Pontifícia Universidade Católica do Rio de Janeiro

Renato José Quiliche Altamirano

A supervised learning approach to predict household aid demand for recurrent clime-related disasters in Peru

Dissertação do Mestrado

Dissertation presented to the Programa de Pós-graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção.

Advisor: Prof. Adriana Leiras
Co-advisor: Prof. Fernanda Araújo Baião Amorim

Rio de Janeiro
September 2023

Renato José Quiliche Altamirano

A supervised learning approach to predict household aid demand for recurrent clime-related disasters in Peru

Dissertation presented to the Programa de Pós-graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção. Approved by the Examination Committee:

Prof. Adriana Leiras
Advisor
Departamento de Engenharia Industrial - PUC-Rio

Prof. Fernanda Araújo Baião Amorim
Co-advisor
Departamento de Engenharia Industrial - PUC-Rio

Prof. Paula Medina Maçaira Louro
Departamento de Engenharia Industrial - PUC-Rio

Prof. Mario Chong
UP

Rio de Janeiro, September 4th, 2023.

Renato José Quiliche Altamirano

Economista pela Pontifícia Universidade Católica do Perú (PUCP) (2019), Mestrando em Engenharia Industrial pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Bolsista pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Bibliographic data

Quiliche Altamirano, Renato José

A supervised learning approach to predict household aid demand for recurrent clime-related disasters in Peru / Renato José Quiliche Altamirano ; advisor: Adriana Leiras ; co-advisor: Fernanda Araujo Baião Amorim. – 2023.
92 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2023.
Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Aprendizado de máquina supervisionado. 3. Classificador de risco de desastre. 4. Ondas de frio. 5. Inundações. 6. Deslizamentos de terra. I. Leiras, Adriana. II. Amorim, Fernanda Araujo Baião. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 658.5

Acknowledgments

I am grateful to my advisor, Adriana Leiras, for guiding me and helping me become a better professional through her example.

To my co-advisor, Fernanda Baião, who provided me with methodological support and warmth in guiding my academic path.

PUC-Rio was my second home during my master’s degree, and I felt I was in the best learning environment, surrounded by excellent professionals.

I am thankful to all the professors from DEI PUC-Rio and the University of Pacifico Perú who helped me clear up many doubts and direct my research toward a more significant impact.

To my family, Jorge, Carmela, and Romina, who supported and motivated me in this step of my professional development, I am always happy that we are all reunited again.

To my grandmother Inés, who believes unconditionally in my dreams and advises me to stay firm in my decisions.

Lastly, I am grateful to my girlfriend, Gera, who, like her family, always gave me unconditional support and freedom to take the course of my life.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001.

Abstract

Quiliche Altamirano, Renato José; Leiras, Adriana (advisor); Baião Amorim, Fernanda Araújo (co-advisor). Rio de Janeiro, 2023. 92p. **A supervised learning approach to predict household aid demand for recurrent clime-related disasters in Peru.** Dissertação de Mestrado – Departamento de Engenharia Industrial Católica do Rio de Janeiro.

This dissertation presents a data-driven approach to the problem of predicting recurrent disasters in developing countries. Supervised machine learning methods are used to train classifiers that aim to predict whether a household would be affected by recurrent climate threats (one classifier is trained for each natural hazard). The approach developed is valid for recurrent natural hazards affecting a country and allows disaster risk managers to target their operations with more knowledge. In addition, predictive assessment allows managers to understand the drivers of these predictions, leading to proactive policy formulation and operations planning to mitigate risks and prepare communities for recurring disasters.

The proposed methodology was applied to the case study of Peru, where classifiers were trained for cold waves, floods, and landslides. In the case of cold waves, the classifier was 73.82% accurate. The research found that low-income families in rural areas are vulnerable to cold wave-related disasters and need proactive humanitarian intervention. Vulnerable families have poor urban infrastructure, including footpaths, roads, lampposts, and water and drainage networks. The role of health insurance, health status, and education is minor. Households with sick members are more likely to be affected by cold waves. Higher educational attainment of the head of the household is associated with a lower probability of being affected by cold snaps.

In the case of flooding, the classifier is 82.57% accurate. Certain urban conditions, such as access to drinking water, lampposts, and

drainage networks, can make rural households more susceptible to flooding. Owning a computer or laptop decreases the likelihood of being affected by flooding while owning a bicycle and being headed by married individuals increases it. Flooding is more common in less developed urban settlements than isolated rural families.

In the case of landslides, the classifier is 88.85% accurate and follows a different logic than that of floods. The importance of the prediction is more evenly distributed among the features considered when learning the classifier. Thus, the impact of an individual feature on the prediction is small. Long-term wealth is more critical: the probability of being affected by a landslide is lower for families with specific appliances and household building materials. Rural communities are more affected by landslides, especially those located at higher altitudes and greater distances from cities and markets. The average marginal impact of altitude is non-linear.

The classifiers provide an intelligent data-driven method that saves resources by ensuring accuracy. In addition, the research provides guidelines for addressing efficiency in aid distribution, such as facility location formulations and vehicle routing.

The research results have several managerial implications, so the authors call for action from disaster risk managers and other relevant stakeholders. Recurrent disasters challenge all of humanity.

Keywords

Supervised Machine Learning; disaster risk classifier; cold waves; floods; landslides; households' features; logistic regression; random forest; XGBoost.

Resumo

Quiliche Altamirano, Renato José; Leiras, Adriana (orientadora); Baião Amorim, Fernanda Araújo (co-orientador). Rio de Janeiro, 2023. 92p. **Uma abordagem de aprendizado supervisionado para prever a demanda de ajuda familiar para desastres climáticos recorrentes no Peru.** Dissertação de Mestrado - Departamento de Engenharia Industrial da Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação apresenta uma abordagem baseada em dados para o problema de predição de desastres recorrentes em países em desenvolvimento. Métodos de aprendizado de máquina supervisionado são usados para treinar classificadores que visam prever se uma família seria afetada por ameaças climáticas recorrentes (um classificador é treinado para cada perigo natural). A abordagem desenvolvida é válida para perigos naturais recorrentes que afetam um país e permite que os gerentes de risco de desastres direcionem suas operações com mais conhecimento. Além disso, a avaliação preditiva permite que os gerentes entendam os impulsionadores dessas previsões, levando à formulação proativa de políticas e planejamento de operações para mitigar riscos e preparar comunidades para desastres recorrentes.

A metodologia proposta foi aplicada ao estudo de caso do Peru, onde foram treinados classificadores para ondas de frio, inundações e deslizamentos de terra. No caso das ondas de frio, o classificador tem 73,82% de precisão. A pesquisa descobriu que famílias pobres em áreas rurais são vulneráveis a desastres relacionados a ondas de frio e precisam de intervenção humanitária proativa. Famílias vulneráveis têm infraestrutura urbana precária, incluindo trilhas, caminhos, postes de iluminação e redes de água e drenagem. O papel do seguro saúde, estado de saúde e educação é menor. Domicílios com membros doentes levam a maiores probabilidades de serem afetados por ondas de frio. Maior realização educacional do chefe da família está associada a uma menor probabilidade de ser afetado por ondas de frio.

No caso das inundações, o classificador tem 82.57% de precisão. Certas condições urbanas podem tornar as famílias rurais mais suscetíveis a inundações, como acesso à água potável, postes de iluminação e redes de drenagem. Possuir um computador ou laptop diminui a probabilidade de ser afetado por inundações, enquanto possuir uma bicicleta e ser chefiado por indivíduos casados aumenta. Inundações são mais comuns em assentamentos urbanos menos desenvolvidos do que em famílias rurais isoladas.

No caso dos deslizamentos de terra, o classificador tem 88.85% de precisão, e é segue uma lógica diferente do das inundações. A importância na previsão é mais uniformemente distribuída entre as características consideradas no aprendizado do classificador. Assim, o impacto de um recurso individual na previsão é pequeno. A riqueza a longo prazo parece ser mais crítica: a probabilidade de ser afetado por um deslizamento é menor para famílias com certos aparelhos e materiais domésticos de construção. Comunidades rurais são mais afetadas por deslizamentos, especialmente aquelas localizadas em altitudes mais elevadas e maiores distâncias das cidades e mercados. O impacto marginal médio da altitude é não linear.

Os classificadores fornecem um método inteligente baseado em dados que economiza recursos garantindo precisão. Além disso, a pesquisa fornece diretrizes para abordar a eficiência na distribuição da ajuda, como formulações de localização da instalação e roteamento de veículos.

Os resultados da pesquisa têm várias implicações gerenciais, então os autores convocam à ação gestores de risco de desastres e outros interessados relevantes. Desastres recorrentes desafiam toda a humanidade.

Palavras-chave

Aprendizado de máquina supervisionado; classificador de risco de desastre; ondas de frio; inundações; deslizamentos de terra.

Table of contents

1 Introduction **13**

2 A predictive assessment of households' risk against disasters caused by cold waves using supervised learning **17**

2.1. Introduction 17

2.2. Theoretical foundation 20

2.2.1. Disaster risk reduction for clime-related disasters 20

2.2.2. Machine Learning in Disaster Risk Reduction 22

2.3. The case of Puno, Perú 24

2.4. Materials and methods 28

2.4.1. Data collection methods and the classification problem 28

2.4.2. Machine Learning Pipeline 29

2.5. Description of results 36

2.5.1. Descriptive characterization of Puno 36

2.5.2. Model training results 40

2.5.3. Complementary descriptive analysis 42

2.6. Discussion and implications 44

2.6.1. Determinants of cold wave-related disaster risk 44

2.6.2. A proposal for improvement of the model 47

2.6.3. Extra considerations about practical implementations 48

2.7. Chapter conclusions 49

3 Identifying final demand points for aid in the aftermath of sudden-onset clime-related recurrent disasters in Peru using supervised learning **52**

3.1. Introduction 52

3.2. Theoretical framework 54

3.2.1. Characterization of households’ disaster vulnerability 55

3.3. Materials and methods 55

3.3.1. Data gathering methods 56

3.3.2. Data analysis method 58

3.3.2.1. The classification problem 58
3.3.2.2. Machine Learning Pipeline 58
3.3.2.3. Bayesian Optimization Gaussian Process 63
3.4. Description of results 65
3.4.1. Descriptive analytics 65
3.4.2. Model training results 67
3.5. Practical Implications for Humanitarian Operations 73
3.6. Chapter conclusions 75

4 Conclusions and recommendations 77

References 79

Appendix - Elsevier copyright 89

List of Figures

Figure 1: Theoretical representation of Disaster Management Helix optimization. 21
Figure 2: Causes of cold-related disastrous events affecting communities 24
Figure 3: Spatial distribution of households exposed to ELTEs 26
Figure 4: Time series plot for average minimum temperature in Puno 2009-2012 27
Figure 5: Detailed procedure for Machine Learning Model Training 30
Figure 6: Features' correlation heatmap 40
Figure 7: Confusion matrix for Random Forest Classifier (Test set) 41
Figure 8: Confusion matrix for Random Forest Classifier (Train set) 42
Figure 9: Feature's importance from RFC 45
Figure 10: Top 10 Feature's partial dependence on the probability of being at risk 46
Figure 11: Confusion matrix with new prediction threshold (Test set) 48
Figure 12: Hazard zones 57
Figure 13: Machine Learning Pipeline 60
Figure 14: Confusion matrixes for the hold-out test set 67
Figure 15: Elbow plot for model feature importance 69
Figure 16: Feature importance indicators for regional dummies 70
Figure 17: Partial dependence plots for Flood's classifier 72
Figure 18: Partial dependence plots for Landslides classifier 73

List of tables

Table 1 – Thesis Structure 15

Table 2: Multi-dimensional vulnerability features 38

Table 3: Models' performance on the test dataset (Puno, 2021). 40

Table 4: Descriptive analytics of misclassified categories 42

Table 5: Hyperparameter definitions 61

Table 6: Empirical characterization of vulnerability 65

Table 7: Performance metrics in the hold-out test set 68

1
Introduction

During the XXI century, disasters have increasingly impacted humanity by producing poverty, disrupting the economy, causing social conflicts, and harming the environment (Besiou et al., 2021). However, the most critical consequences are human losses and, ultimately, human suffering. The intractable magnitude of such losses highlights the need to manage risk and disasters better.

Developing countries are the most susceptible to losses (Guha-Sapir et al., 2015). Their susceptibility increases because of their resource scarcity and reduced capacity to mitigate the impacts of disasters. Having lower indicators of human development, quality of housing, and available income conditions, households from developing countries have higher expected losses from a disaster.

The case of developing countries is more critical than the rest of the world because they face a dynamic problem (Akram et al., 2021). Such countries face higher losses from disasters that reduce their available resources and capabilities because they must waste them on disaster response and recovery activities (after a disaster strikes, there is no such option to save resources in the response and recovery activities). With lower resources and affected capacities, such countries must face future disasters for which they need more preparation. Hence, this reinforcement loop or a vicious cycle may last indefinitely if impactful interventions are absent (Sodhi, 2016).

Disasters are not natural (Besiou et al., 2021). They occur when natural hazards interact with vulnerable populations. Several regions of developing countries that are vulnerable are also affected by recurrent natural hazards that later turn into recurrent disasters. This dissertation attempts to contribute to solving the problem of recurrent disasters using a data-based approach.

Managing disaster risk is critical to the problem of countries recurrently affected by disasters (Bosher et al., 2021). Efficient and effective use of resources in pre-disaster managerial interventions leads to reduced disaster impact.

Reduced disaster impacts save resources in response and recovery activities. The saved resources may be later used to improve disaster risk management. Then, the vicious cycle turns into a virtuous one.

This research proposes using Machine Learning methods to train supervised classifiers to predict if a household would be affected by natural hazards. This approach is valid for recurrent natural hazards affecting a country. Then, demand can be screened, and disaster risk managers can accurately target their operations. Furthermore, predictive assessment allows managers to understand the drivers of such predictions, leading to proactively formulating policies and planning operations to mitigate risks and prepare communities for recurrent disasters.

The proposed risk screening tool has implications for managers such as those mentioned before, but it also provides granular demand estimations that may drive countrywide multi-hazard risk reduction. In harmony with the theory of disaster risk (Twigg, 2003), data on households' multiple dimensions of vulnerability was collected. Economic, social, health, and geographical vulnerability indicators might determine if a household was affected by a disaster (given that it was exposed to a natural hazard).

The proposed methodology was applied to the case study of Peru. Peru is a developing country that is recurrently affected by disasters, particularly those potentialized by El Niño (extreme precipitations causing floods and landslides) and cold waves in both Coastal and Andean regions. Then, classifiers for cold waves, floods, and landslides are trained. We followed an experimental methodology evaluating the algorithms Logistic Regression, Random Forest Classifier, and XGBoost in several random scenarios and selected the classifier with the highest balanced accuracy.

The main contribution of this dissertation is that it uses multiple dimensions of vulnerability based on households' characteristics to build supervised learning classifiers. Other approaches use data from social media or outsourcing platforms that may not represent poor regions with no access to the internet (Lin et al., 2020). Furthermore, this dissertation provides a framework for automated model updating and re-training with new data under the same domain using hyperparameter optimization techniques (Hutter et al., 2019).

The thesis structure is described in Table 1.

Table 1 – Thesis structure

Main research question	To what extent does multidimensional vulnerability can predict the risk of being affected by recurrent disasters (Cold waves, Floods, or Landslides)?	
Main objective	To train classifiers to predict whether a household would be affected by recurrent natural hazards (Cold waves, Floods, and Landslides)	
Secondary research questions	What are the households' characteristics that make them susceptible to Cold waves-related disasters?	What are the households' characteristics that make them susceptible to Floods and Landslides?
Secondary objectives	To train a classifier for Cold waves-related disasters at the household level of analysis. To interpret and understand the logic of predictions in terms of features. To provide guidelines for the practical usage of such a classifier.	To train classifiers for whether a household would be affected by Floods and Landslides. To interpret and understand the logic of predictions in terms of features. To provide guidelines and methods to use the predictions to take action and plan the humanitarian supply chain.
Methodology	Applied supervised Machine Learning	Applied supervised Machine Learning
Deliverable	Cold wave risk classifiers	Floods risk classifier, Landslides risk classifiers
#	**PAPER 1**	**PAPER 2**

A critical difference in both papers is that the second paper provides the decision-makers with guidelines when facing multi-hazards. Also, the second paper provides a nationwide management perspective. Division in two papers is also justified by division on the sample sizes (Paper 1 contains samples from households in Puno, while Paper 2 uses households from all Peruvian regions). Additional tools and guidelines are shown to deepen the potential of implementation of the results of this research.

An Elsevier copyright Appendix was included to guide authors in the legal reproduction of the contents of the first paper that was submitted to the *International Journal of Disaster Risk Reduction*.

The rest of the dissertation is structured as follows. The case of Cold Waves is presented in Section 2, which shows the first paper under the second round of review at the International Journal of Disaster Risk Reduction. Then, Section 3 presents the case of Floods and Landslides in a second paper, to be submitted after the Master's thesis defense committee considerations.

2
A predictive assessment of households' risk against disasters caused by cold waves using supervised learning

2.1. Introduction

The frequency of clime-related disasters has grown exponentially in the last twenty years (EM-DAT, 2022). This fact may be explained mainly by the increase in global warming and population sizes, which, in turn, pressure on natural resources, generating harmful outcomes for the environment (Keja-Kaereho & Tjizu, 2019). Disasters are not natural, as the same hazard leads to different outcomes in different locations worldwide (Besiou et al., 2021). Disaster risk is the outcome of interactions of hazard, vulnerability, and exposure (UNDRR, 2015; Wright et al., 2020). In consequence, the impact of disasters depends on the degree of vulnerability, which is defined by anthropogenic conditions, the hazard's scale and magnitude, and the exposure level.

Hazards might harm humans, animals, and the environment, destroying a specific geographic position in a period (Preciado, 2015). Although hazards are mostly known to be an occurrence that human beings cannot control, human interaction with the environment has caused an increase in the frequency of climate-related hazards (Shafapourtehrany et al., 2022). The clime has become more extreme. Cold waves are not the exception; people with low incomes or the vulnerable are the most affected. However, the literature on cold waves is scarce despite their harmful consequences on the livelihoods of poor agricultural households (Amirkhani et al., 2022; López-Bueno et al., 2021).

We argue that proactive disaster risk reduction is essential for communities affected by recurrent disasters. Disaster risk management phases are not independent (Besiou et al., 2021). Thus, proactive disaster risk reduction activities are carried out before a disaster strikes to help mitigate risks and create savings that communities may use for further development and building of resilience that is urgent due to the increasing magnitude and frequency of disasters.

The increase in the frequency of cold wave-related disasters during the last century disproportionately affected low-income countries (Amirkhani et al., 2022;

López-Bueno et al., 2021). India, Bangladesh, Poland, and Russia are the most affected countries, harming 1,227 million people and generating 184 thousand deaths since 2000.

Cold waves can trigger disasters that result in the loss of human lives, particularly in cases of high vulnerability. Households with poor infrastructure and limited resources to combat the cold are at a higher risk (López-Bueno et al., 2021). Individuals with a high prevalence of comorbidities, such as cardiovascular diseases, are also at an increased risk (Shaposhnikov and Revich, 2016). As such, deaths caused by cold waves are not natural occurrences but rather the extreme result of cold climate conditions affecting impoverished households with chronic illnesses.

This paper examines the situation in Puno, Peru, one of the poorest regions in the country. A significant proportion of Puno's population lives in rural areas, where they are exposed to cold waves and rely on subsistence agriculture to survive. These cold waves pose a significant challenge to the livelihoods of poor rural households, as they must use their limited resources to cope with the harsh conditions. As a result, these households may struggle to develop resilience and improve their situation. The primary concern in Puno is the impact of cold waves on the ability of poor rural households to sustain their livelihoods.

This paper proposes a method to identify and target households prone to be affected by cold waves. These households need interventions in risk mitigation, disaster prevention, and preparedness. To achieve intelligent and accurate targeting, this paper focuses on a data-centric approach (ENAHO, 2022). Specifically, we aim to predict which households must be prepared for a cold wave-related disaster.

This prediction must be accurate for the at-risk households, representing demand points that must be met. This is because when a predictive model misclassifies positive outcomes, defined as at-risk households, deprivation costs represent demand points that need essential supplies. However, the model needs to be more accurate in their risks, and aid goods are not being supplied (Gutjahr and Fischer, 2018; Holguin-Veras et al., 2013). These cases are named false negatives.

Our proposed model gives greater importance to accurate prediction of disaster risk, even if it implies that some households that do not have risk are being misclassified. Considering these objectives, our methodology uses supervised learning algorithms - Logistic Regression and Random Forest Classifier - with data from the Peruvian National Household Survey for Puno, 2019 to learn a binary classifier that discriminates which households are at risk of being affected by a cold wave-related disaster. Machine Learning would help to build a risk screening tool that can be tuned, in terms of models' hyperparameters, to maximize predictive power considering the importance of false negatives.

Puno, in Peru, is affected by recurrent cold waves. Peruvian's South Andean Region is especially susceptible to these types of hazards. Since 2000, considering world-total historical data on disasters caused by Extreme Low-Temperature Events (ELTEs) registered in EM-DAT (2022), 21.28% have affected this geographic boundary. According to EM-DAT estimations, the most harmful ELTE was recorded in 2004 as a cold wave of -35°C that affected 40.30% of the total population of 15 Peruvian regions. Puno is a rural and low-densely populated region in southeast Peru. Puno is the epicenter of ELTEs affecting PSAR, as 70.00% of events registered in EM-DAT affected Puno from 2003 to 2015. As ELTEs affect a sizeable geographic boundary, estimating the number of affected people and economic losses, for example, could be challenging.

Research on proactive disaster risk reduction would significantly impact Puno because of the high prevalence of agricultural households. These disasters may cause economic losses that impact their long-term wealth. If a community is unprepared to face cold wave-related disasters, it might enter a vicious cycle of cold waves affecting the economy, shaping disaster impacts. This vicious cycle affects the ability to respond and recover from disasters, producing a lower budget to invest in resilience mechanisms (Besiou et al., 2021).

The proactive intervention on Puno may significantly impact the disaster response and recovery. Following Holguin-Veras et al. (2013), resources invested in response and recovery include logistic and deprivation costs. An optimized predictive model would identify which households would be the target of proactive interventions. Puno is a case study characterized by spatially dispersed final demand points and high peaks of deprivations caused by accumulated vulnerabilities (Kim and Sohn, 2018; Quiliche et al., 2021); thus, accurate forecasts are significant. Assessment of delivery strategies, transportation costs, and their balance with deprivation costs are left for future complementary research as the social objective function, which includes logistic costs, is the primary concern of humanitarian logistics.

This paper's contribution is twofold. First, we propose a supervised learning pipeline to produce an accurate classifier based on households' vulnerability conditions. This proposal is grounded in previous research. The pipeline includes hyperparameter optimization to automate the search for hyperparameters, including the adaptation of the pipeline to domain requirements, which, in this case, is equivalent to considering the importance of deprivation costs. Second, we report statistical interpretations of the classifier using the logic of average marginal effects to provide decision-makers with more practical insights.

The learned predictive model is expected to contribute to reducing social costs while considering the importance of deprivation costs (Holguin-Veras et al., 2013). As the focus is on disaster preparedness, the predictive model will identify the final demand points that need the prepositioning of supplies, thus producing information regarding the number of supplies required or the demand for humanitarian aid to perform proactive interventions. In the context of disastrous events, the value of information on where and at which level to preposition supplies is high. Those supplies aim to reduce the expected damages to households' livelihoods strongly linked to agriculture and livestock (Quiliche and Mancilla, 2021).

The remainder of this paper is divided into five Sections. Section 2 describes the main works on SLAs, Machine Learning applications to disaster risk management, and emergency assessment. Section 3 details the case study of Puno, Perú. Section 4 depicts data collection methods, Machine Learning pipeline, and experimental setting. Section 5 brings the main results, descriptive analysis, model performance, and further analysis of results. Section 6 discusses the results and their practical implications, such as statistical interpretation and threshold tunning. Finally, Section 7 brings our conclusions and recommendations for improvements in disaster preparedness strategies and future research avenues.

2.2. Theoretical foundation

2.2.1. Disaster risk reduction for clime-related disasters

The most outstanding theory on disaster risk claims that risk is produced if three elements are combined inside a geographic boundary (Ramos et al., 2010; UNDRR, 2015; Twigg, 2004): i., natural hazard, i.e., the natural phenomenon that may harm communities; ii. exposure, i.e., the condition of an agent within the geographic boundary of being exposed to such natural hazard; and iii. vulnerability, which shapes the consequences of a damaging event on agents. If an agent is resilient to disasters, it would have small losses after a disastrous event. Vulnerability is a set of conditions that an agent possesses, making it more prone to high losses when affected by a hazardous event (Christian et al., 2021; Sahana et al., 2019; Tasnuva et al., 2020; Ullah et al., 2022). Among natural hazards that jeopardize vulnerable communities, clime-related hazards such as rainfalls, heat waves, cold waves, or storms have an impact that covariates with the degree of vulnerability of the agents within the geographic boundary exposed to such hazards (Renteria et al., 2021). Furthermore, these hazards tend to be seasonal and localized in a geographic boundary, and the magnitude of losses can be anticipated by considering vulnerability (Simmons and Sutter, 2014).

The challenge of disaster risk reduction comes from vulnerability shaping the magnitude of the losses related to agents' exposure to natural hazards. Disaster risk can be mitigated by reducing vulnerability, or equivalent, by creating resilience, as stated in the Sendai Framework for Disaster Risk Reduction (Aitsi-Selmi et al., 2015). However, the reduction of vulnerability is a long-term goal. From an economic perspective, communities need resources to face disasters. Then, disaster risk reduction could be incredibly challenging when a community is affected by recurrent disasters. In those cases, the resources allocated to disaster response and recovery are more likely to be higher than those invested in risk mitigation and disaster preparedness. Thus, the total cost of the disaster risk management cycle is steadily high, as illustrated by the red line in Figure 1.

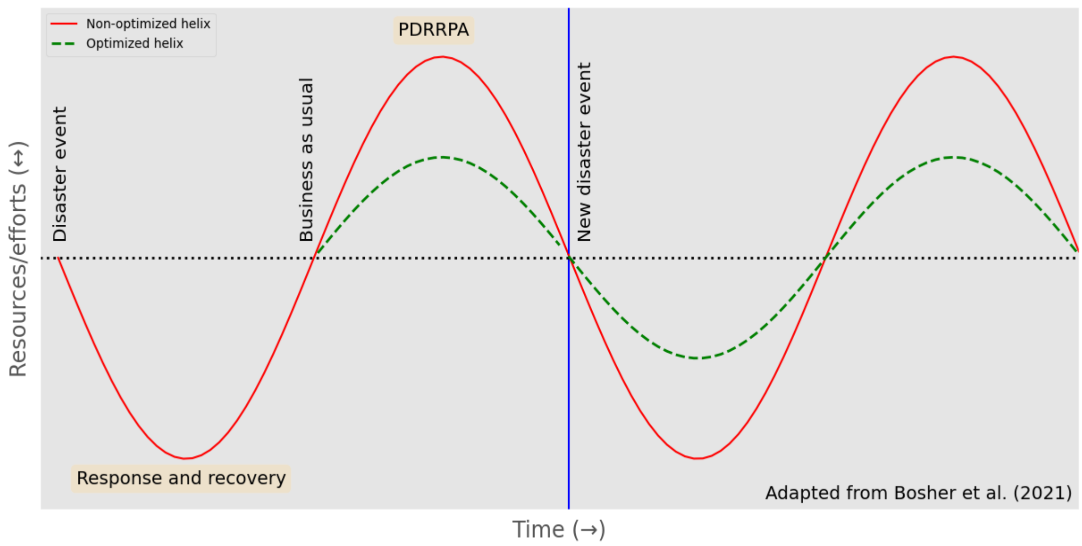

Figure 1: Theoretical representation of Disaster Management Helix optimization.

The helix concept for disaster risk management illustrates the dynamics of disaster risk reduction in the case of recurrent disasters. The long-term is of particular importance. In cases where communities are affected by recurrent disasters, unmitigated disaster losses might harm the overall economic environment by causing infrastructure destruction, systemic agricultural losses, and hazards to public health (Ferreira, 2012; López-Bueno et al., 2021; Quiliche and Mancilla, 2021).

In sum, there are two aspects of disaster risk theory that we want to highlight. The first is that vulnerability is, by definition, the best predictor of disaster risk. Hence, it is crucial to collect data on vulnerability conditions. For instance, Linardos (2022) reported some studies deploying neural networks to collect data on households' equipment and construction materials to predict disasters. The second is that the significant implication of this research is that disaster risk reduction can be done efficiently by having accurate information to plan and implement the pre-disaster risk reduction and preparedness activities such as stock pre-positioning and improved housing infrastructure, among others. If well done, these activities may positively affect the long-term dynamics of the Disaster Management Helix in Figure 1, leading to a more significant cold wave-related risk reduction in Puno.

2.2.2. Machine Learning in Disaster Risk Reduction

Previous studies addressed disaster preparedness with predictive analytics (Davis et al., 2010; Simmons and Sutter, 2014; Van Thang et al., 2022). There are several contributions of Machine Learning to disaster risk management. Lu et al. (2021) performed a comprehensive review of applied Machine Learning in the context of public health emergencies related to disasters. The authors found that the main contribution of Machine Learning is to process information to support decision-making in managing risks by producing forecasts and insights to improve understanding of phenomena. For example, automated models can improve decision-making under time-sensitive conditions by processing big data. In this sense, Machine Learning contributes to multiple edges of information management: demand forecasts may help to reduce material convergence

(Holguin-Veras et al., 2014), stochastic programming in transportation may help to avoid bottlenecks (Alcántara-Ayala, 2019), and so on. Machine Learning to predict and understand complex phenomena helps to mine valuable insights from data (Fayyad and Shapiro, 1996; Tomasini and Van Wassenhove, 2009; Behl and Dutta, 2018). Izquierdo-Horna et al. (2022) applied a hybrid approach to seismic risk assessment in Perú, integrating Random Forest and Hierarchical Analysis to determine seismic risk in Pisco. China is a country known for having densely populated cities. An early-awareness approach based on Machine Learning is beneficial in that context, such as the approach proposed by (Bai et al., 2022), by which a disaster response plan can be executed within a more extended time window before flooding is at its peak.

A critical gap identified in Machine Learning applications for disaster risk reduction is that predictive modeling simplifies vulnerability by economic factors. A multi-dimensional approach must be included to better represent vulnerability (for example, Ahmad and Routray, 2018; Patri et al., 2022). This multi-dimensional vulnerability approach contributes to a better understanding of clime-related disaster risks and improves prediction accuracies (Ramos et al., 2010; Zhao et al., 2022). The applied Machine Learning contributes to accurate estimates of demand for better disaster risk management.

The vulnerability dimensions are composed of endogenous variables—these features might covariate with other predictors not considered in this paper. For example, vulnerable agents tend to be settled in places with high exposure. The classifier is expected to exploit these relationships to produce accurate predictions. According to theory, the essential variables for creating our Machine Learning algorithm are next defined (López-Bueno et al., 2021; Renteria et al., 2021).

Low income and lousy infrastructure are the main drivers of vulnerability to climate-related disasters, according to Tasnuva et al. (2020). Bad outcomes in health, such as a high prevalence of chronic illness, could also be related to a higher vulnerability (Djalante et al., 2020). Specific configurations of socio-economic variables make households especially vulnerable, such as unemployment and low educational achievement. There is evidence that younger and female head of households is related to the probability of being affected by a disaster (Rapeli, 2017). Geographical vulnerability depends on household location, which at the same time is determined by economic vulnerability: households located in vulnerable areas tend to be poor, and this magnifies the vulnerability condition (Mattea, 2019).

2.3. The case of Puno, Perú

This paper analyzes the case of cold wave-related disasters. Cold wave-related disaster risk is sensible to vulnerability. Figure 2 illustrates the triggering process of cold wave-related disasters, from natural hazards to disasters impacting populations (Quiliche and Mancilla, 2021). Losses occur when exposure meets vulnerability (i.e., if an agent had been resilient to cold waves, it would not have been affected by the disaster). Hence, disaster risk reduction is a priority for communities affected by cold waves.

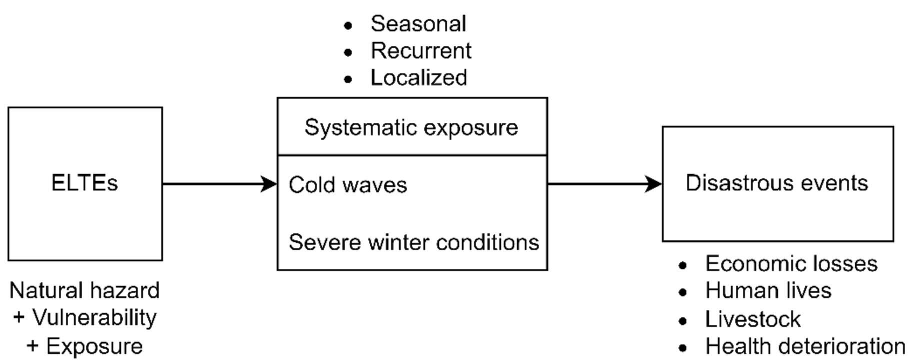

Figure 2: Causes of cold-related disastrous events affecting communities.

The analysis for the Puno case considers the household level. This level of granularity allows the researchers to draw insights into the points of final demand for aid (Eckhardt et al., 2019; Eckhardt et al., 2022; Jardim et al., 2022). Such information is valuable for developing disaster risk reduction strategies.

Vulnerability characterization by socioeconomic features is critical to the proposal of this paper. López-Bueno et al. (2021) performed a statistical analysis of mortality rates in urban and rural areas of Madrid, Spain. The authors conclude that the main risk drivers of mortality rates are socio-economic. Amirkhani et al. (2022) found an interesting pattern for a cross-section of countries worldwide for 1999-2018 using EM-DAT (2022): cold waves and severe winter conditions caused more deaths in middle-income countries than in high-income ones.

The demand is screened using a supervised learning classifier. Hence, a household at risk is a demand point. Then, minimizing False Negatives and False Positives, some misclassification sources is crucial. In Figure 3, the localization of most final demand points (known a priori from past data) is in rural areas outside the principal cities (Gutjahr and Fischer, 2018). The main cities can be identified on the map by conglomerates of households closer than 5km to each other; in the north, several isolated households are located in rural areas and have positive risk classification. False Negatives are households at risk of being affected by a cold wave-related disaster labeled “without risks.” False Negatives would produce deprivation costs because those households need aid, but the model decides they do not (Tomasini and Van Wassenhove, 2009; Eckhardt et al., 2017). A model training was adapted to minimize false negatives to produce accurate classifications with reduced deprivation costs to tackle this obstacle.

The time series of minimum temperatures reported in Figure 4 illustrates the seasonality of the cold waves in Puno (SENAHMI, 2022). Every year, households located within Puno are exposed to cold waves. In July, August, and September, the exposure tends to be higher on average for all the meteorological stations that collect temperature data in Puno.

This recurrent exposure causes frequent disasters affecting most of Puno’s population (Alarcón and Trebejo, 2010). Disaster risk management is different when dealing with recurrent disasters (in contrast to the case of Earthquakes that happen once in a decade in Peru). This paper proposes to estimate the demand for aid in the aftermath of Cold Waves.

In sum, this paper covers the problem of disaster risk reduction for communities with recurrent disasters. Then, it proposes to train a classifier using Machine Learning methods to identify points of final demand and support pre-disaster risk reduction and preparedness activities. Hence, a more significant impact on model implementation is expected in the pre-disaster phase. Nevertheless, the insights may be helpful for post-disaster response and recovery activities, as they also contribute to understanding vulnerability drivers at the household level.

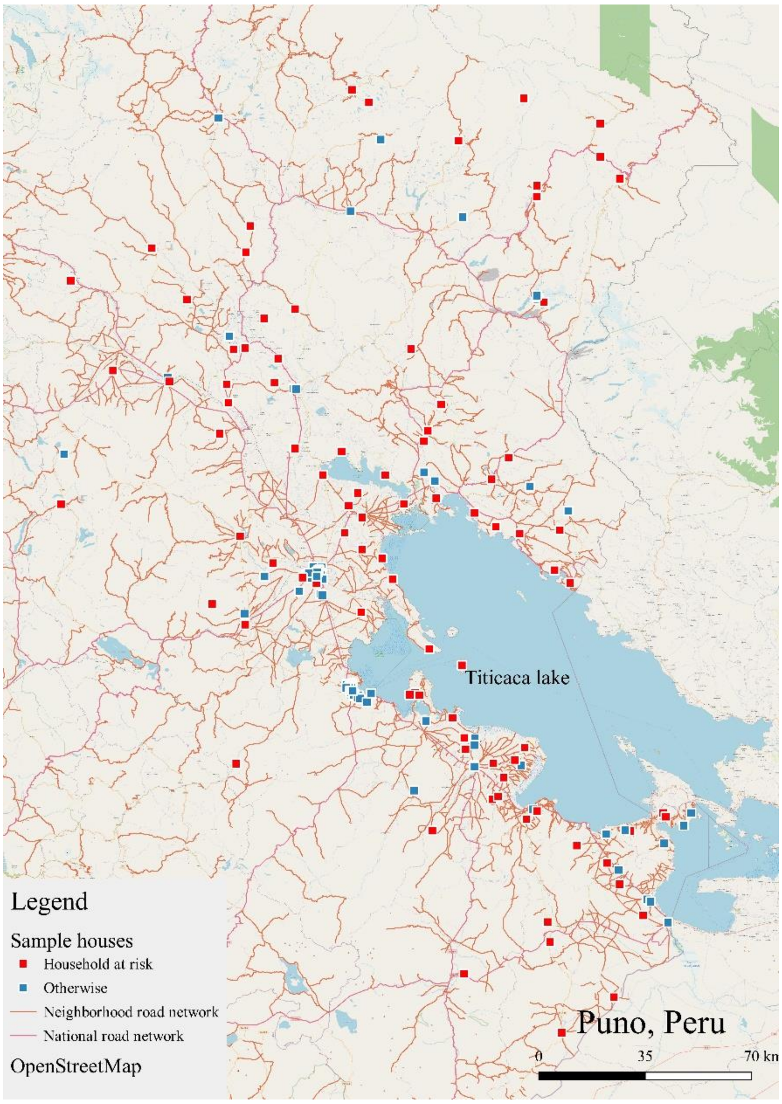

Figure 3: Spatial distribution of households exposed to ELTEs

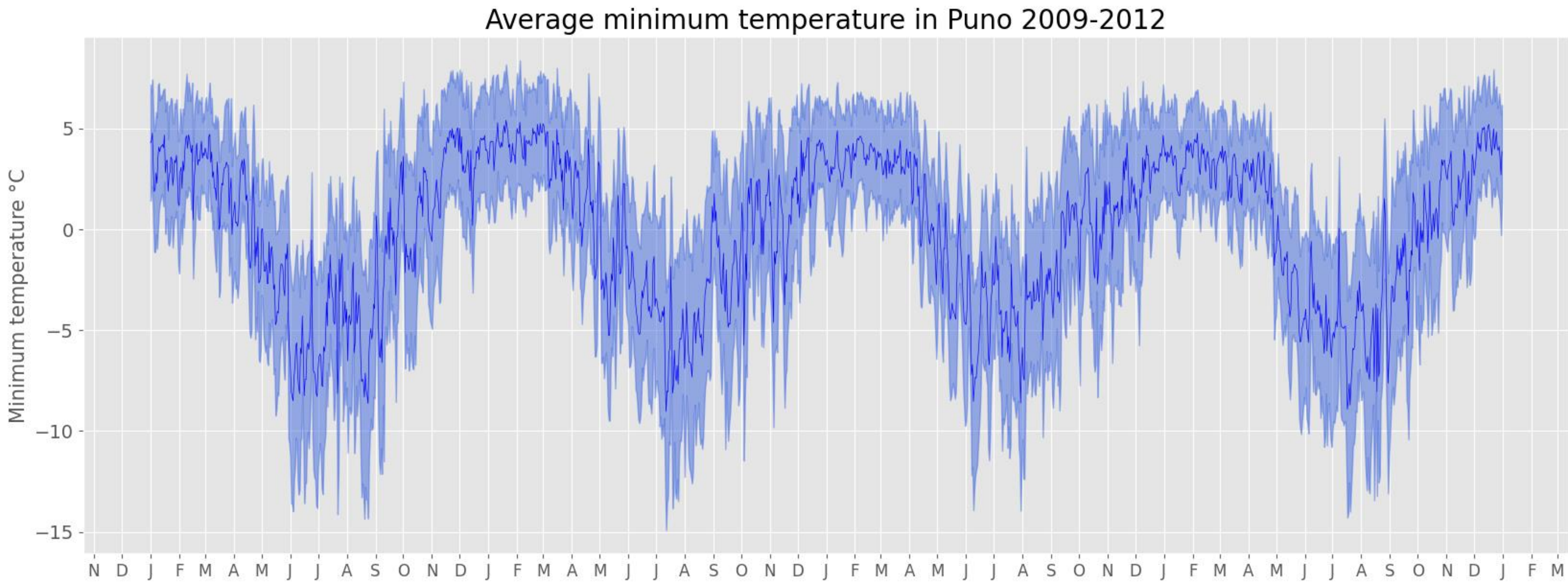

Figure 4: Time series plot for average minimum temperature in Puno 2009-2012

2.4. Materials and methods

2.4.1. Data collection methods and the classification problem

Raw household vulnerability characteristics data were collected from the National Household Survey by Peruvian National Institute of Statistics and Informatics in 2018-2021. Data is available at the national level. The survey's sampling method was stratified over political regions. Thus, the survey is representative of Puno at the regional level. The following survey modules were considered for this analysis: population and housing (modules 100 and 200), education (module 300), health (module 400), employment (module 500), and democracy and transparency (module 612). These modules contain information about the defined dimensions of vulnerability (UNDRR, 2015; Salazar-Briones et al., 2020; Renteria et al., 2021).

The following question is asked to the informers:

In the last 12 months, has your house been affected by natural disasters (drought, storm, plague, flood, etc.)?

The target variable equals one if the respondent said natural disasters had affected their house. In the binary classification jargon, this category is also labeled as positive.

$$1)\mathrm{Y_i} = \begin{cases} 1 \text{ if the household is at risk of being affected by a cold-related disaster} \\ 0 \text{ otherwise} \end{cases}$$

Even though this variable does not provide specific information about the type of disaster, we consider it appropriate to represent risk associated with cold waves because:

1. For the specific case of Puno, there is an overwhelming prevalence of risks related to low temperatures (see Section 2.3 for data analytics support for this proposition). To some extent, every household has a latent degree of cold wave-related disaster risk.
2. The average household's monthly earnings are US$139.53, and the poverty line is estimated at US$104.45 (conversion rate of 1US$ = S/. 3.37). The literature emphasized the importance of economic deprivations driving cold-related disaster risk (López Bueno et al., 2021).

3. Suppose a household is at risk of being affected by a drought, storm, plague, flood, or landslide. In that case, it would likely be at risk of being affected by another climate-related disaster, such as cold waves-related disasters (Rentería et al., 2021). The mechanism explaining this correlation is the vulnerability conditions these households share.

Considering this evidence, it seems reasonable to operationalize the target variable as in Equation 1: equal to one when the household is at risk of being affected by cold waves-related disasters and zero otherwise.

2.4.2. Machine Learning Pipeline

Supervised learning was applied for binary classification, considering that the target variable is categorical but binary encoded (see Equation 1). The methodological approach for model training was based on a standard framework for Machine Learning model training (Giovanelli et al., 2021; Waring et al., 2020). It included three main steps: pre-processing, data processing, and post-processing. Model testing included an experimental validation method for each supervised learning algorithm. The objective was to find the best performing, most explainable, and parsimonious model (Hastie et al., 2001). This model must perform well on unseen data or testing data.

The procedure illustrated in Figure 5 was followed to reach a model with the abovementioned characteristics.

Figure 5 shows an experimental setting different from the classical random train-test split approach. The objective of this experimental setting is to discuss what would have been the outcome of model implementation in 2021 and, hence, shedding light on the practical implications of the implementation of Machine Learning techniques into disaster risk reduction. Additionally, this procedure implies hyperparameters' Optimization, which improves model performance based on an objective function.

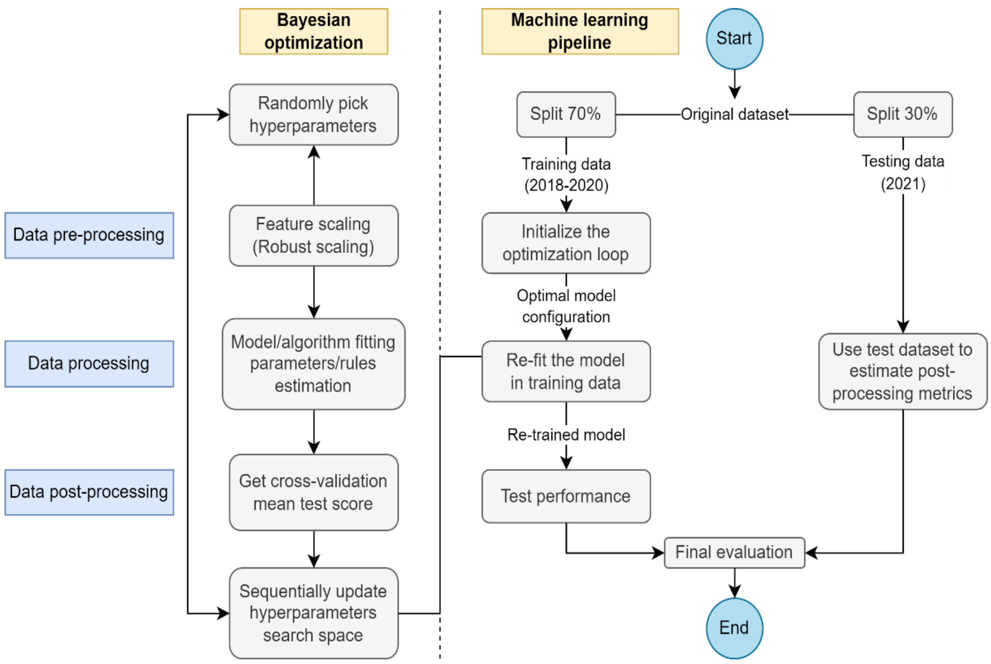

Figure 5: Detailed procedure for Machine Learning model training

Training data (2018-2020) pass through several steps, which include fitting the model to random subsets or batches of the pre-processed data and then doing this multiple times to get a cross-validation score. The Hyperparameter optimization loop is repeated until a certain of iterations are reached. Then, the trained model is used to estimate metrics in test data and evaluate performance. Sections 2.4.2.1 to 2.4.2.3 explain the three phases applied in the optimization loop.

2.4.2.1. Data pre-processing

The feature space extracted from the survey is multi-dimensional. This data feature overcomes the empirical over-simplification of existing disaster vulnerability studies, which are restricted to the socio-economic dimension, thus ignoring the dependence on other factors (Villarroel-Lamb, 2020; Regal, 2021; Szczyrba et al., 2021 are some examples). However, the proposed feature space entails greater empirical complexity as more features are considered for the model training process. Some features may not be robust predictors of the outcome, and supervised learning algorithms must consider a feature selection process (Xu et al., 2019).

The dataset comprises 86 features, of which 84 are binary and two are numeric. The greater the number of features, the more computational time is required for hyperparameter search for each supervised learning algorithm. The literature identifies the three most-used approaches to handle multi-dimensional datasets with supervised learning algorithms: dimensionality reduction, sequential feature selection, and model-based feature selection (Venkatesh & Anuradha, 2019; Pedregosa et al., 2011).

Model-based feature selection was adopted in our procedure, known in the literature as sparse learning (Xu et al., 2019). On the one hand, standard dimensionality reduction techniques, such as Principal Component Analysis, resulted in a significant loss of information that negatively impacted the predictive power of supervised algorithms in preliminary experiments performed with this data. On the other hand, sequential feature selection increased the computational time by increasing the time required to evaluate each hyperparameter configuration and was discarded in our procedure.

Using several dummies may lead to collinearity and a singular matrix. However, this pattern only caused delays in Logistic Regression model training and, in any case, caused non-convergence of the vector of parameters. Thus, it was addressed in Logistic Regression by using a stochastic gradient descent solver that speeded up computations. Collinearity was not a problem for the Random Forest Classifier.

Following packages' documentation guidelines, supervised learning algorithms' performances improve when input features are measured on the same scale (Pedregosa et al., 2011). As shown in Figure 5, the first step in cross-validation iteration is to scale the data. The scaling method is called Robust Scaling, a variation of Standard Scaling that uses median and interquartile ranges for Scaling, thus producing more robust features' standardization (Zheng and Casari, 2018). Missing data was removed before the Robust Scaling.

2.4.2.2. Data processing

Elastic-Net Logistic Regression (ENLR) and Random Forest Classifier (RFC) were selected because of their functionalities regarding features' importance (Micheletti et al., 2014). These algorithms rank the feature's importance and reach the optimal predictive formula as a function of a subset of features, removing large amounts of redundancy and noise in the dataset (Xu et al., 2019).

This paper considered the trade-off between expected performance and interpretability as additional criteria for selecting the best classifier. According to the experimental results of Fauvel et al. (2022) on UCI datasets, ENLR outperformed Support Vector Machines (SVM), Local Cascade, and Multilayer Perceptron (MP). It performed almost as well as Bagging and Boosting and Simple Ensemble Methods that use ensembles of Decision Trees, Gaussian Naïve Bayes, and Stochastic Gradient Descent. On the other hand, RFC outperformed other algorithms, including XGBoost, SVM, Gradient Boosting, Multilayer Perceptron (MP), and ENLR. In the experiments, RFC was the second-best supervised learning algorithm.

ENLR and RFC were selected because other algorithms may perform equally but be more complex to explain their logic to relevant stakeholders. RFC algorithm has some advantages over ENLR. RFC may consider cross-influencing factors in prediction, but ENLR assumes that features are independent of each other and, thus, fails in accounting for cross-influencing factors such as other linear models. ENLR is easier to interpret but sensible to outliers. However, significant outliers were not found in the data, and ENLR could run in tractable time (the stochastic gradient descent method to fit ENLR may facilitate the convergence of ENLR objective function).

- **Elastic-Net Logistic Regression**

Zou and Hastie (2005) proposed for the first time the Elastic-Net regularization technique as a combination of the Least Absolute Shrinkage Selection Operator (LASSO), known as the L1 regularization, and Ridge regression, known as L2 regularization, terms. The adaptation to Logistic Regression was proposed in the literature using different solvers and formulations, but the one used here is based on Pedregosa et al. (2011). The objective function is stated as follows:

$$\min_{\beta,\beta_0} \frac{1-\rho}{2}\beta^{T}\beta + \rho||\beta|| + C\sum_{i}^{N}\log\left(\exp\left(-Y_i\left(x_i^{T}\beta + c\right)\right) + 1\right) \quad (2)$$

Where x_i^T is a data vector corresponding to observation i, Y_i is the respective observation point for target classes. Considering that both optimal Elastic-Net mixing parameter ρ and C inverse of regularization strength are selected based on cross-validation scores, the vector of parameters β is estimated to fit the optimal model to training data, as shown in Figure 5. Equation 2 shows the loss function for ENLR and is minimized through Stochastic Gradient Descent (Bottou, 2010 and Pedregosa et al., 2011) with a learning rate equal to η.

We next define the hyperparameter search for ENLR in Equation 3:

$$ENLR(.) = \begin{cases} \text{Penalty='Elastic-Net'} \\ \eta\text{='Optimal'} \\ \text{C}\sim LOGU(1E^{-2}, 1E^{2}) \\ \rho\sim U(0,1) \end{cases} \quad (3)$$

The procedure in Figure 5 searches for best "C" and ρ considering a priori uniform distributions for both parameters, being "C" defined in logarithmic space.

- **Random Forest classifier**

The algorithm is an ensemble of Decision Trees fitted with the CART algorithm (Jackins et al., 2021) on multiple sub-samples of a dataset. Trees are pruned and then averaged to balance the bias-variance trade-off and maximize the predictive power of the ensemble (Pedregosa et al., 2012). the following steps were followed to train RFCs (Xin and Ren, 2022):

Algorithm 1. Random Forest Classifier

Random Forest Classifier

1. Randomly select a subset of features K_{max}.

2. Randomly sample N observations with replacement.

3. Calculate the first node using the best-split point under criterion CRIT with the obtained subset of data, following the rules defined above (this applies for further nodes):

   3.1. The minimum number of data points placed in a node before the node is split equals Min_{split}.
   3.2. The minimum number of data points allowed in a leaf node equals Min_{leaf}.
   3.3. Perform cost-complexity pruning of lower information-gain nodes according to CPP_{α}.

4. Categorize the node into daughter nodes using the best split with selected criterion CRIT.

5. Categorize more daughter nodes until the tree reaches the defined Max_{depth}.

6. Repeat steps 1 to 5 $N_{estimators}$ times to build the same number of trees, which refers to the size of the ensemble.

7. Build the prediction algorithm by averaging the probabilistic prediction over the ensemble

Authors' adaptation from Jackins et al. (2021).

We next define the hyperparameter search for RFC in Equation 4:

$$RFC(.) = \begin{cases} \mathrm{K_{max}} = 1 \\ \mathrm{CRIT} = UCAT['Gini', 'Entropy'] \\ \mathrm{Min_{split}} \sim U(0.5,1) \\ \mathrm{Min_{leaf}} \sim U(0.5,1) \\ \mathrm{CPP_{\alpha}} \sim U(0,0.1) \\ \mathrm{Max_{depth}} \sim UINT(1,20) \\ \mathrm{N_{estimators}} \sim UINT(0,100) \end{cases} \quad (4)$$

Cross-validation helps to identify the optimal values of CRIT, $\mathrm{Min_{split}}$, $\mathrm{Min_{leaf}}$, $\mathrm{CPP_{\alpha}}$, $\mathrm{Max_{depth}}$, and $\mathrm{N_{estimators}}$. After the cross-validation loop, the optimal ensemble is fitted to training data, as illustrated in Figure 5.

2.4.2.3. Data post-processing

This section describes what happens at the end of every cross-validation loop. Model performance metrics are calculated for each hyperparameter configuration in each iteration, sampled randomly from hyperparameter search spaces defined in Equations 3 and 4. Within the cross-validation loop, a training set is randomly shuffled and split into F folds of equal size; the algorithm is trained with a sample composed of $F - 1$ folds and tested on the remaining. This procedure produces F performance metrics that are averaged to have a point estimate of the performance of the corresponding hyperparameter setting. This procedure is known as K-Fold cross-validation Pedregosa et al., 2012. For robustness purposes, the K-Fold cross-validation method is repeated R times in each iteration, known as Repeated K-Fold cross-validation (Pedregosa et al., 2012).

- **Bayesian Optimization**

In a Grid Search or Random Search scheme, every iteration is independent of the other, and the optimization program would sample $N_{iterations}$. The more hyperparameters to tune, the bigger the required number of iterations. The combinatorics of possible hyperparameter configurations in the RFC algorithm are particularly large. Due to combinatorial search spaces, optimizing hyperparameters is an NP-Hard problem (Yang and Shami, 2020).

Hyperparameter optimization techniques are essential because they improve the performance of ML models. Bayesian Optimization is used as a sequential hyperparameter optimization scheme to overcome the computational complexity inherent in hyperparameter optimization procedures. In the Bayesian method, each cross-validation iteration depends on the previous one. Further theoretical and computational details can be reviewed in Owen (2022).

- **Objective function**

The objective function of Bayesian Optimization is typically defined as the model's accuracy or another performance metric. This paper's objective function is a linear convex combination of Matthews Correlation Coefficient (MCC) and Sensitivity (True Positive Rate).

On the one hand, MCC represents an accurate model regarding both classes (Chico & Jurman, 2021). On the other hand, sensitivity captures the ability of the model to predict positive classes. For ground truth positive classes, this is known as the True Positive Rate (Luque et al., 2019). In Section 3, the importance of deprivation costs was introduced. The definition of this objective function is based on the importance of False Negatives. The True Positive Rate decreases with the increase of False Negatives. Hence, the objective function is shown in Equation 5:

$$Z_m = \lambda . MCC_m + (1-\lambda) . Sensitivity_m, \forall_m, \lambda_\in [0,1] \quad (5)$$

The greater the λ coefficient, the lower the importance of Sensitivity or False Negatives. The value of λ was set to 0.5 in this experiment. The core assumption for this optimization is that there is a trade-off between accuracy of both classes and accuracy of positive class (or Sensitivity). If the model misclassifies positive classes, it labels risk households as non-eligible for humanitarian focalization. Thus, maximizing the $\mathrm{Z_m}$ leads to an accurate and deprivation costs-aware model.

After Bayesian Optimization, the following metrics were calculated to detail model performance for the test dataset:

- **Model performance metrics**

Area Under the ROC Curve (AUC)

This metric represents the distance between the 'no discrimination' classifier (the worst classifier that distributes the predictions over classes uniformly for any probability threshold) and the tested classifier. It is defined in the function of $TruePositiveRate = \frac{TP}{TP+FP}$ and $FalsePositiveRate = \frac{FP}{TP+FP}$ coordinates at various probability threshold settings. The range of this metric varies in the closed interval [0,1], so better classifiers are found when $AUC \to 1$.

Accuracy

The accuracy estimation represents the application of a common heuristic where the diagonal of the confusion matrix is maximized. The formula is given by $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$. The range of this metric varies in the closed interval $[0,1]$, so better classifiers are found when $Accuracy \to 1$.

F1-Score

F1- Score is defined as the harmonic mean of the $Precision = \frac{TP}{TP+FP}$ and $Recall = \frac{TP}{TP+FN}$. The formula is given by $F1 = \frac{TP}{TP+0.5(FP+FN)}$. The range of this metric varies in the closed interval $[0,1]$, so better classifiers are found when $F1 \to 1$.

Matthews Correlation Coefficient

This metric is a correlation coefficient in the [-1,1] interval. The formula is given by $MCC = \frac{TP(TN)-FP(FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FP)}}$. It was selected to choose the best classifier as it tends to co-optimize all elements of the confusion matrix for binary classifications (Luque et al., 2019; Chicco and Jurman, 2020). By maximizing this metric, the classifier minimizes both deprivation and logistic costs.

All the Machine Learning pipeline steps were performed on Python 3.11 programming language using packages Scikit-Learn 1.2.0, Scikit-optimize 0.8.1, Pandas 1.5.2, and NumPy 1.24.1. Data analytics was built using Matplotlib 3.6.2 and Seaborn 0.12.1.

2.5. Description of results

This section presents and describes the main results. First, we characterize the study case using descriptive analytics of features that might predict cold wave-related disaster risk. Second, we present the main results of Machine Learning model training, including model selection. Third, we show descriptive analytics of False Negatives and False Positives to enrich the description of results (Dantas et al., 2021).

2.5.1. Descriptive characterization of Puno

According to historical data, the urban public infrastructure in Puno is poor. Whether households are settled in rural areas, 15.94% have inlaid walls, 52.43% have tracks, 21.93% are paved, and 40.94% are settled near a lighting pole. Regarding ownership, 82.55% of households are owned, but 22.17% have a title of ownership. Housing infrastructure is fragile: 27.49% of households have walls of concrete. Most households in Puno are settled in rural areas (59.04%) at an average altitude of 3880 meters above sea level.

Regarding access to essential services, 34.28% of households are connected to a water and drainage network, and 55.39% have daily access to water for consumption. Nevertheless, access to electricity has improved, with 89.33% of households with electric lighting compared to 74.18% in 2017. Households without electricity use candles (7.14%) or other lighting (3.53%). The main cooking methods are GLP (60.58%) and manure (39.86%). Manure cooking is a characteristic of rural livelihoods (Sagastume-Gutiérrez et al., 2022). Thus, the prevalence of manure cooking is explained by the prevalence of rurality. Regarding access to Information and Communications Technologies, 14.05% of households have internet access, but 83.05% have a cellphone.

Households are equipped with assets like color TVs (47.36%), bicycles (32.05%), motorcycles (24.35%), and DVDs (24.38%). Just 6.54% of households have a particular car, which is explained by the observed poor urban infrastructure. In modern society, ICTs grant opportunities and capabilities for individuals (Oyelami et al., 2022); however, just 18.14% of households have a computer or laptop. Just 8.68% of households have a refrigerator. The annual per capita expenditure approximates short-term household nominal income. The average annual per capita expenditure is US$1634.29. The average expenditure is below Latin America's principal cities, such as Lima, Bogotá, Buenos Aires, and Rio de Janeiro. It is worth mentioning that the mean income is above the median, meaning that more than half of the per capita expenditure distribution is below the average, showing some degree of income inequality.

It is common to find old adults (51 to 65 years old) and old (more than 65 years old) household heads (59.95%). Even though Puno is not densely populated, 38.47% of households are overcrowded, which means they have more inhabitants than bedrooms. 40.68% of households' heads are married. Puno has a poor development of human capital: 19.56% of households' heads are illiterate, 63.02% have no education, and just 2.25% have a postgraduate degree.

Lastly, the population faces a high prevalence of acute illness (96.24%) and chronic illness (87.52%). More than half of the households in the sample have at least one member who searched for medical attention (67.14%), and 73.32% have a subsidized health insurance regime. 32.79% of households have at least one member with one or more disabilities.

Table 2: Multi-dimensional vulnerability features

Category	**Variable**
Household exterior and access to public goods	Households with inlaid walls, households with painted walls, Outside tracks are paved, Outside tracks are terrain, Outside paths, Lighting poles, No public good.
Ownership and physical characteristics	Independent house, the household is a house, the household is totally owned, the household has a title of ownership, Concrete walls, Concrete floor, Concrete roof, Overcrowded bedrooms, No other rooms than bedrooms.
Access and use of essential services	Water network, Potable water, Quality water (chlorine), Daily access to water, Drainage network, Electric lighting, Candle lighting, Other lighting, GLP cooking, Wood cooking, Other cooking, Manure cooking, Phone, cell phone, Cable TV, Internet
Household income and assets	Per capita expenditure, Radio, Color TV, Black-White TV, Sound equipment, DVD, Computer or laptop, Electric iron, Electric blender, Gas stove, Refrigerator, Cloth washing machine, Microwave oven, Sewing machine, Bicycle, Car, Motorcycle, Tricycle
Socio-demographics	The head is employed, The head is a woman, The head is married, The head is literate, The head has no education, The head achieved basic education, The head achieved technic education, The head achieved a college education, The head achieved pos-graduate education, The head is a young adult (17-35), The head is an adult (36-50), The head is an old adult (51-65), The head is old (more than 66)
Health and insurance (for household members)	Illness (last month), Accident (last month), Healthy (last month), Chronic illness, Medical intervention (last month), Contributory health insurance, Subsidized health insurance, Disabilities
Geographical context	The household is located in a rural area, Altitude.

Authors' own elaboration from ENAHO (2023).

Figure 6 shows the correlation heatmap of features listed in Table 2. Statistical correlation between features was estimated using Spearman's Rank-Order Correlation. There are some yellow points in housing variables, such as construction materials. Then, if walls are made of concrete, it is likely that the roof and floor are also made of concrete. Furthermore, economic vulnerability indicators are related to each other; this suggests that measurement is consistent between individual indicators.

However, the high correlation between features is not an obstacle. It is worth mentioning that both ENLR and RFC have mechanisms to handle correlated predictors, so all the variables were kept. Then, the interpretation of the results was made based on post-estimation feature importance.

Following results from the Spearman correlation matrix, households with concrete walls and floors tend to connect to a water and drainage network in urban areas. Rural households have fewer assets, lower educational levels, health access, and lower acute illness prevalence. We next report model training results.

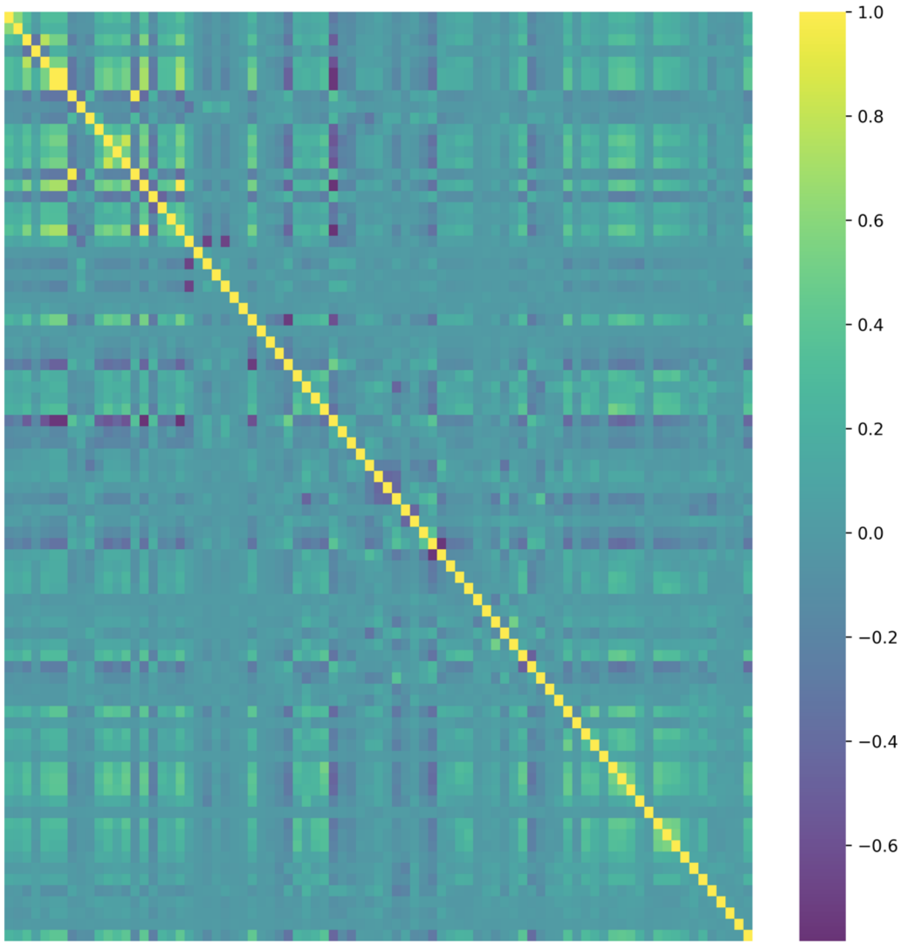

Figure 6: Features' correlation heatmap

2.5.2. Model training results

As optimal hyperparameters were selected based on performance on the training dataset, it is crucial to analyze how trained models perform on unseen data. We use data from 2021 as a test dataset to perform this analysis. Table 3 summarizes the main results regarding model performance.

Table 3: Models' performance on the test dataset (Puno, 2021).

Classifier	ROC-AUC	Accuracy	F1-Score	MCC	Sensitivity
ENLR	73.76	73.5	73.31	47.48	77.75
RFC	74.24	73.82	74.3	48.64	80.9

Authors own elaboration

RFC was selected as the best predictive model for the case of cold wave-related disaster risk in Puno. The RFC produced more accurate results than ENLR and achieved higher sensitivity, making it less prone to misclassify households at risk of being affected by cold wave-related disasters. We next report the optimal hyperparameter configuration in Equation 6:

$$RFC^{*} = \begin{cases} \mathrm{K}_{\mathrm{max}} = 1 \\ \mathrm{CRIT} = Entropy \\ \mathrm{Min}_{\mathrm{split}} = 7 \\ \mathrm{Min}_{\mathrm{leaf}} = 9 \\ \mathrm{CPP}_{\alpha} = 2.47E - 4 \\ \mathrm{Max}_{\mathrm{depth}} = 9 \\ \mathrm{N}_{\mathrm{estimators}} = 40 \end{cases} \tag{6}$$

For reproducible purposes, the trained model was saved to a file to be loaded in software to reproduce the results or to use the model for further practical implementations. We report below the corresponding confusion matrix in Figure 7 for the Test Dataset; we also report the confusion matrix in Figure 8 for the Train Dataset:

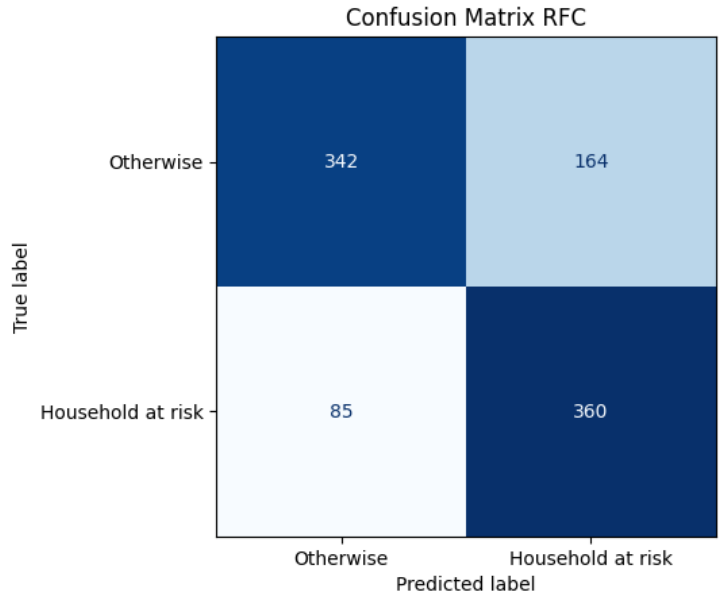

Figure 7: Confusion matrix for Random Forest Classifier (Test set)

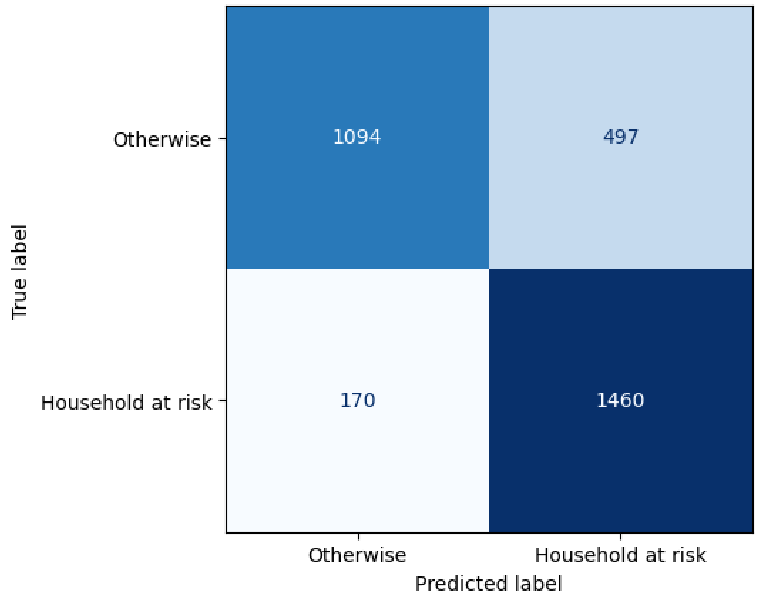

Figure 8: Confusion matrix for Random Forest Classifier (Train set)

As expected, the RFC produced more False Positives than False Negatives. However, negative classes (households that are not at risk) are more frequent than positive classes. The model focuses on positive classes, and the proposed objective function is helping to reduce False Negatives, which is the desired characteristic for the case of disasters.

The critical element for the hyperparameter optimization procedure is the confusion matrix of the predictive models, as logistics costs depend on False Positives and True Positives. True Negatives mean no delivery is required, and deprivation costs arise from False Negatives. Our methodology includes co-optimization of MCC and NPV, where maximization of MCC aims to minimize social costs and maximization of NPV aims to minimize deprivation costs.

2.5.3. Complementary descriptive analysis

We conducted a descriptive analysis of False positives and False Negatives to complement the results above. Table 4 shows the average of each variable across the subpopulations.

Table 4: Descriptive analytics of misclassified categories

Variable	**False positives (N=164)**	**False negatives N=(85)**
Terrain tracks (Yes=1)	30.49% (N=50)	41.18% (N=35)
Paved tracks (Yes=1)	7.32% (N=12)	44.71% (N=38)
Lighting pole (Yes=1)	12.2% (N=20)	85.88% (N=73)
Own house (Yes=1)	88.41% (N=145)	67.06% (N=57)

Variable	False positives (N=164)	False negatives N=(85)
Title of ownership (Yes=1)	10.37% (N=17)	40.0% (N=34)
Concrete walls (Yes=1)	8.54% (N=14)	48.24% (N=41)
Altitude	4001.51	3781.26
Rural (Yes=1)	88.41% (N=145)	17.65% (N=15)
Water network (Yes=1)	7.93% (N=13)	62.35% (N=53)
Drainage network (Yes=1)	7.93% (N=13)	62.35% (N=53)
Electric lighting (Yes=1)	71.34% (N=117)	98.82% (N=84)
Candle lighting (Yes=1)	13.41% (N=22)	1.18% (N=1)
Another lighting (Yes=1)	20.12% (N=33)	0.0% (N=0)
GLP cooking (Yes=1)	15.85% (N=26)	62.35% (N=53)
Manure cooking (Yes=1)	59.15% (N=97)	16.47% (N=14)
Internet (Yes=1)	8.54% (N=14)	27.06% (N=23)
Cellphone (Yes=1)	71.95% (N=118)	95.29% (N=81)
TV color (Yes=1)	18.29% (N=30)	58.82% (N=50)
Bicycle (Yes=1)	25.61% (N=42)	35.29% (N=30)
Motorcycle (Yes=1)	23.17% (N=38)	40.0% (N=34)
DVD (Yes=1)	8.54% (N=14)	23.53% (N=20)
Car (Yes=1)	2.44% (N=4)	10.59% (N=9)
Computer/laptop (Yes=1)	4.27% (N=7)	27.06% (N=23)
Refrigerator (Yes=1)	0.0% (N=0)	8.24% (N=7)
Per capita expenditure	3799.05	5789.32
Young adult (Yes=1)	9.15% (N=15)	18.82% (N=16)
Adult (Yes=1)	22.56% (N=37)	37.65% (N=32)
Old adult (Yes=1)	26.83% (N=44)	31.76% (N=27)
Old (Yes=1)	41.46% (N=68)	11.76% (N=10)
overcrowding (Yes=1)	50.61% (N=83)	35.29% (N=30)
Married (Yes=1)	33.54% (N=55)	31.76% (N=27)
Literacy (Yes=1)	24.39% (N=40)	11.76% (N=10)
No education (Yes=1)	79.27% (N=130)	48.24% (N=41)
Postgraduate education (Yes=1)	0.0% (N=0)	1.18% (N=1)
Illness (Yes=1)	96.34% (N=158)	91.76% (N=78)
Medical attention (Yes=1)	45.73% (N=75)	67.06% (N=57)
Subsidized health insurance (Yes=1)	86.59% (N=142)	69.41% (N=59)
Disabilities (Yes=1)	45.12% (N=74)	25.88% (N=22)

Authors' own elaboration.

The False Positives are households characterized as poor in a multi-dimensional sense. Otherwise, the False Negatives are households with non-poor characteristics. From Table 4, we highlight the following features for False Positives: 7.32% of households have access to paved tracks and 12.20% to lighting poles, 8.54% have concrete walls, 7.93% have water and drainage network, 59.15% cook with manure, 8.54% have internet access, and 0% have a refrigerator. These features suggest that False Positives are poor households. We must consider that 88.41% of them are rural, so for this case, they may have vulnerable conditions but might not be exposed to cold wave-related disasters.

We highlight the following features for False Negatives: 44.71% of households have access to paved tracks and 85.88% to lighting poles, 48.24% have concrete walls, 62.35% have water and drainage network rather than using manure, 62.35% of households cook with GLP, 27.06% have internet access, and 8.24% have a refrigerator. According to this characterization, False Negatives are mostly non-poor households associated with better urban infrastructure. 17.65% of these households are rural. False Negatives might be exposed to cold wave-related disasters but may not have vulnerability conditions.

Regarding educational and health dimensions of vulnerability, False Positives have 31.03% more uneducated household heads than False Negatives and have 21.33% less access to medical attention and 17.18% more households with subsidized health insurance. Finally, on average, False Positives are settled at a higher altitude than False Negatives (220.25 m.a.s.l.) and have lower annual monetary earnings (US$590.58).

2.6. Discussion and implications

This section presents a discussion of the main results and the practical implications of these results for relevant stakeholders and decision-makers.

2.6.1. Determinants of cold wave-related disaster risk

RFC estimated features' importance to understand which features drive cold wave-related disaster risk at the household level. The results for the 15 most important features are shown in Figure 9.

The most important features for prediction were household localization in a rural area (that accounts for the fact that the household is isolated in the space and systematically far away from principal urban settlements) and per capita expenditure (that accounts for short-run household purchase power). Moreover, the feature "Household resides in a rural area" carries twice the significance of per capita expenditure. Access to public goods (measuring the government's presence in public spaces where households are located) was also crucial for cold wave-related disaster risk classification. Other important predictors were altitude (proxies for household exposure to shallow temperature events) and household materials of construction (concrete walls and concrete roofs).

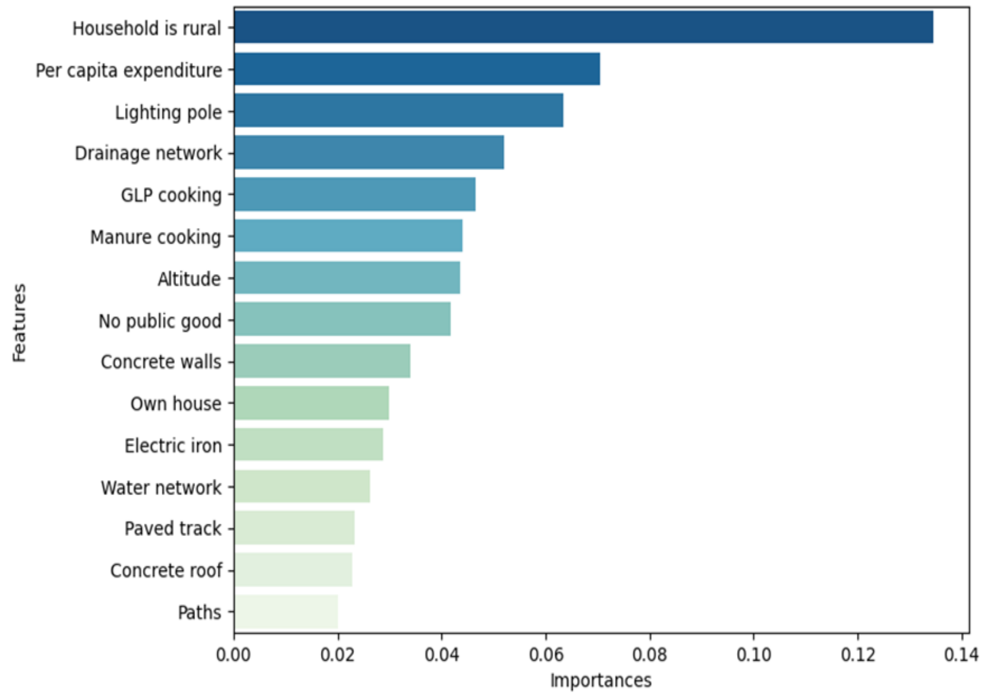

Figure 9: Feature's importance from RFC

Figure 10 estimates the average marginal effect or partial dependence plot for each feature in Figure 9. The axis is continuous for continuous variables and discrete for dummies. According to these results, rural households are 24% more likely to be at risk of being affected by cold wave-related disasters than urban ones. In contrast, having a lighting pole, a drainage network, and cooking by GLP reduces the probability of being at risk by 14%, 11%, and 15%, respectively. The higher the ranking in Figure 9, the greater the robustness of this average estimate. Interestingly, an increase in per capita expenditure lowers the probability of being at risk based on the magnitude of expenditure at different rates. For high-expenditure households, an increase in expenditure is not related to a significant decrease in the probability of being at risk. For poor households, the impact of variations in expenditure is higher. Public goods and concrete on walls and roofs lower the probability of being at risk.

Altitude partial dependence seems constant, but as indicators feature importance indicators suggest, it adds information. This means that for all the households, the probability of case risk does not vary with different altitude levels, but this may not be true for groups of households. As important features may not be significant in their average variation, exploration of heterogeneous effects is needed. However, it must be left for further research.

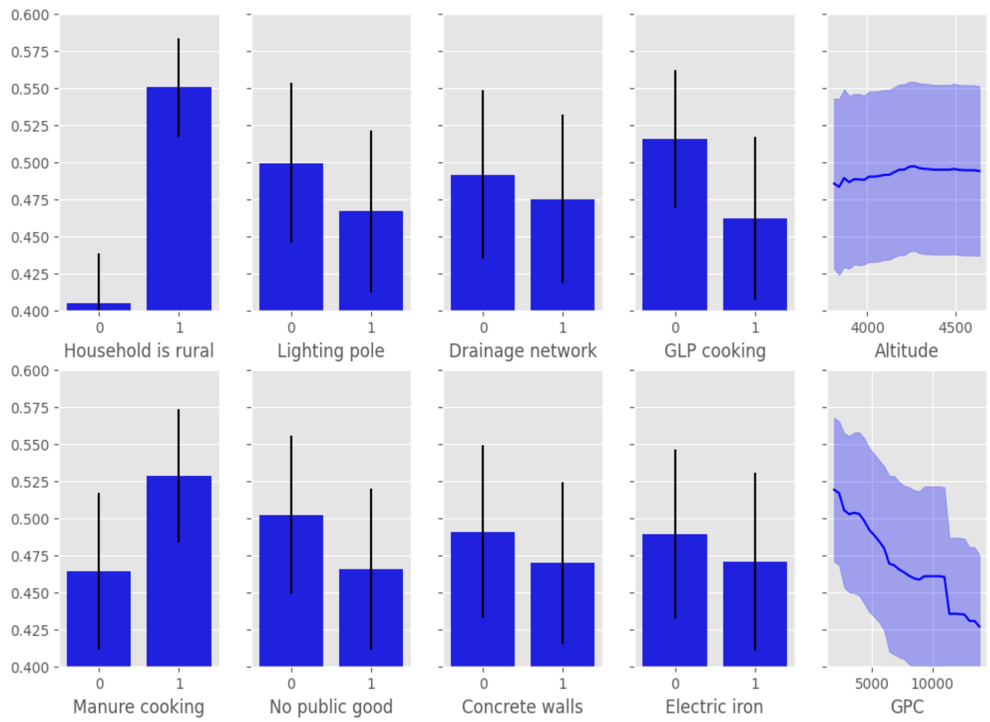

Figure 10: Top 10 Feature partial dependence on the probability of being at risk

In Figure 10, the significance of each predictor is assessed based on the magnitude of its impact on the probability of being at risk and the variance of the partial dependence estimations. The feature "Household is rural" results in discernible differences in the probability of being at risk for the entire sample, indicating its significance. On the other hand, features like "Manure cooking" appear to be influential, but it is essential to quantify the confidence in the disparities between their values. For continuous features, this calculation may be more intricate, and we defer this discussion to future research, with reference to SHAP values for the task.

The RFC estimator is robust to non-linearity, heteroscedasticity, and noise on predictors. As the construction of trees is based on bootstrap methods, the partial dependence estimates are a non-parametric estimator of the impact of exogenous variations on predictors into the target variable, disaster risk.

Considering these results, we conclude that cold wave-related disaster vulnerability is shaped by economic deprivations, geographical localization in rural areas, and the degree of access to public goods in urban environments, including access to essential services. In this sense, to reduce vulnerability, we must act in line with disaster risk reduction main guidelines (Wright et al., 2020): it is necessary to make long-term investments that aim at systematically reducing vulnerabilities to create resilience in communities by achieving economic and urban development of cities.

Resilience is a goal that would be achieved slowly and requires much planning. Puno is a city that was built with scarce resources. Hence, there is an enormous potential for improvement, particularly regarding mitigating disaster risks. It is worth highlighting that in the short term, applied Machine Learning can be used to optimize resource utilization and, in the best of cases, save necessary resources that communities may invest in their future development (Bosher et al., 2022).

2.6.2. A proposal for improvement of the model

The actual model has an accuracy of 73.85% on the test dataset. That means that if the model had been implemented in 2021 and all the demand points had been fulfilled with aid within the context of an intervention, 19.1% of households that would have demanded aid would have been excluded from the targeting. On the other hand, 32.41% of households that were unaffected would have been provided with aid, creating additional costs.

The main pattern regarding False Positives and False Negatives was that poor households without risks were misclassified (False Positives), and non-poor households with risks were labeled non-risky (False Negatives). In this sense, additional costs related to False Positives might not be unjustified, as most households are poor. Considering False Negatives, the average household may be non-poor but still need aid to face cold waves. Considering statistical analysis, we recommend moving the classification threshold of the RFC to balance False Positives and False Negatives and achieve greater accuracy and sensitivity. The following Figure 11 shows the confusion matrix corresponding to a probability threshold of 42% (corresponding to the threshold that maximizes Z_m objective balance between accuracy and sensitivity):

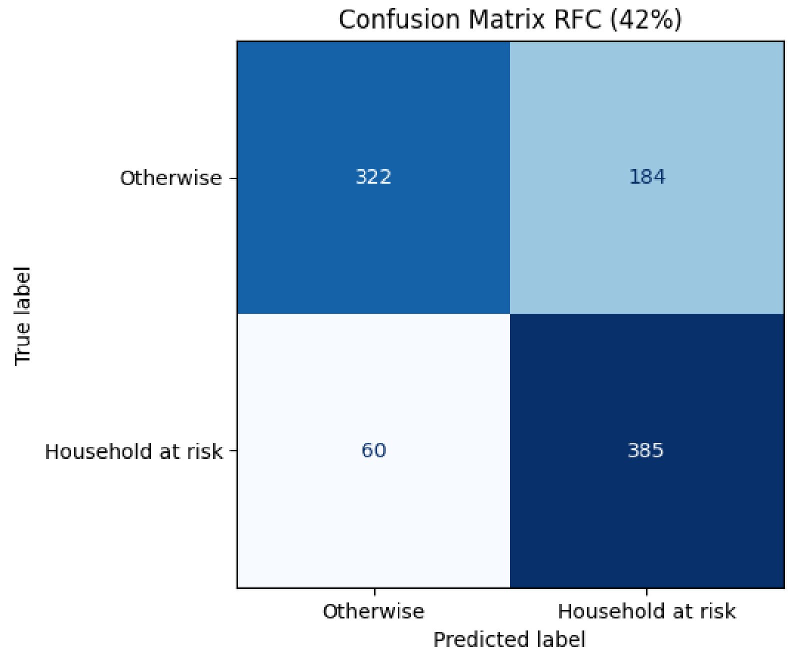

Figure 11: Confusion matrix with new prediction threshold (Test set)

The performance metrics of this confusion matrix are the following: 75.08% ROC-AUC, 74.34% accuracy, 75.94% F1-score, 51.05% MCC, and 86.52% sensitivity. This improvement can have a significant impact on the practice. If a humanitarian intervention had been implemented considering the confusion matrix in Figure 11, 13.48% of households would have generated deprivation costs (a percentual reduction of 29.42%). Although there are more False Positives, most of these households are poor, so aid would attend to other necessities embedded in their multiple vulnerabilities. Any humanitarian project that aims to mitigate the negative impacts of cold wave-related disasters may find this paper helpful since its methodology can be replicated for other case studies.

2.6.3. Extra considerations about practical implementations

Table 4 shows that both False Positives and True Positives are characterized as being poor, rural, and isolated in space. A humanitarian intervention would find reaching households with these characteristics more costly in a real-world scenario. In contrast, False Negatives and True Negatives are households settled in urban areas with transport infrastructure that reduces logistic costs. Although these households are easy to reach, it could be challenging to identify which would be the target of humanitarian intervention. The model in Figure 11 can improve this targeting.

Any practical implementation must consider the guidelines above. The main challenge is to attend to the demand from predicted positives. If the model in Figure 7 is implemented, 32.34% of this demand is expected to be misallocated. However, considering that these households are poor, an excellent strategy is to integrate both the humanitarian intervention that aims to mitigate the impacts of cold wave-related disasters and poverty and short-run hunger interventions. Integrating interventions would lead to a more efficient use of resources, assuming that poor households are vulnerable to food shortages and economic losses during months of cold temperatures in Puno.

2.7. Chapter conclusions

This paper focused on using Machine Learning to build proactive strategies for cold wave-related disaster preparedness in Puno. The aim was on households' disaster risk classification or identification of demand: a predictive classifier was built to identify households that are targets for humanitarian interventions.

Puno has small cities; most of its population is settled in rural areas dispersed in space. The classifier identified the following prediction rules:

Poor households in rural areas are vulnerable to cold wave-related disasters and need proactive humanitarian intervention.

Beyond economic vulnerability, vulnerable households have poor urban infrastructure, including tracks, paths, lighting poles, and water and drainage networks. These features characterize households that are demand points of humanitarian interventions.

The impact of health insurance, health status, and education is minor. Households with unhealthy members have a 0.8% higher probability of being at risk than households with healthy members on average. At the same time, households with graduate members have a 0.6% lower probability of being at risk than other households.

The experimental setting allowed us to select RFC over ENLR as the best classifier, with an MCC of 48.64% and a sensitivity of 80.9% on the test dataset. This result represents a good baseline level for practical implementations because the model's accuracy is relatively high (73.82%), considering that predictions were made with a model trained with past data from 2018-2020. Thus, the model can perform a demand forecast with acceptable accuracy.

After performing a statistical analysis of False Negatives and False Positives, we considered it profitable to modify the probability threshold of the RFC to improve the model's performance. With a threshold of 42% instead of 50%, model accuracy improved to 74.34%, MCC to 51.05%, and sensitivity to 86.52%. This result has several practical implications. First, if this model is implemented, False Negatives would be reduced at the cost of more False Positives. That means that humanitarian operations targeting would improve at the cost of reaching more households that might not need supplies to face cold. The drawback is that such households, known as False Positives, are poor and isolated in space, so most kinds of interventions may find it costly to reach them.

Even though the improved model misclassifies a higher frequency of False Positives, statistical analysis shows that these households have deprivations. Hence, those costs may be justified, especially if the humanitarian intervention is embedded in another, more comprehensive program. This scenario could be the case of a policy to mitigate food and hunger. Using the improved model would enormously impact the Machine Learning-targeted households.

Consequently, Supervised learning offers a data-centered solution to the large-scale problem of deciding where aid must be delivered. This solution is characterized by being detailed and disaggregated at the household level: model predictions can be used to decide which households will require a supply of aid. Decision-makers can implement proactive disaster preparedness strategies such as stock prepositioning, proactive delivery, and gradual delivery based on information drawn from the prediction of trained models (Apte and Yoho, 2011).

This paper confirms previous literature findings regarding cold wave-related disaster risk mitigation. It brings new conclusions: physically vulnerable and economically deprived households are more likely to be affected by a cold-related disaster. The well-known prescription is to create community resilience with solid urban infrastructure, which is difficult to achieve in the short term.

In addition, we suggest using Machine Learning to implement an automated classifier that identifies the demand in the context of uncertainty and intervenes in those demand points to mitigate short-term cold wave risks. This would improve disaster risk reduction decisions.

Better management of disaster risk is related to mitigated response and recovery. The model's implementation gives Puno opportunities to use the saved resources to carry on long-run tasks such as creating resilience.

This paper is not free of limitations. The following limitations were identified:

- Local effects were not estimated; hence, health and education might significantly impact the probability of being affected by cold waves for some households with specific characteristics. A complete analysis was not performed, just an average estimation of marginal effects.
- Although the experimental setting is robust, real-world model implementation is vital to close the gap between academia and practitioners. This paper aimed to provide guidelines and, to the best of our ability, shed light on the uncertainty embedded in practical implementations.
- The model can be further extended to consider more sophisticated predictors such as distance from households to main tracks, livestock, and area of land under cultivation, among others, that may improve the accuracy of the classifier. Measuring experience in cold waves as a proxy for risk resistance or disaster preparedness would improve accuracy (Chen et al., 2022).

Since humanitarian interventions operate with scarce resources and must be optimized regardless of their localization or vulnerability condition, this paper sheds light on practical considerations of applied Machine Learning. One way to measure the contribution is to analyze the accuracy of model forecasts on real data. By doing this, practitioners may observe better Key performance indicators (KPIs), such as emergency response time, normal response time, total coverage, and demand fluctuations (Rejane et al., 2013). Consequently, this paper contributes to closing the gap between academia and practitioners toward an improved disaster risk management system based on data. The Puno community would benefit from the practical implementation of Machine Learning in disaster risk reduction.

3
Identifying final demand points for aid in the aftermath of sudden-onset clime-related recurrent disasters in Peru using supervised learning

3.1. Introduction

There is an increasing concern about natural hazards worldwide, as they are becoming more frequent due to climate change and increasing pressure on natural resources. Less developed countries are more affected by recurrent natural hazards regarding human and economic losses (Guha-Sapir et al., 2015). Losses are linked to disaster risk, which depends on three factors: vulnerability, hazard, and exposure (Twigg, 2004). Consequently, less developed countries have been more affected by natural hazards due to poverty, inadequate infrastructure, and limited resources for disaster preparedness and response (Ghesquiere and Mahul, 2010). Furthermore, recurrent natural hazards tend to produce long-term effects that harm the development of such countries. For instance, Akram et al. (2021) conclude that "disasters impede human development, and their effects are most pronounced in low and lower-middle-income countries."

The disaster risk management lifecycle has five phases: prevention, preparedness, mitigation, response, and recovery. The first three phases are proactive and aim to minimize the impact of an expected disaster. These activities are crucial in the lifecycle of disaster management as they impact the cost and complexity of response and recovery activities. In the case of recurrent disasters, where exposure to natural hazards interacts with population vulnerability, planning for disaster risk management activities becomes particularly important (Bosher et al., 2021). Proactive disaster risk management activities can be seen as an investment that produces returns in the aftermath of disasters.

In the presence of recurrent disasters, there is a reinforcement cycle between expected losses and investment in disaster risk mitigation and preparedness (Sodhi, 2016). This vicious cycle is challenging because investing in risk management activities is necessary to unlock a country's capability to reduce its disaster risk (UNDRR, 2020; ADB, 2015).

Less-developed countries often operate in a context of scarce resources, so decision-making planning is imperative for these cases. Then, decision-makers require high-quality data representing all the problem's edges. In disaster risk management activities, this bias causes an imbalance in demand and supply (for example, wrong stock prepositioning and, thus, material convergence).

The demand for relief supplies must be dynamically estimated from status-quo representative data to improve the disaster risk management activities in pre- and post-disaster areas. However, the availability of data representing the actual status quo is a big concern for decision-making in the aftermath of recurrent natural hazards (Linardos et al., 2022). This paper proposes using supervised learning algorithms to train predictive models to produce dynamic forecasts that can be used to make informed decisions.

The core of the problem, at least in the disaster risk management domain, is that the future distribution of demand is unknown. This paper proposes a novel approach to identify final demand points after sudden-onset clime-related disasters. Thus, the specific objective of this paper is to train supervised learning binary classifiers to identify demand points using households' observable characteristics. The disaster risk management domain demands the predictions to be aware of unmet demand (Silva and Leiras, 2021). For that, the classifiers minimize the false negatives (i.e., households classified as non-prone to disasters when they are prone to them).

This paper assesses the Peruvian case study, a less-developed, low-income country (Annual Disaster Statistical Review, 2014). Considering the 2000-2023 period, floods and landslides were the most frequent causes of disasters that affected communities in Peru, generating up to 4,171,481$ and 3,042,638$ total losses, respectively (EM-DAT, 2023).

The contributions are twofold. First, a supervised learning approach is proposed to infer accurate classifiers for floods and landslides based on households' vulnerability conditions. Such classifiers are flexible to adapt to domain requirements, such as the importance of unmet demand. Given some required characteristics, the models can be re-trained to update their performance or adapt to different realities within the same domain. Second, this paper deepens prediction explanations, which are also scarce in the literature.

In addition to academic contributions, the results contribute to a wide range of public and private stakeholders, helping in public policy design and disaster risk mitigation and preparedness. This paper closes by enumerating the research implications for disaster management that provide decision-makers with recommendations. This paper includes model implementation guidelines that may help them to operate with scarce resources, such as resource allocation, timely response, and improved targeting (Farazmehr and Wu, 2023; IFRC, 2013; OCHA, 2015; WFP, 2018). A spatial decision-support display is built to illustrate the recommended implementation of the model results and define further research agenda.

The rest of the paper is divided into five sections. Section 2 presents the theoretical foundation. Section 3 details the case and describes the materials and methods. Section 4 depicts the main results. Section 5 discusses how the results may be interpreted and used to formulate decisions regarding disaster risk management. Finally, Section 7 brings the conclusions and suggests further research avenues.

3.2. Theoretical framework

Several approaches oriented to prediction and planning for disaster risk management activities aim to exploit the recurrent nature of some disasters to develop early warning systems, automated logistics, demand forecasting, mapping of scenarios, and so on (Yuan and Moyaedi 2020; Zhang et al., 2019; Resch et al., 2018). The predictive assessment is more relevant in a population affected by recurrent disasters.

Although some disasters have a predictable pattern, Machine Learning and artificial intelligence methods are scarce in managing recurrent disasters (Behl and Dutta, 2019). Lin et al. (2020) predict aid demand from crowdsourcing platforms concerned about anticipated earthquake response. French et al. (2023) study the root causes of El Niño related recurrent floods and landslides and conclude that geophysical characteristics interact with exposure and vulnerability of Peru’s population and infrastructure to produce high levels of disaster risk that are recurrent because of institutional factors affecting the management of risks and disasters.

This paper proposes using supervised learning to train classifiers for households affected by floods and landslides. The proposed predictive assessment contributes to disaster risk management in two ways: i. providing intelligent methods to target disaster preparedness activities, and ii. mapping the impact of key features in prediction to derive vulnerability drivers and formulate disaster risk mitigation policies.

3.2.1. Characterization of households' disaster vulnerability

This paper uses Peruvian population-representative stratified random sampling data to train supervised learning classifiers, which is uncommon in the literature. Some powerful approaches extract data from crowdsourcing platforms or social media such as Twitter or Facebook (Zhang et al., 2019; Eckhardt et al., 2021). However, it is crucial to consider alternative approaches that are better suited to regions with a low ratio of Internet access (which is the case for Peru, which has unequal Internet access in their regions). In such realities, overcoming the participation bias is a challenge.

The stratified random sampling method ensures that every household in the sampling area, or the final demand point, has a fair chance of being included in the analysis. Therefore, the sample distribution is representative of the population's actual status quo. Under this approach, big data is not more important than good data (Jayawardene et al., 2021). That means quality must be preferred over quantity if data is used for decision-making. In contrast with information-scrapping approaches (Resch et al., 2018; Lin et al., 2020), classical statistical sampling is grounded in theory, and its use may contribute to further Machine Learning applications in disaster risk management (Jayawardene et al., 2021).

The central theoretical hypothesis is that vulnerability shapes disaster risk (Twigg, 2004). Four general dimensions of vulnerability are considered: economic, health, social, and geographical. The economic dimension measures purchasing power and comprises household income, construction materials, equipment, and access to services (Tasnuva et al., 2020; Pessoa, 2012). The health dimension measures the members' health status, such as chronic and acute illness (Djalante et al., 2020). The social dimension represents the sociodemographic characterization by measuring unemployment, education achievement, sex, age, and marital status (Rapeli, 2017). The geographical dimension captures household location conditions, such as regional dummies and altitude, that make them susceptible to natural hazards (Mattea, 2019; Ullah et al., 2022).

3.3. Materials and methods

In this paper, we focus on the case of Perú, one of the countries most affected by clime-related disasters in Latin America. Between 2000 and 2022, Perú was affected by 47 floods, 17 earthquakes, and 12 landslides (EM-DAT, 2023). Those were the most frequent disasters. While floods caused economic losses of 3,237,000 USD and 823 deaths, landslides caused unrecorded damages and 317 deaths.

Peru has a population of 33.72 million people spread over an area of 1,285,216 km², resulting in a population density of approximately 26.24 inhabitants per km². However, the population density varies significantly within the country. For example, Lima, the capital city, has a population of 11.82 million people living in an area of 2,672 km², resulting in a high population density of approximately 4,425.97 inhabitants per km². In contrast, some regions of Peru are sparsely populated, such as Puno, with a population density of 16.28 inhabitants per km², and Pasco, with 10.03 inhabitants per km².

Natural hazards typically affect rural regions that have a greater poverty rate. For instance, Ayacucho has a poverty rate of 64.8%. It is affected by landslides that happen without media coverage. Then Lambayeque, with a poverty rate of 31.6%, is affected recurrently by floods that destroy houses and livelihoods yearly. Then Lima, with a poverty rate of 18.3%, is affected by earthquakes that, in contrast with floods and landslides, have happened once a decade.

It is a challenge to manage risks in such a heterogeneous country. This paper aims to identify the final demand points for each type of disaster and then prescribe some decisions to operationalize the knowledge produced using supervised learning methods. This inherently data-centric solution adds significant value to society by reducing human suffering and offering tools to improve the quality of decisions taken by stakeholders.

The following subsections detail the data gathering and data analysis methods.

3.3.1. Data gathering methods

This paper uses modules from the National Household Survey (NHS) carried out by the Peruvian National Institute of Statistics and Informatics in 2018-2021. The sampling method was stratified over political regions. Thus, the data provides an excellent quantitative representation of socioeconomic characteristics over urban and rural conglomerates.

The target ground-truth classification labels were built using spatial processing tools. This labeling was done as follows. NHS provides geographical coordinates of households, and the following question is asked to the informers:

In the last 12 months, has your house been affected by natural disasters (drought, storm, plague, flood, etc.)?

Thus, a household is affected by a natural hazard if the household head reports the above question. However, the NHS does not provide details about the type of disaster that affected a household. Hazard zones were built for Floods and Landslides to overcome this limitation using data from Peruvian public entities and Dottori et al. (2016).

Figure 12: Hazard zones

Figure 12a. Hazard zones for Floods and Landslides

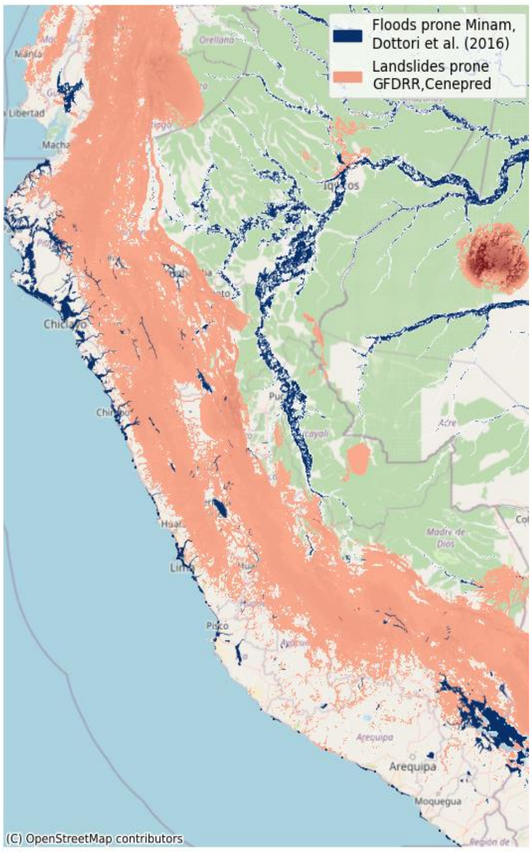

Figure 12b. Ground-truth Affected Households 2021

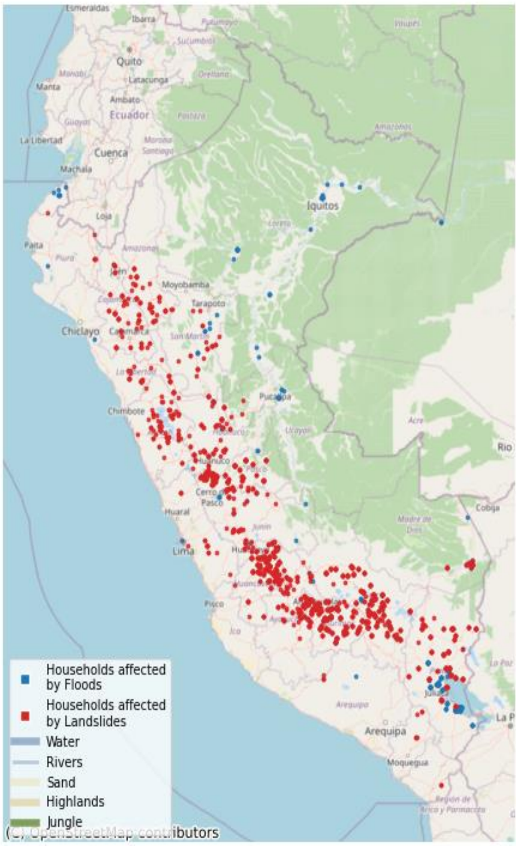

Figure 12a shows Hazard zones that are geographical boundaries recurrently affected by a natural hazard. Data were extracted from the Ministerio del Ambiente (Minam) estimation of flood-prone urban areas, including tsunamis, to build the Flood Hazard layer. Dottori et al. (2016) provided a global flood hazard mapping that was joined with Minam layers to complement rural river-based floods. On the other hand, the Landslide Hazard layer is built upon data from the Global Landslide Hazard Map made by the Global Facility for Disaster Reduction and Recovery (GFDRR) using the mean annual rainfall-triggered landslide hazard assessment for the period 1980-2018. The information from the GFDRR Hazard map was verified by historical data collected by Centro Nacional de Estimación, Prevención y Reducción del Riesgo de Desastres (Cenepred).

Then, a household is affected by a flood/landslide if [1] it was affected by a natural hazard and [2] it is located within the boundaries of **Flood Hazard**/**Landslide Hazard** zones. This is formally defined in Equations 1 and 2

$$\text{Affected by Floods}_i=\begin{cases}1 \text{ if an affected household is located in a Flood Hazard zone}\\0 \text{ otherwise}\end{cases} \quad (1)$$

$$\text{Affected by Landslides}_i=\begin{cases}1 \text{ if an affected household is located in a Landslide Hazard zone}\\0 \text{ otherwise}\end{cases} \quad (2)$$

3.3.2. Data analysis method

3.3.2.1. The classification problem

The ground-truth label construction results in Figure 12b suggest that few households are inside Hazard zones and report being affected by a natural hazard. For floods, 1.20% of households in the sample were affected, and 5.45% for landslides. Hence, the classification problem is imbalanced.

The classification problem is modeled as follows:

$$\Pr(\text{Affected by Floods}_i=1)=F(\text{Vulnerability}_i) \quad (3)$$

$$\Pr(\text{Affected by Landslides}_i=1)=L(\text{Vulnerability}_i) \quad (4)$$

That is, the probability of a household being affected by floods or landslides is calculated as a function of 112 features describing multiple vulnerability dimensions. Equation 5 groups the features into four groups.

$$\text{Vulnerability}_i=[\text{Economic}_i,\text{Social}_i,\text{Health}_i,\ \text{Geographical}_i] \quad (5)$$

This procedure hypothesizes that a household affected by natural hazards is necessarily vulnerable or, at least, there is a strong dependence. Thus, the predictive performance of the algorithms depends on the degree of dependence between vulnerability and the outcome of the household located within a Hazard zone. The literature (Li et al., 2023; Lapietra et al., 2023) supports this hypothesis for the case of multidimensional vulnerability (Equation 5) (UNDRR, 2015). Furthermore, we seek empirical evidence for the same hypothesis in this paper.

3.3.2.2. Machine Learning Pipeline

Both functions F(.) and L(.) from Equations 3 and 4 must be learned from data using supervised learning algorithms. These functions predict the household probability of being affected by Floods or Landslides, respectively. Then, if predicted probabilities for each household $Pr(\text{Affected by Floods}_i=1)$ and $Pr(\text{Affected by Landslides}_i=1)$ are greater than a threshold, defined as 50% by default, the model produces predictions.

However, supervised learning through cross-validation may produce insufficient learning of minority classes or affected households, and learning may be biased towards the majority class or non-affected households (Luque et al., 2019). If this limitation were not addressed, the classifier would not be aware of deprivation costs caused by the misclassification of affected households (Holguin-Veras et al., 2013). That is, affected households are labeled as non-affected or False Negatives. Imbalanced learning techniques are applied in different steps of the machine-learning pipeline to overcome this limitation (Brownlee, 2020).

The Machine Learning pipeline includes pre-processing, processing, and post-processing stages (Waring et al., 2020). The selected train-test split method is the stratified K-fold cross-validation, which keeps the same proportion of class labels in each fold during the training process to overcome bias from class imbalance. A second train-test split method is applied for model out-of-sample validation purposes. Model performance is evaluated in a one-period-ahead hold-out test set. Model parameters and hyperparameters are estimated from 2018-2020 data using stratified K-fold cross-validation and tested for 2021 data using a hold-out test set. The entire procedure is outlined in Figure 13,

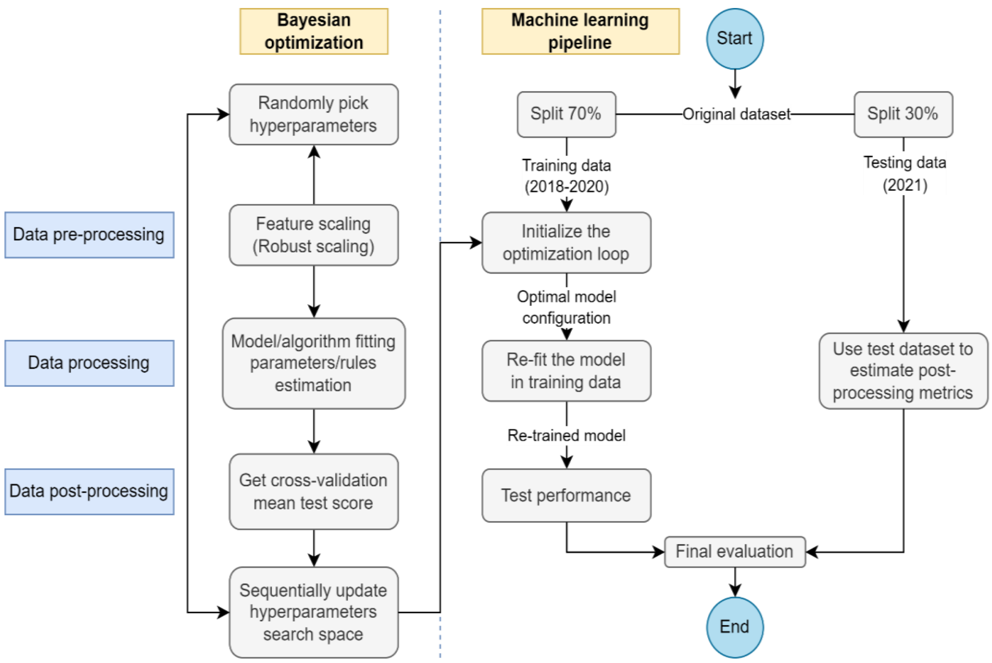

Figure 13: Machine Learning pipeline

Data pre-processing steps include feature scaling or normalization using the robust scaling method (Zheng and Casari, 2018). This is a variation of Standard Scaling that subtracts the median and scales the data to the interquartile range:

$$Z = \frac{X - X_{p50}}{IQR_X} \tag{6}$$

There are several supervised learning algorithms. The common procedure is to test several algorithms and keep the best performing on the experimental grid (Dantas et al., 2021). However, this approach was not considered because of the overwhelming number of experiments that would have to be run to reach the best-performing model (and its best configuration of hyperparameters). Instead of using brute-force approaches (Pedregosa et al., 2012), the Bayesian search for hyperparameter optimization method was applied to the XGBoost supervised learning algorithm.

Data processing considers the XGBoost algorithm for supervised learning for several reasons. This algorithm was selected due to its properties regarding robustness to outliers, noise, and feature selection (Chen et al., 2016). XGBoost stands for eXtreme Gradient Boosting, an ensemble of gradient-boosted trees (More details in Section 4.2). This algorithm is known for excelling in online competitions (Kaggle, 2023). Furthermore, XGBoost trains faster than Random Forests and Neural Networks (Fauvel et al., 2022). XGBoost is also adaptable to a wide spectrum of particularities regarding the domain's nature that conditions the dataset's characteristics. For these reasons, XGBoost was considered the best option to explore.

Equation 5 shows the elements of the objective function of XGBoost:

$$obj(\theta) = L(\theta) + \Omega(\theta) \tag{7a}$$

$$L(\theta) = \sum_i \left[y_i \ln\left(1 + e^{-\hat{y}_i}\right) + (1 - y_i) \ln\left(1 + e^{\hat{y}_i}\right) \right] \tag{7b}$$

$$\Omega(\theta) = L_1(\theta) + L_2(\theta) \tag{7c}$$

XGBoost is an ensemble algorithm composed of individual boosted decision tree classifiers trained with the CART algorithm such as Random Forests (Breiman, 2003). The CART algorithm searches for the best split that produces more loss change, measured by Equation 7a for each tree in the ensemble. Thus, the algorithm estimates weights for each additional tree to minimize the objective function in Equation 7a. Equations 7b shows the log-loss function and 7c, L1, and L2 regularization terms.

Regarding model interpretability, feature selection helps the algorithm learn a parsimonious model in terms of predictors. Model complexity is measured by Ω(θ), which represents regularization terms. Predictions are made using regularized coefficients, so the higher the value of Ω(θ), the more complex the model. By regularizing the boosting process, we address feature selection. Features that do not improve the total loss (as defined in Equation 7a) are considered less important and may be discarded due to the definition of $\Omega(\theta)$.

Additional hyperparameters control different aspects of the algorithmic procedure. The performance of XGBoost is sensible to the choice of these hyperparameters, so these must be fine-tuned to maximize performance. Table 5 defines the role of each parameter in the construction of the XGBoost ensemble:

Table 5: Hyperparameter definitions

Hyperparameter	**Description**
Ensemble	
$Colsample_{tree}$	It is the subsample ratio of columns when constructing each tree.
$Colsample_{level}$	is the subsample ratio of columns for each level.
$Colsample_{node}$	is the subsample ratio of columns for each node
Max_{depth}	Number of splits that define the total depth of each tree in the ensemble

Hyperparameter	Description
$N_{estimators}$	Number of trees in the ensemble
γ	Minimum loss reduction required to make an additional split
η	Make each update more conservative by constraining the loss reduction on each tree (helps in imbalanced learning).
$Subsample$	Bootstrap subsample used in each tree
$Max_{deltastep}$	Make each update more conservative by constraining the loss reduction on each leaf (similar to η.)
$Policy$	Strategy for growing trees in the ensemble
Objective function	
L_1	LASSO-type regularization term
L_2	Ridge-regression regularization term
Scale-pos-weight	The relative cost of prediction error for the minority class

The construction of the ensemble begins with a single decision tree built by subsampling columns according to $Colsample$ parameters (subsampling occurs once for every tree constructed). Trees are grown following rules depicted by parameters γ, that imposes a lower bound to the loss reduction required for further splits, and η and $Max_{deltastep}$ that imposes an upper bound to the loss reduction for each tree and leaf respectively.

As with other tree-based ensemble algorithms, the size of the ensemble is given by the $N_{estimators}$ parameter. The depth of each tree is given by Max_{depth}, considering that, if this parameter is not tuned, trees may grow until prediction is fully accurate, which produces overfitting.

As class labels are imbalanced, it is important to constraint loss reduction to prevent bias from updates learned from majority class labels. However, the "Scale-pos-weight" parameter adds cost-sensitive learning because it re-weights the importance of the minority class in the objective function (Brownlee, 2020) (Equation 7a).

The Policy parameter declares that the XGBoost algorithm must grow trees using an objective function (Equation 7a). However, when the first tree is learned, the algorithm captures what was already learned and then learns one new tree at a time; this is known as the additive learning strategy that makes the XGBoost algorithm scalable to large samples. This is important because of the number of hyperparameters that must be tuned.

The following Equation 8.

$$XGBoost(\mathrm{H}) = \begin{cases} Colsample_{tree} = U[0,1] \\ Colsample_{level} = U[0,1] \\ Colsample_{node} = U[0,1] \\ L_1 = LOGU[E^{-2}, E^{2}] \\ L_2 = LOGU[E^{-2}, E^{2}] \\ \gamma = U[0.1,5] \\ Policy = [\text{'lossguide'}] \\ \eta = U[0.2,0.8] \\ Max_{depth} = INTU[9,25] \\ N_{estimators} = INTU[70,150] \\ Subsample = U[0,1] \\ Max_{deltastep} = INTU[1,20] \\ \text{Scale-pos-weight} = \left[\frac{N_{negatives}}{N_{positives}}\right] \end{cases} \quad (8)$$

Data post-processing includes the Generalized Index of Balanced Accuracy (GIBA) (García et al., 2020) metric for imbalanced learning. Although many other methods exist for learning imbalanced classes, using ad-hoc learning metrics complements cost-sensitive learning. The metric is defined by Equation 9,

$$GIBA_{\alpha}(M) = (1 + \alpha Dom)M \quad (9)$$

In Equation 9, M can be any metric but is defined here as the Geometric Mean Score (GMS) that is the geometric mean of True Positive Rate (TPR) (Sensitivity) and True Negative Rate (TNR) (Specificity): $M =$ GMS = ${}^{TPR}/_{TNR}$. Dominance is represented by $Dom = TPR - TNR$. This weights the metric M to reduce the influence of majority class in $GIBA_{\alpha}$. In this application, we set $\alpha = 0.9$ considering that minority class has high importance. This is supported by domain knowledge (i.e., minority class representing affected households that produce deprivation costs when misclassified) (Shao et al., 2019; Gomes et al., 2021).

3.3.2.3. Bayesian Optimization Gaussian Process

The Bayesian optimization algorithm for hyperparameter search is illustrated in Figure 5 and described in detail in Algorithm 1. Owen (2022) emphasizes the importance of automating hyperparameter tuning using data. Hyperparameter tuning is a methodology commonly used for AutoML (Hutter et al., 2019). This approach allows the Machine Learning developer to achieve the highest performance metric without overfitting. Maximizing the performance metric is crucial for obtaining the best operational results.

The hyperparameters of supervised learning algorithms (parameters that are not inferred from data) must be tuned so that the model can produce better predictions in terms of bias-variance (Owen, 2022). However, the tunning process must consider deprivation costs and sample imbalance while it considers hyperparameter space (that increases with the number of hyperparameters to tune). Hence, finding the set of hyperparameters that maximizes a custom performance metric in a hold-out test set is a combinatorial problem.

In disaster risk management, scarce studies address hyperparameter tunning despite being needed to produce better predictive models (Linardos, 2022). This paper performs hyperparameter tunning and proposes an optimization framework to automate searching for the optimal hyperparameters' configuration. Bayesian Search optimization heuristic was used to develop an automated ML pipeline (Hutter et al., 2019). This development provides valuable methodological contributions to further academic research in disaster risk management.

Bayesian optimization was selected for its simplicity compared to other hyperparameter search methods. A Gaussian Process was used as the surrogate model. The algorithm is explained in detail to provide a clear understanding of the methodology, enabling it to be replicated. The following is a detailed description of the Bayesian optimization algorithm:

Algorithm 2: Bayesian optimization search procedure

Bayesian-optimization search

1. Define the hyperparameter space with the accompanied distributions H
2. Define the objective function, in this case, $GIBA_{0.9}(GMS)$
3. Define the stopping criterion; in this case, the number of iterations is equal to 50
4. Initializes the empty set D. Initialize the sample of several pairs of hyperparameter values and stratified cross-validation scores and store them in D (the sample size is equal to 30)
5. Fit the probabilistic regression model/surrogate model, Gaussian Process (M), using the value pairs in D

5.1. Sample the next set of hyperparameters by utilizing the pairs suggested by the acquisition function, A:
5.2. Perform optimization on the acquisition function, A, with the help of the surrogate model, M, to sample which hyperparameters are to be passed to the acquisition function

6. Get the expected optimal set of hyperparameters based on the acquisition function, A
7. Compute the cross-validation score using the objective function, $GIBA_{0.9}(GMS)$, based on the output from Step 6.
8. Add the hyperparameters and cross-validation score pair from Step 7 and Step 8 to set D.
9. Repeat Steps 6 to 9 until the number of iterations equals 50.
10. Trains on the full training set using the final hyperparameter values.

Authors' own elaboration based on Owen (2022).

3.4. Description of results

3.4.1. Descriptive analytics

This section presents and describes the main results. First, we characterize the study case in terms of features. Table 6 lists the complete set of features considered in the ML pipeline.

Table 6: Empirical characterization of vulnerability

Category	Variable
Household exterior and access to public goods	Households with inlaid walls, households with painted walls, Outside tracks are paved, Outside tracks are terrain, Outside paths, Lighting poles, No public good.
Ownership and physical characteristics	Independent house, the household is a house, the household is totally owned, the household has a title of ownership, Concrete walls, Concrete floor, Concrete roof, Overcrowded bedrooms, No other rooms than bedrooms.
Access and use of essential services	Water network, Potable water, Quality water (chlorine), Daily access to water, Drainage network, Electric lighting, Candle lighting, Other lighting, GLP cooking, Wood cooking, Other cooking, Manure cooking, Phone, Cellphone, Cable TV, Internet
Household income and assets	Per capita expenditure, Radio, Color TV, Black-White TV, Sound equipment, DVD, Computer or laptop, Electric

	iron, Electric blender, Gas stove, Refrigerator, Cloth washing machine, Microwave oven, Sewing machine, Bicycle, Car, Motorcycle, Tricycle
Socio-demographics	The head is employed, The head is a woman, The head is married, The head is literate, The head has no education, The head achieved basic education, The head achieved technic education, The head achieved a college education, The head achieved pos-graduate education, The head is a young adult (17-35), The head is an adult (36-50), The head is an old adult (51-65), The head is old (more than 66)
Health and insurance (for household members)	Illness (last month), Accident (last month), Healthy (last month), Chronic illness, Medical intervention (last month), Contributory health insurance, Subsidized health insurance, Disabilities
Geographical context	The household is located in a rural area, Altitude, Region in which the household is located (25 categories)

There are a total of 112 features. The comprehensive analytics may not be interesting, as households affected by Floods or Landslides represent 1.20% and 5.45% of households in the sample, respectively. Then, descriptive statistics are reported for affected households.

The average household affected by a Flood has an annual per capita expenditure of 1561$ (3.6), is located at 1692 m.a.s.l., and 55.34% of households affected by a Flood are located in rural areas. For the case of the average household affected by Landslides, the annual per capita expenditure is 1257$, located at 3048.503 m.a.s.l., and 74.95% of households are located in rural areas. These numbers show insights regarding poverty and geographical conditioning regarding disaster risk.

Another important insight is related to health. Floods and Landslides seem related to 91.8% and 92.20% of households that report a member having a symptom of an acute illness. Despite being informative, the fact that acute illnesses are probably more a consequence than a cause of disaster risk must be considered.

Although Peru is highly diverse, the disaster risk seems to be concentrated in rural areas. Hazard zones mainly cover territories where impoverished and isolated households are located, far from major cities. Consequently, it can be affirmed that the disaster risk policy in Peru is a policy of territorial planning and rural development.

3.4.2. Model training results

3.4.2.1. Model performance

This section presents results from XGBoost model training and hyperparameter optimization. Figure 14 shows the confusion matrix on the hold-out test set for Floods and Landslides.

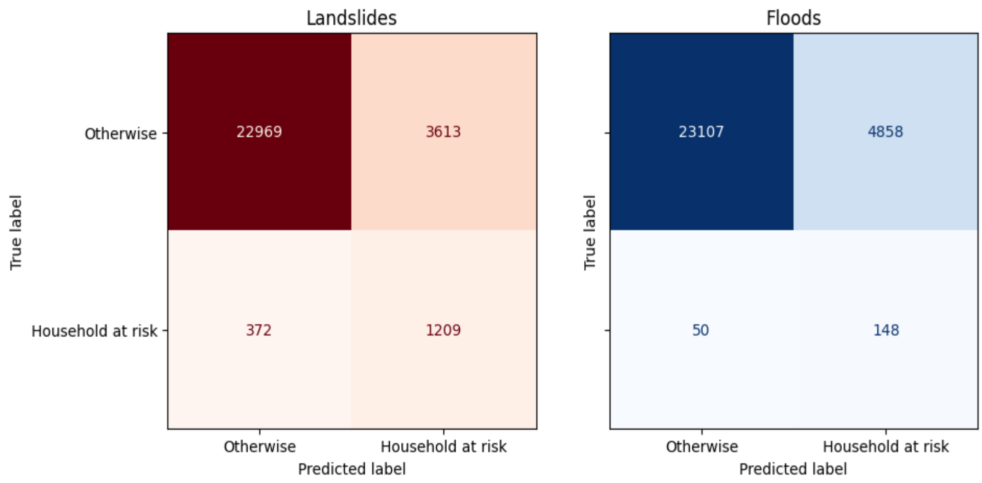

Figure 14: Confusion matrixes for the hold-out test set

The objective function tried to balance the trade-off between the accuracy of the minority class and the total accuracy. This task is NP-Hard, so a hyperparameter search heuristic, such as Bayesian Optimization, was implemented. The confusion matrix shows that the models have good performance for the minority class and medium performance for the majority class. Table 7 shows the performance metrics that measure the success of the training process.

$GIBA_{\alpha}$ controls the optimization process. However, the value of GIBA_α captures the value of the geometric mean score but with lower values when the minority class is misclassified. The metrics tend to have better performance for the Landslides classifier. This may be due to differences in imbalance proportions (the households affected by Floods are a smaller proportion of the total sample and, thus, more imbalanced). The Matthews Correlation Coefficient (MCC) suggests that the Landslides classifier performs significantly better among majority and minority classes.

Table 7: Performance metrics in the hold-out test set

Metric	Floods	Landslides
Accuracy	82.57	88.85
MCC	12.54	38.42
Sensitivity	82.63	86.41
Specificity	74.75	76.47
Geometric mean	78.59	81.29
$GIBA_{\alpha}$	62.0	67.0

With an accuracy higher than 74% for both majority and minority classes, we highlight that the model's performance aligns with the state-of-the-art supervised learning classifiers reported in Linardos (2022). However, the performance is low in comparison with other approaches. Despite this fact, the classifier is evaluated in future data, which conditions accuracy and causes it to decay as it is more complex to predict what will happen in the future.

We affirm that Floods and Landslides classifiers have acceptable performance that can be used for further operations. However, performance can be improved with additional information such as distance from rivers or other water corpses, slope magnitude, distance from mountains, clime, etc. An additional feature collection is left for further research or model implementation.

3.4.2.2. Feature importances and partial dependence

The proposed ML pipeline is expected to output an accurate and parsimonious model. The model is accurate. However, we analyze if it is compounded with several features or can be easily interpreted in terms of a small subset of features.

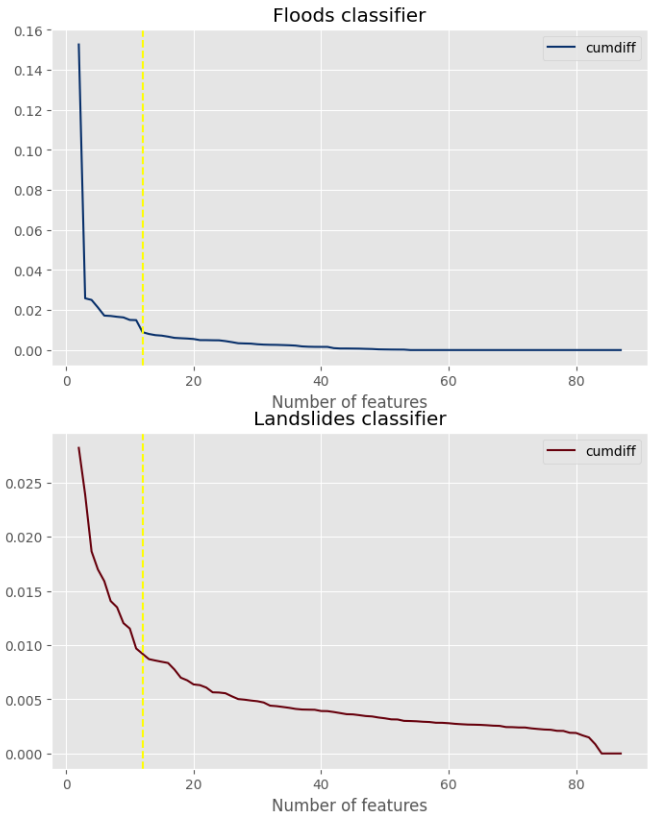

Figure 15: Elbow plot for model feature importance

XGBoost module allows to draw feature importances from information gain. It averages the gain across each ensemble tree, indicating the feature's importance (Pedregosa et al., 2011). Figure 15 shows the marginal average gain for each additional feature (the total gain is the sum of the average gain for all features), with the features being sorted by their importance (regional dummies were omitted from Figure 15 but analyzed separately below).

A threshold of twelve features was selected for analysis. Each additional feature adds a negligible gain beyond this threshold for the Floods classifier. This result suggests that the remaining features are essential as a whole. For the Landslides classifier, a similar pattern is observed where each of the twelve most important features (or more) are important, and the remaining features are essential when grouped.

The regional dummies caused 38.65% of the total gain for the Floods classifier and 48.10% for the Landslides classifier. Thus, households' localization matters in prediction. Figure 16 shows the importance of each classifier's location in a region. In the case of Floods, Puno, Huancavelica, and Ica regions are more prone to flooded households than the rest. On the other hand, Apurimac, Cusco, and Ucayali are the regions with more households affected by Landslides. The region dummies capture the spatial distribution of the affected households.

The case of Puno is remarkable as most households prone to floods lie within its geographic boundaries. Special attention must be given in decision-making to the regions targeted by the classifiers.

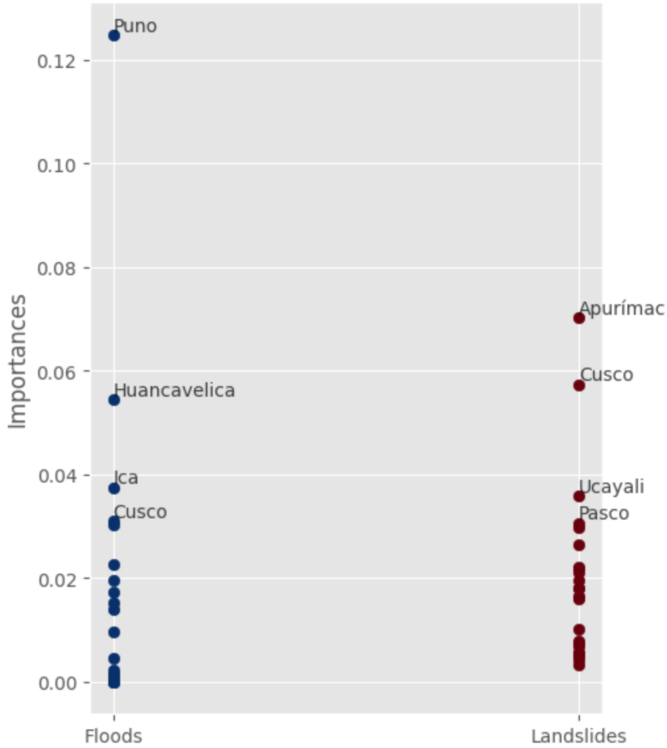

Figure 16: Feature importance indicators for regional dummies

Instead of a parsimonious model, a complex model was inferred from the feature space. However, Figure 15 suggests that the twelve most important features, other than regional dummies, may tell a story about the classification problem. The partial dependence method was selected to analyze the impact of the most important features on predictions. This method creates an interpretable tool that maps the change of predicted probabilities caused by a change in feature values (ceteris-paribus). Figures 17 and 18 show the partial dependence plots for Floods and Landslides classifiers.

3.4.2.3. Interpretation of Floods Classifier

We use several variables to measure the degree of rurality a household is exposed to, including the presence of public goods, cooking with manure, water quality, access to potable water, lighting poles, drainage networks, daily water access, and whether the household is located in a rural area. The "Rural" variable indicates that the household is part of a community with fewer than 2,000 inhabitants. Cooking with manure, for example, suggests a higher degree of spatial isolation and exclusion from the gas or wood cooking fuel market. Households that cook with manure are 13.35% more likely to be affected by floods.

Certain urban conditions can make households more susceptible to Floods. Access to potable water, lighting poles, and drainage networks increases the likelihood of a household being affected by floods. However, daily access to water (Water daily access) of optimal chlorine composition (Water quality chlorine) reduces this probability.

Owning a computer or laptop also decreases the likelihood of being affected by floods. Owning a bicycle may be more common in rural areas or less-developed urban settlements and is associated with an increased probability of being affected by floods. Similarly, households headed by married individuals are also more likely to be affected by floods.

It is important to note that the presence of certain urban public goods (such as paved roads and lighting poles) can increase susceptibility to Floods, as suggested by the "no public good variable". This suggests that floods are more common in less-developed urban settlements than in isolated rural households. Floods are rare in developed urban areas.

Finally, Floods are associated with acute illness, providing valuable information. However, acute illness may not be observable before floods occur since acute illnesses are a consequence rather than a cause. This variable may be excluded from training if the model is intended for use in real-world operations.

The estimated variance of predictions is included in estimations. The variance is lower for “no public good” and “manure cooking” with respect to other features. However, the magnitude of variance indicates that the effect of features is heterogeneous. Individual effects may not be significant, but when several features change together, the probabilities change to produce positive risk classifications. We must leave the track of heterogeneous effects to further research, as it is beyond the scope of research.

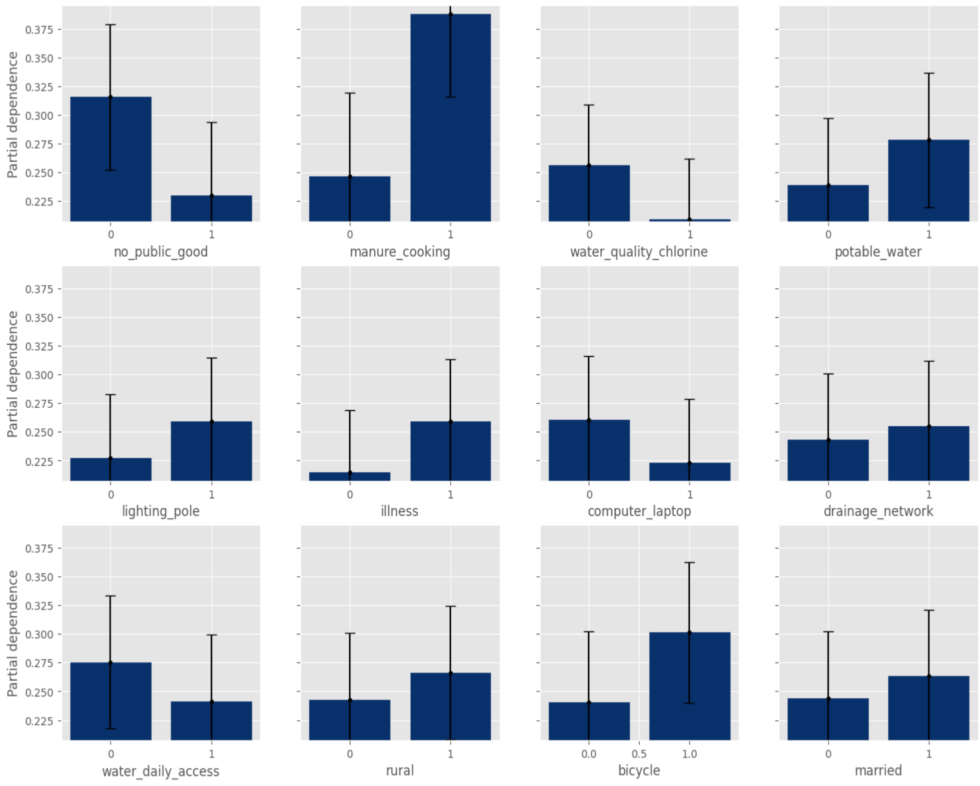

Figure 17: Partial dependence plots for Floods classifier

3.4.2.4. Interpretation of Landslides Classifier

This classifier is different from the Floods one in terms of the marginal impact of each feature. As was depicted before, the total information gain is more evenly distributed between features. Thus, the impact of an individual feature in prediction is small.

In contrast to the case of Floods, long-run wealth seems to be more important as the probability of being affected by a Landslide is lower for households with an Electric iron, Refrigerator, GLP cooking, Computer or laptop, and Washing machine. Furthermore, household construction materials are also important as Landslides susceptibility decreases for households made of Concrete walls, Concrete roofs, and Inlaid walls. In consequence, long-run wealth protects households from being affected by Landslides.

Rural communities are more affected by Landslides. In Rural Perú, finding families that have built their houses with improvised construction materials is a pattern. This pattern is stronger in households located at higher altitudes that, in the case of Peru, are located at larger distances from cities and markets (Gonzales de Olarte, 2021). The altitude feature is continuous, and the average marginal impact is non-linear. From 0 to 900 m.a.s.l., the probability increases from 0% to

25% (and the variance of predicted probabilities is minimal), then the probability oscillates from 901 to 4000 m.a.s.l., having the highest value (27.5%) at 3700 m.a.s.l..

Similar to the case of floods, most variables have high variance estimates; this is also a consequence of the imbalanced nature of class labels. However, “altitude” has a low variance that suggests that it affects the prediction significantly. “Employment” has a variance greater than other features. The high variance of estimates suggests heterogeneous effects. That means that the classification function is strongly non-linear in features.

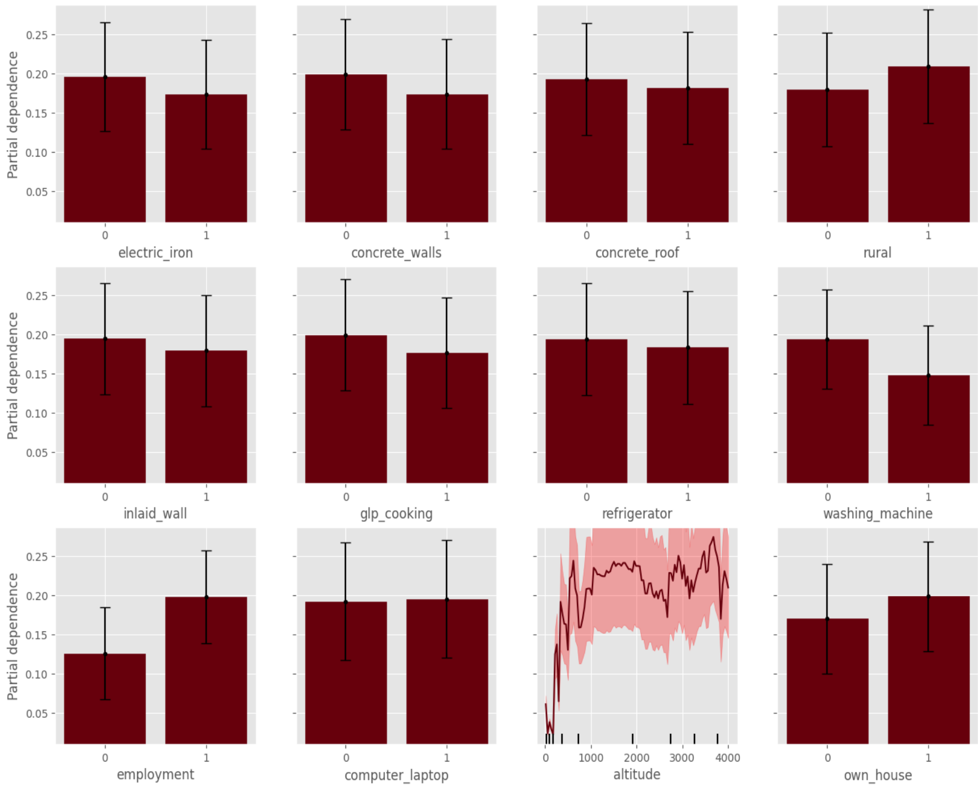

Figure 18: Partial dependence plots for Landslides classifier

3.5. Practical Implications for Humanitarian Operations

This paper delivers two main products oriented to practitioners, managers, operations planners, and policymakers: the classifiers and the interpretation of feature importance. After rigorous research on methodological issues, the authors state that the main objectives of this research were achieved in the case of floods and landslides in Peru.

The floods and landslides classifiers can extrapolate their rules to one-year-ahead data with an accuracy of 78.59% and 81.29% for floods and landslides (corrected for class imbalance), respectively. The proposed ML pipeline is oriented to create high-performing classifiers when class labels are imbalanced. To achieve high accuracy, the ML pipeline is built with XGBoost supervised learning algorithm, cost-sensitive learning, custom performance metric, and Bayesian-optimization hyperparameter search.

Based on the information gain indicator for each feature, the twelve most important features were selected to interpret the rules drawn from XGBoost based on the marginal change in prediction for the average household. In statistical modeling, this method is known as average marginal effects. The computation of partial dependence was done to interpret the impact of the features on prediction and generate interpretable insights.

The humanitarian problem raised by recurrent disasters raises the question: how can human suffering be reduced in crises that combine vulnerability and recurrent exposure to natural hazards (Leiras et al., 2017)? We propose adding value to the formulation of the solution by having an accurate estimate of the future demand distribution. We list the following practical implications of the results for this specific context:

- ✓ This paper describes a blueprint that can produce status-quo representative demand predictions using computational tools. Thus, further research should aim to overcome high computational requirements.
- ✓ The model can be implemented in practice in its current version. However, we recommend improvements to address its probable lack of empathy, transparency, and ethical concerns.
- ✓ Predictions made by the classifiers depend on a probability threshold. The closer the households' predicted probabilities are to the threshold, the more uncertain their outcome.
- ✓ In humanitarian operations, uncertainty must be communicated to stakeholders.
- ✓ Uncertainty should be treated by additional techniques such as threshold tunning, etc.
- ✓ The data was triangulated from different resources without finding inconsistencies. These data collection methods add reliability to this paper's practical implications.

- ✓ The learned classifiers are extrapolated to out-of-sample data to robustly generalize insights, including validating the model through computational experiments.
- ✓ Rural households with poor construction materials located at higher altitudes and deprived of elementary assets are susceptible to Landslides.
- ✓ Rural households in poor and small settlements (manure-cooking households with lighting poles, drainage networks, paved tracks, and paths) with access to potable water are susceptible to Floods.
- ✓ Deploying aid supplies to households with these specific characteristics may create ethical concerns in the practice as households that do not share characteristics would not be targeted for aid.
- ✓ Statistical methods such as proxy-means testing in microeconomic poor-targeting programs such as Conditional Cash Transfer *Ingreso Solidario* in Colombia have proven to improve targeting and reduce ethical concerns.
- ✓ Spatial analysis is essential to make decisions with data. Further facility location or routing formulations are encouraged to prescribe solutions for the humanitarian problem of recurrent disasters.

3.6. Chapter conclusions

This paper has applied supervised Machine Learning to train classifiers to identify final demand points after recurrent disasters. The Peruvian Case study was analyzed. Floods and Landslides were addressed because they were the most critical recurrent disasters that affected households over the entire Peruvian geographical boundaries.

The main contribution of this paper is that it provides the stakeholders with a method to estimate the future distribution of demand for aid with state-of-the-art accuracy. Second, but not less important, is that decision-makers are provided with an interpretation of model predictions that allows them to understand the logic behind the model's outcome. In this sense, the model allows exploiting available data to provide decision-making tools.

The authors believe that data-driven strategies may benefit humanitarian operations and, thus, add significant value to society. However, the fact that this data is just an input for further DRM policies must be highlighted. In this regard, this paper also provides guidelines for potential users. An explicit limitation is that this research was done under scarce collaboration with practitioners, so further research should be focused on generating insights from practical implementations.

Another limitation is that this paper does not contribute to mapping uncertain scenarios. It just provides a point-prediction demand estimation. However, future research may shed light on uncertainty by embedding classifiers into a mathematical optimization framework to prescribe a different decision for each scenario provided by the households' characteristics (Bertsimas and Kallus, 2020). The authors state that research on these formulations may produce even better solutions.

The application of our results might impact the vicious cycle in which the impact of disasters shapes further investments in disaster risk management. The classifiers provide an intelligent data-driven targeting method that saves costs by guaranteeing accuracy. Additionally, the paper provides guidelines to address efficiency in aid distribution, such as facility location and vehicle routing formulations.

Finally, the authors call to action for humanitarian logisticians, disaster risk managers, and other stakeholders. Recurrent disasters are increasingly becoming a challenge for several countries. Furthermore, they challenge the entire humanity. Thus, although research in this area is relevant, real action must be taken for the benefit of the future.

4
Conclusions and recommendations

This research has modeled the Cold waves, Floods, and Landslides disaster risk, considering household characteristics as indicators of multidimensional vulnerability.

In contrast with previous literature, multiple dimensions of vulnerability are being considered, and the predictive approach is complemented by explanatory analytics.

Rural households seem to be the most affected by Cold waves and Landslides in Andean regions located at higher altitudes. Floods affect households that are located next to water corpses at lower altitudes.

Furthermore, households' construction materials were important predictors for the three classifiers. Hence, this research highlights the importance of risk mitigation policies oriented to improve housing to create resilience and mitigate risks.

This research brings insights regarding the statistical association between multidimensional vulnerability and disaster risk reported in the literature. There are several determinants of whether a household is affected by Cold waves, Floods, or Landslides. However, such determinants differ with the type of disaster.

In the case of Cold waves, economic poverty determines the outcome of a household during extreme temperature events. In the case of Floods, affected households are located in semi-urban settlements with inadequate infrastructure. For the case of Landslides, rural poor households located at higher altitudes are the target. Different policies may address the risks of these households.

Recurrent multi-hazards (Floods and Landslides) may affect Peruvian households in the context of El Niño. Risk reduction and disaster preparedness policies must be planned proactively to improve the quality of life and reduce human suffering caused by consequent acute diseases and infrastructure losses.

Implementing the obtained classifiers in actual operations is a challenging task. The first paper concerns how the Cold waves-related disasters classifier can contribute to real operations. It maps the impact of each feature in prediction. It proposes techniques such as statistical analysis and threshold tuning to improve the quality of decisions supported by the usage of the model.

The second paper provided guidelines for practical implementations; however, following them may not necessarily produce a successful implementation. Trial and failure are needed to improve iteratively the model in real-world disaster risk management operations.

Another contribution to the literature is that this research has provided insights into hyperparameter optimization techniques and showed that the learning process can be automated for disaster risk management of recurrent disasters. In short, we developed Machine Learning pipelines that allow computers to learn classifiers by themselves, with reduced human interaction. The authors encourage using such pipelines as they lead to optimal model configuration with data that can be directly observed from reality.

The principal implication of this research is that it provides a management strategy that impacts the pre-disaster management phases. These phases are crucial for the entire lifecycle and the vicious cycle of managing risks and disasters.

Further research must be focused on estimating causal impacts to improve policymaking and proposing location routing problems to plan aid distribution and operate at a lower cost. Several methods may be used for these purposes.

Machine Learning is an applied science, so the best option for further research must be to improve the model's performance. This improvement can be achieved with more features (such as distances from water corpses, the slope in which each household is located, etc.), better Machine Learning pipelines (with other supervised learning algorithms such as neural networks), and the use of exact methods for hyperparameter optimization or improved heuristics. Additionally, we encourage exposing the classifiers to real-world implementations. By doing this and consequently updating the models, they can better adapt to reality.

References

ADB, 2015. Private Sector Disaster Preparedness and Recovery: A Guidebook for Business in Asia. Asian Development Bank. Available online: https://www.adb.org/sites/default/files/publication/168205/private-sector-disaster-preparedness-recovery-guidebook.pdf.

Akram, A., Jamil, F., & Alvi, S. (2022). The effects of natural disasters on human development in developing and developed countries. International Journal of Global Warming, 27(2), 155. https://doi.org/10.1504/IJGW.2022.123279

Alarcón, C., Trebejo, I., & FAO-SENAHMI. (2010). Atlas de helada del Perú. Retrieved from https://repositorio.senamhi.gob.pe/handle/20.500.12542/359

Amirkhani, M., Ghaemimood, S., von Schreeb, J., El-Khatib, Z., & Yaya, S. (2022). Extreme weather events and death based on temperature and CO2 emission – A global retrospective study in 77 low-, middle- and high-income countries from 1999 to 2018. *Preventive Medicine Reports*, 28, 101846. https://doi.org/10.1016/j.pmedr.2022.101846

Bertsimas, D., & Kallus, N. (2020). From Predictive to Prescriptive Analytics. Management Science, 66(3), 1025–1044. https://doi.org/10.1287/mnsc.2018.3253

Besiou, M., & Van Wassenhove, L. N. (2021). System dynamics for humanitarian operations revisited. *Journal of Humanitarian Logistics and Supply Chain Management*, 11(4), 599–608. https://doi.org/10.1108/jhlscm-06-2021-0048

Bosher, L., Chmutina, K., & van Niekerk, D. (2021). Stop going around in circles: towards a reconceptualisation of disaster risk management phases. *Disaster Prevention and Management: An International Journal*, 30(4/5), 525–537. https://doi.org/10.1108/DPM-03-2021-0071

Bottou, L. (2010). Large-Scale Machine Learning with Stochastic Gradient Descent. *In Proceedings of COMPSTAT'2010* (pp. 177–186). Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-2604-3_16

Brownlee, J. (2020). *Imbalanced classification with Python: better metrics, balance skewed classes, cost-sensitive learning*. Machine Learning Mastery.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), 1-13.

Christian, J. I., Basara, J. B., Hunt, E. D., Otkin, J. A., Furtado, J. C., Mishra, V., Xiao, X., & Randall, R. M. (2021). Global distribution, trends, and drivers of flash drought occurrence. *Nature Communications*, 12(1). https://doi.org/10.1038/s41467-021-26692-z

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research*, 12, 2493-2537. Retrieved from https://dl.acm.org/doi/10.5555/1953048.2078195

CRED Centre for Research on the Epidemiology of Disasters (2022). EM-DAT The International Disaster Database [Data set]. Retrieved from https://public.emdat.be

Dantas, L. F., Peres, I. T., Bastos, L.S., Marchesi, J.F., De Souza,G.F., Gelli,J.G.M…& Bozza,F.A.(2021). App-based symptom tracking to optimize SARS-CoV-2 testing strategy using Machine Learning. *PLoS One* ,16(3),e0248920.

Eckhardt, D. ,Leiras, A. & Thomé, A. M. T. (2019). Systematic literature review of methodologies for assessing the costs of disasters. *International Journal of Disaster Risk Reduction* ,33 ,398–416 .https://doi.org/10.1016/j.ijdrr.2018.10.010

Eckhardt,D .,& Leiras,A .,& Thomé,A.M.T.(2022). Using Social Media for Economic Disaster Evaluation: A Systematic Literature Review and Real Case Application. *Natural Hazards Review*, 23(1). https://doi.org/10.1061/(asce)nh.1527-6996.0000539

F. Hutter, L. Kotthoff, & J. Vanschoren (Eds.) (2019). Automated Machine Learning: Methods, Systems and Challenges. In The Springer Series on Challenges in Machine Learning. Springer International Publishing. https://doi.org/10.1007/978-3-030-05318-5

Fauvel,K .,& Fromont,E .,& Masson,V.,& Faverdin,P. ,& Termier, A. (2022).XEM: An explainable-by-design ensemble method for multivariate time series classification. *Data Mining and Knowledge Discovery*, 36(3), 917–957. https://doi.org/10 .1007/s10618-022-00823-6

French, A., Mechler, R., Arestegui, M., MacClune, K., & Cisneros, A. (2020). Root causes of recurrent catastrophe: The political ecology of El Niño-related disasters in Peru. *International Journal of Disaster Risk Reduction*, *47*, 101539.

Ghesquiere, F., & Mahul, O. (2010). Financial protection of disaster risks: is the systemic risk transfer mechanism taking shape? The World Bank Research Observer, 25(1), 77-111.

Giovanelli,J .,& Bilalli,B .,& Abelló Gamazo,A. (2021) .Effective data pre-processing for AutoML. *In Proceedings of the 23rd International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data* (DOLAP): colocated with the 24th International Conference on Extending Database Technology and the 24th International Conference on Database Theory (EDBT/ICDT 2021): Nicosia, Cyprus, March 23 ,2021(pp .1 -10). CEUR-WS.org.

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. Journal of Computational and Graphical Statistics, 24(1), 44-65.

Guha-Sapir, D., Hoyois, P., & Below, R. (2015). Annual Disaster Statistical Review 2014: The numbers and trends. Brussels: Centre for Research on the Epidemiology of Disasters (CRED).

Gutjahr,W.J. & Fischer,S.(2018). Equity and deprivation costs in humanitarian logistics. European Journal of Operational Research, 270(1), 185–197. https://doi.org/10 .1016/j.ejor .2018 .03 .019

Hastie,T., Tibshirani,R., Friedman,J.H., & Friedman,J.H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol.2, pp.1-758). New York: springer.

Holguín-Veras, J., Pérez, N., Jaller, M., Van Wassenhove, L. N., & Aros-Vera, F. (2013). On the appropriate objective function for post-disaster humanitarian logistics models. *In Journal of Operations Management* (Vol. 31, Issue 5, pp. 262–280). Wiley. https://doi.org/10.1016/j.jom.2013.06.002

Hyperopt: A Python library for distributed asynchronous optimization of difficult search spaces, which includes support for both global and local optimization algorithms. Reference: Bergstra, J., Yamins, D., & Cox, D. D. (2013). Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. Proceedings of the 30th International Conference on Machine Learning (ICML-13).

IFRC. (2013). Disaster risk reduction: Saving lives and livelihoods. International Federation of Red Cross and Red Crescent Societies.

Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *The Journal of Supercomputing*, 77, 5198-5219.

Jardim, R., dos Santos, M., Neto, E., Muradas, F. M., Santiago, B., & Moreira, M. (2022). Design of a framework of military defense system for governance of geoinformation. *In Procedia Computer Science* (Vol. 199, pp. 174–181). Elsevier BV. https://doi.org/10.1016/j.procs.2022.01.022

Keja-Kaereho, C., & Tjizu, B. R. (2019). Climate Change and Global Warming in Namibia: Environmental Disasters vs. Human Life and the Economy. *In Management and Economics Research Journal* (Vol. 5, Issue 1, p. 1). HATASO Enterprises LLC. https://doi.org/10.18639/merj.2019.836535

Kim, Y., & Sohn, H.-G. (2018). *Disaster Risk Management in the Republic of Korea.* Springer Singapore. https://doi.org/10.1007/978-981-10-4789-3

Li, X., Caragea, C., Caragea, D., Imran, M., & Ofli, F. (2019). Identifying disaster damage images using a domain adaptation approach. *Proceedings of the International ISCRAM Conference*, 2019-May(May 2019), 633–645.

Lin, A., Wu, H., Liang, G., Cardenas-Tristan, A., Wu, X., Zhao, C., & Li, D. (2020). A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster. International Journal of Disaster Risk Reduction, 49, 101682. https://doi.org/10.1016/j.ijdrr.2020.101682

Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y. (2022). Machine Learning in Disaster Management: Recent Developments in Methods and Applications. Machine Learning and Knowledge Extraction, 4(2), 446–473. https://doi.org/10.3390/make4020020

Lopez, C., Marti, J. R., & Sarkaria, S. (2018). Distributed reinforcement learning in emergency response simulation. IEEE Access, 6, 67261-67276.

López-Bueno, J. A., Navas-Martín, M. Á., Díaz, J., Mirón, I. J., Luna, M. Y., Sánchez-Martínez, G., Culqui, D., & Linares, C. (2021). The effect of cold waves on mortality in urban and rural areas of Madrid. *Environmental Sciences Europe*, 33(1). https://doi.org/10.1186/s12302-021-00512-z

Luque, A., Carrasco, A., Martín, A., & de Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216-231.

Mattea, S. (2019). Exploring spatial sources of preference heterogeneity for landslide protection. *Land Economics*, Vol. 95 No. 3, pp. 333-352.

Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., & Kanevski, M. (2013). Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. *In Mathematical Geosciences* (Vol. 46, Issue 1, pp. 33–57). Springer Science and Business Media LLC. https://doi.org/10.1007/s11004-013-9511-0

OCHA. (2015). Preparedness and contingency planning for humanitarian response. United Nations Office for the Coordination of Humanitarian Affairs.

Optuna: A next-generation hyperparameter optimization framework that is designed to be highly scalable and versatile, with support for various optimization algorithms and ML libraries. Reference: Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19).

Owen, L. (2022). *Hyperparameter tuning with python: boost your Machine Learning model's performance via hyperparameter tuning*. Packt Publishing, 2022

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. The *Journal of Machine Learning Research*, 12, 2825-2830.

Quiliche, R. & Mancilla, L. (2022, January 10). Determinantes de la diversificación de ingresos en los hogares rurales del Perú. http://hdl.handle.net/20.500.12404/21237

Quiliche, R., Rentería-Ramos, R., de Brito Junior, I., Luna, A., & Chong, M. (2021). Using Spatial Patterns of COVID-19 to Build a Framework for Economic Reactivation. In *Sustainability* (Vol. 13, Issue 18, p. 10092). MDPI AG. https://doi.org/10.3390/su131810092

Ramos, M.-H., Mathevet, T., Thielen, J., & Pappenberger, F. (2010). Communicating uncertainty in hydro-meteorological forecasts: mission impossible? In *Meteorological Applications* (Vol. 17, Issue 2, pp. 223–235). Wiley. https://doi.org/10.1002/met.202

Regal Ludowieg, A., Ortega, C., Bronfman, A., Rodriguez Serra, M., & Chong, M. (2022). A methodology for managing public spaces to increase access to essential goods and services by vulnerable populations during the COVID-19 pandemic. *Journal of Humanitarian Logistics and Supply Chain Management*, 12(2), 157–181. https://doi.org/10.1108/JHLSCM-02-2021-0012

'Reisinger, Andy, Mark Howden, Carolina Vera, et al. (2020) The Concept of Risk in the IPCC Sixth Assessment Report: A Summary of Cross-Working Group Discussions. Intergovernmental Panel on Climate Change, Geneva, Switzerland. pp15

García, V., Sánchez, J. S., & Mollineda, R. A. (2012). On the effectiveness of preprocessing methods when dealing with different levels of class imbalance. Knowledge-Based Systems, 25(1), 13–21. doi:10.1016/j.knosys.2011.06.013

Renteria, R., Chong, M., de Brito Junior, I., Luna, A., & Quiliche, R. (2021). An entropy-based approach for disaster risk assessment and humanitarian logistics operations planning in Colombia. *In Journal of Humanitarian Logistics and Supply Chain Management* (Vol. 11, Issue 3, pp. 428–456). Emerald. https://doi.org/10.1108/jhlscm-03-2020-0018

Revich, B. A., & Shaposhnikov, D. A. (2016). Cold waves in southern cities of European Russia and premature mortality. In Studies on *Russian Economic Development* (Vol. 27, Issue 2, pp. 210–215). Pleiades Publishing Ltd. https://doi.org/10.1134/s107570071602012x

Sahana, M., Rehman, S., Paul, A. K., & Sajjad, H. (2019). Assessing socio-economic vulnerability to climate change-induced disasters: evidence from Sundarban Biosphere Reserve, India. In *Geology, Ecology, and Landscapes* (Vol. 5, Issue 1, pp. 40–52). Informa UK Limited. https://doi.org/10.1080/24749508.2019.1700670

Šakić Trogrlić, R., van den Homberg, M., Budimir, M., McQuistan, C., Sneddon, A., & Golding, B. (2022). Early Warning Systems and Their Role in Disaster Risk Reduction. In Towards the "Perfect" Weather Warning. SpringerLink. https://link.springer.com/chapter/10.1007/978-3-030-98989-7_2

Salazar-Briones, C., Ruiz-Gibert, J. M., Lomelí-Banda, M. A., & Mungaray-Moctezuma, A. (2020). An Integrated Urban Flood Vulnerability Index for Sustainable Planning in Arid Zones of Developing Countries. In *Water* (Vol. 12, Issue 2, p. 608). MDPI AG. https://doi.org/10.3390/w12020608

Santos, R., Santos, P., Sharan, P., & Rodriguez, C. (2021). Digital Agglomeration in the Improvement of the Human Development Index in Peru. In *2021 IEEE 9th Region 10 Humanitarian Technology Conference* (R10-HTC). IEEE. https://doi.org/10.1109/r10-htc53172.2021.9641710

Shabani, E. (2023). Site response and liquefaction hazard analysis in Ramsar city, North of Iran. *Copernicus GmbH*. https://doi.org/10.5194/egusphere-egu23-4084

Shafapourtehrany, M., Yariyan, P., Özener, H., Pradhan, B., & Shabani, F. (2022). Evaluating the application of K-mean clustering in Earthquake vulnerability mapping of Istanbul, Turkey. In *International Journal of Disaster Risk Reduction* (Vol. 79, p. 103154). Elsevier BV. https://doi.org/10.1016/j.ijdrr.2022.103154

Shao, J., Wang, X., Liang, C., & Holguín-Veras, J. (2019). Research progress on deprivation costs in humanitarian logistics. International Journal of Disaster Risk Reduction, 101343. doi:10.1016/j.ijdrr.2019.101343

Shao, J., Wang, X., Liang, C., & Holguín-Veras, J. (2020). Research progress on deprivation costs in humanitarian logistics. *International Journal of Disaster Risk Reduction*, 42, 101343. https://doi.org/10.1016/j.ijdrr.2019.101343

Shrivastava, P. (2003). Principles of Emergency Planning and Management. *Risk Management*, 5(2), 67–67. https://doi.org/10.1057/palgrave.rm.8240152

Szczyrba, L., Zhang, Y., Pamukcu, D., Eroglu, D. I., & Weiss, R. (2021). Quantifying the role of vulnerability in hurricane damage via a Machine Learning case study. *Natural Hazards Review*, 22(3), 04021028.

Tasnuva, A., Hossain, Md. R., Salam, R., Islam, A. R. Md. T., Patwary, M. M., & Ibrahim, S. M. (2020). Employing social vulnerability index to assess household social vulnerability of natural hazards: an evidence from southwest coastal Bangladesh. In *Environment, Development and Sustainability* (Vol. 23, Issue 7, pp. 10223–10245). Springer Science and Business Media LLC. https://doi.org/10.1007/s10668-020-01054-9

Tatebe, J., & Mutch, C. (2015). Perspectives on education, children and young people in disaster risk reduction. In *International Journal of Disaster Risk Reduction* (Vol. 14, pp. 108–114). Elsevier Ltd. https://doi.org/10.1016/j.ijdrr.2015.06.011

Tomasini, R., & Wassenhove, L. V. (2009). *Humanitarian Logistics.* Palgrave Macmillan UK. https://doi.org/10.1057/9780230233485

Twigg, J. (2004). *Disaster risk reduction: Mitigation and preparedness in development and emergency programming*. Humanitarian Practice Network, Overseas Development Institute.

Ullah, I., Saleem, F., Iyakaremye, V., Yin, J., Ma, X., Syed, S., Hina, S., Asfaw, T. G., & Omer, A. (2022). Projected Changes in Socioeconomic Exposure to Heatwaves in South Asia Under Changing Climate. In *Earth's Future* (Vol. 10, Issue 2). American Geophysical Union (AGU). https://doi.org/10.1029/2021ef002240

UNDRR, 2020. Global Assessment Report on Disaster Risk Reduction. United Nations Office for Disaster Risk Reduction. Available online: https://gar.unisdr.org/2021.

United Nations Office for Disaster Risk Reduction (2015). The Sendai Framework for Disaster Risk Reduction, United Nation, New York, NY.

Van Wassenhove, L. N. (2006). Humanitarian aid logistics: supply chain management in high gear. *Journal of the Operational Research Society*, 57(5), 475–489. https://doi.org/10.1057/palgrave.jors.2602125

Venkatesh, B., & Anuradha, J. (2019). A Review of Feature Selection and Its Methods. In *Cybernetics and Information Technologies* (Vol. 19, Issue 1, pp. 3–26). Walter de Gruyter GmbH. https://doi.org/10.2478/cait-2019-0001

Villarroel-Lamb, D. (2020). Quantitative risk assessment of coastal erosion in the caribbean region. *Natural Hazards Review*, 21(3), 04020021.

Wang, X., Choi, T. M., Liu, H., & Siu, K. Y. (2018). A novel hybrid ant colony optimization algorithm for emergency transportation problems during post-disaster scenarios. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48(4), 545-556.

Waring, J., Lindvall, C., & Umeton, R. (2020). Automated Machine Learning: Review of the state-of-the-art and opportunities for healthcare. *Artificial intelligence in medicine*, 104, 101822.

WFP. (2018). Forecast-based financing: A new approach to humanitarian action. United Nations World Food Programme.

Wilcox, R. (2017). Introduction to Robust Estimation and Hypothesis Testing (4th ed.). ScienceDirect. https://www.sciencedirect.com/book/9780128047330/introduction-to-robust-estimation-and-hypothesis-testing

Wright, N., Fagan, L., Lapitan, J. M., Kayano, R., Abrahams, J., Huda, Q., & Murray, V. (2020). Health Emergency and Disaster Risk Management: Five Years into Implementation of the Sendai Framework. In *International Journal of Disaster Risk Science* (Vol. 11, Issue 2, pp. 206–217). Springer Science and Business Media LLC. https://doi.org/10.1007/s13753-020-00274-x

Xin, Y., & Ren, X. (2022). Predicting depression among rural and urban disabled elderly in China using a random forest classifier. *BMC Psychiatry*, 22(1), 1-11.

Xu, X., Liang, T., Zhu, J., Zheng, D., & Sun, T. (2019). Review of classical dimensionality reduction and sample selection methods for large-scale data processing. In *Neurocomputing* (Vol. 328, pp. 5–15). Elsevier BV. https://doi.org/10.1016/j.neucom.2018.02.100

Yang, L., & Shami, A. (2020). On hyperparameter optimization of Machine Learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

Zhao, J., Zhang, Q., Wang, D., Wu, W., & Yuan, R. (2022). Machine Learning-Based Evaluation of Susceptibility to Geological Hazards in the Hengduan Mountains Region, China. *International Journal of Disaster Risk Science*, 13(2), 305–316. https://doi.org/10.1007/s13753-022-00401-w

Zheng, A., & Casari, A. (2018). Feature engineering for Machine Learning: principles and techniques for data scientists. O'Reilly Media, Inc.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B* (Statistical Methodology), 67(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Appendix - Elsevier copyright

ELSEVIER

Overview

In order for Elsevier to publish and disseminate research articles, we need certain publishing rights from authors, which are determined by a publishing agreement between the author and Elsevier.
For articles published open access, the authors license exclusive rights in their article to Elsevier where a CC BY-NC-ND end user license is selected, and license non-exclusive rights where a CC BY end user license is selected.
For articles published under the subscription model, the authors typically transfer copyright to Elsevier. In some circumstances, authors may instead grant us (or the learned society for whom we publish) an exclusive license to publish and disseminate their work.
Regardless of whether they choose to publish open access or subscription with Elsevier, authors have many of the same rights under our publishing agreement, which support their need to share, disseminate and maximize the impact of their research.
For open access articles, authors will also have additional rights, depending on the Creative Commons end user license that they select. This Creative Commons license sets out the rights that readers (as well as the authors) have to re-use and share the article. Learn how articles can be re-used and shared under these licenses.
This page aims to summarize authors' rights when publishing with Elsevier; these are explained in more detail in the publishing agreement between the author and Elsevier.
Irrespective of how an article is published, Elsevier is committed to protect and defend authors' works and their reputation. We take allegations of infringement, plagiarism, ethical disputes, and fraud very seriously.

Author rights

The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to publish either open access or subscription. These apply to the corresponding author and all co-authors.

Author rights in Elsevier's proprietary journals	**Published open access**	**Published subscription**
Retain patent and trademark rights	√	√
Retain the rights to use their research data freely without any restriction	√	√
Receive proper attribution and credit for their published work	√	√
Re-use their own material in new works without permission or payment (with full acknowledgement of the original article): 1. Extend an article to book length 2. Include an article in a subsequent compilation of their own work 3. Re-use portions, excerpts, and their own figures or tables in other works.	√	√
Use and share their works for scholarly purposes (with full acknowledgement of the original article): 1. In their own classroom teaching. Electronic and physical distribution of copies is permitted 2. If an author is speaking at a conference, they can present the article and distribute copies to the attendees 3. Distribute the article, including by email, to their students and to research colleagues who they know for their personal use 4. Share and publicize the article via Share Links, which offers 50 days' free access for anyone, without signup or registration 5. Include in a thesis or dissertation (provided this is not published commercially) 6. Share copies of their article privately as part of an invitation-only work group on commercial sites with which the publisher has a hosting agreement	√	√
Publicly share the preprint on any website or repository at any time.	√	√
Publicly share the accepted manuscript on non-commercial sites	√	√ using a CC BY-NC-ND license and usually only after an embargo period (see Sharing Policy for more information)
Publicly share the final published article	√ in line with the author's choice of end user license	×
Retain copyright	√	×

Institution rights

Regardless of how the author chooses to publish with Elsevier, their institution has the right to use articles for classroom teaching and internal training. Articles can be used for these purposes throughout the author's institution, not just by the author:

Institution rights in Elsevier's proprietary journals (providing full acknowledgement of the original article is given)	**All articles**
Copies can be distributed electronically as well as in physical form for classroom teaching and internal training purposes	√
Material can be included in coursework and courseware programs for use within the institution (but not in Massive Open Online Courses)	√
Articles can be included in applications for grant funding	√
Theses and dissertations which contain embedded final published articles as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publication on ScienceDirect	√

Government rights

For US government employees, works created within the scope of their employment are considered to be public domain and Elsevier's publishing agreements do not require a transfer or license of rights for such works.
In the UK and certain commonwealth countries, a work created by a government employee is copyrightable, but the government may own the copyright (Crown copyright). Please find information about UK government employees publishing open access.

Find out more

- Download a sample publishing agreement for articles financed by journal subscriptions in English(opens in new tab/window) and French(opens in new tab/window)
- Download a sample publishing agreement for articles published open access with a commercial user license(opens in new tab/window) (CC BY) and a non-commercial user license(opens in new tab/window). (CC BY-NC-ND)
- For authors who wish to self-archive see our sharing guidelines
- See our author pages for further details about how to promote your article
- See our hosting page for additional information on hosting research published by Elsevier

- For use of Elsevier material not defined here please see our permissions page or visit the Permissions Support Center(opens in new tab/window)
- If an author has become aware of a possible plagiarism, fraud or infringement we recommend contacting their Elsevier publishing contact who can then liaise with our in-house legal department

- If you are publishing in a society or third party owned journal, they may have different publishing agreements. Please see the journal's Guide for Authors for journal specific copyright information