UNIVERSIDAD NACIONAL DE SAN MARTIN

RCSI
https://revistas.unsm.edu.pe/index.php/rcsi
e-ISSN: 2709-992X

Original article
Artículo original
Jul-Dec, 2026

Explainable model of supervised and unsupervised learning to characterize the coverage of social programs in vulnerable territories of Peru

Modelo explicable de aprendizaje supervisado y no supervisado para caracterizar la cobertura de programas sociales en territorios vulnerables del Perú

Marin-Rodriguez, William J.[1]*
Vellón-Flores, Viviana I.[1]
Ausejo-Sánchez, José L.[1]
Solano-Armas, Timoteo[1]
Maguiña-Poma, Yolanda E.[1]
Samanamud-Malca, Sixto[1]

[1]Universidad Nacional José Faustino Sánchez Carrión, Huacho, Perú

Received: 03 Jan. 2026 | **Accepted:** 26 Jun. 2026 | **Published:** 20 Jul. 2026
Corresponding author*: wmarin@unjfsc.edu.pe
How to cite this article: Marin-Rodriguez, W. J., Vellón-Flores, V. I., Ausejo-Sánchez, J. L., Solano-Armas, T., Maguiña-Poma, Y. E. & Samanamud-Malca, S. (2026). Explainable model of supervised and unsupervised learning to characterize the coverage of social programs in vulnerable territories of Peru. *Revista Científica de Sistemas e Informática*, *6*(2), e1266. https://doi.org/10.51252/rcsi.v6i2.1266

ABSTRACT

This study developed and evaluated an explainable supervised and unsupervised learning approach to characterize social program coverage and identify historically vulnerable territories in Peru. A total of 35,952 district-month observations from 1,893 districts and 19 periods were integrated using MIDIS administrative data and poverty maps. Based on 13 coverage indicators, four clustering algorithms and three classifiers were compared through cross-validation and an independent test set. K-means with two clusters was selected as the main solution, achieving a silhouette score of 0.540 and identifying a profile associated with greater coverage and historical vulnerability. For classification, Elastic Net using only coverage indicators achieved an F1-score of 0.655 and, in the independent test set, a sensitivity of 0.829, an F1-score of 0.647, and a ROC-AUC of 0.871, with no significant differences compared with the other models. Explainability analysis highlighted Juntos coverage, estimated FONCODES users, and the number of programs present. The findings show that administrative records can support territorial prioritization when complemented with updated information and institutional review.

Keywords: clustering; social targeting; artificial intelligence; historical poverty; logistic regression

RESUMEN

Este estudio desarrolló y evaluó un enfoque explicable de aprendizaje supervisado y no supervisado para caracterizar la cobertura de programas sociales e identificar territorios históricamente vulnerables del Perú. Se integraron 35 952 observaciones distrito-mes, correspondientes a 1 893 distritos y 19 periodos, con información del MIDIS y mapas de pobreza. A partir de 13 indicadores se compararon cuatro algoritmos de agrupamiento y tres clasificadores mediante validación cruzada y prueba independiente. K-means con dos conglomerados fue seleccionado como solución principal, con una silueta de 0,540, e identificó un perfil de cobertura y vulnerabilidad histórica. En la clasificación, Elastic Net basado únicamente en indicadores de cobertura alcanzó un F1-score de 0,655 y, en la prueba independiente, una sensibilidad de 0,829, un F1-score de 0,647 y un ROC-AUC de 0,871, sin diferencias significativas frente a los demás modelos. La explicabilidad destacó la cobertura de Juntos, los usuarios de FONCODES y el número de programas presentes. Los resultados muestran que estos registros pueden apoyar la priorización territorial, complementados con información actualizada y revisión institucional.

Palabras clave: agrupamiento; focalización social; inteligencia artificial; pobreza histórica; regresión logística

1. INTRODUCTION

Poverty remains one of the main challenges to Peru's territorial development, despite the decline observed in recent years. According to the National Institute of Statistics and Informatics (INEI, 2026), monetary poverty affected 25.7% of the country's population in 2025; however, its incidence reached 35.5% in rural areas and exceeded 40% in departments such as Cajamarca and Loreto. These disparities show that improvements in aggregate indicators have not eliminated the territorial concentration of deprivation and that national averages may conceal markedly heterogeneous realities at the district level. In this context, identifying territories with higher levels of vulnerability is essential for guiding public interventions and allocating social protection resources with greater territorial relevance. Beyond its monetary dimension, poverty also encompasses deprivation related to income, food, basic services, and housing conditions; therefore, its analysis requires a multidimensional perspective (García et al., 2024).

Social programs are key instruments for protecting people living in poverty and expanding their access to basic services and transfers. However, their outcomes depend on targeting mechanisms, the characteristics of the population served, and the structural conditions of each territory (Borga & D'Ambrosio, 2021). In Peru, factors such as education, infrastructure, employment, and institutional constraints produce differentiated effects across urban and rural areas (Hinojosa Pérez et al., 2024; Maco, 2025; Paucara-Charca et al., 2024). Therefore, the presence of social programs and their territorial configuration must be analyzed jointly.

The availability of open administrative data provides an opportunity to advance this type of analysis. The Ministry of Development and Social Inclusion publishes district-level information on the coverage of Cuna Más, Juntos, FONCODES, Pensión 65, Wasi Mikuna, Contigo, and PAIS, expressed through the number of registered users, households, beneficiaries, service instances, institutions, facilities, and projects (MIDIS, 2026). However, the diversity of measurement units, population differences among districts, and coexistence of several programs within the same territory make these records difficult to interpret directly. Consequently, procedures are required to normalize the indicators and examine their relationships simultaneously, enabling administrative information to be transformed into understandable territorial profiles that are useful for decision-making.

In this context, machine learning has created new opportunities to analyze large volumes of social data and identify patterns that would be difficult to detect through conventional descriptive procedures. In Togo, combining mobile phone data with machine-learning models reduced exclusion errors by between 4% and 21% compared with the geographic targeting alternatives considered for distributing humanitarian assistance (Aiken et al., 2022). In Nigeria, high-resolution poverty maps constructed using geospatial information and machine learning reduced inclusion and exclusion errors in the territorial allocation of social transfers (Smythe & Blumenstock, 2022). Similarly, integrating administrative data, imagery, connectivity indicators, and other spatial data sources improved the identification of vulnerable households in urban settings where socioeconomic information was limited (Jung et al., 2026). These advances demonstrate the potential of computational models to complement traditional territorial assessment tools.

Unsupervised and supervised learning provide complementary capabilities for analyzing social program coverage. While clustering algorithms can identify territories with similar patterns

without predefined categories, classification models assess their ability to distinguish vulnerability conditions. However, achieving strong predictive performance is not sufficient in the social domain; it is also necessary to explain which variables influence the estimates and how they relate to the outcomes. The literature indicates that transparency and interpretability remain limited in several computational studies of poverty (Hall et al., 2022). Therefore, the application of artificial intelligence to public policy requires understandable and auditable models that support, rather than replace, the judgment of decision-makers (Lisboa et al., 2023; Papadakis et al., 2024).

Despite these advances, the integrated characterization of multiple social programs at the district level through supervised learning, unsupervised learning, and explainability techniques remains limited. In Peru, studies have focused primarily on specific programs, targeting constraints, or beneficiary classification, without jointly analyzing the territorial patterns of MIDIS services. Therefore, the objective of this study was to develop and evaluate an explainable analytical approach to characterize social program coverage and identify historically vulnerable territories. For this purpose, district-level administrative records were integrated, clustering and classification algorithms were compared, and explainability methods were applied. The proposed approach provides a reproducible methodology for transforming heterogeneous public data into useful evidence to support the territorial assessment and planning of social protection.

2. MATERIALS AND METHODS

2.1 Study design and scope

This was an applied study with a quantitative approach, a non-experimental, retrospective, and ecological design, and an exploratory-analytical scope. The unit of analysis was the district in Peru, identified using the six-digit Geographic Location Code (UBIGEO).

The initial database had a longitudinal district-month structure. For the machine-learning analysis, the monthly records were aggregated at the district level by calculating the temporal mean of the coverage variables, resulting in a single observation for each district.

2.2 Methodological procedure

The study integrated the preparation of official data sources, the construction of territorial indicators, and the analysis of social program coverage patterns. Subsequently, unsupervised and supervised learning techniques were applied to characterize the territories and classify their historical vulnerability. The analysis was complemented by statistical validation, explainability procedures, and the evaluation of classification errors. As shown in Figure 1, these activities were organized into a sequential workflow that maintained data traceability from data collection to model interpretation.

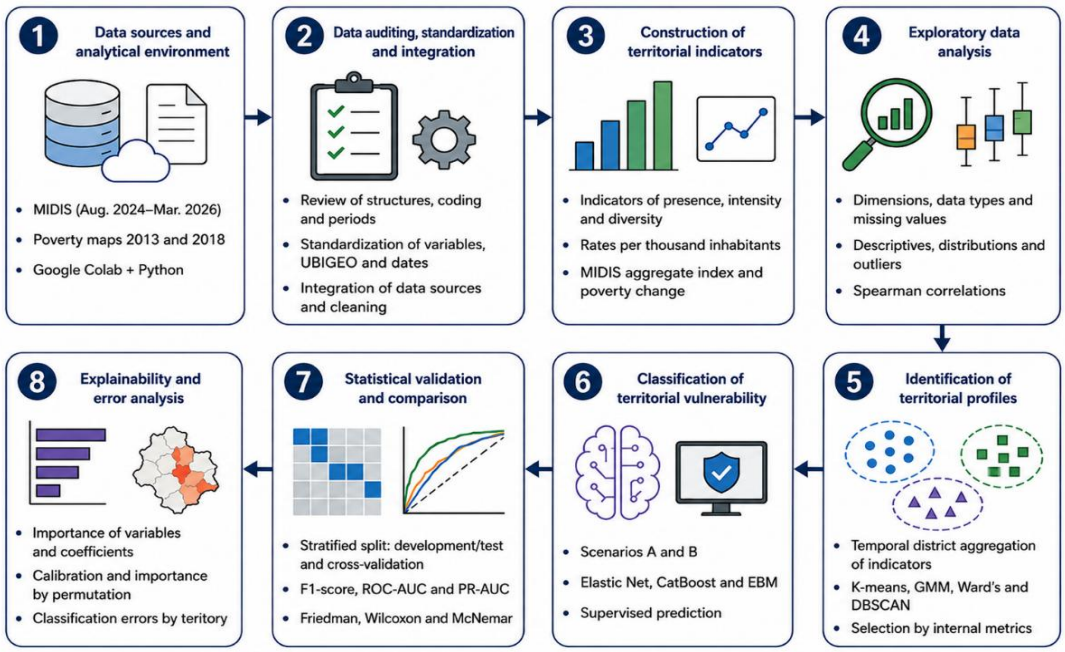

Figure 1. Methodological workflow of the study

Data sources and analytical environment

Two publicly available official data sources were used. The first consisted of 19 monthly CSV files published by the Ministry of Development and Social Inclusion, containing district-level coverage records from August 2024 to March 2026 (MIDIS, 2026). The November 2024 reporting period was excluded because of an inconsistency identified in the date field. The data covered the Cuna Más, Juntos, FONCODES, Pensión 65, Wasi Mikuna, Contigo, and PAIS programs and included records of users, households, beneficiaries, service instances, institutions, Tambos, and projects, according to the operational structure of each program.

The second source consisted of the 2013 and 2018 Provincial and District Poverty Maps, available in Excel format (INEI, 2015, 2020). The following variables were extracted from these files: UBIGEO, department, province, district, estimated population, lower and upper poverty bounds, robust group, and position in the district ranking.

Data processing was performed in Google Colaboratory using Python. pandas and NumPy were used for data manipulation and integration; Matplotlib for visualization; scikit-learn, CatBoost, and InterpretML for machine learning; and SciPy and Statsmodels for statistical analysis.

Data auditing, standardization, and integration

The files were audited based on their structure, encoding, delimiter, dimensions, available columns, missing values, and recorded periods. To ensure that they were read correctly, different encoding and separator combinations were evaluated, and the combination that successfully recovered the tabular structure was selected automatically (Van den Broeck et al., 2005).

Variable names were standardized by converting them to lowercase, removing accent marks and special characters, and replacing spaces with underscores. In addition, an equivalence mapping was implemented to identify changes in program names across periods, such as the transition from Qali Warma to Wasi Mikuna and the naming variants used for the National PAIS Program. The UBIGEO codes were cleaned by retaining only numeric characters and padding them to six digits.

Reporting dates were converted from the YYYYMMDD format, and the coverage variables were converted to numeric format (Harron et al., 2017).

After consolidating the monthly files, the uniqueness of each period–UBIGEO combination was verified. As a data-cleaning rule, when duplicate records were identified, the most recent available observation for each district and period was retained. In the poverty maps, the "Cdr01" worksheet was used for 2013 and the "Anexo1" worksheet for 2018. Headers, invalid codes, records without territorial identification, and observations that did not meet logical population and poverty-percentage criteria were removed.

Average district poverty was calculated as the midpoint of the reported interval:

$$Average\ poverty = \frac{Lower\ poverty\ bound + Upper\ poverty\ bound}{2}$$

The data sources were integrated using UBIGEO, while the MIDIS monthly panel was retained as the main data structure. Missing poverty values were not imputed. For the coverage variables, missing values were replaced with zero as an operational criterion, and an indicator of the original number of missing values was retained to preserve traceability. No temporal interpolation was applied.

Construction of territorial indicators

Indicators of the presence, intensity, and diversity of social program coverage were constructed from the original variables. For each program, a binary variable was generated with a value of one when at least one of its services had a recorded value greater than zero and a value of zero otherwise. The number of programs present and the number of coverage variables with positive values were also calculated for each district and period (Amarante et al., 2026).

To account for population differences among districts, ten coverage variables were expressed per 1,000 inhabitants:

$$Coverage\ rate_j = \frac{Program\ coverage_j}{District\ population} x1000$$

This normalization was applied to Cuna Más day-care services and family support; households enrolled in and receiving payments from Juntos; estimated FONCODES users; Pensión 65 users; children served by Wasi Mikuna; Contigo users; and service instances and beneficiaries of the National PAIS Program.

The ten rates were summed to construct an aggregate MIDIS coverage intensity index per 1,000 inhabitants. This indicator represented the combined magnitude of the administrative records rather than the number of unique beneficiaries, because the same individual could participate in more than one program or service modality (Ferreira et al., 2022). As a contextual variable, the change in district poverty was calculated as follows:

$$Poverty\ change = Average\ poverty\ in\ 2018 - Average\ poverty\ in\ 2013$$

For supervised classification, a district was operationally defined as a territory with high historical vulnerability when its average poverty rate in 2018 was equal to or greater than the 75th percentile of the district-level distribution. The remaining districts were assigned to the lower

relative vulnerability category. Poverty, ranking, and robust-group variables were retained solely to define and interpret the outcome and were not included as predictors.

Exploratory data analysis

The exploratory analysis included a review of the dataset dimensions, data types, missing values, duplicate records, temporal coverage, and information availability by district (Behrens, 1997). For the numerical variables, measures of central tendency, dispersion, and position were calculated, including the mean, median, standard deviation, minimum and maximum values, and the 1st, 5th, 25th, 75th, 95th, and 99th percentiles.

Potential outliers were identified using the interquartile range but were not removed at this stage. The distribution of the number of programs present, the aggregate coverage intensity index, the monthly evolution of the main services, and territorial differences across departments and districts were also examined.

Associations among the coverage indicators were assessed using Spearman's correlation coefficient because of the skewness of their distributions. Their bivariate relationships with average district poverty in 2018 were also analyzed. These associations were interpreted solely for descriptive purposes and not as causal relationships (Schober et al., 2018).

Identification of territorial profiles

For the unsupervised analysis, the monthly records were aggregated at the district level by calculating the temporal mean of 13 indicators: ten coverage rates per 1,000 inhabitants, the number of programs present, the number of coverage variables with positive values, and the aggregate MIDIS coverage intensity index. Missing values were replaced with zero. The variables were then transformed using log1p, winsorized between the 1st and 99th percentiles, and scaled using RobustScaler.

Four clustering algorithms were compared. K-means was evaluated using between two and ten clusters, with 100 initializations and a random seed of 42. The Gaussian mixture model was examined using between two and ten components, a full covariance matrix, 20 initializations, and the same random seed. Agglomerative hierarchical clustering was evaluated over the same range using Ward linkage. For DBSCAN, the eps values were obtained from the 60th to 95th percentiles of the nearest-neighbor distances, while min_samples was assigned values of 5, 13, 16, and 19 (Jain, 2010).

The internal quality of the solutions was evaluated using the silhouette coefficient and the Davies–Bouldin and Calinski–Harabasz indices. For K-means, inertia and the elbow method were also analyzed; for the Gaussian mixture model, the AIC and BIC criteria were assessed; and for DBSCAN, the proportion of observations classified as noise was examined. For the latter algorithm, the metrics were calculated using only observations not classified as noise, and solutions producing between two and ten clusters with no more than 30% noise were retained (Arbelaitz et al., 2013).

The final comparison included the K-means solutions determined by the elbow method and the highest silhouette coefficient, the Gaussian mixture model with the lowest BIC, the agglomerative clustering solution with the highest silhouette coefficient, and the best valid DBSCAN configuration. The following score was calculated for each candidate:

$$P_c = R_S + R_{DB} + R_{CH} + P_R$$

where R_S y R_{CH} represented the ranks assigned in descending order according to the silhouette and Calinski–Harabasz values, respectively, whereas R_{DB} corresponded to the rank assigned in ascending order according to the Davies–Bouldin value. The noise penalty (P_R) was zero for noise proportions of up to 10%, one point when the proportion exceeded 10%, and two points when it exceeded 20%. The solution with the lowest score was selected as the main solution.

Additionally, a four-cluster K-means solution was examined to obtain more detailed territorial profiles. Three-dimensional principal component analysis was used solely for graphical representation. Poverty variables were not included in cluster formation and were subsequently used to characterize the identified profiles.

Classification of territorial vulnerability

Supervised classification used high territorial vulnerability in 2018 as the target variable and the 13 social program coverage indicators as predictors. Before training, an audit was conducted to exclude variables associated with the definition of the target, including poverty, district ranking, robust groups, and vulnerability-derived variables. These variables were retained solely for the subsequent interpretation of the results.

Two scenarios were evaluated. **Scenario A** used only the 13 MIDIS coverage indicators. **Scenario B** additionally incorporated membership in the four clusters obtained through K-means, represented by three binary variables and one reference cluster. To prevent data leakage, preprocessing and K-means fitting were performed exclusively on the training data within each partition (Kapoor & Narayanan, 2023).

Three classifiers were compared in both scenarios:

Elastic Net logistic regression: SAGA solver, L1 ratio of 0.5, C=1, balanced class weights, a maximum of 5,000 iterations, and a random seed of 42 (Zou & Hastie, 2005).

CatBoost: Logloss loss function, AUC evaluation metric, 500 iterations, learning rate of 0.03, maximum depth of 4, L2 regularization of 5, automatic class balancing, and a random seed of 42.

Explainable Boosting Machine: five interaction terms, four outer bags, a maximum of 256 bins, a learning rate of 0.03, a maximum of 2,500 rounds, and parallel processing. Class imbalance was addressed using sample weights.

The estimated probabilities were converted into class labels using a threshold of 0.50.

Statistical validation and comparison

The district-level dataset was stratified into a 75% development set and a 25% independent test set using a random seed of 42. The latter remained isolated throughout model training, cross-validation, and model selection.

Stratified ten-fold cross-validation with random shuffling was applied to the development set, using the same folds for all models and scenarios. Within each fold, zero imputation, the log1p transformation, winsorization, robust scaling, and, in Scenario B, K-means fitting were performed exclusively on the training data. The resulting parameters were subsequently applied to the corresponding validation fold (Varma & Simon, 2006).

Performance was evaluated using accuracy, balanced accuracy, precision, sensitivity, specificity, F1-score, ROC-AUC, PR-AUC, confusion matrix, and training time. Because of class imbalance and

the interest in identifying vulnerable territories, the F1-score was established as the primary metric. For each model–scenario combination, the mean, standard deviation, standard error, and 95% confidence interval across the ten folds were calculated.

The distribution of the F1-score was examined using the Shapiro–Wilk test and quantile–quantile plots. Overall comparisons were performed using the Friedman test, and effect size was estimated using Kendall's W coefficient. As a complementary paired analysis, the Wilcoxon signed-rank test was applied with Holm correction for multiple comparisons and a significance level of 0.05 (García et al., 2010).

The final model was selected based on the lowest mean F1-score rank across the shared folds. The ranks obtained for PR-AUC, ROC-AUC, and balanced accuracy were considered complementary criteria. Each model was then retrained using the entire development set and evaluated once on the independent test set.

Predictions from the independent test set were compared using the exact McNemar test and a paired bootstrap procedure with 1,000 resamples to estimate differences in F1-score and their 95% confidence intervals. The significance values from both comparisons were adjusted using the Holm method.

Explainability and error analysis

Explainability was addressed through both global procedures and model-specific methods. Within the development set, correlations among predictors and variance inflation factors were examined to identify multicollinearity and guide the interpretation of the results.

For Elastic Net logistic regression, standardized coefficients, odds ratios, and approximate marginal effects on the probability of vulnerability were obtained. Their stability was assessed through 200 bootstrap resamples, from which 95% confidence intervals and the proportion of repetitions in which each coefficient retained its sign were calculated.

For CatBoost and the Explainable Boosting Machine, native feature-importance measures were extracted. In addition, permutation importance was calculated for all models using ten repetitions per predictor, with PR-AUC, ROC-AUC, and F1-score as evaluation criteria (Marcinkevičs & Vogt, 2023).

Probability calibration was assessed using the Brier score and calibration curves constructed with ten quantile-based bins (Van Calster et al., 2019). Finally, predictions were classified as true positives, true negatives, false positives, and false negatives. Errors were analyzed both overall and by department and territorial profile; these variables were used solely for interpretive purposes and were not included in model training.

3. RESULTS

3.1 Construction of the territorial dataset

The integration of MIDIS administrative records with the district poverty maps produced a dataset comprising 35,952 district-month observations from 1,893 districts and 19 periods, with no duplicates in the period–UBIGEO combination. For the supervised analysis, 1,872 districts with valid poverty information for 2018 were retained. Average district poverty was 43.95% in 2013 and 34.16% in 2018, while the average change among districts with available information for both

years was −9.31 percentage points. High historical vulnerability was operationally defined using the 75th percentile of the 2018 district poverty distribution, equivalent to 46.80%, as summarized in Table 1.

Table 1. General characteristics of the territorial dataset

Indicator	Result
District-month records	35 952
Districts analyzed	1 893
Available periods	19
Districts with 2018 poverty data	1 872
Districts without 2018 poverty data	21
Districts without 2013 poverty data	90
Average district poverty in 2013	43.95 %
Average district poverty in 2018	34.16 %
Average change, 2013–2018	−9.31 percentage points
Vulnerability cutoff	46.7974 %
Vulnerable districts	468
Districts with lower relative vulnerability	1 404

3.2 Social program coverage

The recorded presence of social programs varied considerably across the districts and periods analyzed. Pensión 65 and Juntos showed the broadest coverage, with positive records in more than 99% of the observations, whereas PAIS and FONCODES exhibited a more concentrated presence, as detailed in Table 2.

Table 2. Presence of social programs in the analytical dataset

Program	Records with presence	Percentage of records	Districts with presence
Pensión 65	35 923	99.92 %	1 891
Juntos	35 789	99.55 %	1 888
Wasi Mikuna	35 127	97.71 %	1 891
Contigo	34 928	97.15 %	1 848
Cuna Más	28 642	79.67 %	1 545
FONCODES	14 950	41.58 %	918
PAIS	7 679	21.36 %	407

At the departmental level, Huancavelica recorded the highest average coverage intensity index, with 710.96 units per 1,000 inhabitants, followed by Ayacucho, Apurímac, and Huánuco, with 683.56, 653.10, and 594.46 units, respectively. The lowest values were recorded in Callao, Ica, and Lambayeque, with 54.68, 144.76, and 179.07 units per 1,000 inhabitants, respectively.

The monthly evolution of the index also showed marked variations, with a minimum of 319.13 units per 1,000 inhabitants in February 2025 and a maximum of 544.34 in April of the same year. These fluctuations were interpreted cautiously, as they may have been related to differences in the updating, accumulation, or reporting of administrative records, as shown in Figure 2.

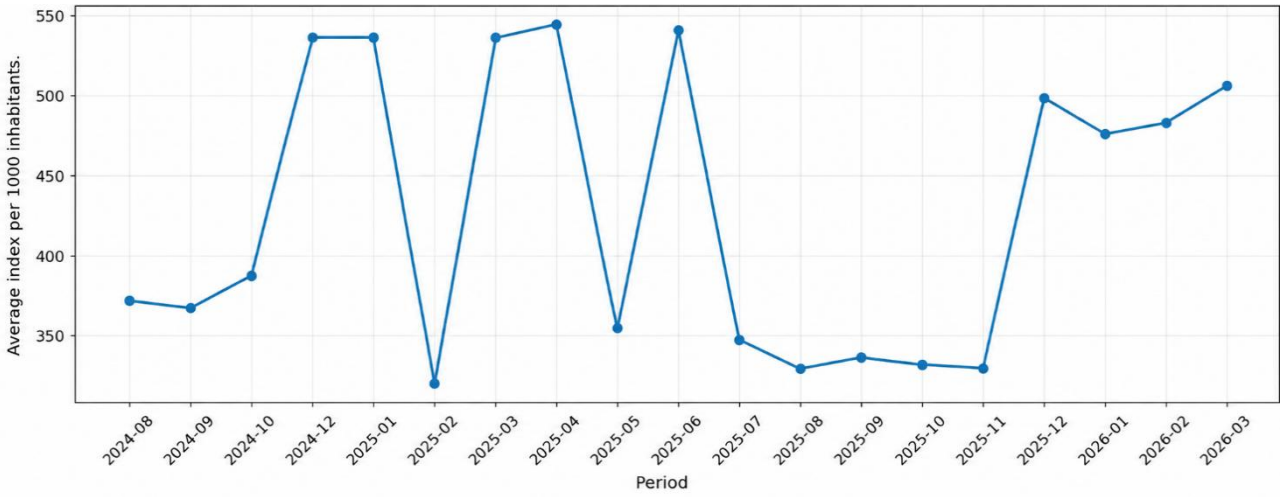

Figure 2. Monthly evolution of the MIDIS coverage intensity index

3.3 Social program coverage and territorial vulnerability

Districts classified as having high historical vulnerability exhibited greater intensity and diversity of social program coverage. The aggregate MIDIS coverage intensity index reached 621.10 units per 1,000 inhabitants, compared with 369.76 in districts with lower relative vulnerability. Likewise, the average number of programs present was 6.05 and 5.15, respectively.

As shown in Table 3, district poverty in 2018 was positively associated with most coverage indicators. The highest coefficients corresponded to households receiving payments from and enrolled in Juntos, with values of 0.672 and 0.664, respectively, followed by the aggregate MIDIS coverage intensity index, with 0.599. In contrast, Cuna Más day-care services showed the only negative association, with a coefficient of −0.261. These results primarily reflected the concentration of social programs in territories with greater historical vulnerability rather than a causal effect of coverage on poverty.

Table 3. Correlations between coverage indicators and district poverty in 2018

Variable	Spearman correlation
Households receiving payments from Juntos	0.672
Households enrolled in Juntos	0.664
Aggregate MIDIS coverage index	0.599
Cuna Más family support	0.565
Number of programs present	0.549
Number of coverage variables present	0.480
Contigo users	0.469
Pensión 65 users	0.448
PAIS beneficiaries	0.301
Estimated FONCODES users	0.286
Children served by Wasi Mikuna	0.256
Cuna Más day-care services	−0.261

3.4 Territorial profiles of social program coverage

The comparison of the algorithms showed that DBSCAN achieved the highest silhouette coefficient, at 0.606; however, it classified 27.05% of the districts as noise. In contrast, K-means with two clusters obtained a silhouette coefficient of 0.540, a Davies–Bouldin index of 0.696, and a Calinski–Harabasz index of 1,946.77. Although its silhouette value was lower than that of DBSCAN, this solution assigned all districts to a profile and provided a better balance among

separation, compactness, and interpretability. Therefore, K-means with two clusters was selected as the main solution for the unsupervised analysis.

The first three principal components jointly explained 85.13% of the variability in the coverage indicators: 56.14% in the first component, 20.97% in the second, and 8.01% in the third. The distribution of the districts within this three-dimensional space showed a consistent distinction between the two territorial profiles identified, as illustrated in Figure 3.

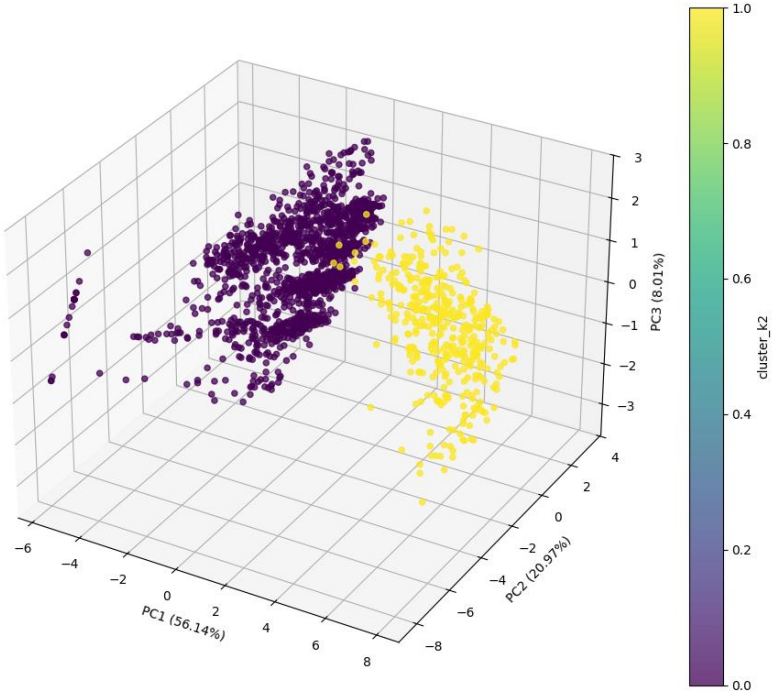

Figure 1. Three-dimensional distribution of the territorial profiles obtained using K-means (k=2)

The main solution identified one profile characterized by lower coverage intensity and diversity, comprising 1,497 districts, and another characterized by higher intensity and diversity, comprising 396 districts. As shown in Table 4, the second profile recorded a MIDIS coverage intensity index more than twice that observed in the first profile, together with higher average numbers of programs and coverage variables with positive values. It also comprised less populated districts with higher historical poverty levels.

Table 4. Characteristics of the territorial profiles obtained using K-means

Indicator	Profile 1: lower vulnerability and coverage	Profile 2: higher vulnerability and coverage
Districts	1 497	396
Percentage of districts	79.08 %	20.92 %
Average poverty in 2018	31.38 %	44.53 %
Average poverty in 2013	40.35 %	57.04 %
Poverty change	−8.46 puntos	−12.40 puntos
Average population in 2020	20 155	7 254
MIDIS coverage index	347.88	729.89
Programs present	5.05	6.59
Coverage variables present	7.46	11.28

The 13.16-percentage-point difference in average poverty in 2018 between the two profiles was particularly relevant because this variable was not used to form the clusters. Consequently, administrative coverage patterns made it possible to distinguish territories associated with different levels of historical vulnerability. Nevertheless, the greater coverage observed in the

second profile should not be interpreted as a causal effect of social programs on poverty, but rather as an expression of their targeting toward historically more disadvantaged districts.

Additionally, the four-cluster K-means solution achieved a silhouette coefficient of 0.362 and made it possible to identify more specific territorial patterns. However, because of its lower internal separation, it was retained only as a complementary segmentation to evaluate its contribution in one of the supervised classification scenarios.

3.5 Performance of the supervised models

Territorial vulnerability was classified under two scenarios. Scenario A used the 13 social program coverage indicators, whereas Scenario B additionally incorporated three binary variables derived from the four-cluster K-means segmentation. As shown in Table 5, Elastic Net logistic regression under Scenario A achieved the highest mean F1-score in cross-validation, at 0.655, and the lowest mean rank, at 2.30. The remaining combinations obtained F1-scores ranging from 0.645 to 0.654, indicating only small numerical differences.

The Friedman test did not identify statistically significant overall differences among the models and scenarios, $x^2(5) = 10{,}310$; $p = 0.0669$, with a Kendall's W coefficient of 0.206. Pairwise comparisons using the Wilcoxon signed-rank test with Holm correction were also not significant. Consequently, Elastic Net under Scenario A was selected because it achieved the best mean rank and the highest F1-score in cross-validation, although this did not imply statistical superiority over the other models.

Table 5. Model performance in cross-validation

Scenario and model	F1-score	Balanced accuracy	Sensitivity	ROC-AUC	PR-AUC	Mean rank
Elastic Net, Scenario A	0.655	0.796	0.818	0.878	0.666	2.30
Elastic Net, Scenario B	0.654	0.795	0.821	0.877	0.666	2.80
EBM, Scenario A	0.652	0.792	0.807	0.881	0.707	3.40
CatBoost, Scenario B	0.653	0.785	0.752	0.874	0.707	3.80
EBM, Scenario B	0.645	0.787	0.801	0.879	0.702	4.25
CatBoost, Scenario A	0.649	0.782	0.752	0.875	0.718	4.45

In the independent test set, whose results are presented in Table 6, Elastic Net under Scenario B achieved the highest numerical F1-score, at 0.656, whereas CatBoost under Scenario A obtained the highest PR-AUC, at 0.683. The selected model, Elastic Net under Scenario A, maintained competitive performance, with an F1-score of 0.647, sensitivity of 0.829, balanced accuracy of 0.792, and ROC-AUC of 0.871. Incorporating the clusters did not produce consistent improvements across the three classifiers; therefore, their usefulness was mainly limited to territorial characterization. The McNemar tests and paired bootstrap comparisons, both adjusted using the Holm method, likewise found no significant differences among the model predictions.

Table 6. Model performance in the independent test set

Model	Accuracy	Balanced accuracy	Precision	Sensitivity	F1-score	ROC-AUC	PR-AUC
Elastic Net, Scenario B	0.778	0.801	0.535	0.846	0.656	0.871	0.638
EBM, Scenario A	0.782	0.795	0.542	0.821	0.653	0.862	0.643
EBM, Scenario B	0.782	0.792	0.543	0.812	0.651	0.857	0.640
Elastic Net, Scenario A	0.774	0.792	0.530	0.829	0.647	0.871	0.639
CatBoost, Scenario B	0.786	0.783	0.552	0.778	0.645	0.864	0.674
CatBoost, Scenario A	0.780	0.776	0.542	0.769	0.636	0.867	0.683

3.6 Explainability, calibration, and error analysis

Explainability and error analysis were conducted for the Elastic Net logistic regression model under Scenario A, which had previously been selected through cross-validation. As shown in Figure 4, the model correctly classified 97 of the 117 vulnerable districts in the independent test set and failed to detect 20. Among the 351 districts with lower relative vulnerability, it correctly identified 265 and generated 86 false positives. The sensitivity of 0.829 showed that the model identified approximately 83% of the vulnerable territories; however, the precision of 0.530 indicated that some of the flagged districts would require additional review.

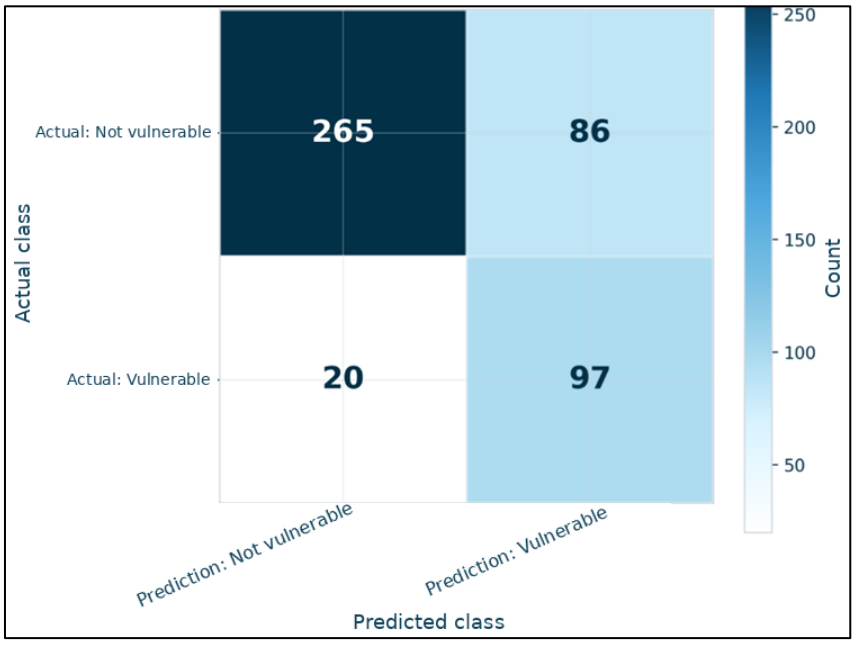

Figure 2. Confusion matrix of the Elastic Net model in the independent test set

The overall contribution of the predictors was examined through permutation importance, using PR-AUC as the reference metric. As shown in Figure 5, households receiving payments from Juntos had the highest importance, with a mean decrease of 0.152 when their values were permuted. They were followed by estimated FONCODES users, with 0.142, and the number of programs present, with 0.102. These indicators constituted the main administrative signals used by the model to distinguish historically vulnerable territories.

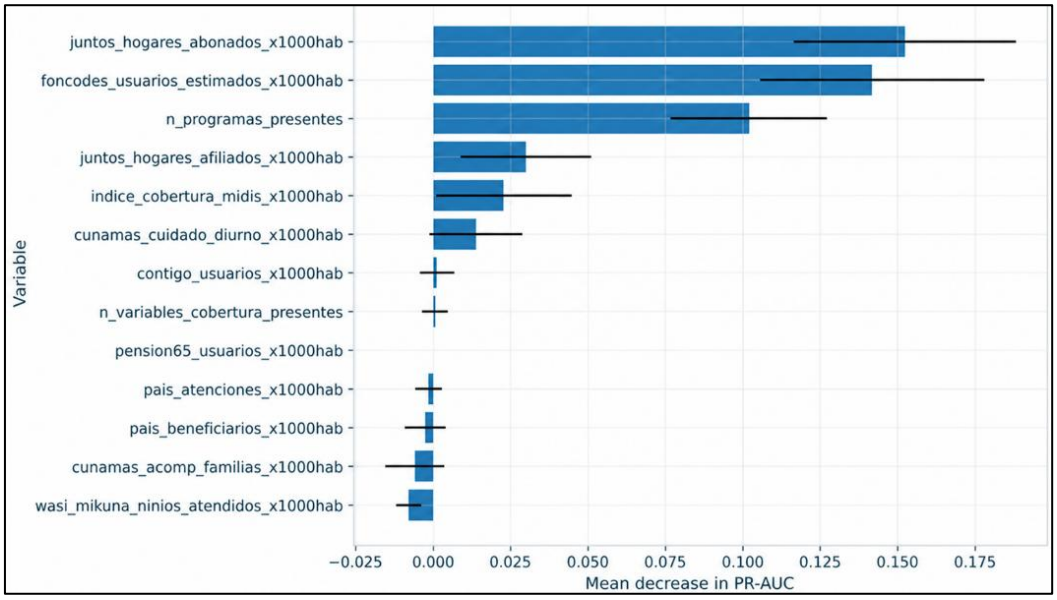

Figure 3. Permutation importance of the Elastic Net model variables with respect to PR-AUC

The stability analysis based on 200 bootstrap resamples showed consistently positive signs for households receiving payments from and enrolled in Juntos, estimated FONCODES users, and the number of programs present. In contrast, Cuna Más day-care services showed a consistently negative sign. These relationships reflected internal associations within the model rather than causal effects of social programs on poverty.

High multicollinearity was also identified among operational variables belonging to the same program. Households enrolled in and receiving payments from Juntos had variance inflation factors of 375.05 and 372.76, respectively, whereas PAIS service instances and beneficiaries reached values of 218.83 and 212.18. Although Elastic Net regularization reduced the instability associated with these dependencies, the individual coefficients were interpreted jointly with the bootstrap results and permutation importance.

Regarding calibration, CatBoost achieved the lowest Brier scores in both scenarios, at 0.143, followed by Elastic Net, at 0.152, and EBM, with values close to 0.154. Therefore, CatBoost produced slightly better-calibrated probabilities, although it did not achieve the lowest mean rank or the highest F1-score during cross-validation. Overall, Elastic Net under Scenario A maintained an appropriate balance among discriminative ability, sensitivity, and interpretability, without showing statistically significant superiority over the other configurations evaluated.

3.7 Discusión

The results showed that administrative coverage records contain useful information for distinguishing territories with different levels of historical vulnerability. Districts with higher poverty levels exhibited greater intensity and diversity of social interventions. This pattern should be interpreted as an expression of territorial targeting rather than as an effect of social programs on poverty. Altındağ et al. (2021) demonstrated that administrative data can support social targeting when subjected to appropriate design and validation procedures.

K-means clustering identified two distinct profiles, even though poverty was not included in cluster formation. The profile with greater coverage also exhibited higher historical poverty levels and a lower average population. This finding indicates that the operational patterns of the programs partially captured the territorial structure of vulnerability. Nevertheless, these profiles

do not replace an up-to-date, multidimensional assessment, because the territories prioritized may vary depending on the welfare criterion applied (Henderson & Follett, 2022).

Elastic Net logistic regression based exclusively on coverage indicators was selected because of its cross-validation performance and maintained competitive results in the independent test set. The absence of significant differences compared with CatBoost and EBM showed that greater algorithmic complexity did not guarantee better results. Li et al. (2022) likewise found that incorporating additional variables or more complex models does not necessarily improve poverty classification. Furthermore, the preference for a parsimonious alternative is consistent with Asri et al. (2022), who highlighted the operational usefulness of simpler and more verifiable targeting criteria.

The model's high sensitivity demonstrated a strong ability to identify vulnerable districts, although its moderate precision generated a considerable number of false positives. This behavior may be useful when the objective is to reduce exclusion errors, provided that the predictions are institutionally reviewed. Poulin et al. (2022) noted that machine-learning-based targeting methods should be assessed not only in terms of performance, but also according to their costs, transparency, and operational consequences. Therefore, the model should be used to support territorial prioritization rather than as an automated mechanism for allocating benefits.

The explainability analysis identified Juntos coverage, estimated FONCODES users, and program diversity as the main signals. However, the high degree of multicollinearity prevented these variables from being interpreted as independent or causal effects. Although CatBoost showed slightly better calibration, Elastic Net provided an appropriate balance among performance, sensitivity, and interpretability. In line with Amarasinghe et al. (2023), explanations applied to public policy should respond to specific decision-making needs and be complemented by up-to-date information and expert judgment.

CONCLUSIONS

The integration of MIDIS administrative records with district poverty maps made it possible to identify territorial profiles associated with historical vulnerability and to develop an explainable model for their classification. K-means distinguished two differentiated profiles, whereas Elastic Net based exclusively on coverage indicators was selected because of its balance among performance, sensitivity, and interpretability, without showing significant differences compared with more complex models. Incorporating the clusters did not consistently improve prediction; therefore, their main contribution was descriptive. Consequently, the model can support territorial prioritization and review, provided that it is complemented with up-to-date socioeconomic information and is not used as an automated mechanism for allocating benefits.

FINANCING

The authors received no funding to conduct this study.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest related to the subject matter of this work.

AUTHORSHIP CONTRIBUTION

Conceptualization: Marin-Rodriguez, W. J.; Vellón-Flores, V. I. Data curation: Marin-Rodriguez, W. J.; Ausejo-Sánchez, J. L.; Solano-Armas, T. Formal analysis: Marin-Rodriguez, W. J.; Vellón-Flores, V. I.; Ausejo-Sánchez, J. L. Investigation: Marin-Rodriguez, W. J.; Vellón-Flores, V. I.; Solano-Armas, T.; Maguiña-Poma, Y. E. Methodology: Marin-Rodriguez, W. J.; Vellón-Flores, V. I.; Samanamud-Malca, S. Project administration: Marin-Rodriguez, W. J.; Samanamud-Malca, S. Software: Marin-Rodriguez, W. J.; Ausejo-Sánchez, J. L. Supervision: Maguiña-Poma, Y. E.; Samanamud-Malca, S. Validation: Vellón-Flores, V. I.; Solano-Armas, T.; Maguiña-Poma, Y. E. Visualization: Marin-Rodriguez, W. J.; Ausejo-Sánchez, J. L. Writing – original draft: Marin-Rodriguez, W. J.; Vellón-Flores, V. I. Writing – review and editing: Marin-Rodriguez, W. J.; Vellón-Flores, V. I.; Ausejo-Sánchez, J. L.; Solano-Armas, T.; Maguiña-Poma, Y. E.; Samanamud-Malca, S.

REFERENCES

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. E. (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature*, *603*(7903), 864–870. https://doi.org/10.1038/s41586-022-04484-9

Altındağ, O., O'Connell, S. D., Şaşmaz, A., Balcıoğlu, Z., Cadoni, P., Jerneck, M., & Foong, A. K. (2021). Targeting humanitarian aid using administrative data: Model design and validation. *Journal of Development Economics*, *148*, 102564. https://doi.org/10.1016/j.jdeveco.2020.102564

Amarante, V., Rossel, C., & Sánchez Laguardia, G. (2026). Social Protection Coverage in Latin America: Aggregate and Individual-level Transformations. *Journal of Human Development and Capabilities*, 1–27. https://doi.org/10.1080/19452829.2025.2610195

Amarasinghe, K., Rodolfa, K. T., Lamba, H., & Ghani, R. (2023). Explainable machine learning for public policy: Use cases, gaps, and research directions. *Data & Policy*, *5*, e5. https://doi.org/10.1017/dap.2023.2

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, *46*(1), 243–256. https://doi.org/10.1016/j.patcog.2012.07.021

Asri, V., Michaelowa, K., Panda, S., & Paul, S. B. (2022). The pursuit of simplicity: Can simplifying eligibility criteria improve social pension targeting? *Journal of Economic Behavior & Organization*, *200*, 820–846. https://doi.org/10.1016/j.jebo.2022.06.003

Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods*, *2*(2), 131–160. https://doi.org/10.1037/1082-989X.2.2.131

Borga, L. G., & D'Ambrosio, C. (2021). Social protection and multidimensional poverty: Lessons from Ethiopia, India and Peru. *World Development*, *147*, 105634. https://doi.org/10.1016/j.worlddev.2021.105634

Ferreira, L. Z., Utazi, C. E., Huicho, L., Nilsen, K., Hartwig, F. P., Tatem, A. J., & Barros, A. J. D. (2022). Geographic inequalities in health intervention coverage – mapping the composite coverage index in Peru using geospatial modelling. *BMC Public Health*, *22*(1), 2104. https://doi.org/10.1186/s12889-022-14371-7

García, S., Fernández, A., Luengo, J., & Herrera, F. (2010). Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences*, *180*(10), 2044–2064. https://doi.org/10.1016/j.ins.2009.12.010

García, V. H. B., Martínez, F. de M. G., Llamas Félix, B. I., & Esparza, R. M. V. (2024). Medición de la pobreza multidimensional en México mediante un análisis bibliométrico y de ecuaciones estructurales. *Iberoamerican Journal of Science Measurement and Communication*, *4*(2), 1–23. https://doi.org/10.47909/ijsmc.1354

Hall, O., Ohlsson, M., & Rögnvaldsson, T. (2022). A review of explainable AI in the satellite data, deep machine learning, and human poverty domain. *Patterns*, *3*(10), 100600. https://doi.org/10.1016/j.patter.2022.100600

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L., & Goldstein, H. (2017). Challenges in administrative data linkage for research. *Big Data & Society*, *4*(2), 205395171774567. https://doi.org/10.1177/2053951717745678

Henderson, H., & Follett, L. (2022). Targeting social safety net programs on human capabilities. *World Development*, *151*, 105741. https://doi.org/10.1016/j.worlddev.2021.105741

Hinojosa Pérez, J. A., Avalos, H. R. B., Salazar, I. Y. V., & Carrasco Mamani, S. C. (2024). Social Programs and Socioeconomic Variables: Their Impact on Peruvian Regional Poverty (2013–2022). *Economies*, *12*(8), 197. https://doi.org/10.3390/economies12080197

INEI. (2015). *Mapa de Pobreza Provincial y Distrital 2013*. Gob.pe. https://www.gob.pe/institucion/inei/informes-publicaciones/3204925-mapa-de-pobreza-provincial-y-distrital-2013

INEI. (2020). *Mapa de Pobreza Provincial y Distrital 2018*. Gob.pe. https://www.gob.pe/institucion/inei/informes-publicaciones/3204872-mapa-de-pobreza-provincial-y-distrital-2018

INEI. (2026). *INEI: Pobreza monetaria alcanzó al 25,7% de la población en el Perú durante el año 2025*. Gob.pe. https://www.gob.pe/institucion/inei/noticias/1387366-inei-pobreza-monetaria-alcanzo-al-25-7-de-la-poblacion-en-el-peru-durante-el-ano-2025

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Jung, W., Benotsmane, R., Stoeffler, Q., Kim, A. H., Ghadimi, S., Hosseini, M., Ntarlagiannis, D., Ammari, T., Lu, Y., & Steiner, J. (2026). Contextualized poverty targeting with multimodal spatial data and machine learning in Brazzaville, Congo. *Cities*, *170*, 106429. https://doi.org/10.1016/j.cities.2025.106429

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, *4*(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

Li, Q., Yu, S., Échevin, D., & Fan, M. (2022). Is poverty predictable with machine learning? A study of DHS data from Kyrgyzstan. *Socio-Economic Planning Sciences*, *81*, 101195. https://doi.org/10.1016/j.seps.2021.101195

Lisboa, P. J. G., Saralajew, S., Vellido, A., Fernández-Domenech, R., & Villmann, T. (2023). The coming of age of interpretable and explainable machine learning models. *Neurocomputing*, *535*, 25–39. https://doi.org/10.1016/j.neucom.2023.02.040

Maco, V. (2025). Lock-in from within: Challenges to expanding cash transfer programs in Peru. *Journal of International and Comparative Social Policy*, *41*(3), 335–349. https://doi.org/10.1017/ics.2025.10066

Marcinkevičs, R., & Vogt, J. E. (2023). Interpretable and explainable machine learning: A methods-centric overview with concrete examples. *WIREs Data Mining and Knowledge Discovery*, *13*(3). https://doi.org/10.1002/widm.1493

MIDIS. (2026). *Cobertura de los Programas Sociales adscritos al MIDIS- [Ministerio de Desarrollo e*

Inclusión Social - MIDIS]. Gob.pe. https://www.datosabiertos.gob.pe/dataset/cobertura-de-los-programas-sociales-adscritos-al-midis-ministerio-de-desarrollo-e-inclusión

Papadakis, T., Christou, I. T., Ipektsidis, C., Soldatos, J., & Amicone, A. (2024). Explainable and transparent artificial intelligence for public policymaking. *Data & Policy*, *6*, e10. https://doi.org/10.1017/dap.2024.3

Paucara-Charca, R., Almanza Huamán, L. M., Szczcpansky Grobas, D., Landa Almanza, A. I., & Muñiz Laura, K. A. (2024). La eficacia de los efectos políticos en el desarrollo alcance y resultados de las políticas regionales orientadas para mitigar el impacto de la explotación laboral en menores de edad en especial en el sector minero de La Pampa Madre de Dios 2018. *Revista Amazónica De Ciencias Sociales*, *2*(2), 12–21. https://doi.org/10.55873/racs.v2i2.262

Poulin, C., Trimmer, J., Press-Williams, J., Yachori, B., Khush, R., Peletz, R., & Delaire, C. (2022). Performance of a novel machine learning-based proxy means test in comparison to other methods for targeting pro-poor water subsidies in Ghana. *Development Engineering*, *7*, 100098. https://doi.org/10.1016/j.deveng.2022.100098

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Smythe, I. S., & Blumenstock, J. E. (2022). Geographic microtargeting of social assistance with high-resolution poverty maps. *Proceedings of the National Academy of Sciences*, *119*(32). https://doi.org/10.1073/pnas.2120025119

Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: the Achilles heel of predictive analytics. *BMC Medicine*, *17*(1), 230. https://doi.org/10.1186/s12916-019-1466-7

Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R., & Herbst, K. (2005). Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. *PLoS Medicine*, *2*(10), e267. https://doi.org/10.1371/journal.pmed.0020267

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(1), 91. https://doi.org/10.1186/1471-2105-7-91

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