.112373 (cit. on pp. 85, 88, 106, 110, 120, 123, 125).

[DM11] K. B. Dear and A. J. McMichael. "The health impacts of cold homes and fuel poverty". In: *Bmj* 342 (2011) (cit. on p. 109).

[DL13] K. Decancq and M. A. Lugo. "Weights in multidimensional indices of wellbeing: An overview". In: *Econometric Reviews* 32.1 (2013), pp. 7–34. DOI: 10.1080/07474938.2012.690641 (cit. on pp. 91, 112, 127).

[DZD22] C. Deng, Y. Zhou, and Z. Dou. "Improving personalized search with dual-feedback network". In: *Proceedings of the 15th ACM International Conference on Web Search and Data Mining (WSDM)*. Virtual Event, USA, 2022, pp. 210–218. DOI: 10.1145/3488560.3498447 (cit. on pp. 11, 13, 14, 16, 22).

[Den+20] L. Deng et al. "Public attitudes towards microplastics: perceptions, behaviors and policy implications". In: *Resour. Conserv. Recycl.* 163 (2020), p. 105096. DOI: 10.1016/j.resconrec.2020.105096 (cit. on pp. 53, 56, 60).

[DPX19] S. Dhongde, P. K. Pattanaik, and Y. Xu. "Well-being, deprivation, and the Great Recession in the U.S.: A study in a multidimensional framework". In: *Review of Income and Wealth* 65.S1 (2019), S281–S306. DOI: 10.1111/roiw.12411 (cit. on pp. 90, 112).

[Dhu+18] A. Dhurandhar et al. "Explanations based on the missing: Towards Contrastive Explanations with Pertinent Negatives". In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NeurIPS)*. Vol. 31. Advances in Neural Information Processing Systems. Montréal, Canada: Curran Associates Inc., 2018. DOI: 10.48550/arXiv.1802.07623 (cit. on p. 115).

[Di +18] T. Di Noia et al. "Using ontology-based data summarization to develop semantics-aware recommender systems". In: *Proceedings of the 15th European Semantic Web Conference (ESWC)*. Heraklion, Greece, 2018, pp. 128–144. DOI: 10.1007/978-3-319-93417-4_9 (cit. on p. 20).

[Dil+19] L. Dilkes-Hoffman et al. "Public attitudes towards plastics". In: *Resour. Conserv. Recycl.* 147 (2019), pp. 227–235. DOI: 10.1016/j.resconrec.2019.05.005 (cit. on pp. 37, 38, 55).

[Dol+14] T. Doliwa et al. "Recursive teaching dimension, VC-dimension and sample compression". In: *The Journal of Machine Learning Research* 15.1 (2014), pp. 3107–3131. DOI: 10.5555/2627435.2697064 (cit. on p. 142).

[DP73] D. H. Douglas and T. K. Peucker. "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature". In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 10.2 (1973), pp. 112–122. DOI: 10.3138/FM57-6770-U75U-7727 (cit. on p. 144).

[DJ21] K. Drescher and B. Janzen. "Determinants, persistence, and dynamics of energy poverty: An empirical assessment using German household survey data". In: *Energy Economics* 102 (2021), p. 105433. DOI: 10.1016/j.eneco.2021.105433 (cit. on pp. 88, 125, 135, 136).

[Ehr+16] F. E. Ehrmann et al. "GPS and injury prevention in professional soccer". In: *J. Strength Conditioning Res.* 30.2 (2016-02), pp. 360–367. DOI: 10.1519/JSC.0000000000001093 (cit. on p. 63).

[EHA12] M. Eitz, J. Hays, and M. Alexa. "How do Humans Sketch Objects?" In: *Proceedings of the 2012 ACM Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH)*. Vol. 31. Los Angeles, CA, USA, 2012, 1(44)–10(44). DOI: 10.1145/2185520.2185540 (cit. on p. 141).

[Eke+19] C. I. Eke et al. "A survey of user profiling: State-of-the-art, challenges, and solutions". In: *IEEE Access* 7 (2019), pp. 144907–144924. DOI: 10.1109/access.2019.2944243 (cit. on pp. 1, 3, 9, 11, 13–15, 18, 20).

[EHW11a] J. Ekstrand, M. Hägglund, and M. Waldén. "Injury incidence and injury patterns in professional football: The UEFA injury study". In: *Br. J. Sports Med.* 45.7 (2011-05), pp. 553–558. DOI: 10.1136/bjsm.2009.060582 (cit. on p. 63).

[EHW11b] J. Ekstrand, M. Hägglund, and M. Waldén. "Epidemiology of muscle injuries in professional football (soccer)". In: *Am. J. Sports Med.* 19.6 (2011-02), pp. 1226–1232. DOI: 10.1177/0363546510395879 (cit. on p. 63).

[EWH16] J. Ekstrand, M. Waldén, and M. Hägglund. "Hamstring injuries have increased by 4% annually in men's professional football, since 2001: A 13-year longitudinal analysis of the UEFA elite club injury study". In: *Br. J. Sports Med.* 50.12 (2016-01), pp. 731–737. DOI: 10.1136/bjsports-2015-095359 (cit. on p. 63).

[ENA19] B. N. El Houda, B. Nadjia, and M. Abdelkrim. "Queries-based profile evolution using genetic algorithm". In: *Proceedings of the 16th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)*. Abu Dhabi, United Arab Emirates, 2019. DOI: 10.1109/aiccsa47632.2019.9035351 (cit. on p. 24).

[Est21] D. .-. D. R. de Estatística da Madeira. *Censos 2021 - Resultados Preliminares Região Autónoma da Madeira*. 2021 (cit. on p. 52).

[Eur23] European Commission. *Commission Recommendation (EU) 2023/2407 of 20 October 2023 on Energy Poverty*. Official Journal of the European Union. 2023 (cit. on p. 122).

[FP06] H. Fan and M. S. Poole. "What is personalization? Perspectives on the design and implementation of personalization in information systems". In: *Journal of Organizational Computing and Electronic Commerce* 16.3-4 (2006), pp. 179–202. DOI: 10.1207/s15327744joce1603&4_2 (cit. on p. 32).

[Far+18a] M. Farid et al. "User profiling approaches, modeling, and personalization". In: *Proceedings of the 11th International Conference on Informatics and Systems (INFOS)*. Cairo, Egypt, 2018. DOI: 10.2139/ssrn.3389811 (cit. on pp. 9, 18, 19).

[Far+18b] G. Farnadi et al. "User profiling through deep multimodal fusion". In: *Proceedings of the 11th ACM International Conference on Web Search and Data Mining (WSDM)*. Marina Del Rey, CA, USA, 2018, pp. 171–179. DOI: 10.1145/3159652.3159691 (cit. on p. 18).

[FL07] C. Felden and M. Linden. "Ontology-based user profiling". In: *Proceedings of the 10th International Conference on Business Information Systems (BIS)*. Poznań, Poland, 2007, pp. 314–327. DOI: 10.1007/978-3-540-72035-5_24 (cit. on p. 16).

[Fel+06] A. Felfernig et al. "An integrated environment for the development of knowledge-based recommender applications". In: *International Journal of Electronic Commerce* 11.2 (2006), pp. 11–34. DOI: 10.1017/cbo9780511763113 (cit. on pp. 29, 30).

[FH18] E. Fermé and S. O. Hansson. *Belief Change: Introduction and Overview*. Springer Briefs in Computer Science Series. Cham, Switzerland: Springer, 2018 (cit. on p. 25).

[Fer+24] E. Fermé et al. "Knowledge-driven profile dynamics". In: *Artificial Intelligence* 331 (2024), 1–30, article no.: 104117. DOI: 10.1016/j.artint.2024.104117 (cit. on pp. 2, 11, 25).

[Fer+21] J. Ferreira et al. "Perception of citizens regarding marine litter impacts: collaborative methodologies in island fishing communities of Cape Verde". In: *J. Mar. Sci. Eng.* 9 (2021), p. 306. DOI: 10.3390/jmse9030306 (cit. on pp. 38, 57).

[Fil+21] W. Filho et al. "An assessment of attitudes towards plastics and bioplastics in Europe". In: *Sci. Total Environ.* 755 (2021), p. 142732. DOI: 10.1016/j.scitotenv.2020.142732 (cit. on pp. 38, 60).

[Fos+18] M. Fossi et al. "Impacts of marine litter on cetaceans: a focus on plastic pollution". In: *Marine Mammal Ecotoxicology. Impacts of Multiple Stressors on Population Health*. Ed. by M. Fossi and C. Panti. Academic Press, 2018, pp. 147–184. DOI: 10.1016/B978-0-12-812144-3.00006-1 (cit. on p. 37).

[Fre+25a] D. Freitas et al. "Relative Drawing Identification Complexity is Invariant to Modality in Vision-Language Models". In: *Proceedings of the 28th European Conference on Artificial Intelligence (ECAI)*. Bologna, Italy, 2025, pp. 4507–4514 (cit. on pp. 6, 202).

[FFB25] D. N. Freitas, E. Fermé, and S. Budría. "A counterfactual approach to energy poverty mitigation: A case study for Australia (Preliminary report)". In: *Proceedings of the Workshop on Foundations and Future of Change in Artificial Intelligence (FCAI)*. Bologna, Italy, 2025, pp. 27–43 (cit. on p. 6).

[FVF25] D. N. Freitas, K. Varela, and E. Fermé. *User profiling and its dynamics: A narrative review*. Accepted at The European Journal on Artificial Intelligence. To appear online. 2025 (cit. on pp. 5, 198).

[Fre+25b] D. N. Freitas et al. *Code for the paper: "Relative Drawing Identification Complexity is Invariant to Modality in Vision-Language Models"*. Available at: https://doi.org/10.5281/zenodo.16762246. 2025 (cit. on p. 152).

[Fre+25c] D. N. Freitas et al. "Predicting noncontact injuries of professional football players using machine learning". In: *PloS one* 20.1 (2025), e0315481 (cit. on pp. 6, 200).

[FSA99] Y. Freund, R. Schapire, and N. Abe. "A short introduction to boosting". In: *J. Jpn. Soc. Artif. Intell.* 14.5 (1999-09), pp. 771–780 (cit. on p. 72).

[FN20] J. Frias and R. Nash. *Perceptions about Marine Anthropogenic Litter and Microplastic Pollution in Ireland – Synopsis of the Online Survey*. 2020 (cit. on p. 54).

[FFT22] J. M. Fry, L. Farrell, and J. B. Temple. "Energy poverty and retirement income sources in Australia". In: *Energy Economics* 106 (2022), p. 105793. DOI: 10.1016/j.eneco.2021.105793 (cit. on pp. 85, 88, 89, 109, 110).

[Fu+19] W. Fu et al. "Deeply Fusing Reviews and Contents for Cold Start Users in Cross-Domain Recommendation Systems". In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence and 31st Innovative Applications of Artificial Intelligence Conference and 9th AAAI Symposium on Educational Advances in Artificial Intelligence*. Honolulu, HI, USA, 2019, pp. 94–101. DOI: 10.1609/aaai.v33i01.330194 (cit. on p. 18).

[Ful+06] C. W. Fuller et al. "Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries". In: *Br. J. Sports Med.* 40.3 (2006-02), pp. 193–201. DOI: 10.1136/bjsm.2005.025270 (cit. on p. 63).

[Für10] J. Fürnkranz. "Decision tree". In: *Encyclopedia of Machine Learning*. Ed. by C. Sammut and G. Webb. Springer, Boston, 2010, pp. 263–267. DOI: 10.1007/978-0-387-30164-8_204 (cit. on p. 41).

[Gab16] T. J. Gabbett. "The training—injury prevention paradox: Should athletes be training smarter *and* harder?" In: *Br. J. Sports Med.* 50.5 (2016-01), pp. 273–280. DOI: 10.1136/bjsports-2015-095788 (cit. on p. 79).

[Gab20] T. J. Gabbett. "Debunking the myths about training load, injury and performance: Empirical evidence, hot topics and recommendations for practitioners". In: *Br. J. Sports Med.* 54.1 (2020-12), pp. 58–66. DOI: 10.1136/bjsports-2018-099784 (cit. on p. 63).

[Gab+17] T. J. Gabbett et al. "The athlete monitoring cycle: A practical guide to interpreting and applying training monitoring data". In: *Br. J. Sports Med.* 51.20 (2017-10), pp. 1451–1452. DOI: 10.1136/bjsports-2016-097298 (cit. on p. 63).

[Gal+12] M. Galar et al. "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42.4 (2012), pp. 463–484. DOI: 10.1109/TSMCC.2011.2161285 (cit. on pp. 93, 113).

[GHM15] F. Galgani, G. Hanke, and T. Maes. "Global distribution, composition and abundance of marine litter". In: (2015). Ed. by M. Bergmann, L. Gutow, and M. Klages, pp. 29–56 (cit. on p. 37).

[GT15] S. Gall and R. Thompson. "The impact of debris on marine life". In: *Mar. Pollut. Bull.* 92 (2015), pp. 170–179. DOI: 10.1016/j.marpolbul.2014.12.041 (cit. on p. 37).

[Gao+21] C. Gao et al. "Advances and challenges in conversational recommender systems: A survey". In: *AI Open* 2 (2021), pp. 100–126. DOI: 10.1016/j.aiopen.2021.06.002 (cit. on p. 14).

[Gao+14] Q. Gao et al. "A multi-agent context-based personalized user preference profile construction approach". In: *Proceedings of the 14th International Symposium on Advanced Intelligent Systems (ISIS)*. Daejeon, South Korea, 2014, pp. 55–69. DOI: 10.1007/978-3-319-05573-2_6 (cit. on pp. 2, 19, 25).

[Gau+07] S. Gauch et al. "User profiles for personalized information access". In: *The Adaptive Web: Methods and Strategies of Web Personalization*. Vol. 4321. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2007, pp. 54–89. DOI: 10.1007/978-3-540-72079-9_2 (cit. on pp. 13, 19).

[GJA24] S. Gawusu, S. A. Jamatutu, and A. Ahmed. "Predictive modeling of energy poverty with machine learning ensembles: Strategic insights from socioeconomic determinants for effective policy implementation". In: *International Journal of Energy Research* 2024.1 (2024), p. 9411326. DOI: 10.1155/2024/9411326 (cit. on pp. 86, 89, 111, 123, 126).

[Gei+23] R. Geirhos et al. *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness*. 2023. DOI: 10.48550/arXiv.1811.12231. arXiv: 1811.12231 [cs.CV] (cit. on p. 140).

[Gen+22] S. Geng et al. "Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5)". In: *Proceedings of the 16th ACM Conference on Recommender Systems*. RecSys '22. Seattle, WA, USA: Association for Computing Machinery, 2022, pp. 299–315. DOI: 10.1145/3523227.3546767 (cit. on p. 17).

[Ges+19] I. Gestoso et al. “Plasticrusts: a new potential threat in the Anthropocene’s rocky shores”. In: *Science of The Total Environment* 687 (2019), pp. 413–415. DOI: 10.1016/j.scitotenv.2019.06.123 (cit. on p. 53).

[Gjy+20] L. Gjyli et al. “Marine litter on the Albanian coastline: baseline information for improved management”. In: *Ocean Coast Manag.* 187 (2020), p. 105108. DOI: 10.1016/j.ocecoaman.2020.105108 (cit. on p. 59).

[GPM20] A. Gkargkavouzi, S. Paraskevopoulos, and S. Matsiori. “Public perceptions of the marine environment and behavioral intentions to preserve it: the case of three coastal cities in Greece”. In: *Mar. Pol.* 111 (2020), p. 103727. DOI: 10.1016/j.marpol.2019.103727 (cit. on pp. 37, 55, 199).

[Goo10] Google. *Google Ngram Viewer*. https://books.google.com/ngrams/. Accessed: 2024-07-09. 2010 (cit. on p. 146).

[Gra+24] A. Grattafiori et al. *The Llama 3 Herd of Models*. 2024. DOI: 10.48550/arXiv.2407.21783. arXiv: 2407.21783 [cs.AI] (cit. on p. 140).

[GPN14] A. Green, A. Putschew, and T. Nehls. “Littered cigarette butts as a source of nicotine in urban waters”. In: *J. Hydrol.* 519 (2014), pp. 3466–3474. DOI: 10.1016/j.jhydrol.2014.05.046 (cit. on p. 53).

[GNR23] W. Günther, U. Ninad, and J. Runge. *Causal discovery for time series from multiple datasets with latent contexts*. 2023. arXiv: 2306.12896 [stat.ME]. URL: https://arxiv.org/abs/2306.12896 (cit. on pp. 123, 128).

[HE17] D. Ha and D. Eck. *A Neural Representation of Sketch Drawings*. 2017. DOI: 10.48550/arXiv.1704.03477. arXiv: 1704.03477 [cs.NE] (cit. on pp. 141, 142).

[Häg+05] M. Hägglund et al. “Methods for epidemiological study of injuries to professional football players: Developing the UEFA model”. In: *Br. J. Sports Med.* 39.6 (2005-05), pp. 340–346. DOI: 10.1136/bjsm.2005.018267 (cit. on p. 63).

[Häg+13] M. Hägglund et al. “Injuries affect team performance negatively in professional football: An 11-year follow-up of the UEFA champions league injury study”. In: *Br. J. Sports Med.* 47.12 (2013-07), pp. 738–742. DOI: 10.1136/bjsports-2013-092215 (cit. on p. 63).

[Hah20] J. Hahladakis. “Delineating the global plastic marine litter challenge: clarifying the misconceptions”. In: *Environ. Monit. Assess.* 192 (2020), p. 267. DOI: 10.1007/s10661-020-8202-9 (cit. on p. 37).

[HK23] G. Halkos and I. Kostakis. “Exploring the persistence and transience of energy poverty: Evidence from a Greek household survey”. In: *Energy Efficiency* 16.6 (2023), p. 50. DOI: 10.1007/s12053-023-10137-1 (cit. on p. 88).

[Hal14] S. L. Halson. "Monitoring training load to understand fatigue in athletes". In: *Sports Med.* 44.2 (2014-09), pp. 139–147. DOI: 10.1007/s40279-014-0253-z (cit. on pp. 63, 66).

[Har+21] B. Hardesty et al. "Socioeconomics effects on global hotspots of common debris items on land and the seafloor". In: *Global Environ. Change* 71 (2021), p. 102360. DOI: 10.1016/j.gloenvcha.2021.102360 (cit. on p. 52).

[HZZ21] M. N. Harris, X. Zhao, and E. Zucchelli. "Ageing Workforces, Ill-health and Multi-state Labour Market Transitions". In: *Oxford Bulletin of Economics and Statistics* 83.1 (2021), pp. 199–227. DOI: 10.1111/obes.12379 (cit. on pp. 127, 128).

[Har13] B. Hartley. *D2.1 Baseline Evaluation of Stakeholder Perceptions and Attitudes towards Issues Surrounding Marine Litter - V0, vol. 6*. Tech. rep. 2013 (cit. on pp. 39, 54–56).

[HTP15] B. Hartley, R. Thompson, and S. Pahl. "Marine litter education boosts children's understanding and self-reported actions". In: *Mar. Pollut. Bull.* 90 (2015), pp. 209–217. DOI: 10.1016/j.marpolbul.2014.10.049 (cit. on p. 56).

[Har+18a] B. Hartley et al. "Exploring public views on marine litter in Europe: perceived causes, consequences and pathways to change". In: *Mar. Pollut. Bull.* 133 (2018), pp. 945–955. DOI: 10.1016/j.marpolbul.2018.05.061 (cit. on pp. 38, 56).

[Har+18b] B. Hartley et al. "Turning the tide on trash: empowering European educators and school students to tackle marine litter". In: *Mar. Pol.* (2018), pp. 227–234. DOI: 10.1016/j.marpol.2018.02.002 (cit. on p. 38).

[HF15] A. Hawalah and M. Fasli. "Dynamic user profiles for web personalisation". In: *Expert Systems with Applications* 42.5 (2015), pp. 2547–2569. DOI: 10.1016/j.eswa.2014.10.032 (cit. on pp. 1, 2, 9, 17, 19, 21, 22, 32).

[HPV16] C. He, D. Parra, and K. Verbert. "Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities". In: *Expert Systems with Applications* 56 (2016), pp. 9–27. DOI: 10.1016/j.eswa.2016.02.013 (cit. on p. 14).

[Hec+18] N. Heck et al. "Management preferences and attitudes regarding environmental impacts from seawater desalination: insights from a small coastal community". In: *Ocean & Coastal Management* 163 (2018), pp. 22–29. DOI: 10.1016/j.ocecoaman.2018.05.024 (cit. on p. 55).

[HC20] M. Hermida and S. Costa. "Between tradition and taste: fish consumption habits in a small Portuguese archipelago". In: *J. Aquat. Food Prod. Technol.* 29 (2020), pp. 335–349. DOI: 10.1080/10498850.2020.1734892 (cit. on p. 39).

[Hos+23] S. Hosan et al. "Evaluating the mediating role of energy subsidies on social well-being and energy poverty alleviation in Bangladesh". In: *Energy Research & Social Science* 100 (2023), p. 103088. DOI: 10.1016/j.erss.2023.103088 (cit. on pp. 88, 110).

[Hou+22] Y. Hou et al. "Towards Universal Sequence Representation Learning for Recommender Systems". In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. KDD '22. Washington DC, USA: Association for Computing Machinery, 2022, pp. 585–593. DOI: 10.1145/3534678.3539381 (cit. on p. 19).

[HDZ22] W. van Hove, F. Dalla Longa, and B. van der Zwaan. "Identifying predictors for energy poverty in Europe using machine learning". In: *Energy and Buildings* 264 (2022), p. 112064. DOI: 10.1016/j.enbuild.2022.112064 (cit. on pp. 86, 88, 106, 110, 120).

[Høy+92] K. Høy et al. "European soccer injuries". In: *Am. J. Sports Med.* 20.3 (1992-05), pp. 318–322. DOI: 10.1177/036354659202000314 (cit. on p. 78).

[Hug+17] T. Hughes et al. "Prognostic factors for specific lower extremity and spinal musculoskeletal injuries identified through medical screening and training load monitoring in professional football (soccer): A systematic review". In: *BMJ Open Sport Exercise Med.* 3.1 (2017-09), e000263:1–e000263:18. DOI: 10.1136/bmjsem-2017-000263 (cit. on p. 79).

[IM22] M. Igawa and S. Managi. "Energy poverty and income inequality: An economic analysis of 37 countries". In: *Applied Energy* 306, Part B (2022), p. 118076. DOI: 10.1016/j.apenergy.2021.118076 (cit. on pp. 87, 110, 125).

[ILO12] ILO - International Labour Organization. *International Standard Classification of Occupations*. Geneva, 2012 (cit. on p. 40).

[INE11] INE - Instituto Nacional de Estatística. *Classificação Portuguesa das Profissões*. Lisbon, 2011 (cit. on p. 40).

[ICF17] M. Iñiguez, J. Conesa, and A. Fullana. "Microplastics in Spanish table salt". In: *Sci. Rep.* 7 (2017), p. 8620. DOI: 10.1038/s41598-017-09128-x (cit. on p. 53).

[Int24] International Energy Agency. *World Energy Outlook 2024*. (Accessed on 10 November 2024). 2024. URL: https://www.iea.org/reports/world-energy-outlook-2024 (cit. on pp. 84, 109, 122).

[IFO15] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh. "Recommendation systems: Principles, methods and evaluation". In: *Egyptian Informatics Journal* 16.3 (2015), pp. 261–273. DOI: 10.1016/j.eij.2015.06.005 (cit. on pp. 1, 9, 11, 12, 14, 19, 20, 27–29).

[JS11] N. Japkowicz and M. Shah. *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge, MA, EUA: Cambridge University Press, 2011-08. ISBN: 978-0-52-119600-0 (cit. on p. 71).

[Jas+18] A. Jaspers et al. "Examination of the external and internal load indicators' association with overuse injuries in professional soccer players". In: *J. Sci. Med. Sport* 21.6 (2018-06), pp. 579–585. DOI: 10.1016/j.jsams.2017.10.005 (cit. on p. 63).

[Jia+25] B. Jiang et al. *Know Me, Respond to Me: Benchmarking LLMs for Dynamic User Profiling and Personalized Responses at Scale*. 2025. arXiv: 2504.14225 [cs.CL] (cit. on pp. 31, 33).

[JRR10] A. M. Jones, N. Rice, and J. Roberts. "Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS". In: *Economic Modelling* 27.4 (2010), pp. 866–880. DOI: 10.1016/j.econmod.2009.10.001 (cit. on p. 127).

[Jon+16] J. Jongejan et al. *The Quick, Draw! - A.I.* https://quickdraw.withgoogle.com/. Accessed: 2024-07-08. 2016 (cit. on p. 142).

[Jos+15] A. Joshi et al. "Likert scale: explored and explained". In: *Br. J. Appl. Sci. Technol.* 7 (2015), pp. 396–403. DOI: 10.9734/BJAST/2015/14975 (cit. on p. 40).

[Kab20] A. T. Kabakus. "A Novel Sketch Recognition Model based on Convolutional Neural Networks". In: *Proceedings of the 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. Ankara, Turkey, 2020, pp. 1–6. DOI: 10.1109/HORA49412.2020.9152911 (cit. on p. 141).

[KŚ20] L. Karpinska and S. Śmiech. "Invisible energy poverty? Analysing housing costs in Central and Eastern Europe". In: *Energy Research & Social Science* 70 (2020), p. 101670. DOI: 10.1016/j.erss.2020.101670 (cit. on pp. 88, 110, 125, 135, 137).

[KŚ21] L. Karpinska and S. Śmiech. "Breaking the cycle of energy poverty. Will Poland make it?" In: *Energy Economics* 94 (2021), p. 105063. DOI: 10.1016/j.eneco.2020.105063 (cit. on p. 135).

[KSC24] U. Khalid, M. Shafiullah, and S. M. Chaudhry. "Does conflict aggravate energy poverty?" In: *Energy Policy* 194 (2024), p. 114317. DOI: 10.1016/j.enpol.2024.114317 (cit. on p. 125).

[Kie+17] T. Kiessling et al. "Who cares about dirty beaches? Evaluating environmental awareness and action on coastal litter in Chile". In: *Ocean Coast Manag.* 137 (2017), pp. 82–95. DOI: 10.1016/j.ocecoaman.2016.11.029 (cit. on pp. 38, 58).

[KB00] M. King and G. Bruner. "Social desirability bias: a neglected aspect of validity testing". In: *Psychol. Market.* 17 (2000), pp. 79–103 (cit. on p. 59).

[KB15] D. P. Kingma and J. Ba. "Adam: A method for stochastic optimization". In: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. San Diego, CA, USA, 2015. DOI: 10.48550/arXiv.1412.6980 (cit. on p. 72).

[Kir+23] H. R. Kirk et al. *Personalisation within bounds: A risk taxonomy and policy framework for the alignment of large language models with personalised feedback*. 2023. DOI: 10.1038/s42256-024-00820-y. arXiv: 2303.05453 [cs.CL] (cit. on pp. 31, 33).

[KUO23] E. Kocak, E. E. Ulug, and B. Oralhan. "The impact of electricity from renewable and non-renewable sources on energy poverty and greenhouse gas emissions (GHGs): Empirical evidence and policy implications". In: *Energy* 272 (2023), p. 127125. DOI: 10.1016/j.energy.2023.127125 (cit. on pp. 87, 110, 125).

[KR25] Y. Kolati and C. Raghutla. "Do renewable energy sources and energy infrastructure contribute to mitigating energy poverty? Exploring uncharted dynamics". In: *Energy Strategy Reviews* 58 (2025), p. 101687 (cit. on pp. 87, 110).

[KA22] I. Koomson and S. Awaworyi Churchill. "Employment precarity and energy poverty in post-apartheid South Africa: Exploring the racial and ethnic dimensions". In: *Energy Economics* 110 (2022), p. 106026. DOI: 10.1016/j.eneco.2022.106026 (cit. on pp. 85, 109).

[Kru97] B. Krulwich. "Lifestyle finder: Intelligent user profiling using large-scale demographic data". In: *AI Magazine* 18.2 (1997), pp. 37–37. DOI: 10.1609/aimag.v18i2.1292 (cit. on p. 13).

[KJ19] M. Kuhn and K. Johnson. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Chapman and Hall/CRC Data Science Series. Boca Raton, FL, EUA: CRC Press, 2019-08. ISBN: 978-1-13-807922-9 (cit. on p. 69).

[KRF15] S. Kühn, E. Rebolledo, and J. Franeker. "Deleterious effects of litter on marine life". In: *Marine Anthropogenic Litter*. Ed. by M. Bergmann, L. Gutow, and M. Klages. Springer, 2015, pp. 75–116 (cit. on p. 37).

[Kus+20] I. Kusumawati et al. "Enhancing millennial awareness towards marine litter through environmental education". In: *E3S Web Conf.* Vol. 147. 2020, p. 02019. DOI: 10.1051/e3sconf/202014702019 (cit. on p. 57).

[LC15] E. Lacroix and C. Chaton. "Fuel poverty as a major determinant of perceived health: The case of France". In: *Public Health* 129.5 (2015), pp. 517–524. DOI: 10.1016/j.puhe.2015.02.007 (cit. on p. 135).

[Lam+20] A. Lamb et al. "SketchTransfer: A New Dataset for Exploring Detail-Invariance and the Abstractions Learned by Deep Networks". In: *Proceedings of the 2020 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Snowmass Village, CO, USA, 2020, pp. 952–961. DOI: 10.1109/WACV45572.2020.9093327 (cit. on p. 143).

[LS12] R. Y. K. Lau and L. Song. "Belief revision for intelligent web service recommendation". In: *Computer and Information Science 2012*. Vol. 429. Studies in Computational Intelligence. Berlin, Germany: Springer, 2012, pp. 53–66 (cit. on p. 25).

[Lau02] R. Y. Lau. "The state of the art in adaptive information agents". In: *International Journal on Artificial Intelligence Tools* 11.1 (2002), pp. 19–61. DOI: 10.1142/S0218213002000770 (cit. on pp. 14, 25).

[LN12] C. Lauschke and E. Ntoutsi. "Monitoring user evolution in Twitter". In: *Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. Istanbul, Turkey, 2012, pp. 972–977 (cit. on pp. 15, 21).

[Le+09] D.-L. Le et al. "Building learner profile in adaptive e-learning systems". In: *Proceedings of the 4th International Conference on e-Learning (ICEL)*. Toronto, Canada, 2009, pp. 15–17. DOI: 10.32508/stdj.v14i1.1881 (cit. on pp. 9, 33).

[Le+10] D.-L. Le et al. "A survey of applying user profile in the adaptive instructional systems". In: *Proceedings of the 5th International Conference on e-Learning (ICEL)*. Penang, Malaysia, 2010, pp. 207–218. DOI: 10.32508/stdj.v14i1.1881 (cit. on p. 33).

[Lee+23] Y.-G. Lee et al. "SHAP value-based feature importance analysis for short-term load forecasting". In: *Journal of Electrical Engineering & Technology* 18.1 (2023), pp. 579–588. DOI: 10.1007/s42835-022-01161-9 (cit. on p. 85).

[LS08] E. E. Lehmann and G. G. Schulze. "What does it take to be a star? The role of performance and the media for German soccer players". In: *Appl. Econ. Q.* 54.1 (2008-03), pp. 59–70. DOI: 10.3790/aeq.54.1.59 (cit. on p. 63).

[LL09] K. W.-T. Leung and D. L. Lee. "Deriving concept-based user profiles from search engine logs". In: *IEEE Transactions on Knowledge and Data Engineering* 22.7 (2009), pp. 969–982. DOI: 10.1109/tkde.2009.144 (cit. on p. 13).

[Li+07] L. Li et al. "Dynamic adaptation strategies for long-term and short-term user profile to personalize search". In: *Proceedings of the Joint 9th Asia-Pacific Web Conference (APWeb)*. Huang Shan, China, 2007, pp. 228–240. DOI: 10.1007/978-3-540-72524-4_26 (cit. on p. 22).

[LZ20] S. Li and H. Zhao. "A Survey on Representation Learning for User Modeling". In: *Proceedings of the 29th International Joint Conference on Artificial Intelligence, IJCAI-20*. International Joint Conferences on Artificial Intelligence Organization, 2020, pp. 4997–5003. DOI: 10.24963/ijcai.2020/695 (cit. on p. 15).

[LLY25] Y. Li, Y. Liu, and M. Yu. "Consumer segmentation with large language models". In: *Journal of Retailing and Consumer Services* 82 (2025), 1–13, article no.: 104078. DOI: 10.1016/j.jretconser.2024.104078 (cit. on pp. 31, 33).

[LZE15] D. Liang, M. Zhan, and D. P. Ellis. "Content-aware collaborative music recommendation using pre-trained neural networks". In: *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*. Malaga, Spain, 2015, pp. 295–301 (cit. on p. 17).

[LO21] B. Lin and M. A. Okyere. "Does energy poverty affect the well-being of people: Evidence from Ghana". In: *Sustainable Production and Consumption* 28 (2021), pp. 675–685. DOI: 10.1016/j.spc.2021.06.031 (cit. on p. 87).

[Liu+23a] Q. Liu et al. "Federated User Modeling from Hierarchical Information". In: *ACM Trans. Inf. Syst.* 41.2 (2023). DOI: 10.1145/3442381.3449926 (cit. on p. 12).

[Liu+18] S. Liu et al. "AK tourism: A property graph ontology-based tourism recommender system". In: *Proceedings of the Knowledge Management International Conference (KMICe)*. Miri Sarawak, Malaysia, 2018, pp. 83–88 (cit. on pp. 16, 29).

[Liu+19] S. Liu et al. "User-sideo co-attention network for personalized micro-video recommendation". In: *Proceedings of the The Web Conference (WWW)*. San Francisco, CA, USA, 2019, pp. 3020–3026. DOI: 10.1145/3308558.3313513 (cit. on p. 17).

[LWZ09] X.-Y. Liu, J. Wu, and Z.-H. Zhou. "Exploratory undersampling for class-imbalance learning". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39.2 (2009), pp. 539–550. DOI: 10.1109/TSMCB.2008.2007853 (cit. on pp. 93, 113).

[Liu+23b] Y. Liu et al. "Urban knowledge graph aided mobile user profiling". In: *ACM Transactions on Knowledge Discovery from Data* 18.1 (2023-10). DOI: 10.1145/3596604 (cit. on p. 21).

[Llo82] S. Lloyd. "Least squares quantization in PCM". In: *IEEE Trans. Inf. Theor.* 28 (1982), pp. 129–137. DOI: 10.1109/TIT.1982.1056489 (cit. on p. 41).

[LMA19] M. Locritani, S. Merlino, and M. Abbate. "Assessing the citizen science approach as tool to increase awareness on the marine litter problem". In: *Mar. Pollut. Bull.* 140 (2019), pp. 320–329. DOI: 10.1016/j.marpolbul.2019.01.023 (cit. on p. 56).

[LFF18] B. Long, J. E. Fan, and M. C. Frank. "Drawings as a window into developmental changes in object representations". In: *Proceedings of the 40th Annual Conference of the Cognitive Science Society*. Madison, WI, USA, 2018, pp. 708–713. DOI: 10.1167/18.10.398 (cit. on p. 138).

[LMR09] V. López-Jaquero, F. Montero, and F. Real. "Designing user interface adaptation rules with T: XML". In: *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI)*. Sanibel Island, FL, USA, 2009, pp. 383–388. DOI: 10.1145/1502650.1502705 (cit. on p. 24).

[Lóp+18] A. López-Valenciano et al. "A preventive model for muscle injuries: A novel approach based on learning algorithms". In: *Med. Sci. Sports Exercise* 50.5 (2018-05), pp. 915–927. DOI: 10.1249/MSS.0000000000001535 (cit. on pp. 63, 72).

[Lop+19] P. Lops et al. *Semantics in Adaptive and Personalised Systems: Methods, Tools and Applications*. Cham, Switzerland: Springer, 2019. DOI: 10.1007/978-3-030-05618-6 (cit. on pp. 8, 13, 18).

[Lot+18] H. Lotze et al. "Public perceptions of marine threats and protection from around the world". In: *Ocean Coast Manag.* 152 (2018), pp. 14–22. DOI: 10.1016/j.ocecoaman.2017.11.004 (cit. on p. 39).

[LYR24] S. Lu, X. Yu, and J. Ren. "Identifying driving factors of energy poverty and their interaction mechanism based on the BPNN-WINGS method: Household-level evidence from China". In: *Journal of Cleaner Production* 452 (2024), p. 142194. DOI: 10.1016/j.jclepro.2024.142194 (cit. on pp. 125, 136).

[Lu+16] Z. Lu et al. "Collaborative evolution for user profiling in recommender systems". In: *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. New York, NY, USA, 2016, pp. 3804–3810. DOI: 10.5555/3061053.3061151 (cit. on p. 29).

[LL17] S. M. Lundberg and S.-I. Lee. "A unified approach to interpreting model predictions". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*. Vol. 30. Advances in Neural Information Processing Systems. Long Beach, California, USA: Curran Associates Inc., 2017, pp. 4768–4777. DOI: 10.48550/arXiv.1705.07874 (cit. on pp. 2, 4, 85, 94, 123).

[LXS22] S. Luo, Y. Xiao, and L. Song. "Personalized Federated Recommendation via Joint Representation Learning, User Clustering, and Model Adaptation". In: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. CIKM '22. Atlanta, GA, USA: Association for Computing Machinery, 2022, pp. 4289–4293. DOI: 10.1109/icicsp55539.2022.10050588 (cit. on p. 12).

[MH13] Y. Ma and H. He. *Imbalanced Learning: Foundations, Algorithms, and Applications*. Piscataway, New Jersey: John Wiley and Sons, 2013-07. ISBN: 978-1-11-807462-6 (cit. on p. 70).

[Mah+24] K. Mahmud Sujon et al. "When to use standardization and normalization: Empirical evidence from machine learning models and XAI". In: *IEEE Access* 12 (2024), pp. 135300–135314. DOI: 10.1109/ACCESS.2024.3462434 (cit. on p. 93).

[Mai+97] T. Mainieri et al. "Green buying: the influence of environmental concern on consumer behavior". In: *J. Soc. Psychol.* 137 (1997), pp. 189–204. DOI: 10.1080/00224549709595430 (cit. on p. 56).

[Mal+17a] S. Malone et al. "High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football". In: *J. Sci. Med. Sport* 20.3 (2017-08), pp. 250–254. DOI: 10.1016/j.jsams.2016.08.005 (cit. on pp. 63, 79).

[Mal+17b] S. Malone et al. "The acute:chonic workload ratio in relation to injury risk in professional soccer". In: *J. Sci. Med. Sport* 20.6 (2017-06), pp. 561–565. DOI: 10.1016/j.jsams.2016.10.014 (cit. on p. 67).

[Mal+18] S. Malone et al. "High-speed running and sprinting as an injury risk factor in soccer: Can well-developed physical qualities reduce the risk?" In: *J. Sci. Med. Sport* 21.3 (2018-03), pp. 257–262. DOI: 10.1016/j.jsams.2017.05.016 (cit. on p. 63).

[MM24] B. Manasi and J. P. Mukhopadhyay. "Definition, measurement and determinants of energy poverty: Empirical evidence from Indian households". In: *Energy for Sustainable Development* 79 (2024), p. 101383. DOI: 10.1016/j.esd.2024.101383 (cit. on pp. 88, 89, 110).

[Mar+19] R. Marchand et al. "Examining the relationship between energy poverty and measures of deprivation". In: *Energy Policy* 130 (2019), pp. 206–217. DOI: 10.1016/j.enpol.2019.03.026 (cit. on p. 136).

[Mar+07] G. Maria et al. "Creating an ontology for the user profile: Method and applications". In: *Proceedings of the 1st International Conference on Research Challenges in Information Science (RCIS)*. Ouarzazate, Morocco, 2007, pp. 407–412. DOI: 10.4018/9781605660325.ch009 (cit. on p. 24).

[MIM10] L. Marin, D. Isern, and A. Moreno. "A generic user profile adaptation framework". In: *Artificial Intelligence Research and Development*. Vol. 220. Frontiers in Artificial Intelligence and Applications. Amsterdam, Netherlands: IOS Press, 2010, pp. 143–152. DOI: 10.3233/978-1-60750-643-0-143 (cit. on pp. 2, 13, 14, 22, 24).

[MIM13] L. Marin, D. Isern, and A. Moreno. "Dynamic adaptation of numerical attributes in a user profile". In: *Applied Intelligence* 39.2 (2013), pp. 421–437. DOI: 10.1007/s10489-012-0421-5 (cit. on pp. 22, 24).

[MMI14] L. Marin, A. Moreno, and D. Isern. "Automatic preference learning on numeric and multi-valued categorical attributes". In: *Knowledge-Based Systems* 56 (2014), pp. 201–215. DOI: 10.1016/j.knosys.2013.11.012 (cit. on pp. 22, 28).

[Mar+16] Martín Abadi et al. *TensorFlow: Large-scale machine learning on heterogeneous systems*. Preprint at https://arxiv.org/abs/1603.04467. 2016-03 (cit. on p. 71).

[MBP17] G. Martinho, N. Balaia, and A. Pires. "The Portuguese plastic carrier bag tax: the effects on consumers' behavior". In: *Waste Manag.* 61 (2017), pp. 3–12. DOI: 10.1016/j.wasman.2017.01.023 (cit. on p. 56).

[McC+14] A. McCall et al. "Risk factors, testing and preventative strategies for non-contact injuries in professional football: Current perceptions and practices of 44 teams from various premier leagues". In: *Br. J. Sports Med.* 48.18 (2014-08), pp. 1352–1357. DOI: 10.1136/bjsports-2014-093439 (cit. on p. 63).

[MF10] E. McKinley and S. Fletcher. "Individual responsibility for the oceans? An evaluation of marine citizenship by UK marine practitioners". In: *Ocean Coast Manag.* 53 (2010), pp. 379–384. DOI: 10.1016/j.ocecoaman.2010.04.012 (cit. on p. 55).

[McK10] W. McKinney. "Data structures for statistical computing in Python". In: *Proceedings of the 9th Python in Science Conference (SciPy)*. Austin, TX, USA, 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a (cit. on p. 71).

[Meh+15] M. Mehrpoor et al. "Using process ontologies to contextualize recommender systems in engineering projects for knowledge access improvement". In: *Proceedings of the 16th European Conference on Knowledge Management (ECKM)*. Udine, Italy, 2015, pp. 524–531. DOI: 10.1080/15228053.2017.1313557 (cit. on p. 29).

[Mes+20] C. Mestanza-Ramón et al. "Coastal scenic evaluation of continental Ecuador and Galapagos islands: human impacts and management issues". In: *J. Mar. Sci. Eng.* (2020). DOI: 10.3390/JMSE8060468 (cit. on p. 54).

[Mil95] G. A. Miller. "WordNet: A lexical database for English". In: *Communications of the ACM* 38.11 (1995), pp. 39–41. DOI: 10.1145/219717.219748 (cit. on p. 19).

[Moh21] G. Mohan. "Young, poor, and sick: The public health threat of energy poverty for children in Ireland". In: *Energy Research & Social Science* 71 (2021), p. 101822. DOI: 10.1016/j.erss.2020.101822 (cit. on p. 109).

[MHB19] K. E. Mokhtari, B. P. Higdon, and A. Başar. "Interpreting financial time series with SHAP values". In: *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering (CASCON)*. Toronto, Ontario, Canada: IBM Corp., 2019, pp. 166–172 (cit. on p. 85).

[MRČ20] M. Mokos, G. Realdon, and I. Čižmek. "How to increase ocean literacy for future ocean sustainability? The influence of non-formal marine science education". In: *Sustainability* 12 (2020), p. 10647. DOI: 10.3390/su122410647 (cit. on p. 57).

[MY16] S. Moran and A. Yehudayoff. "Sample Compression Schemes for VC Classes". In: *Journal of the ACM* 63.3 (2016), 1(21)–10(21). DOI: 10.1145/2890490 (cit. on p. 142).

[MMR18] S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García. "Comparison of SFS and mRMR for oximetry feature selection in obstructive sleep apnea detection". In: *Neural Comput. Appl.* 32.20 (2018-04), pp. 15711–15731. DOI: 10.1007/s00521-018-3455-8 (cit. on p. 69).

[MST20] R. K. Mothilal, A. Sharma, and C. Tan. "Explaining machine learning classifiers through diverse counterfactual explanations". In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 2020, pp. 607–617 (cit. on pp. 2, 4, 114, 115).

[MLB10] J. Mouat, R. Lozano, and H. Bateson. *Economic Impacts of Marine Litter*. 2010 (cit. on pp. 37, 54).

[Mou97] A. Moukas. "User modeling in a multiagent evolving system". In: *Proceedings of the 6th International Conference on User Modeling (UM)*. Sardinia, Italy, 1997 (cit. on pp. 9, 32).

[MWB23] M. Mukelabai, K. Wijayantha, and R. Blanchard. "Using machine learning to expound energy poverty in the global south: Understanding and predicting access to cooking with clean energy". In: *Energy and AI* 14 (2023), p. 100290. DOI: 10.1016/j.egyai.2023.100290 (cit. on p. 126).

[Mun+19] M. Mundt et al. "Prediction of lower limb joint angles and moments during gait using artificial neural networks". In: *Med. Biol. Eng. Comput.* 58.1 (2019-12), pp. 211–225. DOI: 10.1007/s11517-019-02061-3 (cit. on p. 72).

[Mur12] K. Murphy. *Machine Learning: A Probabilistic Perspective*. Adaptive Computation and Machine Learning series. Cambridge, Massuchusetts: MIT Press, 2012. ISBN: 978-0-262-01802-9 (cit. on pp. 94, 113).

[Nag+18] A. Naglah et al. "Athlete-customized injury prediction using training load statistical records and machine learning". In: *Proceedings of the 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. Louisville, KY,USA, 2018, pp. 459–464. DOI: 10.1109/ISSPIT.2018.8642739 (cit. on pp. 64, 80).

[Nau+19] M. Naumov et al. *Deep Learning Recommendation Model for Personalization and Recommendation Systems*. 2019. DOI: 10.1145/3394486.3403059. arXiv: 1906.00091 [cs.IR] (cit. on p. 28).

[Naw21] S. Nawaz. "Energy poverty, climate shocks, and health deprivations". In: *Energy Economics* 100 (2021), p. 105338. DOI: 10.1016/j.eneco.2021.105338 (cit. on pp. 109, 124, 135).

[New+15] S. Newman et al. "The economics of marine litter". In: *Marine Anthropogenic Litter*. Ed. by M. Bergmann, L. Gutow, and M. Klages. Springer, 2015, pp. 367–394 (cit. on p. 37).

[NL24] H. T. Nguyen-Phung and H. Le. "Elderly well-being amidst energy poverty: Exploring the health, social, and economic impacts in Vietnam". In: *Energy Research & Social Science* 118 (2024), p. 103762. DOI: 10.1016/j.erss.2024.103762 (cit. on pp. 85, 109).

[Nov+09] T. Novotny et al. "Cigarette butts and the case for an environmental policy on hazardous cigarette waste". In: *Int. J. Environ. Res. Publ. Health* 6 (2009), pp. 1691–1705. DOI: 10.3390/ijerph6051691 (cit. on p. 53).

[NBM12] P. Nussbaumer, M. Bazilian, and V. Modi. "Measuring energy poverty: Focusing on what matters". In: *Renewable and Sustainable Energy Reviews* 16.1 (2012), pp. 231–243. DOI: 10.1016/j.rser.2011.07.150 (cit. on pp. 110–112).

[Obe+18] C. Obeid et al. "Ontology-based recommender system in higher education". In: *Proceedings of the The Web Conference (WWW)*. Lyon, France, 2018, pp. 1031–1034. DOI: 10.1145/3184558.3191533 (cit. on pp. 16, 17).

[OEC23] OECD. *OECD Economic Outlook, Volume 2023 Issue 2*. (Accessed on 1 December 2024). 2023. URL: https://doi.org/10.1787/7a5f73ce-en (cit. on p. 86).

[Oka+02] K. Okamoto et al. "Comparability of epidemiological information between self- and interviewer-administered questionnaires". In: *J. Clin. Epidemiol.* 55 (2002), pp. 505–511. DOI: 10.1016/S0895-4356(01)00515-7 (cit. on p. 59).

[Oli+20] J. L. Oliver et al. "Using machine learning to improve our understanding of injury risk and prediction in elite male youth football players". In: *J. Sci. Med. Sport* 23.11 (2020-11), pp. 1044–1048. DOI: 10.1016/j.jsams.2020.04.021 (cit. on pp. 63, 64, 67).

[Oli+21] L. Oliveras et al. "The Association of Energy Poverty with Health and Wellbeing in Children in a Mediterranean City". In: *International Journal of Environmental Research and Public Health* 18.11 (2021). DOI: 10.3390/ijerph18115961 (cit. on p. 124).

[Ols+20] J. Olsen et al. "Marine litter: institutionalization of attitudes and practices among Fishers in Northern Norway". In: *Mar. Pol.* 121 (2020), p. 104211. DOI: 10.1016/j.marpol.2020.104211 (cit. on p. 37).

[Oos+22] L. Oosterhout et al. "Public perceptions of marine plastic litter: a comparative study across European countries and seas". In: *Front. Mar. Sci.* 8 (2022), p. 784829. DOI: 10.3389/fmars.2021.784829 (cit. on p. 37).

[Ope24] OpenAI. *API Reference*. https://platform.openai.com/docs/api-reference. Accessed: 2024-07-09. 2024 (cit. on p. 145).

[Oum19] S. Oum. "Energy poverty in the Lao PDR and its impacts on education and health". In: *Energy Policy* 132 (2019), pp. 247–253. DOI: 10.1016/j.enpol.2019.05.030 (cit. on p. 124).

[PWT17] S. Pahl, K. Wyles, and R. Thompson. "Channelling passion for the ocean towards plastic pollution". In: *Nat. Human Behav.* 1 (2017), pp. 697–699. DOI: 10.1038/s41562-017-0204-4 (cit. on p. 53).

[PCS13] F. A. P. de Paiva, J. A. F. Costa, and C. R. M. Silva. "A hierarchical architecture for ontology-based recommender systems". In: *Proceedings of the 2013 BRICS Congress on Computational Intelligence and 11st Brazilian Congress on Computational Intelligence (BRICS-CCI & CBIC)*. Ipojuca, Brazil, 2013, pp. 362–367. DOI: 10.1109/brics-cci-cbic.2013.67 (cit. on p. 18).

[PBL21] L. Pan, A. Biru, and S. Lettu. "Energy poverty and public health: Global evidence". In: *Energy Economics* 101 (2021), p. 105423. DOI: 10.1016/j.eneco.2021.105423 (cit. on pp. 123, 124, 135).

[Ped+11] F. Pedregosa et al. "Scikit-learn: Machine learning in Python". In: *J. Mach. Learn. Res.* 12.85 (2011-11), pp. 2825–2830 (cit. on p. 71).

[PLD05] H. Peng, F. Long, and C. Ding. "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy". In: *IEEE Trans Pattern Anal Mach Intell* 27.8 (2005-08), pp. 1226–1238. DOI: 10.1109/TPAMI.2005.159 (cit. on pp. 64, 68, 69).

[Pfi+16] D. Pfirrmann et al. "Analysis of injury incidences in male professional adult and elite youth soccer players: A systematic review". In: *J. Athletic Training* 51.5 (2016-05), pp. 410–424. DOI: 10.4085/1062-6050-51.6.03 (cit. on p. 78).

[PVR15] E. Phimister, E. Vera-Toscano, and D. Roberts. "The Dynamics of Energy Poverty: Evidence from Spain". In: *Economics of Energy & Environmental Policy* 4.1 (2015), pp. 153–166. DOI: 10.5547/2160-5890.4.1.ephi (cit. on p. 135).

[PK19] H. Phoumin and F. Kimura. "Cambodia's energy poverty and its effects on social wellbeing: Empirical evidence and policy implications". In: *Energy Policy* 132 (2019), pp. 283–289. DOI: 10.1016/j.enpol.2019.05.032 (cit. on pp. 112, 124).

[Pla+99] J. Platt et al. "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods". In: *Adv. Large Margin Classifiers* 10.3 (1999-03), pp. 61–74 (cit. on p. 72).

[Pon+24] T. M. Pondie et al. "Energy poverty and respiratory health in Sub-Saharan Africa: Effects and transmission channels". In: *Energy* 297 (2024), p. 131158. DOI: 10.1016/j.energy.2024.131158 (cit. on pp. 85, 87, 92, 109, 110, 112, 123–125).

[Por+22] L. Portz et al. "Where does marine litter hide? The Providencia and Santa Catalina Island problem, SEAFLOWER Reserve (Colombia)". In: *Sci. Total Environ.* 813 (2022), p. 151878. DOI: 10.1016/j.scitotenv.2021.151878 (cit. on p. 38).

[Pot+17] J. Potting et al. *Circular economy: measuring innovation in the product chain*. 2017. URL: https://www.pbl.nl/sites/default/files/downloads/pbl-2016-circular-economy-measuring-innovation-in-product-chains-2544.pdf (cit. on p. 59).

[PH11] T. Potts and E. Hastings. *Marine Litter Issues, Impacts and Actions*. 2011. URL: https://www.gov.scot/binaries/content/documents/govscot/publications/impact-assessment/2012/09/marine-litter-issues-impacts-actions/documents/00402421-pdf/00402421-pdf/govscot%3Adocument/00402421.pdf (cit. on pp. 37, 54).

[Pou+23] R. Pourreza et al. "Painter: Teaching Auto-regressive Language Models to Draw Sketches". In: *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. Paris, France, 2023, pp. 305–314. DOI: 10.1109/ICCVW60793.2023.00038 (cit. on p. 142).

[PSC05] S. Prada, M. Silva, and J. Cruz. "Groundwater behaviour in Madeira, volcanic island (Portugal)". In: *Hydrogeol. J.* 13 (2005), pp. 800–812. DOI: 10.1007/s10040-005-0448-3 (cit. on p. 53).

[PAS22] K. Prakash, S. Awaworyi Churchill, and R. Smyth. “Are you puffing your children’s future away? Energy poverty and childhood exposure to passive smoking”. In: *Economic Modelling* 114 (2022), p. 105937. DOI: 10.1016/j.econmod.2022.105937 (cit. on pp. 88, 89, 110).

[Pra+19] J. Prata et al. “Solutions and integrated strategies for the control and mitigation of plastic and microplastic pollution”. In: *Int. J. Environ. Res. Publ. Health* 16 (2019), p. 2411. DOI: 10.3390/ijerph16132411 (cit. on p. 59).

[PJC15] P. Price, R. Jhangiani, and I. Chiang. “Survey research: constructing survey questionnaires”. In: *Research Methods of Psychology*. 2015, pp. 172–179 (cit. on p. 60).

[Pro22] D. Proctol. *Coal Use Rises, Prices Soar as War Impacts Energy Markets*. (Accessed on 2 October 2024). 2022. URL: https://www.powermag.com/coal-use-rises-prices-soar-as-war-impacts-energy-markets/ (cit. on p. 86).

[Pro+25] N. J. Prottasha et al. *User Profile with Large Language Models: Construction, Updating, and Benchmarking*. 2025. DOI: 10.1038/s41598-024-75599-4. arXiv: 2502.10660 [cs.CL] (cit. on p. 30).

[Pro00] F. Provost. “Machine learning from imbalanced data sets 101”. In: *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI)’2000 Workshop on Imbalanced Data Sets*. Vol. 68. 2000. Austin, TX, USA: AAAI Press, 2000, pp. 1–3 (cit. on p. 93).

[Puk14] N. Pukkhem. “LORecommendNet: An ontology-based representation of learning object recommendation”. In: *Proceedings of the 10th International Conference on Computing and Information Technology (IC2IT)*. Phuket, Thailand, 2014, pp. 293–303. DOI: 10.1007/978-3-319-06538-0_29 (cit. on p. 24).

[PBD23] E. Purificato, L. Boratto, and E. W. De Luca. “Leveraging graph neural networks for user profiling: Recent advances and open challenges”. In: *Proceedings of the 32nd ACM international conference on information and knowledge management*. 2023, pp. 5216–5219. DOI: 10.1145/3583780.3615292 (cit. on p. 12).

[PD15] S. Pye and A. Dobbins. *Energy poverty and vulnerable consumers in the energy sector across the EU: Analysis of policies and measures*. Tech. rep. INSIGHT_E Consortium, 2015 (cit. on p. 109).

[Qi+20] T. Qi et al. “Privacy-preserving news recommendation model learning”. In: *arXiv preprint arXiv:2003.09592* (2020). DOI: 10.18653/v1/2020.findings-emnlp.128 (cit. on p. 12).

[Qi+22] T. Qi et al. "News recommendation with candidate-aware user modeling". In: *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*. 2022, pp. 1917–1921. DOI: 10.1145/3477495.3531778 (cit. on p. 20).

[Qia+25] M. Qiang et al. "Exploring Unified Training Framework for Multimodal User Profiling". In: *Proceedings of the 31st International Conference on Computational Linguistics*. Abu Dhabi, UAE: Association for Computational Linguistics, 2025, pp. 1699–1710 (cit. on pp. 18, 32).

[Rag+21] A. Ragusa et al. "Plasticenta: first evidence of microplastics in human placenta". In: *Environ. Int.* 146 (2021), p. 106274. DOI: 10.1016/j.envint.2020.106274 (cit. on p. 53).

[Ram72] U. Ramer. "An Iterative Procedure for the Polygonal Approximation of Plane Curves". In: *Computer Graphics and Image Processing* 1.3 (1972), pp. 244–256. DOI: 10.1016/S0146-664X(72)80017-0 (cit. on p. 144).

[RJ14a] C. Rana and S. K. Jain. "An evolutionary clustering algorithm based on temporal features for dynamic recommender systems". In: *Swarm and Evolutionary Computation* 14 (2014), pp. 21–30. DOI: 10.1016/j.swevo.2013.08.003 (cit. on pp. 24, 28).

[RJ14b] C. Rana and S. K. Jain. "An extended evolutionary clustering algorithm for an adaptive recommender system". In: *Social Network Analysis and Mining* 4.1 (2014), 1–13, article no.: 164. DOI: 10.1007/s13278-014-0164-x (cit. on pp. 20, 28).

[RJ15] C. Rana and S. K. Jain. "A study of the dynamic features of recommender systems". In: *Artificial Intelligence Review* 43.1 (2015), pp. 141–153. DOI: 10.1007/s10462-012-9359-6 (cit. on pp. 14, 22).

[RWA18] N. Rangel-Buitrago, A. Williams, and G. Anfuso. "Killing the goose with the golden eggs: litter effects on scenic quality of the Caribbean coast of Colombia". In: *Mar. Pollut. Bull.* 127 (2018), pp. 22–38. DOI: 10.1016/j.marpolbul.2017.11.023 (cit. on p. 54).

[Ran+20] N. Rangel-Buitrago et al. "Curbing the inexorable rising in marine litter: an overview". In: *Ocean Coast Manag.* 188 (2020), p. 105133. DOI: 10.1016/j.ocecoaman.2020.105133 (cit. on p. 38).

[Ran+22] N. Rangel-Buitrago et al. "Litter in coastal and marine environments". In: *Mar. Pollut. Bull.* 177 (2022), p. 113546. DOI: 10.1016/j.marpolbul.2022.113546 (cit. on p. 59).

[Rav+24] M. Ravaut et al. *How Much are LLMs Contaminated? A Comprehensive Survey and the LLMSanitize Library*. 2024. DOI: 10.48550/arXiv.2404.00699. arXiv: 2404.00699 [cs.CL] (cit. on p. 145).

[Rav+21] L. Ravi et al. “An intelligent location recommender system utilising multi-agent induced cognitive behavioural model”. In: *Enterprise Information Systems* 15.10 (2021), pp. 1376–1394. DOI: 10.1080/17517575.2020.1812003 (cit. on pp. 14, 22, 29).

[Ray+19] F. Rayon-Viña et al. “Marine litter and public involvement in beach cleaning: disentangling perception and awareness among adults and children, Bay of Biscay, Spain”. In: *Mar. Pollut. Bull.* 141 (2019), pp. 112–118. DOI: 10.1016/j.marpolbul.2019.02.034 (cit. on p. 37).

[RL08] L. Razmerita and M. D. Lytras. “Ontology-based user modelling personalization: Analyzing the requirements of a semantic learning portal”. In: *Proceedings of the 1st World Summit on the Knowledge Society (WSKS)*. Athens, Greece, 2008, pp. 354–363. DOI: 10.1007/978-3-540-87781-3_39 (cit. on pp. 12, 16, 33).

[Rec+14] S. Rech et al. “Rivers as a source of marine litter – a study from the SE Pacific”. In: *Mar. Pollut. Bull.* 82 (2014), pp. 66–75. DOI: 10.1016/j.marpolbul.2014.03.019 (cit. on p. 53).

[RRG20] C. Ritter, P. Raposeiro, and V. Gonçalves. “Diatom diversity and distribution in Madeira Island streams (Portugal)”. In: *Biodivers. Data J.* 8 (2020), p. 59813. DOI: 10.3897/BDJ.8.e59813 (cit. on p. 53).

[RPC21] A. Rossi, L. Pappalardo, and P. Cintia. “A narrative review for a machine learning application in sports: An example based on injury forecasting in soccer”. In: *Sports* 10.1 (2021-12), 5:1–5:16. DOI: 10.3390/sports10010005 (cit. on p. 72).

[Ros+18] A. Rossi et al. “Effective injury forecasting in soccer with GPS training data and machine learning”. In: *PLoS ONE* 13.7 (2018-07), pp. 1–15. DOI: 10.1371/journal.pone.0201264 (cit. on pp. 64, 79, 80).

[Ros+22] A. Rossi et al. “Blood sample profile helps to injury forecasting in elite soccer players”. In: *Sport Sci. Health* (2022-05). DOI: 10.1007/s11332-022-00932-1 (cit. on pp. 64, 79–81).

[Rud+18a] J. D. Ruddy et al. “Predictive modeling of hamstring strain injuries in elite Australian footballers”. In: *Med. Sci. Sports Exercise* 50.5 (2018-05), pp. 906–914. DOI: 10.1249/MSS.0000000000001527 (cit. on p. 72).

[Rud+18b] J. D. Ruddy et al. “Running exposure is associated with the risk of hamstring strain injury in elite Australian footballers”. In: *Br. J. Sports Med.* 52.14 (2018-06), pp. 919–928. DOI: 10.1136/bjsports-2016-096777 (cit. on p. 64).

[Rui+21] I. Ruiz-Pérez et al. "A field-based approach to determine soft tissue injury risk in elite futsal using novel machine learning techniques". In: *Front. Psychol.* 12 (2021-02), 610210:1–610210:15. DOI: 10.3389/fpsyg.2021.610210 (cit. on pp. 63, 79).

[Run+19] J. Runge et al. "Detecting and quantifying causal associations in large nonlinear time series datasets". In: *Science Advances* 5.11 (2019), eaau4996. DOI: 10.1126/sciadv.aau4996 (cit. on pp. 123, 128).

[Ruo+13] T. Ruotsalo et al. "SMARTMUSEUM: A mobile recommender system for the Web of Data". In: *Journal of Web Semantics* 20 (2013), pp. 50–67. DOI: 10.2139/ssrn.3199062 (cit. on pp. 16, 24, 30).

[Sab+25] M. Sabouri et al. *Towards Explainable Temporal User Profiling with LLMs*. 2025. DOI: 10.1145/3631700.3664869. arXiv: 2505.00886 [cs.IR] (cit. on pp. 31, 33).

[Sal+24] A. Salemi et al. *LaMP: When Large Language Models Meet Personalization*. 2024. DOI: 10.18653/v1/2024.acl-long.399. arXiv: 2304.11406 [cs.CL] (cit. on pp. 32, 33).

[SA19] H. Samin and T. Azim. "Knowledge based recommender system for academia using machine learning: A case study on higher education landscape of Pakistan". In: *IEEE Access* 7 (2019), pp. 67081–67093. DOI: 10.1109/access.2019.2912012 (cit. on pp. 16, 28).

[SLT09] M. Saunders, P. Lewis, and A. Thornhill. *Research Methods for Business Students*. fifth. Pearson, Itália, 2009 (cit. on p. 39).

[SMG06] R. Schaefer, W. Mueller, and J. Groppe. "Profile processing and evolution for smart environments". In: *Proceedings of the 3rd International Conference on Ubiquitous Intelligence and Computing (UIC)*. Wuhan, China, 2006, pp. 746–755. DOI: 10.1007/11833529_76 (cit. on p. 24).

[ST14] R. G. Schneider and T. Tuytelaars. "How do Humans Sketch Objects?" In: *Proceedings of the 2014 ACM Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH) Asia*. Vol. 33. Shenzhen, China, 2014, 1(174)–9(174). DOI: 10.1145/2661229.2661231 (cit. on p. 141).

[Sei+10] C. Seiffert et al. "RUSBoost: A hybrid approach to alleviating class imbalance". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part A (Systems and Humans)* 40.1 (2010), pp. 185–197. DOI: 10.1109/TSMCA.2009.2029559 (cit. on pp. 93, 113).

[Sha+24a] M. S. Shabbir et al. "Energy deprivation to financial prosperity: Unveiling multidimensional energy Poverty's influence". In: *Energy Strategy Reviews* 54 (2024), p. 101473 (cit. on pp. 85, 87, 109, 110).

[ST23] S. Shapira and N. Teschner. "No heat, no eat:(Dis) entangling insecurities and their implications for health and well-being". In: *Social Science & Medicine* 336 (2023), p. 116252 (cit. on p. 109).

[Sha+24b] P. Sharma et al. *A Vision Check-up for Language Models*. 2024. DOI: 10.48550/arXiv.2401.01862. arXiv: 2401.01862 [cs.CV] (cit. on p. 141).

[SR07] S. Sheavly and K. Register. "Marine debris & plastics: environmental concerns, sources, impacts and solutions". In: *J. Polym. Environ.* 15 (2007), pp. 301–305. DOI: 10.1007/s10924-007-0074-3 (cit. on p. 37).

[She+24] X. Shen et al. "PMG: Personalized Multimodal Generation with Large Language Models". In: *Proceedings of the 2024 ACM Web Conference (WWWW)*. Singapore, 2024, pp. 3833–3843. DOI: 10.1145/3589334.3645633 (cit. on p. 32).

[SJC20] H. Sheridan, K. Johnson, and A. Capper. "Analysis of international, European and Scot's law governing marine litter and integration of policy within regional marine plans". In: *Ocean Coast Manag.* 187 (2020), p. 105119. DOI: 10.1016/j.ocecoaman.2020.105119 (cit. on p. 59).

[Sil+23] F. L. da Silva et al. "A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities". In: *Education and information technologies* 28.3 (2023), pp. 3289–3328. DOI: 10.1007/s10639-022-11341-9 (cit. on p. 26).

[SM23] P. Simshauser and W. Miller. "On the impact of targeted and universal electricity concessions policy on fuel poverty in the NEM's Queensland region". In: *Economic Analysis and Policy* 80 (2023), pp. 1848–1857. DOI: 10.1016/j.eap.2023.11.015 (cit. on p. 86).

[Sim23] P. Simshauser. "The 2022 energy crisis: Fuel poverty and the impact of policy interventions in Australia's National Electricity Market". In: *Energy Economics* 121 (2023), p. 106660. DOI: 10.1016/j.eneco.2023.106660 (cit. on p. 86).

[SGC12] C. Slavin, A. Grage, and M. Campbell. "Linking social drivers of marine debris with actual marine debris on beaches". In: *Marine Pollution Bulletin* 64 (2012), pp. 1580–1588. DOI: 10.1016/j.marpolbul.2012.05.018 (cit. on p. 38).

[SFL24] J. Sokołowski, J. Frankowski, and P. Lewandowski. "Energy poverty, housing conditions, and self-assessed health: evidence from Poland". In: *Housing Studies* 39.9 (2024), pp. 2325–2354 (cit. on p. 109).

[SRL23] C. Spandagos, M. A. T. Reaños, and M. Á. Lynch. "Energy poverty prediction and effective targeting for just transitions with machine learning". In: *Energy Economics* 128 (2023), p. 107131. DOI: 10.1016/j.eneco.2023.107131 (cit. on pp. 86, 88, 89, 106, 111, 120, 123, 125, 136).

[SGS01] P. Spirtes, C. N. Glymour, and R. Scheines. *Causation, Prediction, and Search*. 2nd ed. Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press, 2001. ISBN: 978-0-262-28415-8. DOI: 10.7551/mitpress/1754.001.0001 (cit. on p. 128).

[Sri+14] N. Srivastava et al. "Dropout: A simple way to prevent neural networks from overfitting". In: *J. Mach. Learn. Res.* 15.56 (2014-06), pp. 1929–1958 (cit. on p. 72).

[Ste+21] I. Stepin et al. "A Survey of Contrastive and Counterfactual Explanation Generation Methods for Explainable Artificial Intelligence". In: *IEEE Access* 9 (2021), pp. 11974–12001. DOI: 10.1109/ACCESS.2021.3051315 (cit. on p. 115).

[SC12] V. Subramaniyaswamy and S. Chenthur Pandian. "Effective tag recommendation system based on topic ontology using Wikipedia and WordNet". In: *International Journal of Intelligent Systems* 27.12 (2012), pp. 1034–1048. DOI: 10.1002/int.21560 (cit. on p. 28).

[SWW21] J. H. Sullivan, M. Warkentin, and L. Wallace. "So many ways for assessing outliers: What really works and does it matter?" In: *Journal of Business Research* 132 (2021), pp. 530–543. DOI: 10.1016/j.jbusres.2021.03.066 (cit. on pp. 93, 113).

[SKR23] C. Sun, A. Khan, and Y. Ren. "Empowering Progress: Education, innovations and financial development in the battle against energy poverty". In: *Journal of Cleaner Production* 425 (2023), p. 138941. DOI: 10.1016/j.jclepro.2023.138941 (cit. on p. 125).

[Sun+17] Y. Sun et al. "Understanding consumers' intention to use plastic bags: using an extended theory of planned behaviour model". In: *Natural Hazards* 89 (2017), pp. 1327–1342. DOI: 10.1007/s11069-017-3022-0 (cit. on p. 56).

[Sun+07] Y. Sun et al. "Cost-sensitive boosting for classification of imbalanced data". In: *Pattern Recognit.* 40.12 (2007-12), pp. 3358–3378. DOI: 10.1016/j.patcog.2007.04.009 (cit. on p. 79).

[SF22] Z. Sun and C. Firestone. "Seeing and speaking: How verbal "description length" encodes visual complexity". In: *Journal of Experimental Psychology: General* 151.1 (2022), pp. 82–96. DOI: 10.1037/xge0001076 (cit. on p. 139).

[SDD08] M. Sutterer, O. Droegehorn, and K. David. “User profile selection by means of ontology reasoning”. In: *Proceedings of the 4th Advanced International Conference on Telecommunications (AICT)*. Athens, Greece, 2008, pp. 299–304. DOI: 10.1109/aict.2008.47 (cit. on p. 25).

[SM22] S. A. Sy and L. Mokaddem. “Energy poverty in developing countries: A review of the concept and its measurements”. In: *Energy Research & Social Science* 89 (2022), p. 102562. DOI: 10.1016/j.erss.2022.102562 (cit. on pp. 87, 89, 110).

[Sye23] M. N. Syed. “Feature selection in machine learning via variable neighborhood search”. In: *Optimization Letters* 17.9 (2023), pp. 2321–2345 (cit. on p. 3).

[TKL19] B. Tabuenca, M. Kalz, and A. L"ohr. “Massive open online education for environmental activism: the worldwide problem of marine litter”. In: *Sustainability* 11 (2019), p. 2860. DOI: 10.3390/su11102860 (cit. on p. 57).

[Tah16] H. Taherdoost. “Sampling Methods in Research Methodology: How to Choose a Sampling Technique for Research”. In: *International Journal of Academic Research in Management* 5.2 (2016), pp. 18–27. DOI: 10.2139/ssrn.3205035 (cit. on p. 144).

[Tai+21] C.-Y. Tai et al. “User-centric path reasoning towards explainable recommendation”. In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (ICTIR)*. Virtual Event, Canada, 2021, pp. 879–889. DOI: 10.1145/3404835.3462847 (cit. on p. 31).

[TT11] K. Tanaka-Ishii and H. Terada. “Word Familiarity and Frequency”. In: *Studia Linguistica* 65.1 (2011), pp. 96–116. DOI: 10.1111/j.1467-9582.2010.01176.x (cit. on p. 146).

[Tan+10] J. Tang et al. “A combination approach to web user profiling”. In: *ACM Transactions on Knowledge Discovery from Data* 5.1 (2010), 1–44, article no.: 2. DOI: 10.1145/1870096.1870098 (cit. on pp. 2, 19).

[TNM18] J. K. Tarus, Z. Niu, and G. Mustafa. “Knowledge-based recommendation: A review of ontology-based recommender systems for e-learning”. In: *Artificial Intelligence Review* 50.1 (2018), pp. 21–48. DOI: 10.1007/s10462-017-9539-5 (cit. on pp. 16, 17, 28–30).

[TNY17] J. K. Tarus, Z. Niu, and A. Yousif. “A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining”. In: *Future Generation Computer Systems* 72 (2017), pp. 37–48. DOI: 10.1016/j.future.2017.02.049 (cit. on p. 16).

[Tch+10] D. Tchuente et al. “Visualizing the evolution of users’ profiles from online social networks”. In: *Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. Odense, Denmark, 2010, pp. 370–374. DOI: 10.1109/asonam.2010.79 (cit. on p. 22).

[Tea+24] G. Team et al. *Gemini 1.5: Unlocking Multimodal Understanding Across Millions of Tokens of Context*. 2024. DOI: 10.48550/arXiv.2403.05530. arXiv: 2403.05530 [cs.AI] (cit. on p. 140).

[Tek+21] M. Tekman et al. *The amount and distribution of litter and microplastic*. https://litterbase.awi.de/litter_graph. 2021 (cit. on p. 53).

[THF19] J. A. Telle, J. Hernández-Orallo, and C. Ferri. “The Teaching Size: Computable Teachers and Learners for Universal Languages”. In: *Machine Learning* 108.8–9 (2019), pp. 1653–1675. DOI: 10.1007/s10994-019-05821-2 (cit. on p. 142).

[TVB09] J. Thomsen, Y. Vanrompay, and Y. Berbers. “Evolution of context-aware user profiles”. In: *Proceedings of the 2009 International Conference on Ultra Modern Telecommunications & Workshops (ICUMT)*. St. Petersburg, Russia, 2009. DOI: 10.1109/icumt.2009.5345395 (cit. on pp. 2, 18, 19, 24).

[Tho+19] H. R. Thornton et al. “Interunit reliability and effect of data-processing methods of global positioning systems”. In: *Int. J. Sports Physiol. Perform.* 14.4 (2019-04), pp. 432–438. DOI: 10.1123/ijspp.2018-0273 (cit. on p. 66).

[Tia+24] C. Tian et al. “Privacy-preserving Cross-domain Recommendation with Federated Graph Learning”. In: *ACM Trans. Inf. Syst.* 42.5 (2024). DOI: 10.1145/3653448 (cit. on p. 12).

[TM07] N. Tintarev and J. Masthoff. “A survey of explanations in recommender systems”. In: *Proceedings of the 23rd IEEE International Conference on Data Engineering Workshop (ICDEW)*. Istanbul, Turkey, 2007, pp. 801–810. DOI: 10.1007/978-0-387-85820-3_15 (cit. on pp. 4, 30).

[TL17] S. Tonin and G. Lucaroni. “Understanding social knowledge, attitudes and perceptions towards marine biodiversity: the case of tegnùe in Italy”. In: *Ocean & Coastal Management* 140 (2017), pp. 68–78. DOI: 10.1016/j.ocecoaman.2017.02.019 (cit. on p. 60).

[Tor+19] H. Torres et al. “Examining youth perceptions and social contexts of litter to improve marine debris environmental education”. In: *Environmental Education Research* 25 (2019), pp. 1400–1415. DOI: 10.1080/13504622.2019.1633274 (cit. on pp. 56, 57).

[UB02] S. Ujjin and P. J. Bentley. “Learning user preferences using evolution”. In: *Proceedings of the 4th Asia-Pacific Conference on Simulated Evolution and Learning (SEAL)*. Singapore, 2002 (cit. on p. 28).

[Uma+24] M. Umair et al. “Does the disparity between rural and urban incomes affect rural energy poverty?” In: *Energy Strategy Reviews* 56 (2024), p. 101584 (cit. on pp. 88, 110).

[UNE09] UNEP - United Nations Environment Programme. *Marine Litter: A Global Challenge*. Nairobi, 2009 (cit. on p. 37).

[Val+20] E. Vallance et al. “Combining internal-and external-training-loads to predict non-contact injuries in soccer”. In: *Appl. Sci.* 10.15 (2020-07), 5261:1–5261:20. DOI: 10.3390/app10155261 (cit. on pp. 64, 79–81).

[Van+21] H. Van Eetvelde et al. “Machine learning methods in sport injury prediction and prevention: A systematic review”. In: *J. Exp. Orthop.* 8.1 (2021-04), pp. 1–15. DOI: 10.1186/s40634-021-00346-x (cit. on pp. 63, 72).

[VFA12] M. C. Varley, I. H. Fairweather, and R. J. Aughey. “Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion”. In: *J. Sports Sci.* 30.2 (2012-11), pp. 121–127. DOI: 10.1080/02640414.2011.627941 (cit. on p. 66).

[Vei+16] J. Veiga et al. “Identifying Sources of Marine Litter”. In: *MSFD GES TG Marine Litter Thematic Report* (2016) (cit. on p. 37).

[Ver+24] S. Verma et al. “Counterfactual Explanations and Algorithmic Recourses for Machine Learning: A Review”. In: *ACM Computing Surveys* 56.12 (2024), p. 312. DOI: 10.1145/3677119 (cit. on pp. 113, 115).

[VML18] F. D. Villa, B. R. Mandelbaum, and L. J. Lemak. “The effect of playing position on injury risk in male soccer players: Systematic review of the literature and risk considerations for each playing position”. In: *Am. J. Orthop.* 47.10 (2018-10). DOI: 10.12788/ajo.2018.0092 (cit. on p. 78).

[Vin+24] Y. Vinker et al. *SketchAgent: Language-Driven Sequential Sketch Generation*. 2024. DOI: 10.48550/2411.17673. arXiv: 2411.17673 [cs.CV] (cit. on p. 142).

[WHE05] M. Walden, M. Hägglund, and J. Ekstrand. “Injuries in Swedish elite football-a prospective study on injury definitions, risk for injury and injury pattern during 2001”. In: *Scand. J. Med. Sci. Sports* 15.2 (2005-03), pp. 118–125. DOI: 10.1111/j.1600-0838.2004.00393.x (cit. on p. 79).

[WZZ+24] G. Wan, J. Zhang, T. Zeng, et al. “Multidimensional energy poverty and its urban-rural and regional disparities: Evidence from China”. In: *Journal of Cleaner Production* 466 (2024), p. 142874. DOI: 10.1016/j.jclepro.2024.142874 (cit. on p. 125).

[Wan+20] P. Wang et al. "Incremental mobile user profiling: Reinforcement learning with spatial knowledge graph for modeling event streams". In: *Proceedings of the 26th ACM International Conference on Knowledge Discovery & Data Mining (SIGKDD)*. Virtual Event, CA, USA: Association for Computing Machinery, 2020, pp. 853–861. DOI: 10.1145/3394486.3403128 (cit. on p. 21).

[Wan+24] W. Wang et al. "Learnable Item Tokenization for Generative Recommendation". In: *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. CIKM '24. Boise, ID, USA: Association for Computing Machinery, 2024, pp. 2400–2409. DOI: 10.1145/3627673.3679569 (cit. on p. 17).

[Wen+18] J. Wen et al. "Visual background recommendation for dance performances using deep matrix factorization". In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 14.1 (2018), 1–19, article no.: 11. DOI: 10.1145/3152463 (cit. on p. 18).

[WC08] S.-S. Weng and H.-L. Chang. "Using ontology network analysis for research document recommendation". In: *Expert Systems with Applications* 34.3 (2008), pp. 1857–1869. DOI: 10.1016/j.eswa.2007.02.023 (cit. on p. 16).

[WO18] S. Werner and A. O'Brien. "Marine litter". In: *Handbook on Marine Environment Protection: Science, Impacts and Sustainable Management*. Ed. by M. Salomon and T. Markus. Cham: Springer International Publishing, 2018, pp. 447–461. DOI: 10.1007/978-3-319-60156-4_23 (cit. on p. 37).

[Wer+16] S. Werner et al. "Harm Caused by Marine Litter". In: *MSFD GES TG Marine Litter - Thematic Report* (2016). DOI: 10.2788/690366 (cit. on p. 54).

[Wic+22] C. Wichmann et al. "Promoting pro-environmental behavior through citizen science? A case study with Chilean schoolchildren on marine plastic pollution". In: *Marine Policy* 141 (2022), p. 105035. DOI: 10.1016/j.marpol.2022.105035 (cit. on pp. 38, 58).

[Wil22] N. Willand. "Opportunity, ideal or distraction? Exploring stakeholder perceptions of tackling energy poverty and vulnerability among older Australians". In: *Energy Research & Social Science* 94 (2022), p. 102852. DOI: 10.1016/j.erss.2022.102852 (cit. on p. 86).

[WGT05] A. Williams, M. Gregory, and D. Tudor. "Marine Debris – onshore, off shore, seafloor litter". In: *Encyclopedia of Coastal Science*. Ed. by M. Schwartz. The Netherlands: Springer, 2005, pp. 623–628 (cit. on p. 54).

[WR19] A. Williams and N. Rangel-Buitrago. "Marine litter: solutions for a major environmental problem". In: *Journal of Coastal Research* 35 (2019), pp. 648–663. DOI: 10.2112/JCOASTRES-D-18-00096.1 (cit. on pp. 58, 59).

[Wu+22] L. Wu et al. "A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation". In: *IEEE Transactions on Knowledge and Data Engineering* 35.5 (2022), pp. 4425–4445. DOI: 10.1109/tkde.2022.3145690 (cit. on pp. 9, 15, 16, 24, 28).

[Wu+24] L. Wu et al. "A survey on large language models for recommendation". In: *World Wide Web* 27.5 (2024), 1–31, article no.: 60. DOI: 10.1007/s11280-024-01291-2 (cit. on pp. 31, 33).

[Wue11] K. Wuensch. "Chi-square tests". In: *International Encyclopedia of Statistical Science*. Ed. by M. Lovric. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 252–253. DOI: 10.1007/978-3-642-04898-2_173 (cit. on p. 41).

[Xia+10] L. Xiang et al. "Temporal recommendation on graphs via long- and short-term preference fusion". In: *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. Washington, DC, USA, 2010, pp. 723–732. DOI: 10.1145/1835804.1835896 (cit. on p. 18).

[Xie+20] R. Xie et al. "Deep feedback network for recommendation". In: *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*. Yokohama, Japan, 2020, pp. 2519–2525. DOI: 10.24963/ijcai.2020/349 (cit. on pp. 13, 14, 18).

[Xu+10] K. Xu et al. "Context-aware content filtering and presentation for pervasive and mobile information systems". In: *Proceedings of the 1st International ICST Conference on Ambient Media and Systems (AMBI-SYS)*. Quebec, Canada, 2010, 1–8, article no.: 20. DOI: 10.4108/icst.ambisys2008.2907 (cit. on p. 25).

[XJB22] P. Xu, C. K. Joshi, and X. Bresson. "Multigraph Transformer for Free-Hand Sketch Recognition". In: *IEEE Transactions on Neural Networks and Learning Systems* 33.10 (2022), pp. 5150–5161. DOI: 10.1109/TNNLS.2021.3069230 (cit. on p. 141).

[Yan+24] F. Yang et al. "A systematic literature review of deep learning approaches for sketch-based image retrieval: datasets, metrics, and future directions". In: *IEEE Access* 12 (2024), pp. 14847–14869. DOI: 10.1109/ACCESS.2024.3357939 (cit. on p. 141).

[YDH21] Z. Yang, S. Dong, and J. Hu. "GFE: General Knowledge Enhanced framework for explainable sequential recommendation". In: *Knowledge-Based Systems* 230.3 (2021), 1–14, article no.: 107375. DOI: 10.1016/j.knosys.2021.107375 (cit. on pp. 18, 31).

[Yin+15] H. Yin et al. "Dynamic user modeling in social media systems". In: *ACM Transactions on Information Systems* 33.3 (2015), 1–44, article no.: 10. DOI: 10.1145/2699670 (cit. on pp. 9, 22, 25).

[YB19] J. Yu and W. Bi. “Evolution of marine environmental governance policy in China”. In: *Sustainability* 11 (2019), p. 5076. DOI: 10.3390/su11185076 (cit. on p. 59).

[YL04] L. Yu and H. Liu. “Efficient feature selection via analysis of relevance and redundancy”. In: *The Journal of Machine Learning Research* 5 (2004), pp. 1205–1224 (cit. on p. 3).

[Yu+15] Q. Yu et al. *Sketch-a-Net That Beats Humans*. 2015. DOI: 10.48550/arXiv.1501.07873. arXiv: 1501.07873 [cs.CV] (cit. on p. 141).

[ZA20] M. Zambrano-Monserrate and M. Alejandra Ruano. “Do you need a bag? Analyzing the consumption behavior of plastic bags of households in Ecuador”. In: *Resources, Conservation and Recycling* 152 (2020), p. 104489. DOI: 10.1016/j.resconrec.2019.104489 (cit. on p. 56).

[Zan+22] T. Zang et al. “A survey on cross-domain recommendation: taxonomies, methods, and future directions”. In: *ACM Transactions on Information Systems* 41.2 (2022), pp. 1–39. DOI: 10.21203/rs.3.rs-1955580/v1 (cit. on p. 27).

[Zha+24] J.-C. Zhang et al. “A review of recommender systems based on knowledge graph embedding”. In: *Expert Systems with Applications* 250 (2024), p. 123876. DOI: 10.1016/j.eswa.2024.123876 (cit. on pp. 20, 21).

[Zha+25] J. Zhang et al. “Envisioning Recommendations on an LLM-Based Agent Platform”. In: *Communications of the ACM* 68.5 (2025-04), pp. 48–57. DOI: 10.1145/3699952 (cit. on p. 32).

[ZAK21] Q. Zhang, S. Appau, and P. L. Kodom. “Energy poverty, children’s wellbeing and the mediating role of academic performance: Evidence from China”. In: *Energy Economics* 97 (2021), p. 105206. DOI: 10.1016/j.eneco.2021.105206 (cit. on pp. 87, 110, 123).

[Zha+19] S. Zhang et al. “Deep learning based recommender system: A survey and new perspectives”. In: *ACM Computing Surveys* 52.1 (2019), 1–38, article no.: 5. DOI: 10.1145/3285029 (cit. on p. 17).

[Zha+20] X. Zhang et al. “A Hybrid convolutional neural network for sketch recognition”. In: *Pattern Recognition Letters* 130 (2020), pp. 73–82. DOI: 10.1016/j.patrec.2019.01.006 (cit. on p. 141).

[Zha17] X. Zhang. “Support Vector Machines”. In: *Encyclopedia of Machine Learning and Data Mining*. Boston, MA, USA: Springer US, 2017, pp. 1214–1220. ISBN: 978-1-4899-7687-1. DOI: 10.1007/978-1-4899-7687-1_810 (cit. on p. 72).

[ZC20] Y. Zhang and X. Chen. “Explainable Recommendation: A Survey and New Perspectives”. In: *Foundations and Trends® in Information Retrieval* 14.1 (2020), pp. 1–101. DOI: doi.org/10.1561/1500000066 (cit. on pp. 30, 31).

[Zha+21] Z. Zhang et al. "Household multidimensional energy poverty and its impacts on physical and mental health". In: *Energy Policy* 156 (2021), p. 112381. DOI: 10.1016/j.enpol.2021.112381 (cit. on pp. 87, 110, 124, 125, 135).

[Zha+16] Z. Zhang et al. "A weighted adaptation method on learning user preference profile". In: *Knowledge-Based Systems* 112 (2016), pp. 114–126. DOI: 10.1016/j.knosys.2016.09.003 (cit. on pp. 13, 14, 19, 24).

[Zha+17] H. Zhao et al. "Meta-graph based recommendation fusion over heterogeneous information networks". In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. Halifax, NS, Canada, 2017, pp. 635–644. DOI: 10.1145/3097983.3098063 (cit. on p. 18).

[ZW25] L.-Y. Zhou and Y.-Y. Wang. "Simulation of personalized english learning path recommendation system based on knowledge graph and deep reinforcement learning". In: *Scientific Reports* 15.1 (2025-10), p. 34554. DOI: 10.1038/s41598-025-17918-x (cit. on p. 21).

[Zho+20] S. Zhou et al. "Interactive recommender system via knowledge graph-enhanced reinforcement learning". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (ICTIR)*. Virtual Event, China, 2020, pp. 179–188. DOI: 10.1145/3397271.3401174 (cit. on p. 14).

[Zho+12] X. Zhou et al. "The state-of-the-art in personalized recommender systems for social networking". In: *Artificial Intelligence Review* 37.2 (2012), pp. 119–132. DOI: 10.1007/s10462-011-9222-1 (cit. on p. 12).

[Zhu+21] F. Zhu et al. "Cross-Domain Recommendation: Challenges, Progress, and Prospects". In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence, IJCAI-21*. International Joint Conferences on Artificial Intelligence Organization, 2021, pp. 4721–4728. DOI: 10.24963/ijcai.2021/639 (cit. on p. 27).

[Zhu+18] X. Zhu et al. *An Overview of Machine Teaching*. 2018. DOI: 10.48550/arXiv.1801.05927. arXiv: 1801.05927 [cs.LG] (cit. on pp. 2, 140, 142).

[ZHZ12] E. Zucchelli, M. Harris, and X. Zhao. *Ill-health and transitions to part-time work and self-employment among older workers*. Health, Econometrics and Data Group (HEDG) Working Papers. HEDG, c/o Department of Economics, University of York, 2012. URL: https://EconPapers.repec.org/RePEc:yor:hectdg:12/04 (cit. on p. 136).

APPENDIX A

Supplementary Material for the Narrative Review of User Profiling and its Dynamics

The supplementary material for the article *User profiling and its dynamics: A narrative review* [FVF25] can be accessed at the following link: https://doi.org/10.5281/zenodo.17957816.

APPENDIX B

Supplementary Material for the Study on User Profiles of Marine Litter Literacy

The supplementary material for the article *Public perceptions, knowledge, responsibilities, and behavioral intentions regarding marine litter: Identifying profiles of small oceanic island inhabitants* [GPM20] can be accessed at the following link: `https://doi.org/10.5281/zenodo.14965846`.

APPENDIX C

Supplementary Material for the Study on Profiles of Football Players

The supplementary material for the article *Predicting noncontact injuries of professional football players using machine learning* [Fre+25c] can be accessed at the following link: https://doi.org/10.5281/zenodo.14965970.

APPENDIX D

Supplementary Material for the Studies Related to Energy Poverty

The supplementary material for all the works related to energy poverty (Chapters 5 to 7) can be accessed at the following link: `https://doi.org/10.5281/zenodo.17211103`.

APPENDIX E

Supplementary Material for the Study on Relative Drawing Identification Complexity

The supplementary material for the article *Relative drawing identification complexity is invariant to modality in vision-language models* [Fre+25a] can be accessed at the following link: `https://doi.org/10.5281/zenodo.17214663`.

REPÚBLICA PORTUGUESA
CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR
Fundação para a Ciência e a Tecnologia