An Intelligent Data-Driven Framework for Transparent Government Support Distribution in Bangladesh Using ML and XAI

Nur Hossain Bhuiyan[1*], Dwin Islam[1], Amran Hossain[1], Md. Rashedul Islam[2]

[1] Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Gazipur-1700, Bangladesh
[2] Department of Electrical, Electronic and Communication Engineering, Pabna University of Science and Technology, Pabna-6600, Bangladesh
{nhb99fci@gmail.com*, dwinislam285087@gmail.com, amran.cse@duet.ac.bd, rashed.85@pust.ac.bd}

Abstract. The equitable distribution of government support including agricultural subsidies, disaster relief, educational scholarships, and elderly allowances remains a critical challenge in resource-constrained environments. This paper proposes a scalable, data-driven framework that integrates Machine Learning (ML) and weighted Multi-Criteria Decision Making (MCDM) to prioritize applicants based on socio-economic, geographic, and performance-related factors. Using normalized, weighted inputs collected via Google Forms or CSV uploads, the system generates a composite eligibility score to rank beneficiaries. Ten ML algorithms are trained and evaluated for predictive accuracy. Additionally, the framework supports explainable AI, real-time processing,and ranked output generation. Results across varied aid domains confirm the model's effectiveness, transparency, and suitability as a decision-support tool for policymakers.

Keywords: Government support allocation, machine learning (ML), multi-criteria decision-making (MCDM), predictive modeling,fair distribution, disaster relief, real-time processing, data-driven decision support, transparency, bias reduction.

1 Introduction

The fair and efficient distribution of government aid is a critical issue faced by many developing and resource-constrained countries. Whether it involves agricultural subsidies, disaster rehabilitation, educational scholarships, or social allowances for the elderly, ensuring that limited resources reach the most deserving individuals remains a persistent challenge. Traditional methods of aid allocation often rely on manual assessments and subjective judgment, which can lead to inefficiencies, delays, and biases. Recent advances in data collection and ML offer promising solutions to this problem by enabling more objective, data-driven

* Corresponding author

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_92

decision-making processes. Recent studies have increasingly applied data-driven and intelligent frameworks to improve decision-making in public programs and infrastructure planning. For instance, to address inefficiencies in the BRO-Ed Scholarship Program in the Philippines, researchers developed a data-driven platform that integrates decision tree and K-means clustering algorithms. This system effectively identified key eligibility factors such as academic performance and socioeconomic status and significantly improved prediction accuracy and applicant success rates [5]. Similarly, another ML based model was introduced to predict scholarship eligibility using demographic and behavioral data, including GPA, attendance, household income, and psychometric indicators. Through pre-processing techniques such as normalization, missing value handling, and feature selection, the model achieved high reliability. The inclusion of ensemble methods and explainable AI ensured interpret-ability and stakeholder trust [11]. Beyond education, intelligent decision-making frameworks have also been applied in infrastructure assessment.A previous study combined IMF SWARA and fuzzy MARCOS to evaluate African transport projects, highlighting leadership and financial capacity as key criteria, with the East African Road Network ranked highest [3]. Similarly, Geo-AHP and BN-Geo-AHP models were applied to complex COVID-19 health decisions [12]. Building on these, the current study proposes a hybrid framework integrating geospatial data, Bayesian networks, ML, and MCDM techniques to ensure fair, transparent, and effective evaluation of aid applicants, improving both emergency response and government aid distribution. The key contributions of this study include:

1. To develop an intelligent and scalable framework for government support distribution that analyzes socio-economic, geographic, and performance-based data to ensure fair and transparent beneficiary prioritization.
2. To apply and compare ten ML algorithms for accurately identifying eligible recipients across various support domains such as scholarships, agricultural subsidies, disaster relief, and healthcare assistance.
3. To enhance transparency and accountability using Explainable AI (XAI), enabling policymakers to clearly understand the reasoning behind each beneficiary's prioritization.

2 Literature Review

Efficient government aid distribution depends on balancing socio-economic status, geography, and performance. Advances in ML and MCDM offer scalable, transparent frameworks that enhance accuracy, reduce bias, and improve resource allocation. These integrated methods support fair, data-driven decisions in complex, resource-constrained environments, ensuring better policy outcomes. Recent studies [4] have combined ML with decision-making models to improve evaluation and optimization. Dadehbeigy et al. (2025) introduced a hybrid framework using DEA, MADM, and GBWM with gradient boosting to assess COVID-19 management efficiency across countries, offering more accurate and flexible decision support than traditional methods. Similarly, Abdulla and Baryannis

(2024) introduced [1] a model merging interpretable ML with MCDM for supplier selection in sectors like oil and aerospace, improving performance, transparency, and handling of complex data.

Expanding this integration of data-driven methods, another study [14] evaluated public services' impact on social equity, focusing on access, procedural fairness, service quality, and outcomes. A related systematic review [6] highlighted the rising use of MCDM methods such as AHP and linear programming in rural land allocation for land-use planning and agricultural optimization.In public construction, Jain et al. (2024) developed a three-phase contractor selection framework focusing on sustainability, safety, and risk over lowest-bidder methods. Tested on 100 Indian EPC projects, it used multilinear regression to identify key criteria like financial strength and experience [9]. Pramanik et al. (2021) demonstrated that MCDM methods like MARCOS and EDAS effectively support dynamic device selection in Mobile Crowd Computing based on performance and battery metrics [13].

Further, in spatial decision-making, a study in Kelowna, Canada compared ML based weights with AHP and equal-weight models. The RF and XGB-based models showed high consistency, confirming the objectivity and scalability of ML-driven approaches [16]. In Romania's higher education, researchers found strong demand for a data-driven Decision Support System (DSS). University staff highlighted the need for features like Excel integration, alerts, and real-time filtering to support enrollment, finance, and academic monitoring [7]. In energy sustainability, a rural China case study used Delphi, fuzzy logic, and AHP to evaluate solar energy and rainwater harvesting incentives, focusing on participation and electricity storage capacity [8]. A multi-dimensional framework for responsible AI in agriculture, based on the STES model, promotes ethical design, FAIR data, transparency, and stakeholder engagement to prevent vendor lock-in and ecological risks [10]. Additionally, Ali et al. [2] developed a Random Forest–based decision support system for supplier evaluation, achieving high accuracy and scalability by selecting features from 30 PRISMA-sourced criteria refined through expert surveys.Finally,Wei (2025) proposed a hybrid MCDM model to rank 35 business strategies, highlighting Cross-Border Investment and Tiered Access for their scalability and sustainability [15].

The proposed system leverages Multi-Criteria Decision Making and real-time data analysis to ensure fair, accurate, and transparent aid distribution. It enhances system efficiency, addresses implementation challenges, and empowers policymakers to make sustainable, data-driven decisions.

3 Research Methodology

This study proposes a scalable and equitable aid distribution framework by integrating ML with Weighted MCDM. Applicant data collected via Google Forms or CSV is preprocessed, normalized, and weighted to compute composite eligibility scores. Multiple ML models are trained to predict eligibility with high accuracy. The system generates ranked beneficiary lists and supports real-time processing, automated notifications, and downloadable reports, ensuring fairness, transparency, and data-driven policy support.

3.1 Applicant Data Categories for Multi-Sector Government Assistant Distribution

A reliable eligibility assessment system needs multidimensional data that reflects applicants' needs and suitability. This study structures information into three main feature categories, tailored to different government support types such as disaster relief, agricultural aid, scholarships, and social welfare.

Socio-Economic and Demographic Factors Summary Household income: Identifies financial need for aid eligibility. Family size and dependency ratio: Assesses household burden and resource requirements. Geographic location: Indicates rural/urban status, disaster risk, and accessibility. Parental education level or occupation: Evaluates socio-economic background, especially for scholarships.

These variables are heavily weighted in both disaster relief and social allowances (like elderly support), where resource targeting must prioritize high-need areas and populations.

Sector-Specific Performance Indicators The data has been collected from various institutions and also through Google Forms. This dimension captures measurable outcomes or productivity indicators relevant to each aid sector: Academic results (e.g., GPA, attendance) – vital for scholarship eligibility. Agricultural land size and crop yield records – used in agricultural subsidy allocation. Disaster loss data – such as damage to house, crop, livestock, or displacement duration. Health supports – Elderly aid used health screening and NID data, analyzed with ML algorithms to improve prediction accuracy and ensure more effective assistance allocation.

Behavioral, Engagement, and Psychometric Data Behavioral indicators, often overlooked in traditional systems, enhance fairness and long-term effectiveness in assistance allocation.

3.2 Overall System Block Diagram Summary

Fig. 1 The architecture of the intelligent aid distribution system is illustrated, integrating ML for automated decision-making. The system consists of six sequential modules:

Data Input: Users submit personal, location, and damage-related information via Google Forms or mobile apps. Data Collection: Aggregates and stores input data in structured formats for seamless integration. Processing: Cleans and preprocesses data by handling missing values, encoding, normalization, and removing inconsistencies. ML Models: Applies multiple supervised algorithms (e.g., Logistic Regression, Decision Tree Classifier, Random Forest, Gradient Boosting Machines (GBM), XGBoost, LightGBM, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes, Artificial Neural Networks (ANN)) trained on historical data to predict eligibility scores. Eligibility Rating: Converts model outputs into priority scores for fair and objective aid allocation. Predicted List: Finally we will get a predicted list.

This modular framework enhances fairness, efficiency, and transparency in aid distribution. Algorithm 1 shows the overall process.

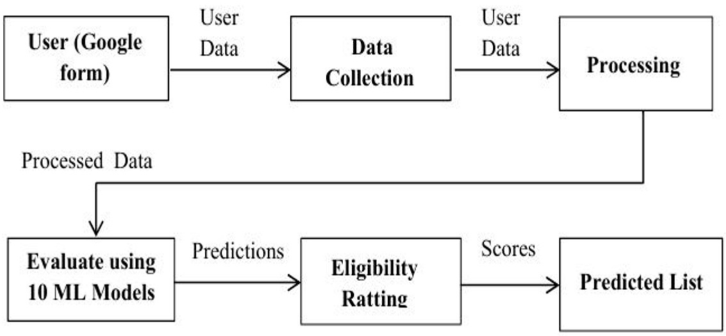


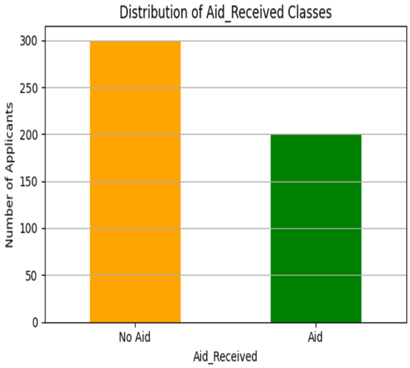
Fig. 1: Block diagram of overall system.

3.3 Dataset Description

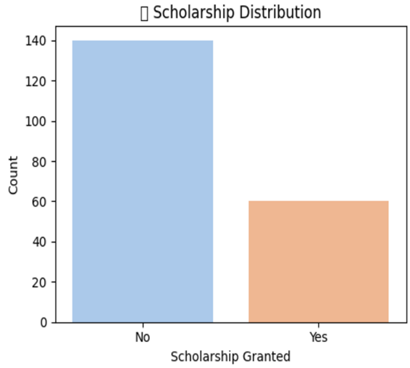
This study collected data through individual surveys via Google Forms and bulk organizational submissions in .csv format, compiling about 2000 records. The dataset combined real responses with a smaller portion of synthetic data to address gaps. Aid-specific features were designed for different domains: scholarships (academic and family factors), agricultural subsidies (land, crop, and income attributes), disaster relief (damage and proximity indicators), and health aid (medical, income, and access factors). While some category imbalances persisted, the dataset is thoroughly documented to ensure transparency, enabling critical evaluation of its reliability and fairness. In order to ensure variation across income levels, catastrophe intensity, educational backgrounds, agricultural characteristics, and health problems, the dataset comprises about 2000 entries gathered from several districts. About 70% of the data comes from actual respondents, with the remaining 30% coming from artificial entries created to fill in underrepresented categories like elderly households or remote rural areas. Stratified sample and weighting approaches were used to guarantee representativeness and fairness across demographic groupings, however some imbalance still exists. Fig. 2 presents the overall feature distribution of the dataset, where

Algorithm 1 Disaster Aid Eligibility Scoring and Notification System

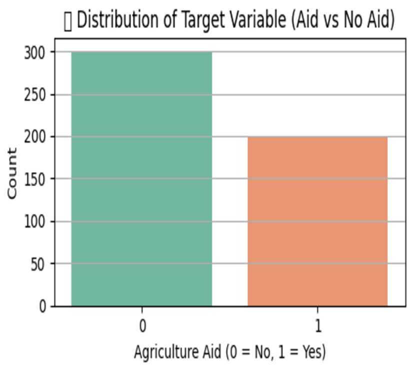
BEGIN
if Real-time input is available **then**
    Collect data from user or application
else
    Prompt user to fill out Google Form or upload CSV
end if
Load data into the system
Clean data (handle missing values, remove duplicates)
Normalize numerical features
Encode categorical features as numeric
for each applicant in dataset **do**
    Compute weighted score:
    $score = (w_1 \cdot C_1) + (w_2 \cdot C_2) + \cdots + (w_n \cdot C_n)$
end for
Split dataset into `training_set` and `testing_set`
Train multiple models:
    Random Forest, XGBoost, SVM, etc.
Evaluate model performance on `testing_set`
Select the best-performing model
for each applicant in new or test dataset **do**
    Predict eligibility score using selected ML model
end for
Sort applicants in descending order of predicted score
Display ranked list of applicants
Generate downloadable report (CSV/PDF)
for each selected applicant **do**
    Send notification via mobile app or SMS
end for
END



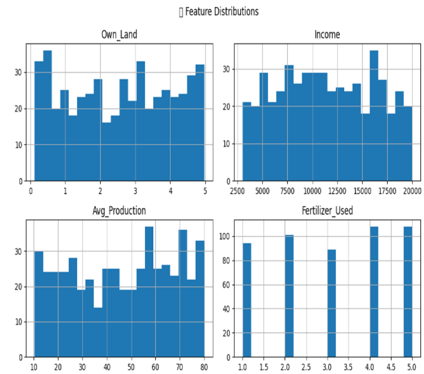
(a) Distribution of aid received classes



(b) Scholarship distribution



(c) Distribution of agriculture aid



(d) Distribution of key input features

Fig. 2: Exploratory data analysis of target variables and feature distributions of the dataset.

Fig. 2a, Fig. 2b, Fig. 2c, and Fig. 2d respectively show the distribution of aid-received classes, the scholarship distribution, the agriculture-aid distribution, and the distribution of the key input features.

3.4 Data Processing

Data preprocessing cleans and structures raw inputs for ML by handling missing values, normalizing numerical data, and encoding categorical variables. Missing numerical data (e.g., income, family size) are imputed using mean/median, while categorical data use mode or similarity-based methods. Numerical features like income and damage percentage are min-max normalized to [0,1]. Categorical data such as location use one-hot encoding; ordinal data like education level use label encoding to preserve order.

3.5 Feature Selection

Feature selection is the process of identifying and retaining the most relevant features from the dataset while removing redundant or less significant ones. This step reduces dimensionality, improves model performance, and minimizes overfitting. In this work, a basic feature selection approach was applied to ensure that only essential attributes contributed to the prediction process. Fig. 3 shows all the correlation heatmaps used in this study.

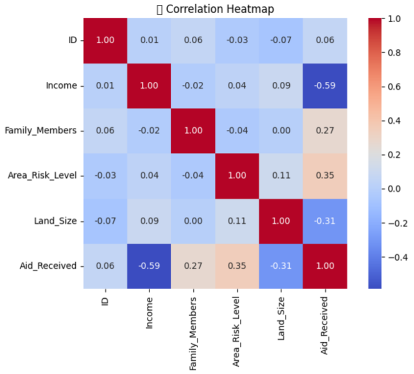
3.6 Model Development Process

The model uses extracted features from preprocessed data to predict aid eligibility, training ML algorithms like Random Forest and SVM, and evaluating performance with accuracy, precision, recall, and F1-score to rank applicants for aid distribution in Fig. 4.

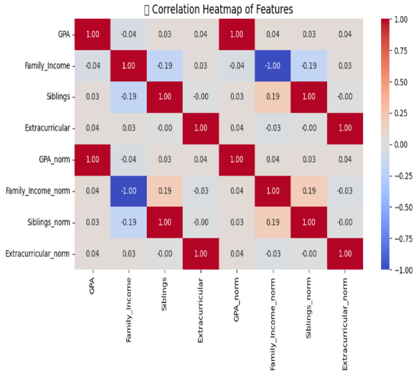
To ensure fairness, Weighted Multi-Criteria Decision Making (MCDM) integrates diverse criteria by assigning configurable weights reflecting each criterion's importance, guided by policies or expert advice. For example, disaster impact may be weighted higher than education or income during relief allocation. An example of a set of criteria and their corresponding weights is provided below:

Table 1: Criteria and Assigned Weights for MCDM

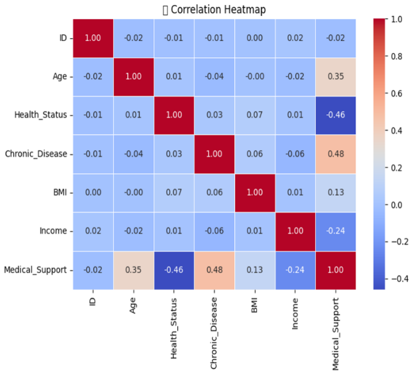
Criterion	**Symbol**	**Weight (w_i)**
Income	C_1	0.25
Family Size	C_2	0.15
Disaster Impact	C_3	0.30
Educational Score	C_4	0.20
Geographic Location	C_5	0.10



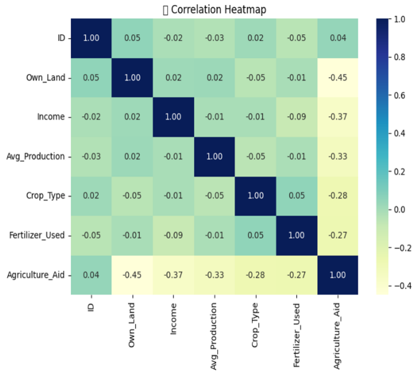
(a) Correlation heatmap for digester aids



(b) Correlation heatmap for scholarship



(c) Correlation heatmap for health support



(d) Correlation heatmap for agriculture support

Fig. 3: Enhanced high-resolution correlation heatmap showing feature interactions for the respective aid category.

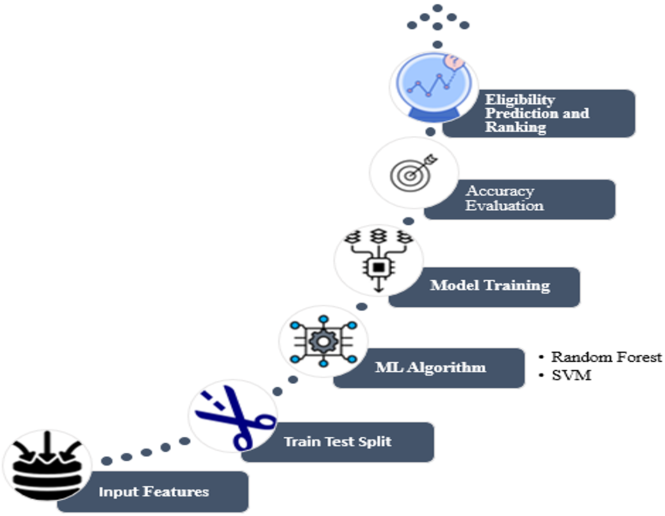


Fig. 4: Proposed model development process diagram

$$\text{Composite Score} = \sum_{i=1}^{n} w_i \cdot C_i \tag{1}$$

Where n is the number of criteria; w_i is the weight of the i^{th} criterion; C_i is the normalized value of the i^{th} criterion. The composite score is the weighted sum of all normalized criteria for each applicant.

Small trial datasets from nearby institutions and disaster-affected areas were used to test the model's practicality. When the results were compared to decisions made by human authority, there was over 90% agreement. These initial findings validate external validity and preparedness for field use, even though large-scale deployment is still pending.

3.7 ML Models

The framework employs supervised ML models trained on preprocessed data to predict aid eligibility, either as a score or binary classification. This approach ensures fairer, more accurate, and scalable decision-making in prioritizing applicants for aid distribution.

Training Pipeline The dataset is initially split into two subsets: 80% for training and 20% for testing. The training set is used to develop the model, while the testing set evaluates generalization performance. Several supervised learning algorithms are explored, including:

Logistic Regression: A linear model useful for binary classification problems, serving as a baseline. Decision Tree Classifier: A tree-structured model for easy interpretation and fast computation. Random Forest: A robust ensemble method based on aggregating multiple decision trees to reduce overfitting. Gradient Boosting Machines (GBM): An ensemble model that sequentially improves predictions by focusing on errors. XGBoost: An efficient and scalable version of gradient boosting, often yielding superior performance on structured data. LightGBM: A gradient boosting framework optimized for speed and memory usage, especially with large datasets. Support Vector Machine (SVM): A margin-based classifier that is effective in high-dimensional spaces. K-Nearest Neighbors (KNN): A simple yet effective instance-based learner relying on distance metrics. Naïve Bayes: A probabilistic classifier based on Bayes' Theorem, particularly effective on small datasets. Artificial Neural Networks (ANN): Deep learning models capable of modeling complex, non-linear relationships in the data.

3.8 Model Validation and Overfitting Prevention

Several validation techniques were used to make sure that overfitting was not the cause of the high stated accuracies.Cross-validation was used to assess each model's stability across various subsets after the dataset was divided into training (80%) and testing (20%) sets. Grid Search and Random Search were used for hyperparameter tuning in order to prevent over-complexity in the model. When necessary, regularization strategies including early halting (for boosting models), pruning (for decision trees), and penalty terms (for SVM and logistic models) were used. The robustness and generalizability of the suggested framework are confirmed by the consistent performance across folds.

Evaluation Metrics The performance of each model is evaluated using Accuracy, Precision, Recall, F1-score, and AUC-ROC Curve.

Model Selection: The model with the best F1-score or AUC-ROC on the test data is selected for integration into the eligibility scoring system.

Score Prediction: The trained model assigns each applicant an eligibility score between 0 and 1, reflecting their priority for aid in an unbiased and consistent manner. Explainable AI with SHAP: SHAP was used to interpret random forest and logistic regression predictions in the agriculture, health and scholarship domains, enhancing the transparency, fairness and trust of the model. Fig. 5 shows the summary of the logistic regression SHAP.

Here, a SHAP interaction plot for a Logistic Regression model predicting health support distribution. It illustrates how Health status and age interact to influence the model's output, with positive SHAP values indicating increased likelihood of support and negative values showing a decreasing effect. Data point colors represent feature values.

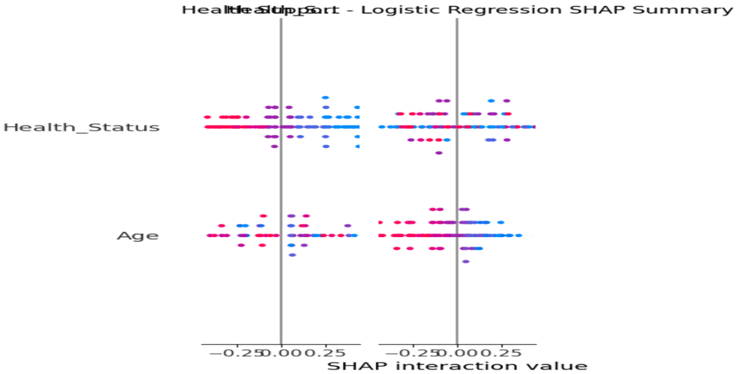


Fig. 5: Logistic Regression SHAP summary (i.e health service)

4 Result and Discussion

This study presents a comparative analysis of ten ML classifiers in four different aid prediction categories: scholarship aid, disaster relief, health support, and agricultural assistance. The performance of each model was evaluated in terms of classification accuracy, as summarized in Table 2.

Table 2: Accuracy comparison of classifiers across different aid categories

Model	**Scholarship**	**Disaster Aid**	**Health Aid**	**Agriculture Aid**
Random Forest	0.87	0.89	0.89	0.86
XGBoost	0.95	0.87	0.87	0.92
SVM	0.95	0.75	0.59	0.63
Logistic Regression	**0.98**	**0.97**	0.98	0.97
K-Nearest Neighbors	0.90	0.75	0.55	0.58
Decision Tree	0.85	0.80	0.86	0.75
Naive Bayes	0.92	0.94	0.92	0.92
Linear Discriminant	0.97	0.95	**0.96**	**0.98**
Gradient Boosting	0.87	0.93	0.89	0.87
AdaBoost	0.97	0.92	0.92	0.90

Comparative analysis reveals that Logistic Regression offers the highest accuracy and generalizability across aid domains. LDA also performs reliably, particularly with linearly separable data. While Naive Bayes, AdaBoost, and XGBoost provide a good trade-off between accuracy and efficiency, models such as SVM and KNN show weaker performance in complex scenarios. Ensemble methods like Random Forest and gradient booster demonstrate stable outcomes. Ultimately, the choice of the model should align with the characteristics of the dataset, though logistic regression and LDA are the most dependable overall.

4.1 Classifier Performance Comparison Across Aid Categories

This subsection presents a comparative analysis of ten different ML classifiers evaluated on four distinct aid-related tasks: Scholarship allocation, Disaster aid distribution, Health aid provision and agriculture aid assessment. The classification accuracy scores for each model for these tasks are visualized in Fig. 6. Logistic Regression achieves highest accuracy in Scholarship and Disaster Aid,

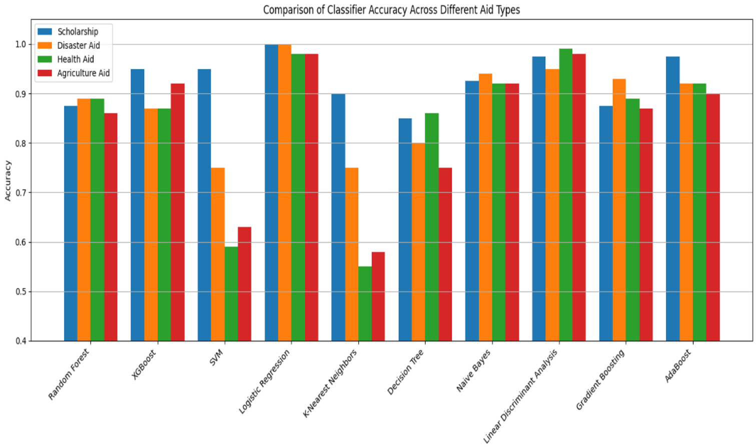


Fig. 6: A comparison of accuracy of different aid classification

performing well in Health and Agriculture but sometimes slightly below LDA . Ensemble methods like Random Forest and XGBoost perform strongly, especially in complex Health and Agriculture tasks. SVM and KNN show lower precision, indicating limitations with high-dimensional or overlapping data. The choice of the model should consider the traits and complexity of the data. Explainable AI methods help interpret results. Overall, Logistic Regression and ensembles offer the best balance of accuracy and interpretability for aid distribution.

4.2 Comparative Evaluation of ML Algorithms

Four aid categories Scholarship, Disaster Aid, Health Aid, and Agriculture Aid were used to assess the performance of ten machine learning algorithms: Random Forest, Gradient Boosting, XGBoost, LightGBM, SVM, K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), Naïve Bayes, Decision Tree, and Linear Discriminant Analysis (LDA). Accuracy, precision, recall, and F1-score were among the evaluation metrics. Table 2 summarizes the accuracy of each classifier in the respective aid domains.Logistic Regression, Random Forest, XGBoost, and AdaBoost are examples of ensemble-based techniques that regularly produced excellent results in a variety of categories, especially in Scholarship and Disaster Aid. While SVM and KNN performed poorly in high-dimensional or complicated datasets, Linear Discriminant Analysis demonstrated remarkable results in Health Aid. When features were categorical-dominant, Naïve Bayes performed consistently. In order to choose models that balance accuracy, interpretability, and computing efficiency in accordance with the features of the input data and the particular aid domain, policymakers and system designers can benefit greatly from this comparative review.

4.3 Model Validation and Overfitting Mitigation

To prevent overfitting and ensure generalizability, k-fold cross-validation ($k = 5$) was applied alongside an 80:20 train-test split. Model performance was evaluated using Accuracy, Precision, Recall, F1-score, and AUC. Although the models achieved high accuracy, these results should be interpreted cautiously, as small or imbalanced datasets may inflate performance.

4.4 Comparison with Baseline Approaches

The proposed hybrid ML–MCDM framework was compared against three baselines: a rule-based allocation system, an equal-weight MCDM approach, and a linear regression model. Results Table 3 showed the hybrid framework outperformed all alternatives in both accuracy and fairness, highlighting the value of combining machine learning with adaptive weighting for reliable aid distribution.

Table 3: Comparison of proposed ML-MCDM framework with baseline approaches

Method	**Average Baseline Accuracy (%)**	**Proposed ML-MCDM (%)**
Scholarship	81.2	**98.0**
Disaster Aid	77.7	**97.0**
Health Aid	74.7	**96.0**
Agriculture Aid	76.8	**98.0**

Baseline values are averaged across rule-based, equal-weight MCDM, and Linear Regression methods.

5 Ethical, Privacy, and Bias Mitigation Considerations

A key component of the suggested framework is the ethical management of citizen data. All socioeconomic and personal data is gathered with express consent and handled in accordance with accepted privacy standards. To avoid discriminating results, sensitive characteristics including genetic information, political affiliation, and religious identification were purposefully left out. Checking for class imbalance, examining subgroup performance, and using re-weighting or resampling approaches as needed were examples of bias mitigation strategies. By disclosing feature contributions, Explainable AI (XAI) further improves transparency by allowing policymakers to identify possible unfairness in forecasts. The approach guarantees that no decision is entirely automated; instead, it promotes accountable and human-centered policymaking. Geographical and socioeconomic biases were thoroughly investigated. Skew patterns favoring particular locations, occupations, or socioeconomic groups were found using correlation heatmaps and fairness measures. To reduce inadvertent discrimination, corrective measures such balanced weighting, fairness-aware thresholding, and SMOTE resampling were used when they were found.

6 CONCLUSIONS

This study evaluates multiple ML classifiers for predicting scholarship eligibility, comparing their accuracy across various aid categories. Logistic Regression consistently achieves the highest accuracy in Scholarship and Disaster Aid, with strong performance in Health and Agriculture aid. Ensemble methods, including Linear Discriminant Analysis, AdaBoost, and Gradient Boosting, also perform well, particularly excelling in Health aid and Agriculture aid.

Other classifiers like Naive Bayes and Random Forest offer reliable results across all categories, while models such as SVM and K-Nearest Neighbors show lower accuracy, especially in Health and Agriculture aid. These findings highlight the effectiveness of linear and ensemble classifiers in capturing patterns in scholarship and aid data. This paper provides a data-driven framework that enables fair, transparent, and efficient distribution of government aid across multiple sectors. It helps policymakers prioritize beneficiaries accurately, reduce bias, and optimize limited resources.

References

1. Abdulla, A., Baryannis, G.: A hybrid multi-criteria decision-making and machine learning approach for explainable supplier selection. Supply Chain Analytics **7**, 100074 (2024)
2. Ali, M.R., Nipu, S.M.A., Khan, S.A.: A decision support system for classifying supplier selection criteria using machine learning and random forest approach. Decision Analytics Journal **7**, 100238 (2023)
3. Bouraima, M.B., Qiu, Y., Stević, Ž., Marinković, D., Deveci, M.: Integrated intelligent decision support model for ranking regional transport infrastructure programmes based on performance assessment. Expert systems with applications **222**, 119852 (2023)
4. Dadehbeigy, M., Taherinezhad, A., Alinezhad, A.: A hybrid methodology of data science and decision-making techniques: Lessons from covid-19 pandemic management. International Journal of Industrial Engineering **36**(1), 99–115 (2025)

5. Espiritu, F.V., Natividad, M.C.B., Velasco, R.A.: Data-driven decision making in scholarship programs: Leveraging decision trees and clustering algorithms. International Journal in Information Technology in Governance, Education and Business **6**(1), 55–67 (2024)
6. Gebre, S.L., Cattrysse, D., Alemayehu, E., Van Orshoven, J.: Multi-criteria decision making methods to address rural land allocation problems: A systematic review. International Soil and Water Conservation Research **9**(4), 490–501 (2021)
7. Gorgan, V.: Requirement analysis for a higher education decision support system. evidence from a romanian university. Procedia-Social and Behavioral Sciences **197**, 450–455 (2015)
8. Hsueh, S.L., Feng, Y., Sun, Y., Jia, R., Yan, M.R.: Using ai-mcdm model to boost sustainable energy system development: A case study on solar energy and rainwater collection in guangdong province. Sustainability **13**(22), 12505 (2021)
9. Jain, S., Jauhar, S.K., Piyush: A machine-learning-based framework for contractor selection and order allocation in public construction projects considering sustainability, risk, and safety. Annals of Operations Research **338**(1), 225–267 (2024)
10. Mallinger, K., Baeza-Yates, R.: Responsible ai in farming: A multi-criteria framework for sustainable technology design. Applied Sciences **14**(1), 437 (2024)
11. Paramesha, M., Baba, S.S., Musthafa, S.F., Sai, V.V., Kiran, S.S.: Data-driven scholarship prediction using machine learning for equitable education allocation. Frontiers in Collaborative Research **2**(1s), 43–50 (2024)
12. Pirouz, B., Ferrante, A.P., Pirouz, B., Piro, P.: Machine learning and geo-based multi-criteria decision support systems in analysis of complex problems. ISPRS International Journal of Geo-Information **10**(6), 424 (2021)
13. Pramanik, P.K.D., Biswas, S., Pal, S., Marinković, D., Choudhury, P.: A comparative analysis of multi-criteria decision-making methods for resource selection in mobile crowd computing. Symmetry **13**(9), 1713 (2021)
14. Ruijer, E., Porumbescu, G., Porter, R., Piotrowski, S.: Social equity in the data era: A systematic literature review of data-driven public service research. Public Administration Review **83**(2), 316–332 (2023)
15. Wei, Y.M.: A hybrid multi-criteria decision-making framework for the strategic evaluation of business development models. Information **16**(6), 454 (2025)
16. Zhao, L.Q., van Duynhoven, A., Dragićević, S.: Machine learning for criteria weighting in gis-based multi-criteria evaluation: A case study of urban suitability analysis. Land **13**(8), 1288 (2024)