This section presents the main results. First, we characterize the study case in terms of descriptive statistics of features that might predict cold wave-related disaster risk. Second, we present the main results of Machine Learning model training including model selection

**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% of them have inlaid walls, 52.43% have tracks of which 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 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.

In terms of access to basic services, 34.28% of households are connected to a water and drainage network, 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 uses candle (7.14%) or other lighting (3.53%). The main cooking methods are cooking by GLP (60.58%) and cooking by 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 cellphone.

Households are equipped with assets like color TV (47.36%), bicycle (32.05%), motorcycle (24.35%) and DVD (24.38%). Just 6.54% of households have a particular car, which is explained by the observed poor urban infrastructure. In the 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, that means they have more inhabitants than bedrooms. The 40.68% of households’ heads are married. Puno has a poor development of human capital: the 19.56% of households’ heads are illiterate, 63.02% have no education, and just 2.25% have a postgraduate degree.

Last but not least, 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.** Multidimensional vulnerability features

|  |  |
| --- | --- |
| **Category** | **Variable** |
| Household exterior and access to public goods | Household with inlaid walls, Household with painted walls, Outside tracks are paved, Outside tracks are terrain, Outside paths, Lighting pole, 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 basic services | Water network, Potable water, Quality water (chlorine), Daily access to water, Drainage network, Electric lighting, Candle lighting, Other lighting, GLP cooking, Wood cooking, Other cooking, Manure cooking, Phone, Cellphone, Cable TV, Internet |
| Household income and assets | Per capita expenditure, Radio, Color TV, Black-White TV, Sound equipment, DVD, Computer or laptop, Electric iron, Electric blender, Gas stove, Refrigerator, Cloth washing machine, Microwave oven, Sewing machine, Bicycle, Car, Motorcycle, Tricycle |
| Socio-demographics | The head is employed, The head is a woman, The head is married, The head is literate, The head has no education, The head achieved basic education, The head achieved technic education, The head achieved 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 | 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 high correlation between features, however it is worth mentioning that both ENLR and RFC have mechanisms to handle correlated predictors.

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

Graphical user interface, chart

Description automatically generated

**Figure 5.** Features’ correlation heatmap

**Machine learning model training results**

As optimal hyperparameters were selected based on performance on training dataset, it is important to analyze how trained models perform on unseen data. We use data from 2021 as test dataset to perform this analysis. Table 2 summarizes the main results regarding model performance.

**Table 2.** Models’ performance on test dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **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 it achieved higher Sensitivity that makes 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:

(5)

For reproductible purposes, the trained model was saved to a file, so that it can be loaded in software to reproduce the results or use the model for practical implementations or further research.

We report below the corresponding confusion matrix in Figure 6:

Chart, treemap chart

Description automatically generated

**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 is clearly focusing 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

**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% |
| Other 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 that are characterized as poor in a multidimensional 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 access to internet, 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 vulnerability 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 cooks with GLP, 27.06% have access to internet, and 8.24% have a refrigerator. According to this characterization, False Negatives are mostly non-poor households that are associated with better urban infrastructure. 17.65% of these households are rural. False Negatives might be exposed to cold wave-related disasters, but they 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).

**Discussion, implications, and proposals**

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

**Determinants of cold wave-related disaster risk**

From RFC, features’ importance was estimated to understand which features drives cold wave-related disaster risk at the household level. The results for the 15 most important features are shown in Figure 7 below:

**Chart

Description automatically generated**

**Figure 7.** Feature’s importance from RFC

The insights are clear: 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 household is isolated on the space and systematically far away from principal urban settlements). Access to public goods (that can be measuring the presence of the government on public spaces were households are located) was also important for cold wave-related disaster risk classification. Other important predictors were altitude (proxies for household exposure to extreme low temperature events), and households’ materials of constructions (concrete walls and concrete roof).

Figure 8 shows an estimate of 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 a 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. It is interesting that an increase in per capita expenditure lowers the probability of being at risk at different rates based on the magnitude of expenditure. For instance, for high expenditure households, an increase in expenditure is not related to a large decrease in probability of being at risk. For poor households, the impact of variations in expenditure is higher. Having public goods and concrete on walls and roofs lowers the probability of being at risk.

**Chart, line chart

Description automatically generated**

**Figure 8.** Feature’s average marginal effect on probability

The RFC estimator is robust to non-linearity, heteroscedasticity and noise on predictors. As 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 target variable that is disaster risk.

Considering these results, we arrive to the conclusion 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, that includes access to basic 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 aims at systematically reduce vulnerabilities to create resilience in communities by achieving economic and urban development of population.

Development is a goal that would be achieved at a slow rate, and it 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 mitigation of risks related to disasters. It is worth highlighting the fact that in the short-term, that is the important term for this analysis, applied Machine Learning can be used to optimize resource utilization and, in the best of cases, save important resources that communities may invest in their future development (Bosher et al., 2022).

**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 has been implemented in 2021 and all the demand points would have 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, hence creating additional costs.

Regarding False Positives and False Negatives, the main pattern was the following: poor households without risks that were misclassified (False Positives) and non-poor households with risks that were labelled as non-risky (False Negatives). In this sense, additional costs related to False Positives might not be unjustified, as the majority of households are poor. Considering False Negatives, the average household may be non-poor but may still need aid to face cold waves. Considering statistical analysis, we recommend moving the classification threshold of the RFC in order 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%:

Chart, treemap chart

Description automatically generated

**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 an important impact on the practice. If a humanitarian intervention would been implemented considering confusion matrix in Figure 9, 13.48% of households would have generated deprivation costs. Although there are more False Positives, the majority 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 useful, since methodology can be replicated for other case studies.

**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. In a real-world scenario, a humanitarian intervention would find it more costly to reach households with these characteristics. In contrast, False Negatives and True Negatives are households that are settled in urban areas, with transport infrastructure that reduces logistic costs. Although these households are easy to reach, it could be difficult to identify which of them would be the target of a 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 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, a good strategy to follow is to integrate both the humanitarian intervention that aims to mitigate the impacts from cold wave-related disasters and poverty and hunger short-run 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 extremely low temperatures in Puno.

**Conclusions, recommendations and future research**

This paper focused on the use of 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 the majority of its population is settled in rural areas that are dispersed in space. The classifier identified the following prediction rules:

* Poor households that are settled in rural areas are vulnerable to cold wave-related disasters and, hence, need proactive humanitarian intervention.
* Beyond economic vulnerability, vulnerable households have poor urban infrastructure including tracks, paths, lighting poles, and water and drainage networks. These features characterize households that are demand points of humanitarian interventions.
* The impact of health insurance, health status, and education is minor. Households with unhealthy members have 0.8% higher probability of being at risk than households with healthy members on average. While households with graduate members have 0.6% lower probability of being at risk than other households on average.

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

After performing statistical analysis of False Negatives and False Positives, we considered it profitable to modify the probability threshold of the RFC to improve 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 has several practical implications. First, if this model is implemented, False Negatives would be reduced at the cost of more False Positives, this 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 bigger, program. This could be the case of a policy to mitigate food and hunger. The use of the improved model would have an enormous impact on the Machine Learning-targeted households.

In consequence, Machine Learning offers a data-centered solution to the large-scale problem of deciding where aid must be delivered. This solution is characterized for being detailed and disaggregated at the household level: model predictions can be used to decide which households will require supply of aid. Decision-makers can implement proactive disaster preparedness strategies such as stock pre-positioning, 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 economic deprived households have a greater probability of being affected by a cold-related disaster. The well-known prescription is to create community resilience with strong urban infrastructure, which is difficult to achieve in the short term. In addition to this, we suggest using Machine Learning to implement an automatized classifier that identifies the demand in context of uncertainty and intervene 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 save resources that, otherwise, would have been wasted in unmitigated response and recovery, this provides Puno with opportunities to use the saved resources to carry on long run tasks such as the creation of 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 probability of being at risk of being affected by cold waves for some households with certain characteristics, a full analysis was not performed, just an average estimation of marginal effects.
* Although experimental setting is robust, real-world model implementation is still key to close the gap between academia and practitioners. This paper aimed to provide guidelines and, to the best of our ability, shed light into the uncertainty embedded in practical implementations.
* The model can be further extended to consider more sophisticated predictors such as distance from households to principal 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 practical implementation of Machine Learning in terms of disaster risk reduction. Considering that humanitarian interventions operate with scarce resources and need to be optimized regardless of their localization or vulnerability condition, this paper sheds light into practical considerations of applied Machine Learning. Hence, it contributes to close the gap between academia and practitioners towards an improved disaster risk management system that is based on data.