**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 classifier: Logistic Regression and Random Forest Classifier. Hyperparameters of such models were optimized through a heuristic applied to results of Random Search Cross-Validation, such as the configuration to maximize the model’s ability to produce accurate predictions and to minimize false negatives and, hence, deprivation costs. In the test dataset, Logistic Regression achieved a Matthews Correlation Coefficient of 63.19% and a Negative Predictive Value of 84.15%, while Random Forest Classifier achieved 62.47% and 85.71%, respectively. In the optimal setting, Negative Predictive Value, which controls the false negatives, increased without trading off a significant amount of model performance as suggested by MCC and other metrics. Considering experiments with different sizes of test datasets, the optimal Random Forest Classifier outperformed the optimal Logistic Regression classifier. 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.

# Introduction

Clime-related hazards are more frequent in this century than in the previous one (EM-DAT, 2022). This 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 of 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 large 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 key 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 pre-positioning 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 and 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.

# Theoretical foundation

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

## 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.

Chart, line chart, histogram

Description automatically generated

**Figure 1.** Theoretical representation Disaster Management Helix optimization.

The disaster risk management helix concept 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.

## 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 the process 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 in (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 is that predictive modeling, in the cases where it incorporates vulnerability, characterizes vulnerability by economic factors. A multidimensional approach needs to be included to better represent vulnerability. This multidimensional 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 multidimensional vulnerability (for example, Ahmad and Routray, 2018; Patri et al., 2022). The dimensions of vulnerability are composed of variables with endogenous nature—these features likely 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 multidimensional 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 socioeconomic 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 basic 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.

Graphical user interface, application

Description automatically generated

**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 socioeconomic. They estimate an index of socioeconomic 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 the period 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.

# 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 make the decision to diversify their income sources (coming from crops, livestock, among other by-products) considering the risk of not being able to guarantee their own 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 during 1969-2010. 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 settled and 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. This 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). To overcome this obstacle, a model training was adapted to minimize false negatives to produce accurate classifications with reduced deprivation costs.

Map

Description automatically generated

**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 in order to identify points of final demand and to support pre-disaster risk reduction and preparedness activities. Hence, in the pre-disaster phase, a more significant impact on model implementation is expected. 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.

|  |  |
| --- | --- |
| Chart  Description automatically generated | Chart  Description automatically generated |
| Chart  Description automatically generated | Chart  Description automatically generated |

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

# Materials and methods

## Data collection methods and the classification problem

Raw data on households’ vulnerability characteristics was 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 his/her house has been affected by natural disasters. In the binary classification jargon this category is also labeled as positive.

(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. If a household is at risk of being affected by a drought, storm, plague, flood or landslide, then it is likely that it would 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.

## 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 standard framework to Machine Learning model training (Giovanelli et al., 2021 and Waring et al., 2020) and included three main steps: data pre-processing, data processing and data 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 above-mentioned characteristics.

Diagram, text

Description automatically generated

**Figure 4.** Machine Learning model training procedure

Figure 4 shows an experimental setting that is 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 into practical implications of implementation of Machine Learning techniques into disaster risk reduction. Additionally, this procedure implies hyperparameters’ optimization, that improves model performance on the basis of an objective function.

### Data pre-processing

The feature space extracted from the survey is multi-dimensional. This 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 2 are numeric. The greater the number of features, the more the computational time required for hyperparameter search for each supervised learning algorithm. The literature identifies three most-used approaches to handle multidimensional 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, that is 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 on 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 for evaluation of each hyperparameter configuration and was therefore 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 is shown in Figure 4, the first step in each iteration of cross-validation is to scale the data. The scaling method is known as Robust Scaling, which is 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.

### 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).

As mentioned before, this paper considered expected performance and interpretability as additional criteria for selection of techniques. According to experimental results of Fauvel et al., (2022) on UCI datasets, ENLR outperformed Support Vector Machines (SVM), Local Cascade, and Multilayer Perceptron (MP), and it performed almost as good as Bagging and Boosting, and Simple Ensemble Methods that uses 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 briefly the mathematical and algorithmic formulation of ENLR and RFC.

* **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:

(2)

Where is a data vector corresponding to observation , is the respective observation point for target classes. Considering that both optimal Elastic-Net mixing parameter and inverse of regularization strength is selected based on cross-validation scores, the vector of parameters is estimated fitting 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 learning rate equal to .

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

(3)

The procedure in Figure 4 search for best and considering a priori uniform distributions for both parameters, being defined in logarithmic space.

* **Random Forest classifier**

The algorithm is an ensemble of Decision Trees fitted with 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). To train RFCs, the following steps were followed (Xin and Ren, 2022):

**Algorithm 1.** Random Forest Classifier

|  |
| --- |
| **Random Forest Classifier** |
| 1. Randomly select a subset of features . |
| 2. Randomly sample observations with replacement. |
| 3. Calculate the first node using the best split point under criterion with the obtained subset of data, following the rules defined above (this applies for further nodes): |
| * 1. The minimum number of data points placed in a node before the node is split is equal to |
| * 1. The minimum number of data points allowed in a leaf node is equal to   2. Perform cost-complexity pruning of lower information-gain nodes according to . |
| 4. Categorize the node into daughter nodes using the best split with selected criterion . |
| 5. Categorize more daughter nodes until the tree reaches the defined |
| 6. Repeat steps 1 to 5 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:

(4)

Cross-validation helps to identify the optimal values of , , , , , and . After cross-validation loop, optimal ensemble is fitted to training data, as illustrated in Figure 4.

### 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, that is sampled at random from hyperparameter search spaces defined in Equations 3 and 4. Within cross-validation loop, a training set is randomly shuffled and split into folds of equal size, algorithm is trained with sample composed of folds and tested on the remaining. This produces 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 times in each iteration, this is 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 each other, and optimization program would sample . The more hyperparameters to tune, the bigger the required number of iterations. In RFC algorithm, the combinatorics of possible hyperparameters configurations is particularly large. Due to combinatorial search spaces, optimization of hyperparameters is a NP-Hard problem (Yang and Shami, 2020).

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

* **Objective function**

Regarding the objective function of Bayesian Optimization, it is typically defined as model’s accuracy or other performance metric. In this paper, the 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), and, 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. True Positive Rate decreases with the increase of False Negatives. Hence, the objective function is the following:

If the model misclassifies positive classes, it is labeling risk households as non-eligible for humanitarian focalization. Thus, maximization of leads to 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 ‘no discrimination’ classifier (worst classifier that distributes uniformly the predictions over classes for any probability threshold) and tested classifier. It is defined in function of and coordinates at various probability threshold settings. The range of this metric varies in the closed interval so better classifiers are found when .

**Accuracy**

The estimation of accuracy represents the application of common heuristic where diagonal of confusion-matrix is maximized. The formula is given by . The range of this metric varies in the closed interval so better classifiers are found when .

**F1-Score**

F1- Score is defined as the harmonic mean of the and . The formula is given by . The range of this metric varies in the closed interval so better classifiers are found when .

**Matthews Correlation Coefficient**

This metric is in essence a correlation coefficient that lies in the [-1,1] interval. The formula is given by . 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 is minimizing both deprivation costs 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.

# Results

Section 5.1 presents…

## 5.1 Descriptive statistics

Table 5 shows descriptive statistics for categorical features (dummy-encoded) and Table 6 for numerical features. Additional pre-processing techniques were applied, in this case, categorical features with a frequency lower than 2% of samples were discarded to improve the results of supervised learning algorithms. Small frequencies in categorical features led to null models in the train-test split phase of the training process. As 10 folds were selected for cross-validation, the train-test split procedure entails a high probability of produce a split with a categorical feature equal to zero, which is the same as not considering it at all. Statistical analysis is recommended to investigate the importance of such features as they could be important, regarding disaster risk, or they could be noise.

Table 5. Descriptive statistics

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Variable** | **Count** | **Percent** |
| Household exterior and access to public goods | Household with inlaid walls | 194 | 17.54% |
| Household with painted walls | 156 | 14.10% |
| Outside tracks are paved | 259 | 23.42% |
| Outside tracks are terrain | 311 | 28.12% |
| Outside paths | 226 | 20.43% |
| Lighting pole | 442 | 39.96% |
| No public good | 442 | 39.96% |
| Ownership and physical characteristics | Independent house | 944 | 85.35% |
| Household is a house | 955 | 86.35% |
| Household is totally owned | 913 | 82.55% |
| Tittle of ownership | 231 | 20.89% |
| Concrete walls | 287 | 25.95% |
| Concrete floor | 361 | 32.64% |
| Concrete roof | 228 | 20.61% |
| Overcrowded bedrooms | 374 | 33.82% |
| No other rooms than bedrooms | 124 | 11.21% |
| Access and use of basic services | Water network | 382 | 34.54% |
| Potable water | 531 | 48.01% |
| Quality water (chlorine) | 122 | 11.03% |
| Daily access to water | 668 | 60.40% |
| Drainage network | 382 | 34.54% |
| Electric lighting | 1017 | 91.95% |
| Candle lighting | 51 | 4.61% |
| Other lighting | 49 | 4.43% |
| GLP cooking | 485 | 43.85% |
| Wood cooking | 61 | 5.52% |
| Other cooking | 141 | 12.75% |
| Manure cooking | 417 | 37.70% |
| Phone | 34 | 3.07% |
| Cellphone | 918 | 83.00% |
| Cable TV | 114 | 10.31% |
| Internet | 142 | 12.84% |
| Household assets | Radio | 890 | 80.47% |
| Color TV | 536 | 48.46% |
| Black-White TV | 132 | 11.93% |
| Sound equipment | 92 | 8.32% |
| DVD | 312 | 28.21% |
| Computer or laptop | 194 | 17.54% |
| Electric iron | 264 | 23.87% |
| Electric blender | 336 | 30.38% |
| Gas stove | 985 | 89.06% |
| Refrigerator | 112 | 10.13% |
| Cloth washing machine | 61 | 5.52% |
| Microwave oven | 41 | 3.71% |
| Sewing machine | 74 | 6.69% |
| Bicycle | 307 | 27.76% |
| Car | 82 | 7.41% |
| Motorcycle | 255 | 23.06% |
| Tricycle | 86 | 7.78% |
| Socio-demographics | The head is employed | 959 | 86.71% |
| The head is a woman | 328 | 29.66% |
| The head is married | 482 | 43.58% |
| The head is literate | 217 | 19.62% |
| The head has no education | 708 | 64.01% |
| The head achieved basic education | 264 | 23.87% |
| The head achieved technic education | 53 | 4.79% |
| The head achieved college education | 51 | 4.61% |
| The head achieved pos-graduate education | 30 | 2.71% |
| The head is a young adult (17 to 35 years) | 103 | 9.31% |
| The head is an adult (36 to 50 years) | 316 | 28.57% |
| The head is an old adult (51 to 65 years) | 361 | 32.64% |
| The head is old (more than 66 years) | 326 | 29.48% |
| Health and insurance  (for household members) | Illness (last month) | 1082 | 97.83% |
| Accident (last month) | 247 | 22.33% |
| Healthy (last month) | 305 | 27.58% |
| Chronic illness | 968 | 87.52% |
| Medical intervention (last month) | 739 | 66.82% |
| Contributory health insurance | 198 | 17.90% |
| Subsidized health insurance | 803 | 72.60% |
| Disabilities | 351 | 31.74% |
| Geographical context | Household is located in a rural area | 670 | 60.58% |

Table 6. Continuous features descriptive statistics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **P50** | **Std** | **Min** | **Max** |
| Household per capita expenditure | 5642.5 | 4307.2 | 4727.5 | 832.4 | 69328.4 |
| Household altitude (m.u.s.l.) | 3836.4 | 3860.0 | 437.2 | 1529.0 | 4835.0 |

The population of Puno is composed of owned households (82.55%), and they cook using GLP (43.85%) and manure (37.70%). The prevalence of manure cooking is explained by the prevalence of rurality (60.58%), as GLP logistics can be challenging. It is important to notice that only 60.40% of households have daily access to water. Given this context and the high exposure to ELEs, it is theoretically logical that the population faces a high prevalence of respiratory illness. However, the categorical features give information at a general level: illness (97.83%), and chronic illness (87.52%). In rural regions over the world, it is common to find that the population has health problems (Lopez-Bueno et al., 2022). More than half of the households in the sample have at least one member that searched for medical attention (66.82%), and 72.60% of households have subsidized health insurance. Figure 8 shows correlations of features with at least another feature with a correlation higher than 70%. According to Figure 8, the most correlated features are ‘rural’, ‘concrete walls’, ‘concrete floor’, ‘drainage network’, ‘water network’, ‘paved tracks’, and ‘paths’. These features can potentially be endogenous, and further statistical modeling is needed to draw robust insights into the relationship between these variables and disaster risk. Correlation between features produces multicollinearity, addressed by elastic-net regularization for ENLR and tree-based permutation, a predictive score or importance in each tree on the ensemble for RFC. RFC can provide insights about feature importance based on multiple permutations.

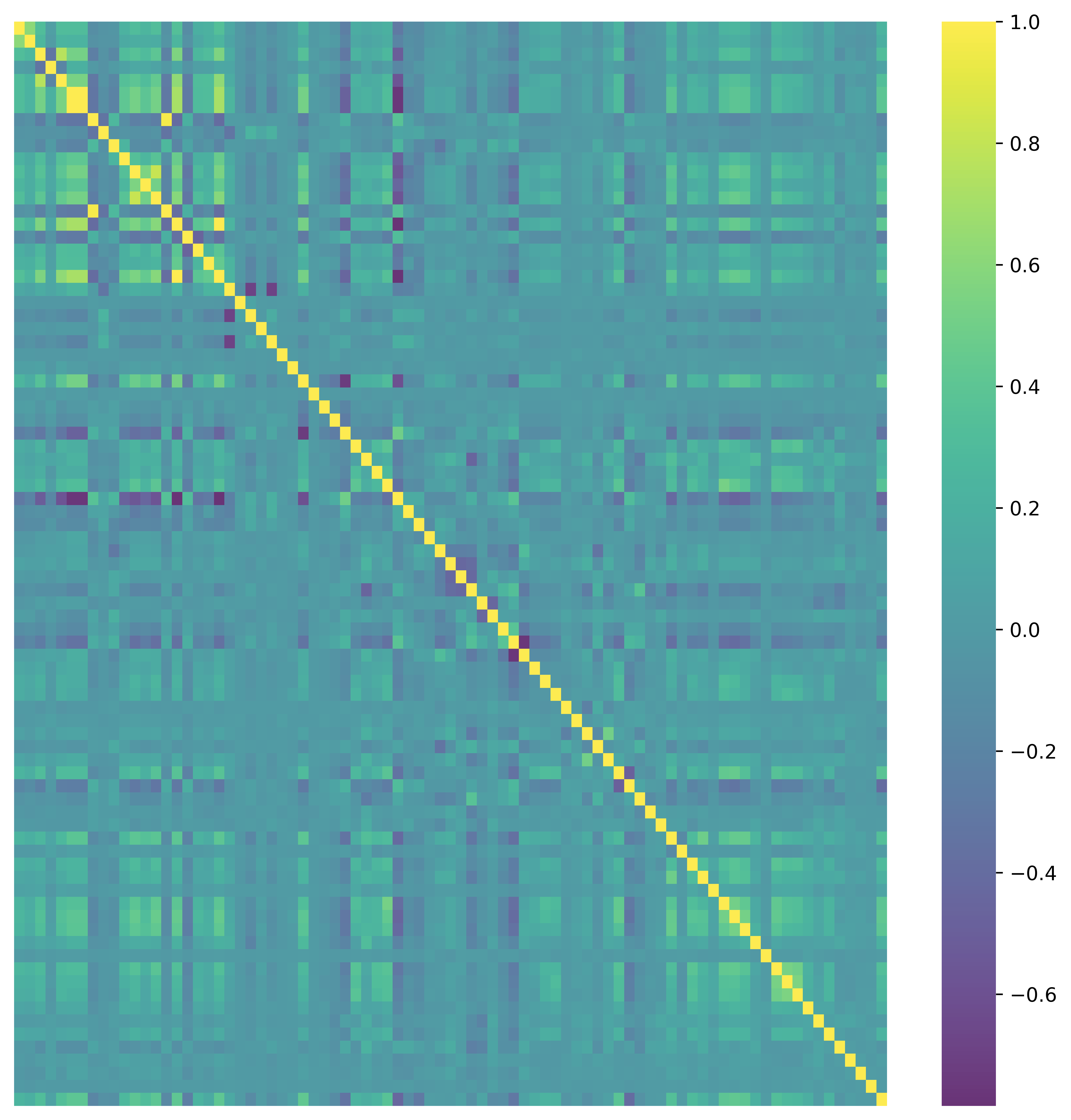


Figure 8: Features' correlation heatmapGráfico, Gráfico de mapa de árvore

Descrição gerada automaticamente

Regarding numerical variables, the annual per capita expenditure measures short-term household nominal income. The average annual per capita expenditure is S/. 5642.5 nuevos soles from 2017 which is equivalent to 1433$ US dollars at current exchange. The average income is below Latin America's principal cities, such as Lima, Bogotá, Buenos Aires, and Rio de Janeiro. Also, for Puno, the mean income is above the median, meaning that more than half of the per capita expenditure distribution is below the average.

Income altitude

## Hyperparameter optimization

Regarding the hyperparameter optimization approach, the best hyperparameters were selected based on experimental results using a repeated stratified cross-validation scheme. Table 7 summarizes the results. One characteristic of the proposed solution is that it guarantees a certain robustness of hyperparameters’ configuration as it is based on multiple experiments (k=10) and repetitions (n=2). After NPV optimization, there is no change in ENLR hyperparameters. For RFC, the new parameters are numerically close to the best results of Random Search CV that optimizes MCC alone. For RFC, the change in NPV is greater than the change in MCC. This is important because MCC is the metric that governs classifier performance for binary classification problems (Chicco and Jurman, 2021). A higher AUC suggests that the new model may be more robust to different probability thresholds for prediction. As the data is balanced, the diminution in Accuracy is explained by the reduction in MCC.

Table 7. Hyperparameter configuration before and after NPV optimization

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Classifier** | **Before NPV optimization** | | | **After NPV optimization** | |
| **ENLR** | **Parameter** | **Value** |  | **Parameter** | **Value** |
| C | 0.100 |  | C | 0.100 |
| l1\_ratio | 0.138 |  | l1\_ratio | 0.138 |
| **Metric** | **Value** |  | **Metric** | **Value** |
| MCC | 54.58 |  | MCC | 54.58 |
| NPV | 80.02 |  | NPV | 80.02 |
| AUC | 82.09 |  | AUC | 82.09 |
| Accuracy | 77.76 |  | Accuracy | 77.76 |
| F1-Score | 81.65 |  | F1-Score | 81.65 |
| **RFC** | **Parameter** | **Value** |  | **Parameter** | **Value** |
| criterion | 'entropy' |  | criterion | 'entropy' |
| max\_depth | 9 |  | max\_depth | 9 |
| max\_features | 0.142 |  | max\_features | 0.141 |
| min\_samples\_leaf | 0.002 |  | min\_samples\_leaf | 0.0001 |
| min\_samples\_split | 0.012 |  | min\_samples\_split | 0.002 |
| n\_estimators | 66 |  | n\_estimators | 87 |
| **Metric** | **Value** |  | **Metric** | **Value** |
| MCC | 56.50 |  | MCC | 56.37 |
| NPV | 81.87 |  | NPV | 82.26 |
| AUC | 82.83 |  | AUC | 82.92 |
| Accuracy | 78.62 |  | Accuracy | 78.53 |
| F1-Score | 82.44 |  | F1-Score | 82.45 |

Metric’s values are computed as the average among folds and repeats within cross-validation scheme.

Figures 9 and 10 show the distribution of MCC and NPV metrics for both ENLR and RFC best hyperparameters’ configuration based on the algorithm in Equation 2. Experimental results show a relatively low variability of MCC and NPV across repeats. However, between the folds, there is an important amount of variability. This finding suggests that the trained model is producing variable results among the data. Considering that the data is a sample drawn from population, this imply that the subset of data that is producing low performance on MCC and NPV could be better modelled by another supervised algorithm. The positive fact is that variability between folds is a pattern, it exists for all the possible configurations of hyperparameters. Future research must seek to minimize the variability between folds, and some algorithms may pay higher attention to mechanisms to minimize this variability.

**Figure 9.** Summary of cross-validation estimates of MCC and NPV for Logistic Regression

|  |  |
| --- | --- |
| **Figure9a.** Boxplot of MCC over repeats ENLR | **Figure9b.** Boxplot of NPV over repeats ENLR |
| Chart, box and whisker chart  Description automatically generated | Chart, box and whisker chart  Description automatically generated |

**Figure 10.** Summary of cross-validation estimates of MCC and NPV for Random Forest Classifier

|  |  |
| --- | --- |
| **Figure.** Boxplot of MCC over repeats RFC | **Figure.** Boxplot of NPV over repeats RFC |
| Chart, box and whisker chart  Description automatically generated | Chart, box and whisker chart  Description automatically generated |

Figures 11 and 12 show the confusion matrix for each algorithm fitted on the best hyperparameters’ configuration obtained through the proposed algorithm. The ‘test\_size’ parameter was fixed to 20%. Thus, the model is trained on 80% of the sample and tested on the other 20%. For these additional experiments, ENLR achieved a MCC of 63.19% and a NPV of 84.15%. On the other hand, RFC achieved 62.47% and 85.71%, respectively.

Nevertheless, these results could be tricky due to the randomness of the split. To overcome this challenge, we repeated the experiments for different values of ‘test\_size’. Results are summarized in Figures 13 and 14. Based on this set of experiments, RFC has a higher chance of producing high results with different sizes of train and test subsets. This implies that RFC is producing systematically better predictions than ENLR, and, thus is more likely to perform better on real-world applications.

Regarding domain-based hyperparameter optimization (algorithm in Equation 2), the proposed approach led to results with minimum false negatives, as shown in confusion matrices (Figures 11 and 12). Experiments with different test sizes show better results in NPV metrics for RFC (Figures 13 and 14). It is worth mentioning that although differences between the MCC optimization alone and co-optimization of MCC and NPV are small, in real-world applications, this would make a difference. When the dataset is a sample of a population, assuming that it is representative, implementing the model with thousands of inhabitants would lead to important savings in terms of deprivation costs that are important to mitigate risks over time for recurrent disasters. In this case, the sample was designed with the objective of representativity of the population, and it also has been previously used to draw insights for policymaking with important impacts (Proaño and Bernabe, 2018).

**Figure 11.** Confusion matrix holdout cross-validation ()

Chart, treemap chart

Description automatically generated

**Figure 12.** Confusion matrix holdout cross-validation ()

Chart, treemap chart

Description automatically generated

Figures 13 and 14 show that, for different test sizes, RFC can adapt better to unseen data as it produces low-variance performance metrics. Following this line, ENLR is more likely to not perform as well as RFC for bigger test sizes. This fact makes a difference in practice, as the cost of misclassifying households at risk of disaster is high because of the potential peaks of deprivation that this could imply.

**Figure 13.** Experiments with different test sizes for Logistic Regression

|  |  |
| --- | --- |
| ENLR |  |
| Chart  Description automatically generated | Chart  Description automatically generated |

**Figure 14.** Experiments with different test sizes for Random Forest Classifier

|  |  |
| --- | --- |
| RFC |  |
| Chart  Description automatically generated | Chart  Description automatically generated |

# Discussions

Regarding the results of the training, the proposed strategy for HPO led to good results in terms of performance on test (unseen) data. Furthermore, all the features used for prediction are vulnerability drivers. The main insight of the predictive analysis is that it is plausible to build a good predictive model for disaster risk that is entirely based on vulnerability. This result is important because it states that it is possible to infer where aid is going to be needed whether decision-makers have prior knowledge about geophysical or meteorological characteristics of disasters. It is not true that predictive modeling of cold waves and severe winter conditions based on geophysical-meteorological features is no longer relevant to decision-making, but it is true that proactive measures and interventions to reduce social costs based on open-data empirical modeling would lead to big savings that are important for the disaster risk management cycle, represented as a helix.

The proposed model can be further extended and improved in terms of predictive power, incorporating geophysical-meteorological features such as distance from lakes, rivers, urban settlements. An improvement of predictive power would lead to greater savings, and eventually an optimization of disaster risk management that is focused on proactive PDRRPA. In terms of risk-reduction, we suggest that further statistical analysis and policymaking focus on the most important features that are drawn from model fitting on best hyperparameters configuration. For this case, the features’ importance can be drawn from estimation of RFC on train dataset. The following Figure 15 shows the features’ importance drawn from a single experiment with optimized hyperparameters for RFC:

**Figure 15.** Random Forest Classifier feature importance

Chart, bar chart

Description automatically generated

The insights are clear: most important features for prediction were per capita expenditure (that accounts for short-run household purchase power), household localization in a rural area (that accounts for the fact that household is isolated on the space and systematically far away from principal urban settlements), altitude (that accounts for household exposure to extreme low temperature events), public goods (that can be measuring the presence of the government on public spaces were households are located) and concrete walls (that is capturing the quality of household construction materials). The other features reported on Figure 15 above tell a similar story. Following these results, we confirm a finding that is in line with disaster risk reduction main guidelines: it is necessary to make long-term investment to systematically reduce vulnerabilities to create resilience in communities by achieving socio-economic development of population. Development is a goal that would be achieved at a slow rate, according to historical data there were few examples of rapid development of communities, but these are considered exceptions (cases of study). For instance, human development index tends to evolute slowly over periods of 6 years (Santos et al., 2021). It is worth highlighting the fact that in the short-term, that is the important term for this analysis, machine learning models can be used to minimize resource utilization and, in the best of cases, save important resources that communities may invest in their future development (Bosher et al., 2022).

# Conclusions, recommendations and future research

The main objective of this paper was to discuss the applicability of machine learning based predictive models to solve a humanitarian logistics problem: the proactive supply of aid to a rural community. Additionally, an alternative hyperparameter optimization strategy, to improve solution considering logistics and deprivation costs as multiple objectives, is presented. This strategy is different from state-of-the-art approaches such as Grid Search, Random Search, Genetic Algorithm and other heuristics proposed to find best hyperparameter configuration. The proposed strategy is summarized as follows: optimize by Random Search Cross-Validation considering MCC as the goal in the training process, then from 5% best found hyperparameter configurations pick up the one that produces the highest NPV. MCC metric is important because it accounts for the classification performance considering equal weight for both positive and negative cases. NPV accounts only for negative cases. The main idea behind this is that the 5% best configurations based only on MCC are very likely to produce a result that minimizes the trade-off between misclassification of negative cases and overall misclassification.

The proposed approach gets better predictive performance for negative cases at the cost of a slightly increase in misclassification of positive cases. For humanitarian logistics domain, misclassification of positive cases implies that aid should be delivered to households that are not at risk of being affected by disasters. However, as the majority of Puno’s territory is exposed to cold waves and severe winter conditions it is probably that all the households in population have at least certain degree of risk of being affected by a cold-related disaster, so the delivery of aid to households labeled as ‘non-risk’ could not be unjustifiably increasing costs. This misclassification produces higher logistic costs, but the key assumption behind this analysis is that the reduction in deprivation costs, that comes from accuracy improvement for negative cases, produces more savings than costs caused by the increase in logistic costs, caused by misclassification of positive cases. Thus, the balance of social costs is positive, and this led to important savings considering the case of study that is characterized by a population suffering from high deprivations. For the case of Puno this approach can potentially led to good results, however, the main assumption is only testable by real-world implementation of trained models. For example, in urban areas the savings of the proposed approach may not be as high as in rural case, as urban household are agglomerated in space.

Machine learning offers a solution to the large-scale problem of deciding where aid must be delivered at a disaggregated level. Model predictions can be used to decide what households would require supply of aid. Decision-makers can implement proactive disaster preparedness strategies such as stock pre-positioning (), proactive delivery, and gradual delivery (Apte and Yoho, 2011) based on information drawn from the prediction of trained models. The models can be applied to census data to estimate the magnitude of savings by generating predictions on disaster risks and building an experimental setting. However, model implementation on a context of a real disaster is advisable, considering the objective of measuring savings caused by proactive disaster preparedness strategies applied based on model predictions. The ideal case is to reach an equilibrium between logistic costs and deprivation costs in real-world outcome. Further research will focus on the aspects of model implementation. For future extensions, the recommended pipeline to use SLAs is to train the model with sample data and test the model with real data. The SLAs used in this paper are not scalable to big data, as training time increases logarithmically with number of samples. Testing other SLAs is recommended for future research, for example XGBoost mixes regularization and ensemble, and it is scalable to big data so a big number of experiments can be performed to reach better solutions regarding predictive power of classification metrics. The actual solution achieved an average MCC of 54.58 for ENLR and 56.50 for RFC, and a NPV of 80.02 and 81.87 respectively.

Regarding disaster risk mitigation, this paper confirms the literature findings about vulnerability and disaster risk. Vulnerable or deprived households, systematically have a greater probability of being affected by a cold-related disaster. The well-known prescription is to create community resilience, which is difficult to achieve in the short term. Instead, we suggest using machine learning to decide where aid must be supplied, considering that humanitarian logisticians operate with scarce resources and need to optimize logistics and provide help to communities regardless of their localization or vulnerability condition. Equality in aid distribution can be achieved at a lower cost if aid is delivered proactively, considering that cold-related disasters are seasonal, recurrent, and localized in Puno.

**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