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

Abstract

This paper trains a household-level disaster risk classifier based on supervised machine learning algorithms for cold wave-related disasters. The households' features considered for this task proxy multiple dimensions of vulnerability to disasters accounting for economic, health, social, and geographical dimensions. These features are theoretically hypothesized to explain disaster risk classification. We test our predictive model based on the case of Puno, Peru, where cold wave-related disasters (e.g., -28°in 2003 and -35° in 2004) are recurrent and overwhelming. Two supervised learning algorithms were tested to build the classifiers: Logistic Regression and Random Forest Classifier. Hyperparameters of such models were optimized through Bayesian Optimization heuristic. Random Forest Classifier outperformed Logistic Regression by 1.16% in MCC and 2.34% in Sensitivity. In the test dataset, Random Forest Classifier achieved MCC of 48.64% and Sensitivity of 80.9%. After statistical analysis and trhreshold tunning, in the optimal setting, both MCC and Sensitivity increased to 51.05% and 86.52% respectively. Feature importance drawn from features' contribution to a reduction in entropy in the construction of the forest suggests that per capita expenditure, household localization in a rural area, altitude, access to public goods, and concrete walls drive the disaster risk classification. Further research must propose strategies to validate the predictive model externally and to analyze the causality of the most important features regarding endogenous disaster risk classification.

1. Introduction

Clime-related hazards are more frequent in this century than in the previous one (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 would lead 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). The impact of disasters depends on the degree of vulnerability, 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 clime-related hazards (Shabani 2022). Vulnerability shapes the damage a natural hazard could cause and is entirely defined by anthropogenic conditions (Bolin, 2006). Exposure is the geographical conditioning of infrastructure, housing, and other tangible assets into hazard-prone areas (Mattea, 2019).

Proactive disaster risk reduction is essential for communities affected by recurrent disasters. However, disaster risk management phases are not independent (Besiou et al., 2021). Thus, proactive disaster risk reduction activities are carried out before a disaster strike 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; Lopez-Bueno et al., 2020). India, Bangladesh, Poland, and Russia are the most-affected countries, harming 1227 million people and generating 184 thousand deaths since 2000. Disasters triggered by cold waves cause losses of human lives in cases of high vulnerability, where households have poor infrastructure and scarce goods to face cold (Lopez-Bueno et al., 2020) or inhabitants have a high prevalence of comorbidities such as cardiovascular diseases (Shaposhnikov and Revich, 2016).

This work focuses on disaster preparedness following a data-centric approach (EM-DAT, 2022). We aim to predict which households need to be prepared for a disaster that cold waves or severe winter conditions can trigger. This prediction must be accurate for the households at risk, representing demand points that must be met. When a predictive model misclassifies positive outcomes, defined as households at risk, deprivation costs are created to represent demand points that need essential supplies, but the model misclassifies 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% of them 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 located in the southeast of 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, it could be challenging to estimate the number of affected people, the economic losses, etc.

Research on proactive disaster risk reduction would significantly impact Puno because of the high prevalence of agricultural households', in which these disasters may cause economic losses that impact their long-run wealth. If a community is not prepared to face cold wave-related disasters, it might enter into a vicious cycle of cold waves affecting the economy. 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 costs 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 of particular importance. Assessment of delivery strategies, transportation costs, and their balance with deprivation costs are left for future research as the objective function is the primary concern of humanitarian logistics.

The contribution of this paper is twofold. First, we introduce vulnerability-based disaster risk prediction, contrasting it with other predictive strategies based on meteorological, geophysical, or geographical modeling. Then, we propose a hyperparameter optimization algorithm based on domain requirements, such as minimizing false negatives. 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. The experimental setting for hyperparameters' Optimization considers confusion matrix metrics by co-optimizing on Matthews Correlation Coefficient (MCC) and Negative Predictive Value (NPV). HPO is usually based on one metric, but our methodology includes sequential Optimization of MCC and NPV, where maximization of MCC aims to minimize social costs and maximization of NPV aims to minimize deprivation costs.

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 be used to identify the final demand points that need 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, as those supplies aim to reduce the expected damages to households' livelihoods that are strongly linked to agriculture and livestock (Quiliche and Mancilla, 2021).

The remaining 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 data collection methods, processing pipeline, and experimental setting. Section 4 brings the results for hyperparameter optimization and deprivation costs. Section 5 discusses the main results. Finally, Section 6 brings our conclusions and recommendations for improvements in disaster preparedness strategies and future research avenues.

2. Theoretical foundation

In this section, we first discuss disaster risk reduction (DRR) concepts. Then, we overview the data science applications in DRR.

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 for a geographic boundary (Mors, 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, then it would have small losses after a disastrous event. Vulnerability is a set of conditions that an agent posse that makes 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., 2021). 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 especially 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.

In this regard, Bosher et al. (2021) state that pre-disaster risk reduction and preparedness activities must aim to reduce the total cost of the disaster risk management lifecycle. If a community successfully builds resilience through proactive disaster risk reduction and preparedness interventions, future disasters will produce lower losses. In such cases, the total cost can be smoothed, as illustrated by the green line in Figure 1.

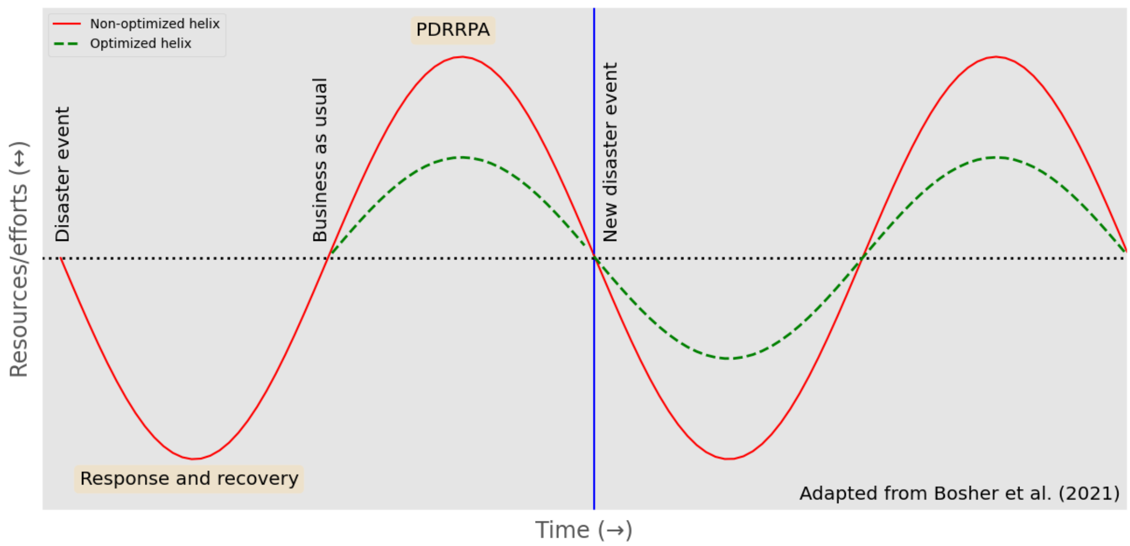

Figure 1. Theoretical representation Disaster Management Helix optimization.

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 having infrastructure destruction, systemic agricultural losses, and hazard to public health (Ferreira, 2012; López-Bueno et al., 2021; Quiliche and Mancilla, 2021).

The first contribution of this paper is that it proposes a solution that is aware of the long-term dynamics of the disaster risk management lifecycle. Implementing a Machine Learning classifier of this type aims to anticipate disaster-related losses to target policies to mitigate risks and prepare agents for upcoming disasters.

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 not only helps to predict but also helps to understand complex phenomena. For instance, data mining applied to disaster risk management is where algorithms find insightful patterns in data that represent chaotic environments characterized by high uncertainties (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, where it incorporates vulnerability, is that predictive modeling characterizes vulnerability by economic factors. A multi-dimensional approach needs to be included to represent vulnerability better. This multi-dimensional vulnerability approach contributes to a better understanding of clime-related disaster risks and improves prediction accuracies (Mors, 2010).

Regarding disaster risk understanding, few studies have considered comprehensive data on multi-dimensional vulnerability (for example, Ahmad and Routray, 2018; Patri et al., 2022). 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. In short, the amount of information that multi-dimensional vulnerability features provide makes it feasible to train an accurate classifier from the vulnerability characterization of agents.

Low-income and lousy infrastructure are the main drivers of vulnerability to clime-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).

Our approach is grounded on previous results summarized in exploratory statistical analysis for clime-related disasters (López-Bueno et al., 2021; Renteria et al., 2021). The probability that an agent would be affected by a natural hazard increases when a set of characteristics are met, such as lack of access to essential services, lack of health, low educational achievement, low social development (Pessoa, 2012), and geographical exposure for the case of disasters (Ullah et al., 2021). In this paper, vulnerability has four dimensions: economic, health, social and geographical. The second contribution of this paper is that it explores an expanded feature space for a vulnerability that considers multiple dimensions that may explain disaster risk.

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. 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.

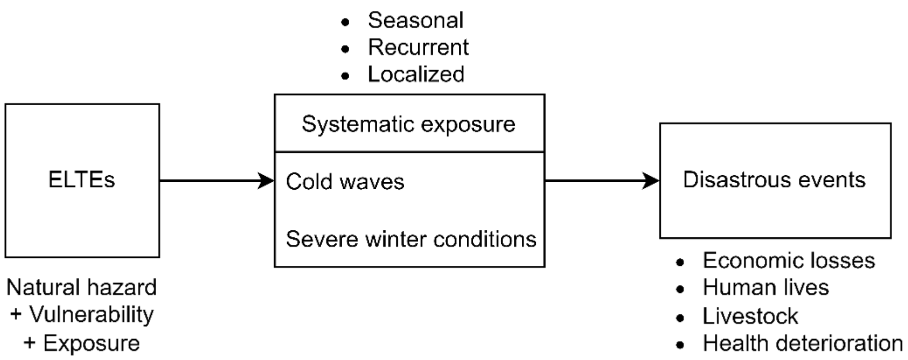

Figure 2. Causes of cold-related disastrous events affecting communities

Regarding cold-wave-related mortality, López-Bueno et al. (2021) performed a statistical analysis of mortality rates in both urban and rural areas of Madrid, Spain. The authors conclude that the main risk drivers of mortality rates are socio-economic. They estimate an index of socio-economic deprivation positively related to mortality rates, controlling for differences between urban and rural municipalities. Amirkhani et al. (2022) found an interesting pattern for a cross-section of countries around the world 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. CO2 emissions are strongly correlated with both frequencies of cold waves and overall temperature variability.

3. The case of Puno, Perú

The third contribution of this paper is that it adapts the standard Machine Learning pipeline to a particular case of study: the Puno region of Peru. The Puno region is affected by cold wave-related disasters. Cold waves reach significant geographic boundaries. In cases where the entire population is exposed to hazardous events, vulnerability differences shape disaster risk. Disasters are more likely to happen where households are more vulnerable. In this case, cold waves produce higher losses for agricultural households or households built with low-quality materials.

Regarding the livelihoods of inhabitants in Peru, Quiliche and Mancilla (2021) stated that rural households decide to diversify their income sources (coming from crops, livestock, among other by-products) considering the risk of being unable to guarantee their subsistence and the reposition of their livelihoods. Rural households must maintain a minimum level of food production, reposition and have a monetary surplus to exchange for health and education services in local markets in contexts of severe deprivations and ELTEs for the case of Puno.

The time series of minimum temperature reported in Figure 3 illustrates the seasonality of the cold waves in Puno. 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.

According to an institutional report published by the Food and Agriculture Organization (Alarcón and Trebejo, 2010), 76.2% of the Puno territory had minimum temperatures from -16 °C to 8 °C for an average of 15 days for June, July, and August. The authors conclude that there was at least one cold wave each year from 1969 to 2010. This finding does not mean that a disastrous event was triggered for each instance. The historical data about disastrous events is limited, but the report states that the hazards are seasonal and recurrent. This fact characterizes disaster risk for households settled over the Puno region: the probability that ELTEs, such as cold waves or severe winter conditions, will trigger disastrous events is considerably high, despite the underreporting of these types of disasters for low-income countries found in EM-DAT (2022) (Amirkhani et al., 2022).

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). Such information is valuable for developing disaster preparedness strategies.

A particular characteristic of the Puno case is that the population is dispersed in space, which matters because the classifier aims to identify the final demand points. In Figure 4, the localization of most final demand points is in rural areas outside the principal cities. The geographic distribution of Puno's households implies increasing logistic costs (Gutjahr and Fischer, 2018). However, misclassifying households at risk of being affected by a cold wave-related disaster 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.

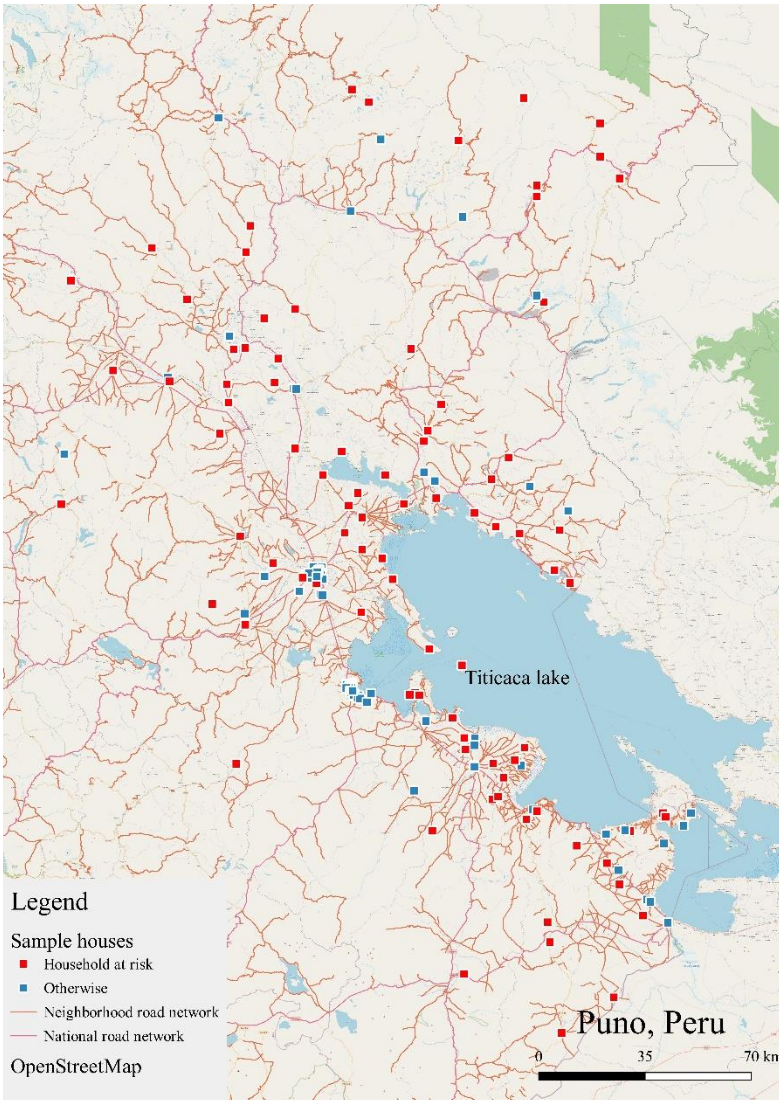

Figure 3. Spatial distribution of households exposed to ELTEs

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 it also contributes to understanding vulnerability drivers at the household level.

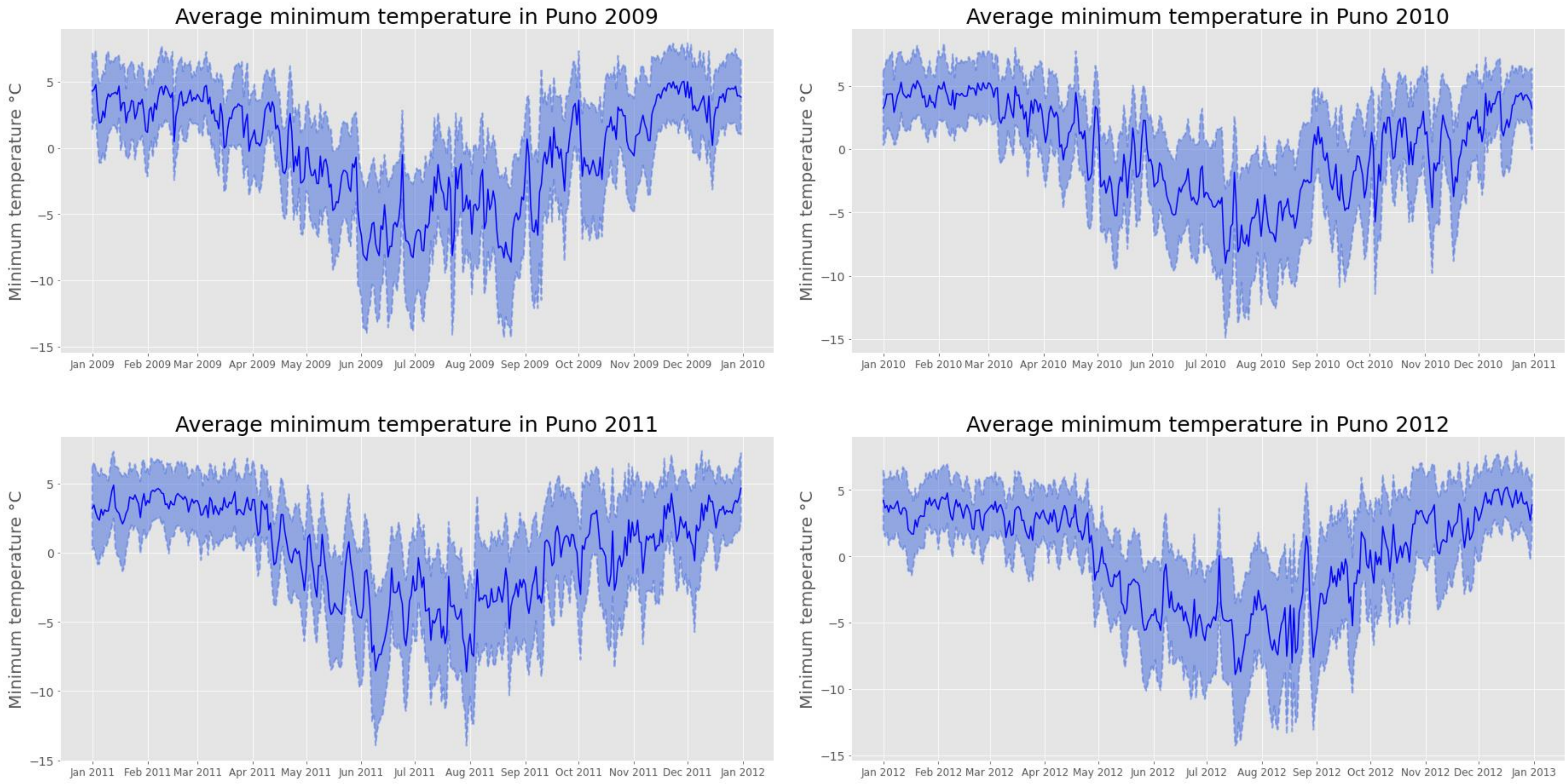

Figure 3. Time series plot for average minimum temperature in Puno 2009-2012

4. Materials and methods

4.1. Data collection methods and the classification problem

Raw data on households' vulnerability characteristics were collected from the National Household Survey carried out by Peruvian's 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 is equal to one If the respondent informed that their house had been affected by natural disasters. In the binary classification jargon, this category is also labeled as positive.

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

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 certain 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 (Lopez Bueno et al., 2020).
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 clime-related disaster, such as cold waves-related disasters (Rentería et al., 2021). The mechanism that explains this correlation is the vulnerability conditions shared by these households.

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.

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 and 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 4 was followed to reach a model with the abovementioned characteristics.

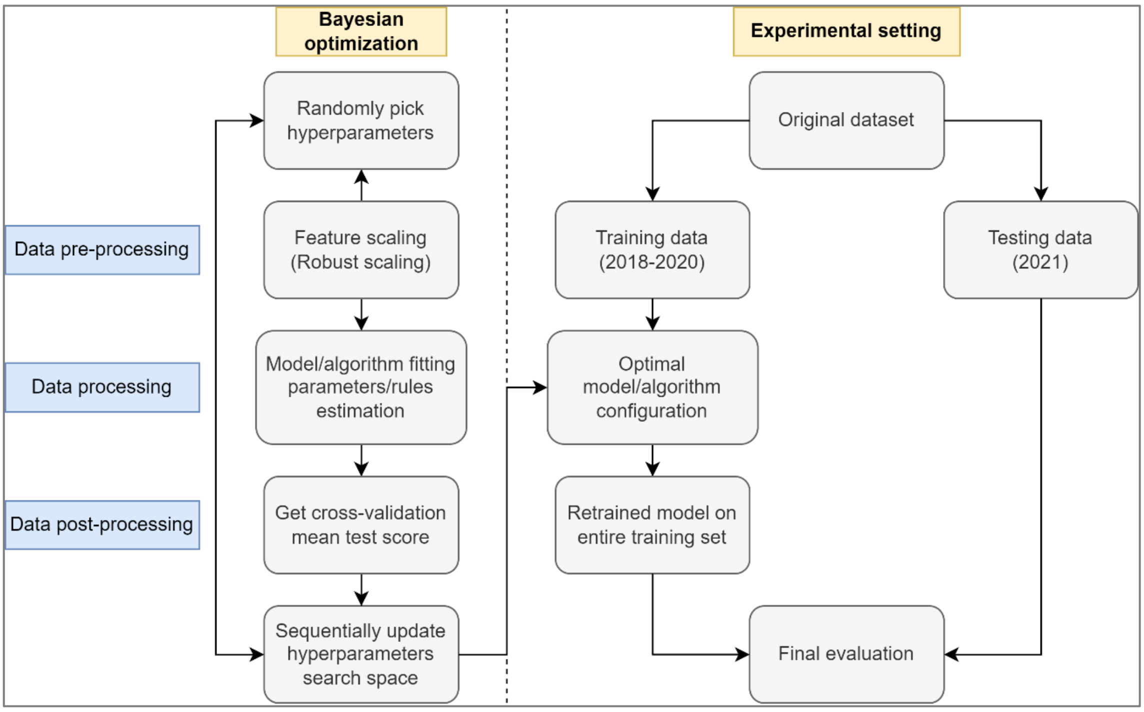

Figure 4. Detailed procedure for Machine Learning model training

Figure 4 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 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.

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 (Xuan 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, and Pedregosa et al., 2011).

Model-based feature selection was adopted in our procedure, known in the literature as sparse learning (Xuan 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 required computational time by increasing the time required to evaluate each hyperparameter configuration and was discarded in our procedure.

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 4, 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.

4.2.2. Data processing

Elastic-Net Logistic Regression (ENLR) and Random Forest Classifier (RFC) were selected because of their functionalities regarding features' importance (Jian et al., 2008; Robert, 2011). 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 (Xuan et al., 2019).

This paper considered 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. Considering these results, we decided to proceed with ENLR and RFC, thus leaving other algorithms for future research. We next explain the mathematical and algorithmic formulation of ENLR and RFC briefly.

- **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:

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

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 4. 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:

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

The procedure in Figure 4 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 $\mathrm{K}_{\mathrm{max}}$. 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 is equal to $\mathrm{Min}_{\mathrm{split}}$. 3.2. The minimum number of data points allowed in a leaf node is equal to $\mathrm{Min}_{\mathrm{leaf}}$. 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 $\mathrm{Max}_{\mathrm{depth}}$. 6. Repeat steps 1 to 5 $\mathrm{N}_{\mathrm{estimators}}$ 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:

$$RFC(.) = \begin{cases} \mathrm{K}_{\mathrm{max}} = 1 \\ \mathrm{CRIT} = UCAT['Gini', 'Entropy'] \\ \mathrm{Min}_{\mathrm{split}}\sim U(0.5,1) \\ \mathrm{Min}_{\mathrm{leaf}}\sim U(0.5,1) \\ \mathrm{CPP}_{\alpha}\sim U(0,0.1) \\ \mathrm{Max}_{\mathrm{depth}}\sim UINT(1,20) \\ \mathrm{N}_{\mathrm{estimators}}\sim UINT(0,100) \end{cases} \tag{4}$$

Cross-validation helps to identify the optimal values of CRIT, $\mathrm{Min}_{\mathrm{split}}$, $\mathrm{Min}_{\mathrm{leaf}}$, CPP_{α}, $\mathrm{Max}_{\mathrm{depth}}$, and $\mathrm{N}_{\mathrm{estimators}}$. After the cross-validation loop, the optimal ensemble is fitted to training data, as illustrated in Figure 4.

4.2.3. Data post-processing

This section describes what happens at the end of every cross-validation loop. In each iteration, model performance metrics are calculated for each hyperparameter configuration, 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 is particularly large. Due to combinatorial search spaces, the Optimization of hyperparameters is an NP-Hard problem (Yang and Shami, 2020).

Hyperparameter optimization techniques are essential because they improve the performance of ML models. To overcome the computational complexity inherent in hyperparameter optimization procedures, Bayesian Optimization is used as a sequential hyperparameter optimization scheme. 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 the following:

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

If the model misclassifies positive classes, it labels risk households as non-eligible for humanitarian focalization. Thus, the maximization of 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 that lies 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.

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., 2019).

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 or not: 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 another 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 TV (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 that 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 1. Multi-dimensional vulnerability features

Category	Variable
Household exterior and access to public goods	Households with inlaid walls, household with painted walls, Outside tracks are paved, Outside tracks are terrain, Outside paths, Lighting poles, No public good
Ownership and physical characteristics	Independent house, household is a house, household is totally owned, Tittle 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, Another 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

Figure 5 shows the correlation heatmap of features listed in Table 1. Statistical correlation between features was estimated using Spearman's Rank-Order Correlation. There is no visual evidence of a high correlation between features; however, it is worth mentioning that both ENLR and RFC have mechanisms to handle correlated predictors.

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

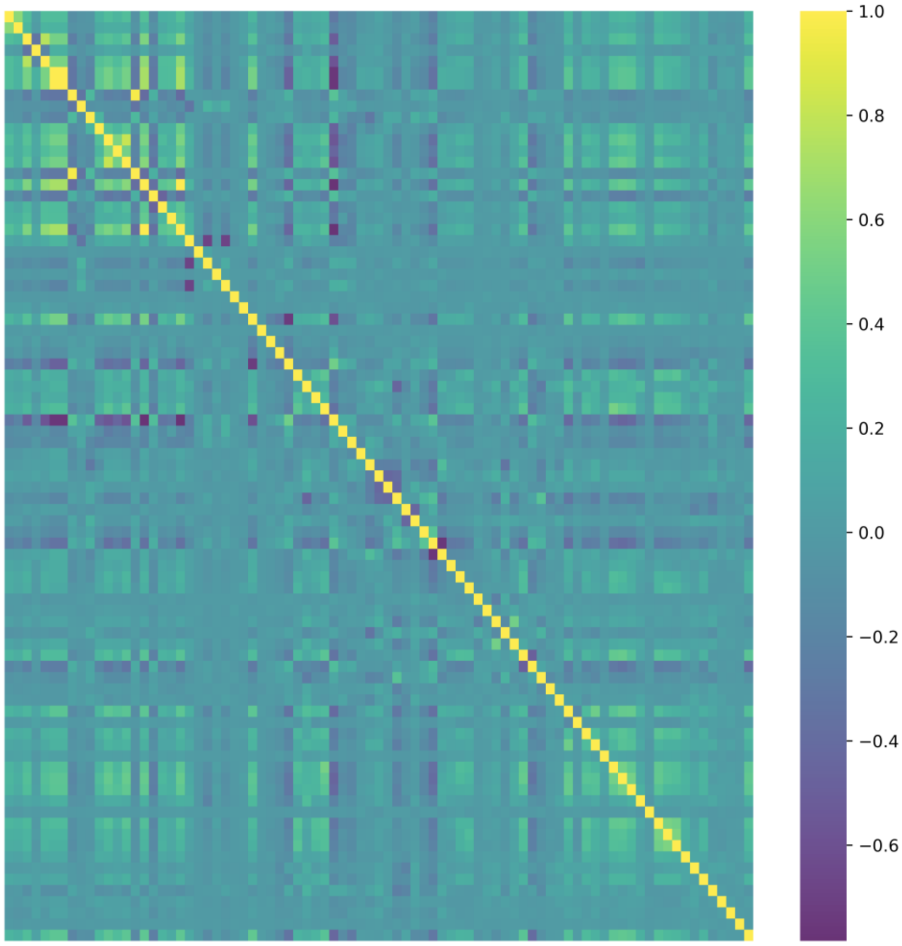

Figure 5. Features' correlation heatmap

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 2 summarizes the main results regarding model performance.

Table 2. Models' performance on the test dataset (Puno, 2021).

	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

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 5:

$$RFC^* = \begin{cases} \mathrm{K_{max}} = 1 \\ \mathrm{CRIT} = Entropy \\ \mathrm{Min_{split}} = 7 \\ \mathrm{Min_{leaf}} = 9 \\ \mathrm{CPP_{\alpha}} = 2.47E-4 \\ \mathrm{Max_{depth}} = 9 \\ \mathrm{N_{estimators}} = 40 \end{cases} \quad (5)$$

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 6:

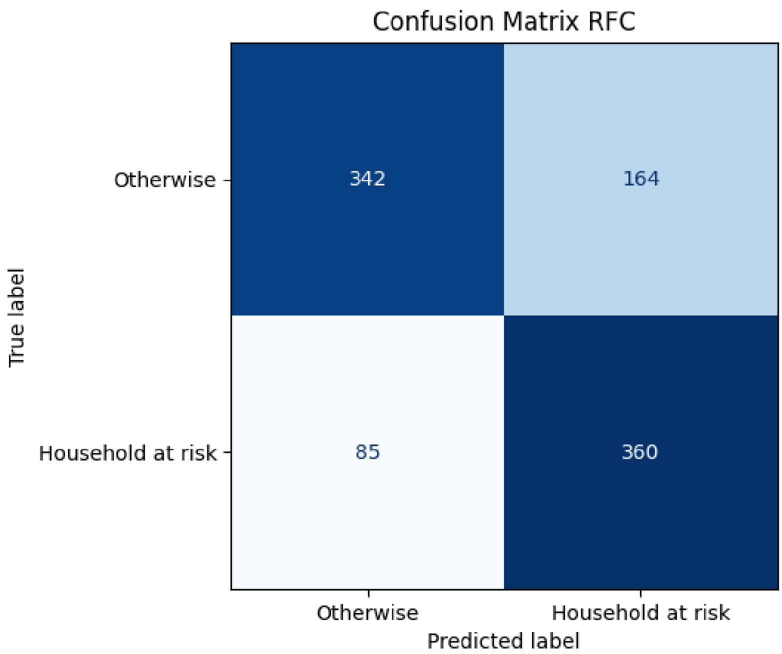

Figure 6. Confusion matrix for Random Forest Classifier

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. We next report complementary results regarding False positives and False Negatives

5.3. Complementary descriptive analysis

We carried out a descriptive analysis of False positives and False Negatives to complement the results above. Table 3 shows the average of each variable across the subpopulations.

Table 3. Descriptive analytics of misclassified categories

Variable	False positives	False negatives
Terrain tracks	30.49%	41.18%
Paved tracks	7.32%	44.71%
Lighting pole	12.20%	85.88%
Own house	88.41%	67.06%
Title of ownership	10.37%	40.00%
Concrete walls	8.54%	48.24%
Altitude	4001.51	3781.26
Rural	88.41%	17.65%
Water network	7.93%	62.35%
Drainage network	7.93%	62.35%
Electric lighting	71.34%	98.82%
Candle lighting	13.41%	1.18%
Another lighting	20.12%	0.00%
GLP cooking	15.85%	62.35%
Manure cooking	59.15%	16.47%
Internet	8.54%	27.06%
Cellphone	71.95%	95.29%
TV color	18.29%	58.82%
Bicycle	25.61%	35.29%
Motorcycle	23.17%	40.00%
DVD	8.54%	23.53%
Car	2.44%	10.59%
Computer/laptop	4.27%	27.06%
Refrigerator	0.00%	8.24%
Per capita expenditure	3799.05	5789.32
Young adult	9.15%	18.82%
Adult	22.56%	37.65%
Old adult	26.83%	31.76%
Old	41.46%	11.76%
overcrowding	50.61%	35.29%
Married	33.54%	31.76%
Literacy	24.39%	11.76%
No education	79.27%	48.24%
Postgraduate education	0.00%	1.18%
Illness	96.34%	91.76%
Medical attention	45.73%	67.06%
Subsidized health insurance	86.59%	69.41%
Disabilities	45.12%	25.88%

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 3, 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% cooks 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 households' 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).

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.

6.1. Determinants of cold wave-related disaster risk

From RFC, features' importance was estimated 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 7 below:

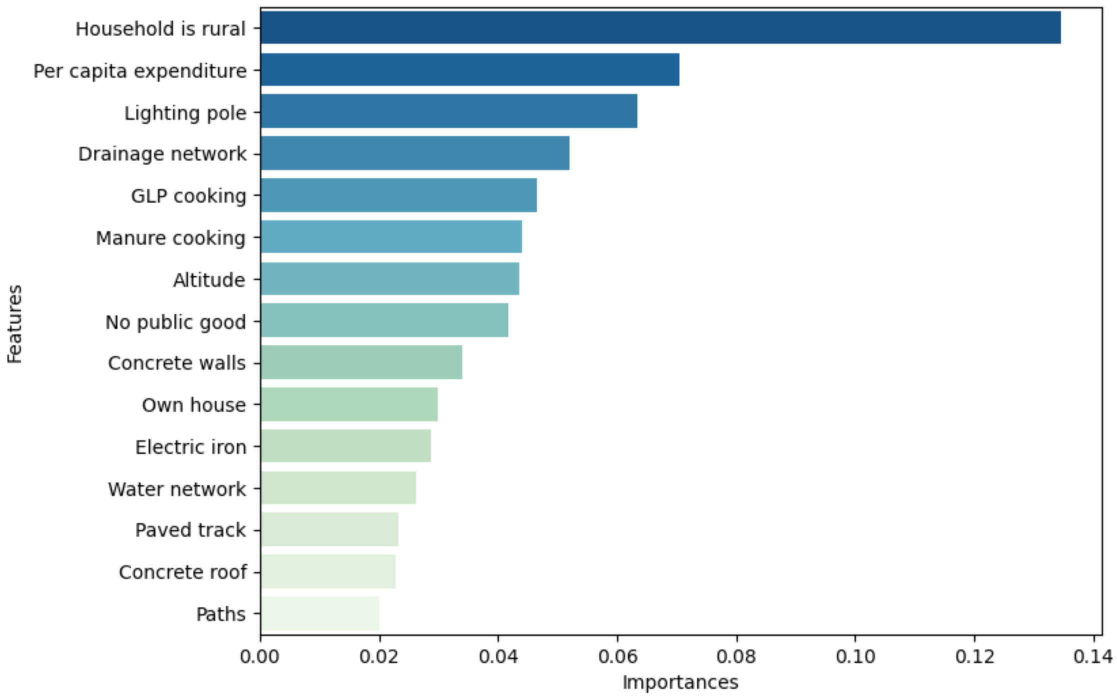

Figure 7. Feature's importance from RFC

The insights are clear: the most important features for prediction were per capita expenditure (that accounts for short-run household purchase power) and 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). 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 significant predictors were altitude (proxies for household exposure to shallow temperature events) and household materials of construction (concrete walls and concrete roofs).

Figure 8 shows an estimate of the average marginal effect or partial dependence plot for each feature in Figure 7. 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 7, 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.

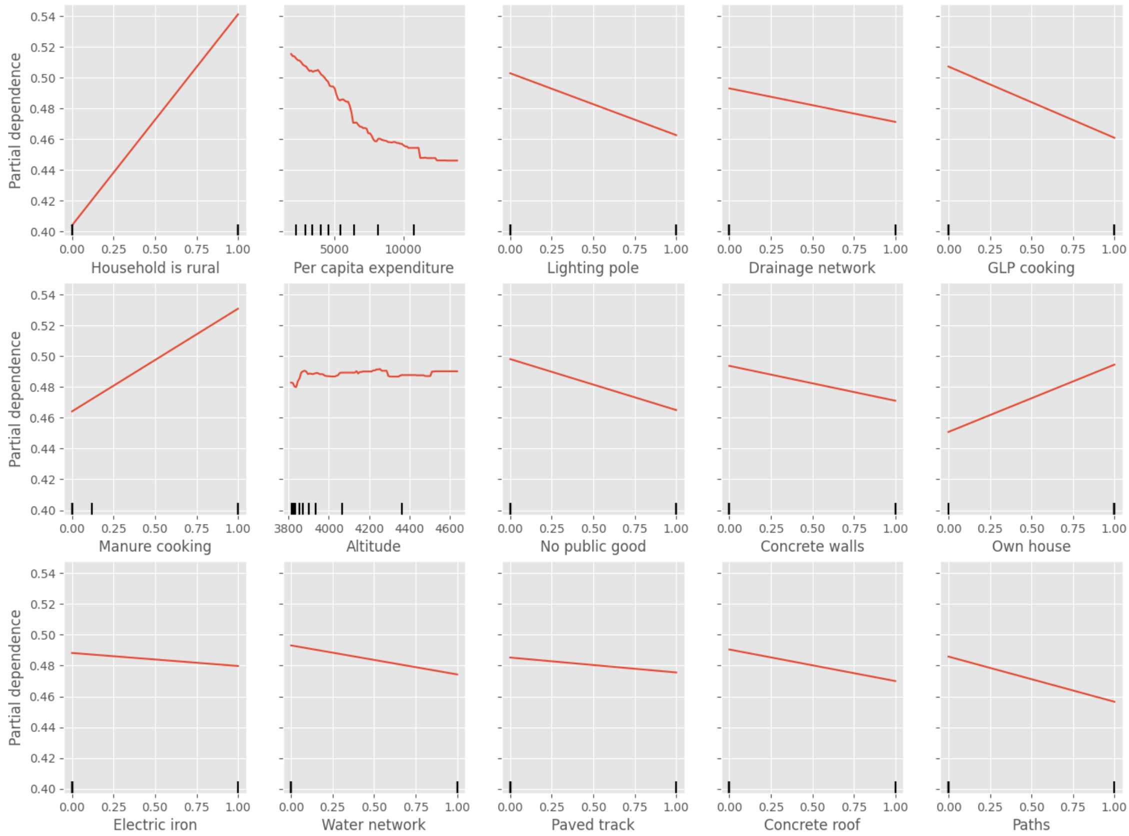

Figure 8. Feature's average marginal effect on the probability of being at risk

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.

Development is a goal that would be achieved at a slow rate and requires a lot of planning. Puno is a city that was built with scarce resources. Hence there is an enormous potential for improvement, particularly regarding mitigating risks related to disasters. It is worth highlighting the fact 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).

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 not affected 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 9 shows the confusion matrix corresponding to a probability threshold of 42%:

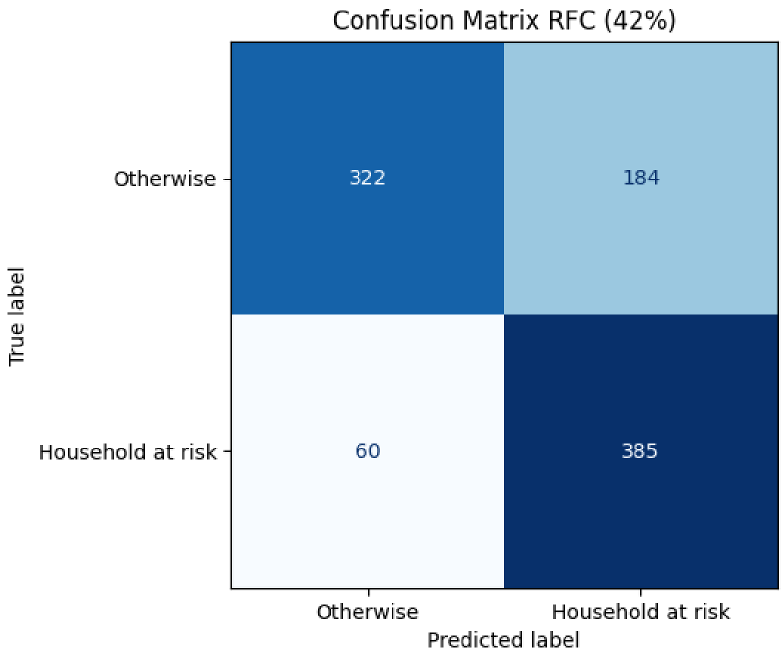

Figure 9. Confusion matrix with new prediction threshold

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 9, 13.48% of households would have generated deprivation costs. 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.

6.3. Extra considerations about practical implementations

From Table 3 and Figure 8, it is known that both False Positives and True Positives are characterized as being poor, rural, and isolated in space. A humanitarian intervention would find it more costly to reach households with these characteristics 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 9 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 9 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. The integration of 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.

7. Conclusions, recommendations, and future research

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 is a region with small cities, and most of its population is settled in rural areas dispersed in space. The classifier identified the following prediction rules:

- Poor households settled in rural areas are vulnerable to cold wave-related disasters and, hence, need proactive humanitarian intervention.
- Beyond economic vulnerability, vulnerable households have the 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 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, maybe 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, Machine 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).

Regarding cold wave-related disasters' risk mitigation, this paper confirms the literature findings: 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 risks related to cold waves in the short term. By reducing the impact of the incoming disaster, Puno's community saves resources that would have been wasted in unmitigated 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 have significant impacts on the probability of being at risk 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 still 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, area of land under cultivation, among others, that may improve the accuracy of the classifier.

This paper concludes that Puno's community would benefit from the practical implementation of Machine Learning in disaster risk reduction. Since humanitarian interventions operate with scarce resources and need to be optimized regardless of their localization or vulnerability condition, this paper sheds light on practical considerations of applied Machine Learning. Hence, it contributes to closing the gap between academia and practitioners toward an improved disaster risk management system based on data.

8. References

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). 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC). IEEE. https://doi.org/10.1109/r10-htc53172.2021.9641710

Ahmed Arafa, A., Radad, M., Badawy, M., & El-Fishawy, N. (2022). Logistic Regression Hyperparameter Optimization for Cancer Classification. *Menoufia Journal of Electronic Engineering Research*, *0*(0), 0–0. https://doi.org/10.21608/mjeer.2021.70512.1034

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

Balcik, B., Beamon, B. M., Krejci, C. C., Muramatsu, K. M., & Ramirez, M. (2010). Coordination in humanitarian relief chains: Practices, challenges and opportunities. *International Journal of Production Economics*, *126*(1), 22–34. https://doi.org/10.1016/j.ijpe.2009.09.008

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

Cantillo, V., Serrano, I., Macea, L. F., & Holguín-Veras, J. (2018). Discrete choice approach for assessing deprivation cost in humanitarian relief operations. *Socio-Economic Planning Sciences*, *63*, 33–46. https://doi.org/10.1016/j.seps.2017.06.004

Daumé, H., & Marcu, D. (2006). Domain adaptation for statistical classifiers. *Journal of Artificial Intelligence Research*, *26*, 101–126. https://doi.org/10.1613/jair.1872

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

Hasebrook, N., Morsbach, F., Kannengießer, N., Franke, J., Hutter, F., & Sunyaev, A. (2022). *Why Do Machine Learning Practitioners Still Use Manual Tuning? A Qualitative Study*. http://arxiv.org/abs/2203.01717

Holguín-Veras, J., Jaller, M., van Wassenhove, L. N., Pérez, N., & Wachtendorf, T. (2014). Material Convergence: Important and Understudied Disaster Phenomenon. *Natural Hazards Review*, *15*(1), 1–12. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000113

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. *Journal of Operations Management*, *31*(5), 262–280. https://doi.org/10.1016/j.jom.2013.06.002

Huang, X., Wu, L., & Ye, Y. (2019). A Review on Dimensionality Reduction Techniques. *International Journal of Pattern Recognition and Artificial Intelligence*, *33*(10), 1950017. https://doi.org/10.1142/S0218001419500174

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

Kim, Y., & Sohn, H.-G. (2018b). *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.

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

Petak, W. J. (1985). Emergency Management: A Challenge for Public Administration. *Public Administration Review*, *45*, 3. https://doi.org/10.2307/3134992

Peterson, T. C., Heim, R. R., Hirsch, R., Kaiser, D. P., Brooks, H., Diffenbaugh, N. S., Dole, R. M., Giovannettone, J. P., Guirguis, K., Karl, T. R., Katz, R. W., Kunkel, K., Lettenmaier, D., McCabe, G. J., Paciorek, C. J., Ryberg, K. R., Schubert, S., Silva, V. B. S., Stewart, B. C., … Wuebbles, D. (2013). Monitoring and understanding changes in heat waves, cold waves, floods, and droughts in the United States: State of knowledge. *Bulletin of the American Meteorological Society*, *94*(6), 821–834. https://doi.org/10.1175/BAMS-D-12-00066.1

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

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. *International Journal of Disaster Risk Reduction*, *79*, 103154. https://doi.org/10.1016/j.ijdrr.2022.103154

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

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

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

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. *International Journal of Disaster Risk Science*, *11*(2), 206–217. https://doi.org/10.1007/s13753-020-00274-x

Yan, X., Xu, K., Feng, W., & Chen, J. (2021). A Rapid Prediction Model of Urban Flood Inundation in a High-Risk Area Coupling Machine Learning and Numerical Simulation Approaches. *International Journal of Disaster Risk Science*, *12*(6), 903–918. https://doi.org/10.1007/s13753-021-00384-0

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

Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. *Journal of Applied Science and Technology Trends*, *1*(2), 56–70. https://doi.org/10.38094/jastt1224

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

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