BAREKENG: Journal of Mathematics and Its Applications
September 2026 Volume 20 Issue 3 Page 2601-2612
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss3pp2601-2612

IMPLEMENTATION OF CUCKOO SEARCH-BASED ENSEMBLE VARIABLE IMPORTANCE IN THE CLASSIFICATION OF NON-CASH FOOD ASSISTANCE (BPNT) RECIPIENTS IN WEST JAVA

Indra Mahib Zuhair Riyanto [1*], Laras Suprapti [2], Salsabila Fayiza [3], Elke Frida Rahmawati [4], Farid Yafi Suwandi [5], Sachnaz Desta Oktarina [6], Rahma Anisa [7]

[1,2,3,4,5,6,7] *Statistics and Data Science Study Program, School of Data Science, Mathematics, and Informatics, IPB University*
Jln. Raya Dramaga, Bogor, 16680, Indonesia

*Corresponding author's e-mail: * **indramahibzuhair@apps.ipb.ac.id***

Article Info

Article History:

Received: 22nd September 2025
Revised: 6th January 2026
Accepted: 17th March 2026
Published: 8th April 2026

Keywords:

Cuckoo search;
Ensemble Variable Importance (EVI);
Machine Learning.

ABSTRACT

The BPNT program is a government initiative to efficiently distribute social assistance to poor households. However, the challenge of achieving accurate recipient identification remains a major obstacle. This research aims to build a classification model for BPNT recipients in West Java using machine learning methods (Random Forest, XGBoost, CatBoost, and LightGBM) and a Cuckoo Search-Based Ensemble Variable Importance (EVI) approach to identify which predictors most strongly affect classification. Class imbalance in the response data was addressed through weighting during model training, and performance was evaluated using balanced accuracy through 10-fold cross-validation. Although all models performed well, the variable importance results varied across models. Using the Random-Key Cuckoo Search algorithm, an EVI ranking was generated that integrated VI rankings from each model, achieving a minimum Spearman correlation of 0.6538. The results show that roof quality, living status, calorie consumption, and per capita expenditure are the main indicators for classifying BPNT recipients. This approach shows great potential to improve modeling interpretability and to provide stronger data-driven support for social policy-making.

How to cite this article:

I. M. Z. Riyanto, L. Suprapti, S. Fayiza, E. F. Rahmawati, F. Y. Suwandi, S. D. Oktarina and R. Anisa., "IMPLEMENTATION OF CUCKOO SEARCH-BASED ENSEMBLE VARIABLE IMPORTANCE IN THE CLASSIFICATION OF NON-CASH FOOD ASSISTANCE (BPNT) RECIPIENTS IN WEST JAVA", *BAREKENG: J. Math. & App.*, vol. 20, no. 3, pp. 2601-2612, Sep, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The Non-Cash Food Assistance Program, or in Bahasa known as Bantuan Pangan Non-Tunai (BPNT), is a development of the previous program, namely Rice for Prosperous Families, which began to be implemented gradually in 2017. The main objective of this program is to improve the efficiency of social assistance distribution through a non-cash system that allows recipient families to purchase food directly. The government is trying to reduce the burden of spending on the poor while encouraging greater financial literacy among lower-income groups. This program also aims to encourage balanced nutrition for each household [1].

The implementation of BPNT in West Java is expected to make a positive contribution to beneficiaries' food security. For example, studies in Bogor City show that although BPNT increases household purchasing capacity, its effectiveness remains closely tied to beneficiaries' understanding of program mechanisms and access to supporting infrastructure [2]. Despite these benefits, some challenges remain, particularly inaccurate beneficiary targeting. Misclassification can create social friction, undermine program credibility, and weaken the effectiveness of resource allocation.

These challenges highlight the need for more robust, data-driven tools to support program evaluation and targeting accuracy. Machine learning algorithms offer a solution for efficiently managing large amounts of diverse data and producing accurate predictions. The application of machine learning algorithms such as Random Forest (RF), XGBoost, CatBoost, and LightGBM has been proven to handle complex data characteristics and achieve good predictive performance. The advantage of this method lies in the model's algorithm, which combines multiple predictions to produce a more stable classification and reduce the risk of overfitting [3].

The biggest challenge in machine learning-based modeling is understanding how each variable affects the classification result. The concept of variable importance (VI) has been widely used by researchers to identify which predictors have a strong impact on classification. Unfortunately, different machine learning algorithms can yield different VIs even when the predictors used are identical. In addition to the previous reason, there are many methods for producing VI, such as Permutation Variable Importance and Shapley Additive Explanation. There is a need for a method that can unify variable importances to better understand the effect of each predictor in classification when different ML or VI algorithms are used.

Research by [4] developed an ensemble variable importance (EVI) technique to unify various VI rankings produced by four different ML algorithms. First, the ML model is built to classify the Food Insecurity Experience Scale with 24 predictors using Random Forest, XGBoost, Neural Network, and Support Vector Machine. Each VI ranking from each model is unified with Cuckoo Search optimization. The EVI results show strong Spearman correlations with the individual variable importance rankings from each model, with coefficients of 0.952, 0.923, 0.941, and 0.942, respectively. The results show EVI's ability to produce stable, interpretable, unified rankings. This methodological framework can also be applied to the BPNT case to systematically identify the key household characteristics that drive program eligibility and classification outcomes.

Despite these advances, several research gaps remain. First, existing studies have largely focused on the classification performance of Indonesia's BPNT program using machine learning, while rarely interpreting results using variable importance (VI) methods, as in [5]. The application of ensemble VI methods for evaluating social assistance targeting can be beneficial because they are constructed from multiple VI models and take into account various algorithms. Second, prior work has rarely tested EVI on real-world, heterogeneous household survey datasets that exhibit data imbalance, regional heterogeneity, and mixed data types, except for [4], which used SUSENAS data for food insecurity classification.

Addressing these gaps, this study develops a machine-learning-based classification model and applies a Cuckoo Search–based Ensemble Variable Importance approach to unify VI rankings across multiple algorithms. The aim is to identify the most influential predictors of BPNT recipient status and to provide a more data-driven foundation for program evaluation. This research offers a novel methodological contribution and generates actionable insights by integrating advanced ML techniques with an optimization-based VI framework. The end result can support more accurate and equitable targeting in Indonesia's food assistance programs.

2. RESEARCH METHODS

2.1 Data Sources

The data used in this research are secondary data obtained from the National Socio-Economic Survey (SUSENAS 2023), conducted by the Central Statistics Bureau (BPS) of West Java Province. The variables used in this research are related to non-cash food assistance (BPNT). The dataset consists of 14 variables and 25,791 observations. The variables used pertain to food security. A detailed description is given in Table 1 below:

Table 1. The Variables Used

Code	Name of Variable	Type
Y	Non-cash food assistance (BPNT) recipient status	Categorical
X_1	Food insecurity concern status	Categorical
X_2	Urban/Rural area	Categorical
X_3	Regency/City area	Categorical
X_4	Homeownership status	Categorical
X_5	Roof quality	Categorical
X_6	Cash assistance (BLT)	Categorical
X_7	Floor area	Numerical
X_8	Expenditure per capita	Numerical
X_9	Average calorie consumption per capita	Numerical
X_{10}	Water source quality	Categorical
X_{11}	Toilet facility quality	Categorical
X_{12}	Credit usage status	Categorical
X_{13}	Household receiving family hope program (PKH) or prosperous family card (KKS)	Categorical

Data source: *BPS 2023*

2.2 Machine Learning Models

Machine learning models are statistical methods for building predictive models from data. Among the most widely used are decision tree-based models and ensemble learning techniques such as Random Forest, Extreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), and Light Gradient Boosting Machine (LightGBM). Random Forest is an ensemble method that employs a bagging approach, where multiple decision trees are constructed from randomly sampled training data, and their predictions are aggregated using majority voting for classification or averaging for regression. This method is known for its stability and ability to handle high-dimensional data [6].

XGBoost is a boosting algorithm designed for high efficiency and performance. Its strengths lie in its ability to regulate model complexity through regularization and its support for datasets containing missing values. This algorithm has been widely applied in various data analysis tasks due to its high predictive accuracy [7]. CatBoost is a boosting algorithm specifically developed to handle categorical data. It uses an ordered boosting technique and stable target statistics to reduce overfitting. Several studies have demonstrated CatBoost's competitive performance, especially when handling datasets with numerous categorical variables [8].

LightGBM is an improved version of XGBoost, offering greater computational efficiency and lower memory usage. This algorithm employs a histogram-based approach and a leaf-wise tree-growth strategy to accelerate training. The optimization strategy allows LightGBM to have good performance while maintaining speed and memory efficiency. However, this approach may also increase the risk of overfitting if appropriate parameter tuning is not applied [6].

2.3 Optuna

The hyperparameters of a machine learning model can affect the model's ability to perform a given task; hyperparameter tuning is crucial to achieving a good model. As machine learning models are developed, the demand for automatic hyperparameter tuning frameworks is increasing as a convenient way to optimize hyperparameters. Optuna is a hyperparameter tuning framework that allows users to construct the hyperparameter space and search it efficiently using an efficient search-and-prune algorithm. Optuna also provides versatility in many tasks of machine learning modelling, making it a good choice when it comes to hyperparameter tuning [9].

2.4 Permutation Variable Importance (PVI)

Permutation Variable Importance (PVI) is a method for evaluating the contribution of individual features to a machine learning model. The PVI algorithm is described in Algorithm 1 [10]. The main advantage of PVI is its model-agnostic nature, allowing it to be applied across various types of algorithms [11]. In addition to its flexibility, this method is intuitive and easy to implement compared to other model-agnostic methods, such as Shapley Additive Explanation (SHAP), which require heavy computation. Research by [12] demonstrated that PVI consistently identifies the most influential features contributing to a model's predictive performance.

Algorithm 1. Permutation Variable Importance algorithm
Input : $x_i, y, f, M; i = 1,2,..p;$
Output: $Score_i = M_i - M_{base}\ ; i = 1,2, \dots, p$
1: Fit the model with all predictors $(x_1, x_2, \dots, x_p)$
2: Compute the performance metrics of the fitted model f and call it M_{base}
3: For $i = 1,2, \dots, p$ do
4: Permute x_i
5: Calculate the performance metrics on permuted data with the fitted model f and call it (M_i)
6: Calculate the VI score: $Score_i = M_i - M_{base}$
7: End for.

2.5 Cuckoo Search Algorithm

The Cuckoo Search (CS) Algorithm is a metaheuristic optimization algorithm inspired by the parasitism behavior of cuckoo birds. It leverages the concept of Levy flight to effectively perform both global and local solution searches. CS was performed by creating a few nests that represent the solution, which were later optimized. Each nest will be evaluated and compared to the new solution generated through the Levy flight. The Levy flight characteristic of frequently taking short steps and occasionally making large jumps helps the algorithm escapes local minima. The worst nest, which represents the poorest solution, is discarded and replaced with a new one. This process is repeated over several iterations until the objective function converges.

Since its first introduction, Cuckoo Search has been widely used to solve various optimization problems in engineering and computer science. Its main advantages lie in its simplicity and the minimal number of parameters required, making it easy to implement across a wide range of applications[13]. Over time, Cuckoo Search has undergone various modifications, particularly for solving certain problems. One such improvement is the integration of the random key encoding scheme, leading to the development of the Random Key Cuckoo Search (RKCS) for discrete problems, which was later used in this research.

Other research [14] on CS shows that it is commonly used in engineering, energy systems, and robotics, demonstrating the algorithm's flexibility and effectiveness in tackling diverse challenges. Additionally, [15] proposed an adaptive model-based variant of Cuckoo Search that dynamically adjusts the search steps, leading to more efficient and convergent optimization. Numerous applications of Cuckoo Search have been explored in areas such as control system tuning, structural fault detection, and controller parameter optimization, further highlighting the algorithm's vast potential in the fields of machine learning and optimization [16].

2.6 Ensemble Variable Importance

One of the key outputs of machine learning algorithms is variable importance (VI), which quantifies the contribution of each predictor variable to the model's performance. Although various supervised machine learning algorithms can be applied to the same dataset using the same VI computation methods, they often produce differing variable importance rankings. These differences can lead to ambiguity in interpreting and selecting important features. To address this issue, it is necessary to harmonize variable importance measures across models. One promising approach is to use an optimization framework, such as simulated annealing or cuckoo search, to find a series of VI rankings that can represent all ML rankings through maximizing the minimum correlation of each VI ranking from different ML with the EVI [17].

Simple aggregation methods, such as majority voting or averaging ranks (mean or median), are often inadequate for several reasons. In majority voting, when the number of models is even, and no clear majority exists at a given rank, the method fails to reach a definitive consensus. Similarly, using the mean or median of ranks may result in non-integer values, which are problematic in ranking contexts where unique, whole-numbered positions are required. Moreover, these methods can lead to duplicate ranks for different features, reducing the interpretability and distinctiveness of the final importance list. In contrast, optimization-based approaches explicitly aim to maximize the unifying process across individual model rankings, yielding a VI ranking that better reflects the collective judgment of all models' VI rankings.

2.7 Data Analysis Stages

The following were the stages of analysis in this research:

1. Preparing the dataset.
2. Conducting data exploration to understand the characteristics of the data.
3. This ratio is chosen because the available dataset is sufficiently large, and the 80:20 split provides a balanced trade-off between supplying enough data for model training while reserving enough independent data to reliably evaluate model performance.
4. Addressing class imbalance in the target variable, which can affect the machine learning modeling process, as models tend to perform better in predicting classes with higher frequencies compared to minority classes [18]. This was handled using the class weight method, a technique in machine learning that assigns greater weight to minority class instances [19].
5. Conducting binary classification modeling using Random Forest, XGBoost, CatBoost, and LightGBM methods, with hyperparameter tuning through the Optuna hyperparameter tuning framework.
6. Applying cross-validation to evaluate the model's performance in making predictions and to assess the predictive accuracy [20]. Balanced accuracy is used as the performance metric to evaluate ensemble model performance and to address imbalanced data [21]. The following balanced accuracy function in Eq. (1) below:
$$Balanced\ accuracy = \frac{1}{2}(Sensitivity + Specificity). \tag{1}$$
7. Identifying variable importance using the permutation feature importance method to determine the contribution of each independent variable to the target variable [22].
8. Applying optimization using the Cuckoo Search Algorithm to find Ensemble Variable Importance Measure (EVI-CS), with the following objective function in Eq. (2) below. The maxmin strategy helps to find a series of VI ranks that maximize the minimum Spearman correlation (ρ) between the ensembled VI rank and each VI rank from 4 different ML algorithms.
$$max_{r_e}\left[min\{\rho(r_e, r_{rf}), \rho(r_e, r_{xgb}), \rho(r_e, r_{cat}), \rho(r_e, r_{lgbm})\}\right], \tag{2}$$
where:
r_e : Ensembe VI rankings
r_{rf} : Random Forest VI rankings
r_{xgb} : XGBoost VI rankings
r_{lgbm} : LightGBM VI rankings
9. Evaluating by examining the correlation between the variable importance from each model and the ensemble variable importance derived from the Cuckoo Search method (EVI-CS).

The flowchart illustrating the research methodology for this research is presented in Fig. 1.

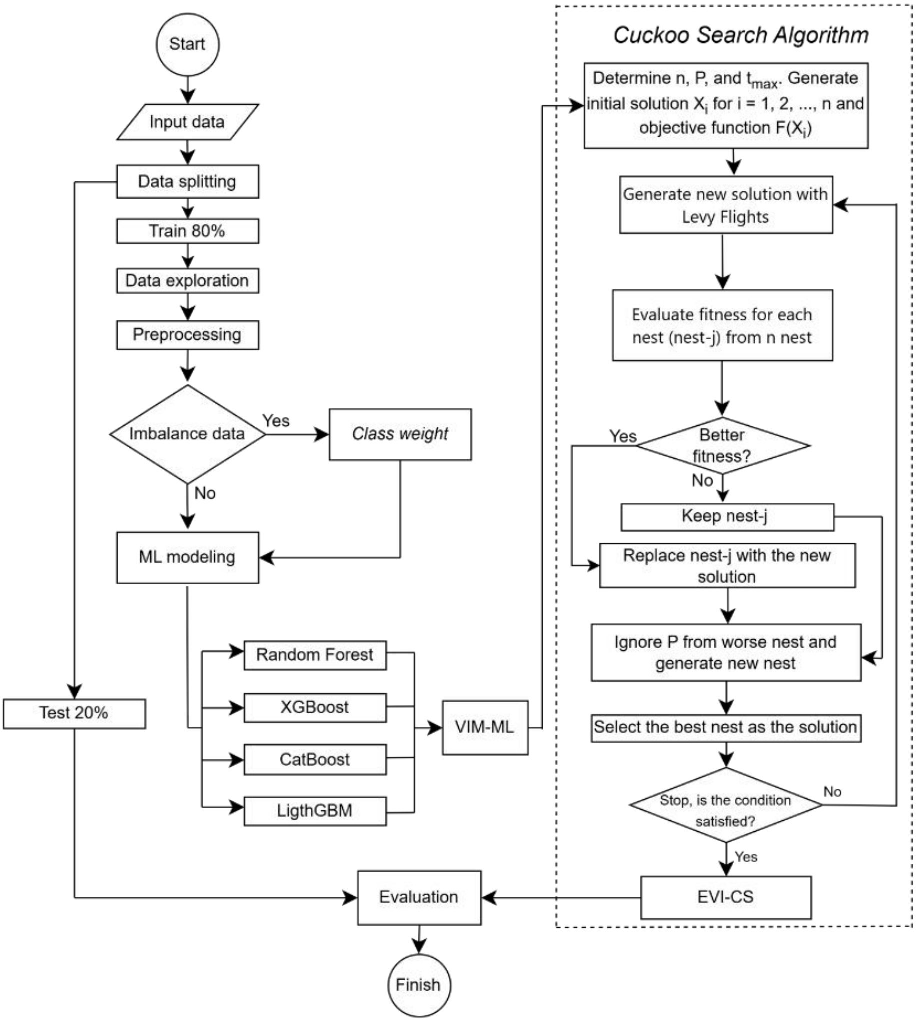

Figure 1. Research Flow Chart

3. RESULTS AND DISCUSSION

3.1 Exploratory Data Analysis

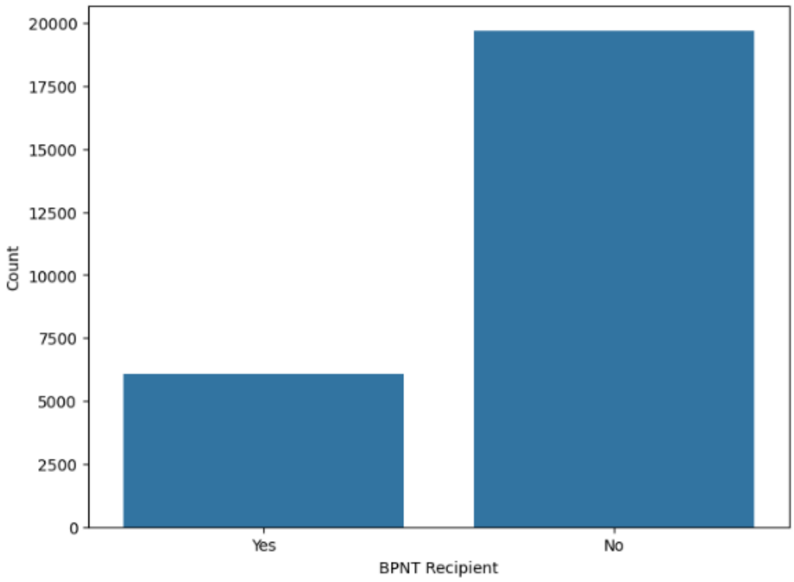

Figure 2. Distribution of Classes in the Response Variable

The distribution of households receiving Non-Cash Food Assistance (BPNT) in West Java Province in 2023 indicates an imbalance in the response variable classes. A total of 6,085 households (23.6%) received BPNT, while 19,706 households (76.4%) were non-recipients. Fig. 2 illustrates the distribution of households across the response variable categories. This class imbalance must be considered during the modeling process, particularly when selecting appropriate methods and model evaluation metrics.

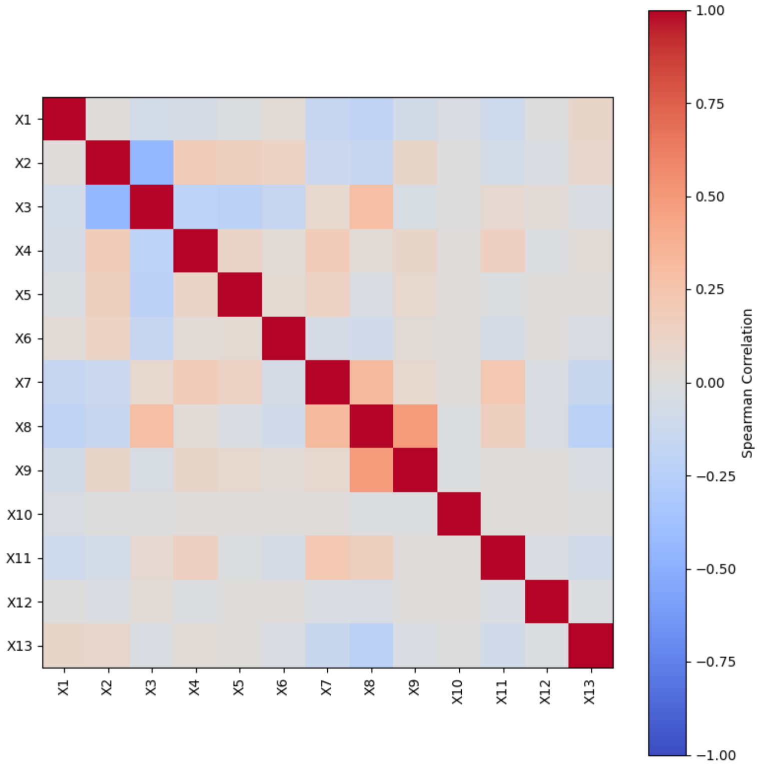

Figure 3. Spearman Correlation Plot of Predictor Variables

Fig. 3 displays the Spearman correlation matrix for the predictor variables included in the analysis. The Spearman rank correlation coefficient was chosen because the data were categorical and may not meet the assumptions of normality or linearity. Overall, the matrix reveals that most pairwise correlations fall within the weak to moderate range, with only a few relationships showing notable strengths. This suggests a relatively low degree of multicollinearity, which is suitable for ensembling variable importances [17].

Among the observed relationships, a moderate positive association is evident between X7 (floor area) and X8 (expenditure per capita), and between X9 (calorie intake per capita) and X8, aligning with economic theory that suggests better economic standing leads to improved nutritional outcomes and life conditions. Other variables show weak positive correlations with each other. The absence of high-magnitude correlations (e.g., >0.8 or <-0.8) across the matrix indicates that the predictors provide largely distinct information, which is desirable in building predictive models and deriving variable importances.

3.2 Machine Learning Modelling

The modelling starts with dividing the dataset into training and testing sets, with 80% allocated for training and 20% for testing. This corresponds to 20,632 observations in the training set and 5,159 observations in the test set. The modelling process was conducted using the training data and involved four widely used machine learning algorithms: Random Forest (RF), XGBoost, CatBoost, and LightGBM. The class imbalance in the response variable, as previously discussed, was addressed by incorporating class weights during model training. These weights were computed based on the ratio between the positive and negative classes.

Each model underwent hyperparameter tuning using the Optuna optimization framework, with the hyperparameter search space defined in Table 2. Furthermore, model evaluation was performed using 10-fold cross-validation, with balanced accuracy Eq. (1) employed as the evaluation metric. This metric was chosen to ensure fair assessment in the presence of class imbalance. Finally, the best-performing models were applied to the test set to assess their generalizability and predictive performance.

Table 2. Hyperparameter Search Space

Model	Hyperparameter	Choice Type	Search Space
RF	n_estimators	Integer	100-1,000
	max_depth	Integer	3-20
	min_samples_split	Integer	2-10
	min_samples_leaf	Integer	1-5
	max_features	Category	Sqrt, log2
CatBoost	iterations	Integer	100-1,000
	depth	Integer	3-10
	learning_rate	Float	0.01-0.3
	l2_leaf_reg	Float	1-10
	bagging_temperature	Float	0-1
XGBoost	n_estimator	Integer	100-1,000
	max_depth	Integer	3-20
	learning_rate	Float	0.01-0.3
	subsample	Float	0.6-1.0
	colsample_bytree	Float	0.6-1.0
	gamma	Float	0-1
	reg_alpha	Float	0-10
	reg_lambda	Float	0-10
LightGBM	n_estimator	Integer	100-1,000
	max_depth	Integer	3-20
	learning_rate	Float	0.01-0.3
	num_leaves	Integer	20-100
	subsample	Float	0.6-1.0
	colsample_bytree	Float	0.6-1.0
	gamma	Float	0-1
	reg_alpha	Float	0-10
	reg_lambda	Float	0-10

The machine learning models used in this study demonstrated relatively similar performance based on the evaluation metrics. A small difference between the evaluation metric values on the training and test sets indicates good classification ability and generalizability. The incorporation of class weights during the modelling process resulted in a slight increase in balanced accuracy on both the training and test datasets across all four models. This suggests that even without explicit handling of class imbalance, the models were already capable of performing classification effectively. Despite the marginal differences in evaluation metrics, the model version incorporating class weights was selected for further analysis to obtain variable importance. A comprehensive summary of model evaluation results is presented in Table 3.

Table 3. Model Evaluation

Modelling	Model	Balanced Accuracy on Training Data (K-fold)	Balanced Accuracy on Testing Data
Without class weights	RF	0.7803 (± 0.0072)	0.7794
	XGBoost	0.7923 (± 0.0097)	0.7959
	CatBoost	0.7989 (± 0.0114)	0.8020
	LightGBM	0.7938 (±0.0135)	0.7942
With class weights	RF	0.8054 (±0.0138)	0.8124
	XGBoost	0.8052 (±0.0141)	0.8110
	CatBoost	0.8055 (±0.0135)	0.8115
	LightGBM	0.8047 (±0.0146)	0.8111

Previous research [4] that employed undersampling techniques for BPNT classification increased the recall of Random Forest and XGBoost models to 80.01% and 74.04%, respectively. The increase is offset by a decrease in overall accuracy (from 55.21% to 60.52%). On the other hand, the model developed in this study demonstrates a more balanced performance. Specifically, the proposed model achieved an average balanced accuracy of 80%, indicating strong recall and a better balance between sensitivity and specificity. This suggests that the current model may offer better classification capability, particularly in handling class imbalance.

The selected model was used to obtain variable importance (VI) rankings for each feature using the Permutation Variable Importance (PVI) algorithm. The VI ranking for each variable was determined based on its PVI score, where a rank of 1 indicates an important feature for classification, and a rank of 13 indicates a less important one. To obtain an ensemble ranking, the RKCS algorithm is used to find a series of rankings that maximizes the minimum Spearman correlation across the VI rankings of the machine learning models. The optimization process was carried out using the following parameters: the number of nests (n_nest) set to 20, the discovery probability (Pa) set to 0.5, and the step size (step_size) set to 0.5. The optimization results indicated convergence after approximately 750 iterations, with the minimum Spearman correlation value reaching 0.6538. The optimization history is illustrated in Fig. 4.

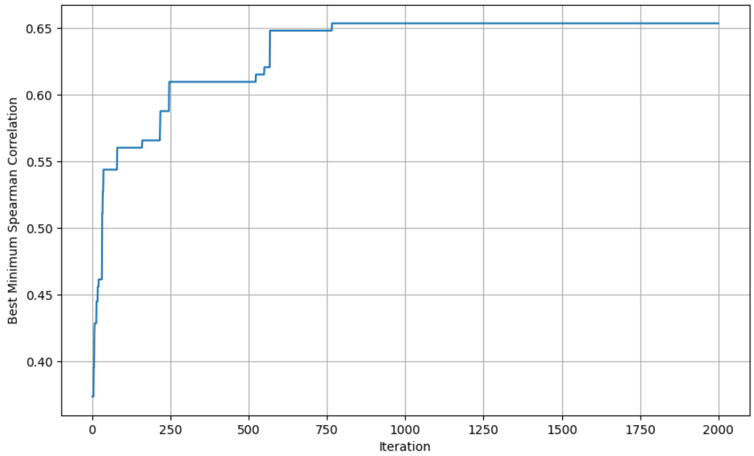

Figure 4. Optimization History of EVI-RKCS

The Ensemble Variable Importance (EVI) ranking, generated using the Random-Key Cuckoo Search (RKCS) algorithm, achieved a Spearman correlation of 0.6538 with the VI ranking from the CatBoost model, 0.6648 with that of the Random Forest (RF) model, 0.6813 with that of the XGBoost model, and 0.8626 with the VI ranking from the LightGBM model. Table 4 demonstrates that although the models exhibit similar predictive performance, the resulting variable importance (VI) rankings may differ. This variation arises from differences in the algorithms underlying each model. The result of the EVI ranking indicates that it was constructed by integrating and balancing the VI rankings from all four machine learning models. The full EVI ranking is also presented in Table 4.

Table 4. Results of Ensemble Variable Importance (EVI) and Individual Model-Based Variable Importance (VI-ML) Rankings

Predictors	VI-ML Rankings				EVI-CS Rankings
	RF	CatBoost	XGBoost	LightGBM	
X1	10	7	10	10	11
X2	12	10	12	12	12
X3	5	12	5	5	9
X4	4	5	4	4	2
X5	3	4	1	1	1
X6	11	3	3	11	8
X7	1	11	11	7	5
X8	9	1	2	3	4
X9	6	2	7	2	3

Predictors	VI-ML Rankings				EVI-CS Rankings
	RF	CatBoost	XGBoost	LightGBM	
X10	2	9	6	6	7
X11	8	6	9	9	6
X12	7	8	8	8	10
X13	13	13	13	13	13
ρ (VI-ML, EVI-RCKS)	0.6648	0.6538	0.6813	0.8626	

Fig. 5 presents a simplified visualization that highlights the capability of the derived Ensemble Variable Importance (EVI) ranking. The EVI result serves as a strong tool for interpreting each variable's importance in classifying BNPT recipients, as it effectively captures the collective pattern of variable importance across the four machine learning models and provides deeper insights. Variables that were consistently categorized as important in the individual VI rankings remained in the important category in the EVI results. The condition can be seen on variables such as X5, which remained in the important category, and X2, which remained in the less important category. For variables with more diverse VI rankings, such as X8, the EVI assigned a rank by balancing the varying importance scores across models.

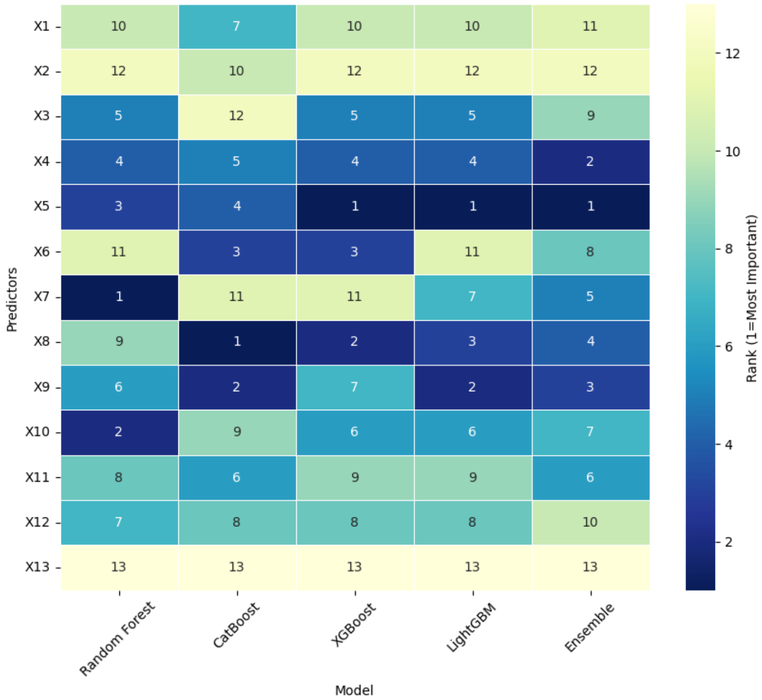

Figure 5. Visualization of EVI-RCKS and VI-ML

The EVI ranking identifies the top four most important variables as: roof quality (X5), homeownership status (X4), average calorie consumption per capita (X9), and average expenditure per capita (X8). These results suggest that household consumption capacity and living conditions are the key factors influencing the classification of BNPT recipients in West Java in 2023. Based on the result, the living conditions represented by X5 and X4 are considered more important than the consumption capacity represented by X9 and X8. While consumption capacity indicators directly capture a household’s economic ability to meet basic needs, they are often difficult to measure accurately in practice because they rely on detailed, self-reported information that may be prone to recall bias or data gaps.

On the other hand, living conditions indicators are practical proxies for determining household conditions. Roof quality and homeownership status can reflect accumulated economic resilience in each household. This makes it a more reliable indicator than other variables. Characteristics such as the physical structure of the house, the quality of the roof, and homeownership status can be directly assessed during field visits, making them practical proxies for household welfare and a tool for evaluating the accuracy of BNPT implementation when detailed consumption data is absent.

4. CONCLUSION

This study successfully implemented the Cuckoo Search-Based Ensemble Variable Importance (EVI) approach to improve the interpretability of the BPNT recipient classification model in West Java in 2023. Although all Machine Learning models used (Random Forest, XGBoost, CatBoost, and LightGBM) showed good classification performance, with small differences in evaluation metrics, the resulting variable importance (VI) differed across models. Through optimization using the Random-Key Cuckoo Search (RKCS) algorithm, an EVI ranking was obtained that integrated the key information from the four models, with a minimum Spearman correlation of 0.6538, demonstrating convergence and stability of the results. Validation of the EVI with VI-ML rank categorization confirmed that the EVI consistently represented important patterns across models. Variables such as roof quality, homeownership status, average calorie consumption, and per capita expenditure were identified as the main factors in determining BPNT recipients. This finding confirms that housing conditions and household consumption power are crucial indicators in social assistance targeting and demonstrates the potential of the ensemble VI approach to improve accuracy and transparency in public policy modeling.

Despite these contributions, the study is not without limitations. The use of cross-sectional SUSENAS 2023 data restricts temporal analysis, and the EVI results remain sensitive to the specific set of machine learning models included in the study. Future research can expand this work by incorporating multi-year or panel data, integrating a wider range of machine learning algorithms or other optimization strategies, and combining EVI with local explanation methods such as SHAP at the household-level. In the end, there is still considerable room for improvement in this study.

Author Contributions

Indra Mahib Zuhair Riyanto: Conceptualization, Data Curation, Methodology, Writing-Original Draft, Software, Validation. Laras Suprapti: Conceptualization, Methodology, Data Exploration, Writing-Original Draft, Validation. Salsabila Fayiza: Methodology, Writing-Original Draft, Validation. Elke Frida Rahmawati: Methodology, Writing-Original Draft, Validation. Farid Yafi Suwandi: Data Exploration, Writing-Original Draft, Validation. Sachnaz Desta Oktarina: Conceptualization, Data Curations, Writing-Review, Validation. Rahma Anisa: Writing-Review, Validation. All authors discussed the results and contributed to the final manuscript.

Funding Statement

The authors declare that no specific funding was received for this research from any public, commercial, or non-profit entity.

Acknowledgment

The authors extend their deepest gratitude to IPB University, particularly the Statistics and Data Science Study Program, for the invaluable academic resources, institutional support, and continuous support that were instrumental in completing this research

Declarations

The authors declare no conflicts of interest to report study.

Declaration of Generative AI and AI-assisted technologies

Generative AI tools (e.g., ChatGPT) were used solely for language refinement (grammar, spelling, and clarity). The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors reviewed and approved all final text.

REFERENCES

[1] I. Hermawan, Izzaty, E. Budiyanti, R. Sari, Y. Sudarwati, and M. Teja, "EFEKTIVITAS PROGRAM BANTUAN PANGAN NONTUNAI DI KOTA YOGYAKARTA," *Jurnal Ekonomi & Kebijakan Publik*, vol. 12, no. 2, pp. 131–145, 2021, doi: https://doi.org/10.22212/jekp.v12i2.2237

[2] A. Nasution, B. Krisnamurthi, and D. Rachmina, "ANALISIS PERMINTAAN PANGAN RUMAH TANGGA PENERIMA MANFAAT BANTUAN PANGAN NON TUNAI (BPNT) DI KOTA BOGOR," *Forum Agribisnis*, vol. 10, no. 1, pp. 1–10, 2020, doi: https://doi.org/10.29244/fagb.10.1.1-10

[3] K. Thirugnanasambandam *et al.*, "OPTIMIZING MULTIMODAL FEATURE SELECTION USING BINARY REINFORCED CUCKOO SEARCH ALGORITHM FOR IMPROVED CLASSIFICATION PERFORMANCE," *PeerJ Comput Sci*, vol. 10, pp. 1–26, 2024, doi: https://doi.org/10.7717/peerj-cs.1816

[4] A. Rusyana, A. H. Wigena, I. M. Sumertajaya, and B. Sartono, "AN OPTIMAL APPROACH TO IDENTIFY THE IMPORTANCE OF VARIABLES IN MACHINE LEARNING USING CUCKOO SEARCH ALGORITHM," *Mathematics and Statistics*, vol. 11, no. 6, pp. 895–909, 2023, doi: https://doi.org/10.13189/ms.2023.110604

[5] A. M. Nazmi and A. Pambudi, "CLASSIFICATION ANALYSIS OF THE ELIGIBILITY OF RECIPIENTS OF NON-CASH FOOD ASSISTANCE AND FAMILY HOPE PROGRAMS IN THE CITY OF SUKABUMI USING THE NAÏVE BAYES CLASSIFIER ALGORITHM," vol. 3, no. 2, pp. 2–6, 2024, doi: https://doi.org/10.59934/jaiea.v3i2.420

[6] K. Ileri, "COMPARATIVE ANALYSIS OF CATBOOST, LIGHTGBM, XGBOOST, RF, AND DT METHODS OPTIMISED WITH PSO TO ESTIMATE THE NUMBER OF K-BARRIERS FOR INTRUSION DETECTION IN WIRELESS SENSOR NETWORKS," *International Journal of Machine Learning and Cybernetics*, no. 0123456789, 2025, doi: https://doi.org/10.1007/s13042-025-02654-5

[7] T. Chen and C. Guestrin, "XGBOOST: A SCALABLE TREE BOOSTING SYSTEM," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-Augu, pp. 785–794, 2016, doi: https://doi.org/10.1145/2939672.2939785

[8] G. Lyu and M. Nakayama, "PREDICTION OF RESPIRATORY FAILURE RISK IN PATIENTS WITH PNEUMONIA IN THE ICU USING ENSEMBLE LEARNING MODELS," *PLoS One*, vol. 18, no. 9 September, pp. 1–16, 2023, doi: https://doi.org/10.1371/journal.pone.0291711

[9] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "OPTUNA: A NEXT-GENERATION HYPERPARAMETER OPTIMIZATION FRAMEWORK," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2623–2631, 2019, doi: https://doi.org/10.1145/3292500.3330701

[10] D. Alvares *et al.*, "A BAYESIAN JOINT MODEL OF MULTIPLE LONGITUDINAL AND CATEGORICAL OUTCOMES WITH APPLICATION TO MULTIPLE MYELOMA USING PERMUTATION-BASED VARIABLE IMPORTANCE," *ArXiv*, vol. 2, no. 1, pp. 1–27, 2025.

[11] C. Molnar, *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable.* 2020.

[12] A. Fisher, C. Rudin, and F. Dominici, "ALL MODELS ARE WRONG, BUT MANY ARE USEFUL: LEARNING A VARIABLE'S IMPORTANCE BY STUDYING AN ENTIRE CLASS OF PREDICTION MODELS SIMULTANEOUSLY," *Journal of Machine Learning Research*, vol. 20, pp. 1–81, 2019.

[13] X.-S. Yang and S. Deb, "CUCKOO SEARCH VIA LEVY FLIGHTS," 2010.

[14] S. Almufti, R. Marqas, R. Asaad, and A. Shaban, "CUCKOO SEARCH ALGORITHM: OVERVIEW, MODIFICATIONS, AND APPLICATIONS," *International Journal of Scientific World*, vol. 11, pp. 1–9, 2025, doi: https://doi.org/10.14419/efkvvd44

[15] Y. Xiong, Z. Zou, and J. Cheng, "CUCKOO SEARCH ALGORITHM BASED ON CLOUD MODEL AND ITS APPLICATION," *Sci Rep*, vol. 13, no. 1, pp. 1–13, 2023. https://doi.org/10.1038/s41598-023-37326-3

[16] M. Guerrero-Luis, F. Valdez, and O. Castillo, "A REVIEW ON THE CUCKOO SEARCH ALGORITHM," pp. 113–124, 2021, doi: https://doi.org/10.1007/978-3-030-68776-2_7

[17] A. Rusnaya, "PENGEMBANGAN ENSEMBLE VARIABLE IMPORTANCE UNTUK BEBERAPA MODEL MACHINE LEARNING MENGGUNAKAN ALGORITMA METAHEURISTIK," 2024.

[18] R. Yulianti, E. A. Ilmani, M. Z. Waliulu, B. Sartono, and A. R. Firdawanti, "PERBANDINGAN ALGORITMA RANDOM FOREST DAN XGBOOST DALAM KLASIFIKASI PENERIMA BANTUAN PANGAN NON-TUNAI (BPNT) DI PROVINSI JAWA BARAT," *Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, vol. 6, no. 1, pp. 21–34, 2025.

[19] H. Akbar and W. Sanjaya, "KAJIAN PERFORMA METODE CLASS WEIGHT RANDOM FOREST PADA KLASIFIKASI IMBALANCE DATA KELAS CURAH HUJAN," *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 3, 2023, doi https://doi.org/10.20885/snati.v3i1.30

[20] H. Azis, P. Purnawansyah, F. Fattah, and I. P. Putri, "PERFORMA KLASIFIKASI K-NN DAN CROSS VALIDATION PADA DATA PASIEN PENGIDAP PENYAKIT JANTUNG," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 81–86, 2020. , doi https://doi.org/10.33096/ilkom.v12i2.507.81-86

[21] A. Khairunnisa, "PERBANDINGAN MODEL RANDOM FOREST DAN XGBOOST UNTUK PREDIKSI KEJAHATAN KESUSILAAN DI PROVINSI JAWA BARAT," *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, p. 202, 2023, doi https://doi.org/10.26798/jiko.v7i2.799

[22] U. A. Syam *et al.*, "EVALUASI KINERJA MODEL RANDOM FOREST DAN LIGHTGBM UNTUK KLASIFIKASI STATUS IMUNISASI HEPATITIS B (HB-0) PADA BALITA," *Jurnal Ilmiah Matematika, Sains, dan Teknologi*, vol. 13, no. 1, pp. 1–8, 2025, doi https://doi.org/10.37905/euler.v13i1.29762