International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 3, March, 2026, pp. 805-817
Journal homepage: http://iieta.org/journals/isi

Enhancing Intelligent Decision Making for Green Economy-Based Social Forestry Development Using K-Means Random Forest Fuzzy Learning Vector Quantization

Pradityo Utomo1*, Rahmanta Setiahadi, Dwi Nor Amadi

1 Engineering Faculty, Merdeka University of Madiun, Madiun 63133, Indonesia
2 Agriculture Faculty, Merdeka University of Madiun, Madiun 63133, Indonesia

Corresponding Author Email: pradityo@unmer-madiun.ac.id

https://doi.org/10.18280/isi.310314

Received: 31 October 2025
Revised: 15 January 2026
Accepted: 20 March 2026
Available online: 31 March 2026

Keywords:
hybrid machine learning, K-Means Random Forest Fuzzy Learning Vector Quantization, intelligent decision making, green economy, social forestry development, variable optimization

ABSTRACT

Social forestry development requires balancing ecological sustainability with community welfare, yet decision-making often relies on subjective assessments without systematic optimization. This study proposes a hybrid machine learning model—K-Means Random Forest Fuzzy Learning Vector Quantization (KM-RF-FLVQ)—that integrates K-Means clustering, Random Forest (RF) feature selection, and Fuzzy Learning Vector Quantization (FLVQ) classification to enhance intelligent decision making for green economy-based social forestry. Survey data were collected from 1,000 respondents living near forest borders in Ngawi Regency, East Java, Indonesia, covering 13 green economy variables. KM-RF-FLVQ first applies K-Means to cluster unlabeled citizen assessments, then RF to rank variable importance, and finally FLVQ for classification. Experiments evaluated three variable sets (9, 11, and 13 variables) under varying iteration counts (250, 500, 750, 1000) and three RF tree configurations (T = 20, 40, 60). Results demonstrate that increasing the number of variables improves model accuracy. The optimal configuration—using all 13 variables, T = 40, and 1000 iterations—achieves the highest accuracy of 91.5%, outperforming configurations with fewer variables (9 variables: 89%; 11 variables: 90.5%). KM-RF-FLVQ also surpasses existing methods, including RF (82.26%) and K-Means FLVQ (88.00%), as reported in comparable studies. These findings confirm that hybrid machine learning with systematic variable optimization can significantly improve intelligent decision-making accuracy for green economy-based social forestry development.

1. INTRODUCTION

Social forestry is a forest management approach that recognizes individuals as critical stakeholders. To efficiently oversee this system, the community organizes itself into various organizations, including agroforestry, agrosilvopasture, agroindustry, and ecotourism [1]. The primary objectives of social forestry management are to improve the welfare of the community and the local economy in the vicinity of the forest [2].

To promote sustainable development, numerous European Union countries are implementing green economy principles [3]. This approach is advantageous in the context of social forestry. Social forestry can help preserve forests and optimize resource use by applying green economy principles. Implementing a green economy offers a range of benefits, including the potential to create jobs, improve community well-being, and enhance sustainability [4]. The implementation of green economy principles in developing countries is well-suited to advancing social forestry [5]. Nevertheless, the social forestry concept has not made the most substantial progress [6-8]. It is important that the Ministry of Forestry has a social forestry concept suitable for decision-making related to social forestry development.

In numerous studies, the concept of social forestry can be developed using intelligent decision-making models [9, 10]. Intelligent decision-making models can be developed using machine learning algorithms [11-21]. These models have the potential to improve efficiency and facilitate agreement formulation [22, 23]. Given that machine learning has the potential to improve model reliability [24-26]. Intelligent decision-making systems can be improved through machine learning. Intelligent decision-making systems based on hybrid machine learning can improve decision quality. This methodology is suitable for managing data complexity in decision-making systems [27-29]. Intricate data, such as evaluative information from communities at the forest periphery, is required for optimal decision-making in social forestry development.

Machine learning algorithms have been demonstrated to be capable of effectively managing forests in previous research. For example, machine learning models can manage forests with an accuracy of over 90% [30, 31]. However, when a forest management approach incorporated social data (citizen surveys), the accuracy rate was 55% [32].

The Fuzzy Learning Vector Quantization (FLVQ) method is employed in machine learning. Specifically, individuals have used the FLVQ approach to manage forests. When applied, FLVQ achieves over 85% accuracy [33, 34]. To achieve optimal results, its accuracy must still be enhanced. Nonetheless, FLVQ has demonstrated greater precision than LVQ in multiple studies [35]. Building upon this, a superior version called K-Means FLVQ has been developed through research. In this method, the K-Means algorithm and FLVQ are combined to organize and categorize objects. The K-Means FLVQ technique achieved 88% accuracy, suggesting that FLVQ's efficacy can be enhanced by K-Means [36].

Additionally, machine learning can be optimized through the Random Forest (RF) method. The RF method has been shown to be effective when combined with K-Means in numerous investigations [37, 38]. The RF method can increase K-Means accuracy by up to 99% [37]. This is particularly noteworthy. Additionally, it has the potential to simplify the classification procedure [39, 40]. RF can select variables [41, 42]. Precision can be improved by selecting variables. This approach may function as an optimization strategy [42, 43].

Optimization has also been demonstrated in hybrid machine learning models. Compared with singular models, hybrid models exhibit superior accuracy [36, 44-48]. Compared with LVQ, fuzzy LVQ achieves superior results [35]. Compared with Neuro Fuzzy and K-Means, K-Means FLVQ exhibits superior performance [47]. The Hybrid RF with CNN outperforms CNN alone [48]. Hybrid models have been demonstrated to be effective for optimization. Additionally, several of them have demonstrated the capacity to improve decision-making systems [36, 49, 50]. However, the impact of variable quantity on model optimization, particularly in hybrid decision-making systems, has not been investigated in prior research.

This research has several objectives: (1) to develop a hybrid optimization model for an intelligent decision-making system framework, particularly for social forestry development; (2) to analyze the impact of variable quantity on the hybrid model for an intelligent decision-making system in social forestry development.

Machine learning and principles from the green economy are employed in the model development. K-Means, FLVQ, and RF were among the machine learning methods employed in this research. It is anticipated that the development of green-economy-based social forestry will be more optimal when K-Means, FLVQ, and RF are combined. Local citizen-based assessment data on forest edges can be clustered using the K-Means algorithm. Next, the RF algorithm can identify the most suitable variables. The classification of FLVQ is performed using optimal variable selection. The classification results inform intelligent decision-making for the development of social forestry within the green economy.

The accuracy of a model can be evaluated by employing the Confusion Matrix approach for model validation. The accuracy value is obtained by comparing the goal data to the total data. The model validation results may be derived from data-suitability outcomes [51-54]. It is anticipated that the results of this research will advance the field of information systems, particularly by enabling more precise and efficient decision-making.

2. LITERATURE REVIEW

2.1 Intelligent decision making system

A decision support system is equivalent to a decision-making system. An evolution of an information system is a decision-making system. This system has the potential to help resolve decision-making challenges. Relevant data is used in the decision-making process. An example pertains to the decision-making process within an organization and a business [55]. Rapid decision-making is facilitated by decision-making systems, which are indispensable in an organization. To resolve issues, it is imperative to provide guidance on which options to pursue [56].

Rapid decision-making is facilitated by decision-making systems, which are indispensable in an organization. To resolve issues, it is imperative to provide guidance on which options to pursue. The development of decision-making systems can address increasingly complex issues. The Intelligent Decision-Making System serves as an example. Informed decision-making can be facilitated by artificial intelligence (AI). Artificial intelligence is implemented through machine learning. Intricate issues can be resolved by incorporating machine learning and decision-making systems [57, 58]. An Intelligent Decision-Making System that is based on hybrid machine learning has the potential to make accurate decisions when confronted with complex variable problems [29]. The system's decision-making outcomes depend on the selection of variables [59].

2.2 K-Means clustering

Citizen assessments generate unlabeled data. The K-Means method can be employed to cluster this unlabeled data. Data patterns that exhibit similar characteristics can serve as a basis for classification. The principal advantage of K-Means is its computational efficiency [60]. The K-Means method is a multi-stage process [36, 61, 62].

- Set K
- Determine the distance x and K with Eq. (1):

$$Dist(x, K) = \sqrt{\sum_{i=1}^{n} (x_i - K_i)^2} \tag{1}$$

- Determine K_{new} with Eq. (2):

$$K_{new} = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i \tag{2}$$

- Back to step 1 with K_{new}
- If vertex point update, then back to step 2. Otherwise, it is finished.

where, K is cluster centre, x is vector data, K_{new} is new cluster centre, n is total vector, i is data index, x_i is vector data with index.

2.3 Random Forest optimization

The RF method is a machine learning approach based on decision trees. This approach is advantageous for optimizing models [63, 64]. A critical component of optimization is identifying the most pertinent variables. The RF method comprises multiple phases [63]:

- Set D and T
- For $t=1$ to T, then:
    a. Determine D_t

b. Training $G_t(x)$ with random feature subset.
- Classify:

$$f(x) = majority\ voting\left(G_t(x)\right) \tag{3}$$

where, D is dataset, T is number of trees, D_t is bootstrap sample, $G_t(x)$ is decision tree, $f(x)$ is final prediction function, and t is iteration.

2.4 Fuzzy Learning Vector Quantization classification

The FLVQ method is a machine learning approach used for classification [36]. It integrates components of LVQ and Fuzzy C-Means [65]. The FLVQ method comprises numerous stages [36, 65, 66]:

- Initializing C, N, m_i, and m_f
- Set $t = 0$
- Set w_0
- Calculate iteration with Eq. (4):

$$t = t + 1 \tag{4}$$

- Calculate m with Eq. (5):

$$m = m_i + t\left[\frac{(m_f - m_i)}{N}\right] \tag{5}$$

- Calculate $\alpha_{ij,t}$ with Eq. (6):

$$\alpha_{ij,t} = \left[\sum_{i=1}^{C}\left(\frac{\|x_i - w_{j,t-1}\|^2}{\|x_i - w_{l,t-1}\|^2}\right)^{1/(m-1)}\right]^{-m} \tag{6}$$

with $1 \le i \le M$ and $1 \le j \le C$.

- Calculate $\eta_{j,t}$ with Eq. (7):

$$\eta_{j,t} = \left(\sum_{i=1}^{M}\alpha_{ij,t}\right)^{-1} \tag{7}$$

with $1 \le j \le C$.

- Calculate $w_{j,t}$ with Eq. (8):

$$w_{j,t} = w_{j,t-1} + \eta_{j,t}\sum_{i=1}^{M}\alpha_{ij,t}\left(x_i - w_{j,t-1}\right) \tag{8}$$

with $1 \le j \le C$.

- Calculate E_t with Eq. (9):

$$E_t = \sum_{j=1}^{C}\|w_{j,t} - w_{j,t-1}\|^2 \tag{9}$$

- If $w < N$ and $E_t > \xi$, back to step 4

where, C is number of class, N is maximum iteration, m_i is initial weighted ranking as a membership function, m is weighting rank, m_f is final weighted ranking as a membership function, t is iteration, w_0 is initial weight, η is learning rate, E_t is error for iteration, w is weight, ξ is error toletance. and α is degree of membership.

2.5 Euclidean Distance

Euclidean Distance is a data recognition technique that utilizes distance computations. The minimal distance value is used in the recognition process. Eq. (10) provides the formula for calculating the Euclidean Distance [67, 68].

$$d_{ij} = \sqrt{\sum_{k=1}^{n}\left(x_{ik} - w_{jk}\right)^2} \tag{10}$$

where, d is distance similarity, x is data value, w is weight, i is data index, and k is reference data index.

2.6 Confusion matrix for multi-label

The effectiveness of a model can be demonstrated through the confusion matrix method. Accuracy is one of the primary metrics for success in this approach, which employs a table for performance evaluation [51, 52]. The confusion matrix can also evaluate the efficacy of multi-class recognition in models that contain more than two classes. Table 1 illustrates the multi-label components of the confusion matrix [69].

Table 1. Confusion matrix

Actual / Predicted		**Predicted Value**			
		Label1	**Label2**	**Label3**	**Label4**
Actual Value	Label1	TP_1			
	Label2		TP_2		
	Label3			TP_3	
	Label4				TP_4

The accuracy of the model can be calculated using Eq. (11) [51, 52].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{11}$$

where, TP is the number of data items correctly identified as positive, TN is the number of data items correctly identified as negative, FP is the number of data items correctly identified as positive when they are actually negative, FN is the number of data items correctly identified as negative when they are actually positive.

3. METHODS

The Methods chapter outlines the process for developing an intelligent decision-making model that supports the advancement of a green economy through social forestry. The stages of model development are illustrated in Figure 1.

The research procedure stages are presented in Figure 1. The initial phase begins with a literature review, the first step in model development. This preliminary phase is crucial for identifying research gaps and collecting relevant literature. After this, the researcher proceeds to the second stage: data collection. This stage utilizes observational methods to collect citizen assessments, which provide valuable input for the intelligence model that informs decision-making. Clustering, variable optimization, and classification comprise the decision-making model. Upon completion of these procedures, the results are assessed to determine the model's validity, a critical step in the validation process.

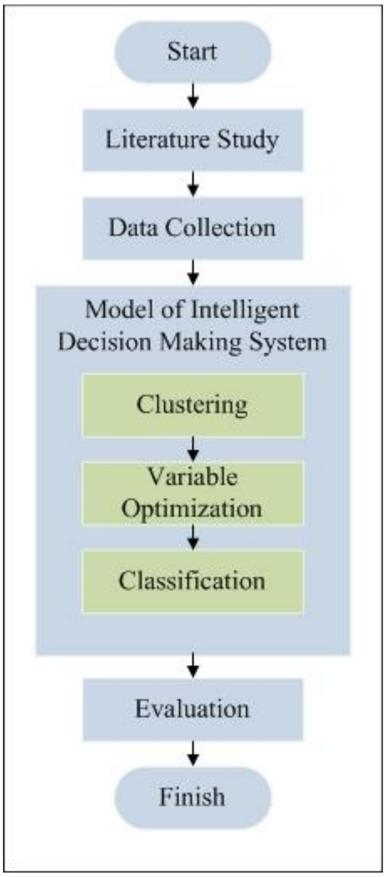

Figure 1. Stages of model development

3.1 Problem identification

This research identifies two key problem formulations based on the findings from the literature study: (1) The development of a hybrid optimization model for a framework for intelligent decision-making systems, particularly for the growth of social forestry; and (2) The examination of the impact of variables varying amounts on the hybrid model for an intelligent decision-making system in social forestry growth.

The research void addressed in this research is underscored by these problem formulations. Previous research has demonstrated the advantages of various machine learning techniques, such as K-Means, Random Forests, and FLVQ. K-Means is effective for aggregating unlabeled data, RF is particularly adept at identifying influential variables, and FLVQ is beneficial for classifying labeled data.

3.2 Dataset

The research is concentrated on Ngawi Regency in East Java, Indonesia, where part of the geographical area is forested. The research required the participation of residents living in the area surrounding the forest border to collect data. Assessing residents' perceptions of green economy variables was a critical component of the research. Observational techniques, particularly questionnaire assessments, were employed to gather this information. The model's effectiveness depends on citizen assessment data. Improving citizen well-being is one of the advantages of this research. As a result, the success of this research depends on citizen feedback.

Additionally, literature review methods were employed to gather information on variables related to the green economy. The variables were chosen based on earlier studies that were related to the growth of social forestry. Additionally, the factors were confirmed by forestry expert Dr. Anang Susanto, S.Hut., M.Sc. The variables must pertain to the notion of the green economy. A literature review has identified the following green economy variables pertinent to social forestry development [3-5]:

1. Distance from village to forest (x_1)
2. Land cover percentage (x_2)
3. Soil fertility (x_3)
4. Citizen income (x_4)
5. Land conflict potential (x_5)
6. Forest history classification (x_6)
7. Main plant type (x_7)
8. Plant cycle (x_8)
9. Land topography (x_9)
10. Livestock potential (x_{10})
11. Home industry (x_{11})
12. potential Ecotourism types (x_{12})
13. Land for Food source potential (x_{13})

The Likert Scale can be used for citizen assessment [70, 71]. Residents living at the edge of the forest evaluated various factors related to social forestry development, assigning ratings on a scale from 1 to 9. The outcomes of these citizen assessments are illustrated in Table 2.

Table 2. Citizen assessment data

Data ID	Variable Code	Value
Data1	x_1	1
	x_2	2
	x_3	1
	x_4	8
	x_5	5
	x_6	2
	x_7	8
	x_8	1
	x_9	3
	x_{10}	3
	x_{11}	3
	x_{12}	3
	x_{13}	2
Data2	x_1	1
	x_2	2
	x_3	2
	x_4	8
	x_5	8
	x_6	5
	x_7	9
	x_8	3
	x_9	2
	x_{10}	3
	x_{11}	2
	x_{12}	3
	x_{13}	1
…	…	…

Table 1 displays the citizen ratings for all the green economy variables associated with social forestry development. In total, the study incorporated 1000 data points to serve as input for the model. The data also has been verified by forestry expert Dr. Anang Susanto, S.Hut., M.Sc. This verification is based on the assessment of the green economy variable by the forest edge residents.

3.3 The proposed method

This study presents an intelligent decision-making model for developing a green economy-based social forestry initiative. The model utilizes hybrid machine learning method. This study introduces the hybrid machine learning illustrated in Figure 2.

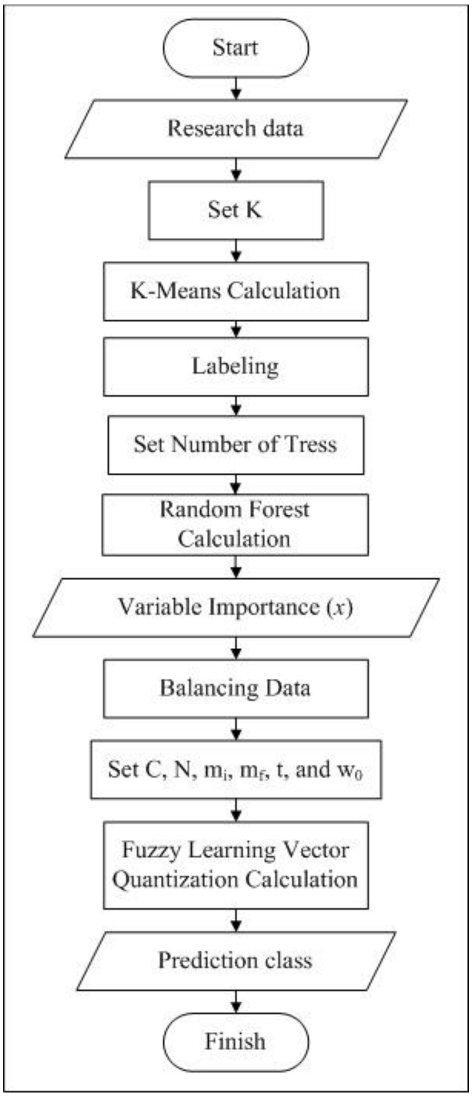

Figure 2. The proposed method

Figure 2 illustrates the flowchart of the proposed method. This proposed method incorporates algorithms from K-Means, RF, and FLVQ. The objective of this study is to achieve optimal results through the combination of these methods. From this point forward, the model will be referred to as K-Means Random Forest Fuzzy Learning Vector Quantization (KM-RF-FLVQ). The algorithm of the KM-RF-FLVQ method is as follows:

- Set *K*
- Determine the distance x and *K* with Eq. (1)
- Determine K_{new} with Eq. (2)
- Back to step 2 with K_{new}
- If vertex point update, then back to step 3. Otherwise, go to step 2
- Labeling
- Set *T*
- For $t_{RF} = 1$ to *T*, then:
  a. Determine $D_{t_{RF}}$
  b. Training $G_t(x)$ with random feature subset.
- Classify with Eq. (3)
- Balancing data
- Initializing *C, N, m_i*, and m_f
- Set $t = 0$
- Set w_0
- Calculate *t* with Eq. (4)
- Calculate *m* with Eq. (5)
- Calculate $\alpha_{ij,t}$ with Eq. (6)
- Calculate $\eta_{j,t}$ with Eq. (7)
- Calculate $w_{j,t}$ with Eq. (8)
- Calculate E_t with Eq. (9)
- If $w < N$ and $E_t > \xi$, back to step 12
- Get optimal prediction class

The KM-RF-FLVQ hybrid method is a component of an intelligent decision-making system model. This model serves as a foundation for advancing social forestry based on the green economy. The model can generate a variety of outputs, including class predictions (regional potential), variable evaluations, and analysis results that demonstrate the impact of the number of variables. The green economy and the intelligent decision-making system model based on KM-RF-FLVQ are illustrated in Figure 3.

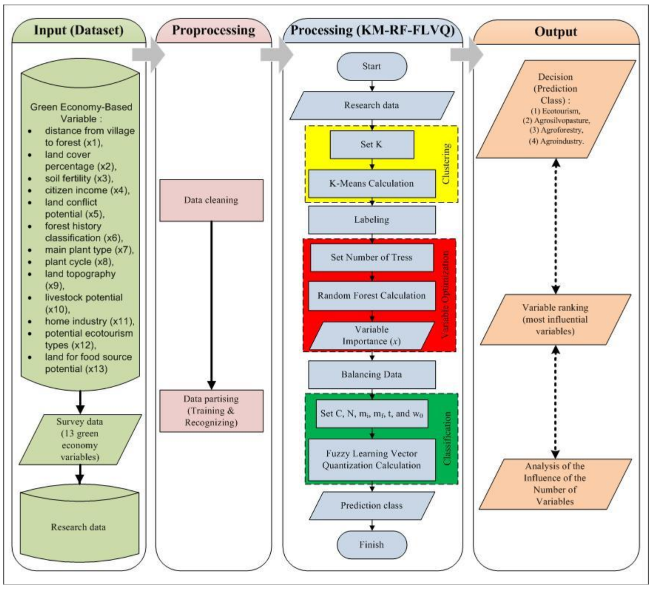

Figure 3. Model of intelligent decision making system based on green economy

A model of an intelligent decision-making system is illustrated in Figure 3. The model serves as a framework for decision-making that is informed by the green economy. Input, preprocessing, processing, and output comprise the model components. The research dataset is the input component. The assessment of forest border residents on 13 green economy-based variables is the subject of the research dataset. The preprocessing phase involves cleaning and disseminating data. Filtering data from unusable data is the process of data cleansing. The data-sharing procedure involves dividing data recognition and data training. Following preprocessing, the subsequent phase is processing. The KM-RF-FLVQ method's calculation procedure is referred to as the processing phase. The sub-parts of this section are clustering, variable optimization, and classification. The final component is output following the processing. The decision, variable ranking, and the results of the variable influence analysis comprise the model output. The decision is the outcome of the KM-RF-FLVQ prediction. (1) Ecotourism, (2) Agrosilvopasture, (3) Agroforestry, and (4) Agroindustry are the components of the prediction that determine the decision.

The government can use model outputs from a green-economy-based intelligent decision-making system to inform its decision-making process. The government's decision involves dividing social forestry areas based on the green economy and regional potential. The Ministry of Forestry is the administration. The division of areas based on potential can enhance citizens' welfare by aligning with regional conditions. For instance, the government may determine that a specific forest area is more conducive to agro-industry because of numerous cottage industries specializing in wood processing. This can be achieved by employing an intelligent decision-making system grounded in the green economy.

4. RESULTS AND DISCUSSION

4.1 Research results

This part shows the outcomes from each step of the model. The first step after gathering all the data was to perform K-means clustering with K = 4 cluster centers. K-Means is a way to provide names to data that doesn't have any. Data for this study were collected from evaluations performed by individuals residing adjacent to the forest boundary. Table 3 shows the results of the K-Means calculations.

Table 3. K-Means calculation results

Data ID	Label
Data1	Cluster_2
Data2	Cluster_2
…	…
Data1000	Cluster_0

The K-Means clustering analysis results are shown in Table 3. The clusters identified in this approach were manually grouped into different types of social forestry. Specifically, Cluster 0 corresponds to Ecotourism, Cluster 1 to Agrosilvopasture, Cluster 2 to Agroforestry, and Cluster 3 to Agroindustry. Forestry experts checked the labels to make sure they were correct. The next step after labeling the data was to improve the RF model.

This study utilized T = 20, T = 40, and T = 60 for the RF computations. RF can rank variables by their importance. In this study, a RF analysis of the labeled data identified variables with a strong effect. Table 4 shows the results of the RF.

The model's most significant variables are ranked in Table 4. In the experiment with T = 20, variable x1 had a significant impact, while variable x10 had a very minor impact. In the experiments conducted with T = 40 and T = 60, variable x1 had the most significant impact, while variable x9 had a very limited impact. The FLVQ classification stage optimises the model by leveraging the influence of these variables.

Before the classification stage, this study balanced the dataset, which comprised a total of 1,000 entries. The clustering and labelling processes led to imbalances in each class. Such imbalanced data can negatively impact the classification process. Out of the 1,000 entries, 600 were used for training, while 200 were reserved for recognition. The training dataset consisted of four classes, with each class having 150 entries. For recognition, each class used 50 entries, leading to a total of 200 recognition entries.

Table 4. Random Forest calculation results

Ranking	Variable Code		
	Experiment T = 20	Experiment T = 40	Experiment T = 60
1	x_1	x_1	x_1
2	x_3	x_3	x_3
3	x_2	x_6	x_6
4	x_6	x_2	x_4
5	x_{13}	x_8	x_8
6	x_7	x_4	x_2
7	x_8	x_{13}	x_{13}
8	x_{12}	x_7	x_7
9	x_5	x_{10}	x_{10}
10	x_{11}	x_{11}	x_{11}
11	x_4	x_{12}	x_5
12	x_9	x_5	x_{12}
13	x_{10}	x_9	x_9

After balancing the data, FLVQ was employed to classify the label samples. The experiment included three variable scenarios. The first scenario used nine variables selected based on the highest RF weights, the second utilized 11 variables with the highest RF weights, and the third involved 13 variables. The goal of testing multiple scenarios were to identify the optimal number of variables for the model. Additionally, FLVQ was classified with a varying number of iterations for each scenario, specifically using 250, 500, 750, and 1,000 iterations. The FLVQ process produced final weights, which served as a reference for the recognition phase. The first experiment used a nine-variable scenario. As evidenced by experiments conducted with T = 20, these variables were x_1, x_3, x_2, x_6, x_{13}, x_7, x_8, x_{12}, and x_5. In addition to the variable scenarios, the first experiment also used variations in the number of iterations. The results of the first experiment are presented in Tables 5 and 6.

Table 5. The results of the first experiment using nine variables, T = 20, 250 iterations, and 500 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	47	1	0	2
	Class2	1	44	2	3
	Class3	2	4	41	3
	Class4	2	2	3	43

Table 5 shows the recognition results of the first scenario. However, the iterations used 250 iterations and 500 iterations. Class1 is Ecotourism, Class2 is Agrosilvopasture, Class3 is Agroforestry, Class4 is Agroindustry. Based on Table 5, TP_{Class1} is 47, TP_{Class2} is 44, TP_{Class3} is 41, TP_{Class4} is 43, TN_{Class1} is 145, TN_{Class2} is 143, TN_{Class3} is 145, TN_{Class4} is 142, FP_{Class1} is 5, FP_{Class2} is 7, FP_{Class3} is 5, FP_{Class4} is 8, FN_{Class1} is 3, FN_{Class2} is 6, FN_{Class3} is 9, and FN_{Class4} is 7. The first scenario experiment also used 750 and 1000 iterations. The results are presented in Table 6.

Table 6 shows the recognition results of the first scenario. However, the iterations used were 750 iterations and 1000 iterations. Based on Table 6, TP_{Class1} is 47, TP_{Class2} is 45, TP_{Class3} is 41, TP_{Class4} is43, TN_{Class1} is 145, TN_{Class2} is 143, TN_{Class3} is 145, TN_{Class4} is 143, FP_{Class1} is 5, FP_{Class2} is 7, FP_{Class3} is 5, FP_{Class4} is 7, FN_{Class1} is 3, FN_{Class2} is5, FN_{Class3} is 9, and

FN_{Class4} is 7. The second experiment used an 11-variable scenario. These variables are x_1, x_3, x_2, x_6, x_{13}, x_7, x_8, x_{12}, x_5, x_{11}, dan x_4. The second experiment used variations of 250 iterations and 500 iterations, the results are presented in Table 7.

Table 6. The results of the first experiment using nine variables, T = 20, 750 iterations, and 1000 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	47	1	0	2
	Class2	1	45	2	2
	Class3	2	4	41	3
	Class4	2	2	3	43

Table 7. The results of the second experiment using 11 variables, T = 20, 250 iterations, and 500 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	48	1	0	1
	Class2	1	46	2	1
	Class3	2	4	42	2
	Class4	2	2	3	43

Table 7 shows the recognition results of the second scenario. However, the iterations used were 250 iterations and 500 iterations. Based on Table 7, TP_{Class1} is 48, TP_{Class2} is 46, TP_{Class3} is 42, TP_{Class4} is 43, TN_{Class1} is 145, TN_{Class2} is 143, TN_{Class3} is 145, TN_{Class4} is 146, FP_{Class1} is 5, FP_{Class2} is 7, FP_{Class3} is 5, FP_{Class4} is 4, FN_{Class1} is2, FN_{Class2} is 4, FN_{Class3} is 8, and FN_{Class4} is 7. The second scenario experiment also used 750 and 1000 iterations. The results are presented in Table 8.

Table 8. The results of the second experiment using 11 variables, 750 iterations, and 1000 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	48	1	0	1
	Class2	1	46	2	1
	Class3	2	3	43	2
	Class4	2	2	3	43

Table 9. The results of the second experiment using 13 variables, T = 20, 250 iterations, 500 iterations, and 750 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	49	1	0	0
	Class2	1	47	2	0
	Class3	2	3	43	2
	Class4	2	2	3	43

Table 8 also shows the recognition results from the second scenario. However, the iterations used were 750 and 1000 iterations. According to Table 8, TP_{Class1} is 48, TP_{Class2} is 46, TP_{Class3} is 43, TP_{Class4} is43, TN_{Class1} is 145, TN_{Class2} is144, TN_{Class3} is 145, TN_{Class4} is 146, FP_{Class1} is 5, FP_{Class2} was 6, FP_{Class3} was 5, FP_{Class4} is 4, FN_{Class1} is 2, FN_{Class2} is 4, FN_{Class3} is 7, dan FN_{Class4} is 7. The third experiment used a scenario with 13 variables. These variables are x_1, x_3, x_2, x_6, x_{13}, x_7, x_8, x_{12}, x_5, x_{11}, x_4, x_9, and x_{10}. The third experiment used variations of 250 iterations, 500 iterations, and 750 iterations, the results are presented in Table 9.

Table 9 shows the recognition results of the third scenario. However, the iterations used were 250 iterations, 500 iterations, and 750 iterations. Based on Table 9, TP_{Class1} is 49, TP_{Class2} is 47, TP_{Class3} is 43, TP_{Class4} is 43, TN_{Class1} is145, TN_{Class2} is 144, TN_{Class3} is 145, TN_{Class4} is 148, FP_{Class1} is 5, FP_{Class2} is 6, FP_{Class3} is5, FP_{Class4} was 2, FN_{Class1} is 1, FN_{Class2} is 3, FN_{Class3} is 7, and FN_{Class4} is 7. The third scenario experiment also used 1000 iterations. The results are presented in Table 10.

Table 10. The results of the second experiment using 13 variables, T = 20, and 1000 iterations

Actual / Predicted		Predicted Value			
		Class1	Class2	Class3	Class4
Actual Value	Class1	49	1	0	0
	Class2	1	47	2	0
	Class3	2	2	44	2
	Class4	2	2	3	43

Table 10 also shows the recognition results of the third scenario. However, the iteration uses 1000 iterations. Based on Table 10, TP_{Class1} is 49, TP_{Class2} is 47, TP_{Class3} is44, TP_{Class4} is 43, TN_{Class1} is 145, TN_{Class2} is 145, TN_{Class3} is 145, TN_{Class4} is 148, FP_{Class1} ish 5, FP_{Class2} is 5, FP_{Class3} is 5, FP_{Class4} is 2, FN_{Class1} is 1, FN_{Class2} is 3, FN_{Class3} is 6, and FN_{Class4} is 7.

Based on experiments using varying numbers of variables and iterations at T = 20, the results have various accuracy. A comparison of the accuracy results from all experiments is shown in Figure 4.

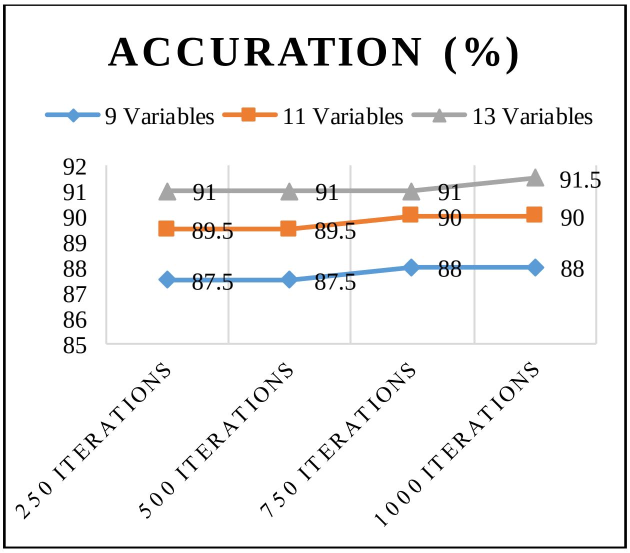

Figure 4. Comparison of accuracy in the T = 20 experiment

Figure 3 presents a comparison of the accuracy achieved in several experiments. When using nine variables, the accuracy was 87% for both 250 and 500 iterations. For 750 and 1000 iterations, the accuracy increased to 88%. By using 11 variables, the accuracy improved to 89.5% for 250 and 500 iterations, and reached 90% for 750 and 1000 iterations. When 13 variables were utilized, the accuracy was consistently 91% for 250, 500, and 750 iterations, and increased slightly to 91.5% for 1000 iterations. Overall, the highest accuracy recorded was 91.5% by using 13 variables.

The experimental scenario of variable variation and number of iterations was also carried out on the experimental results with T = 40. The difference is, the first experiment used nine variables, namely x_1, x_3, x_6, x_2, x_8, x_4, x_{13}, x_7, dan x_{10}. The second experiment used 11 variables, namely x_1, x_3, x_6, x_2, x_8, x_4, x_{13}, x_7, x_{10}, x_{11}, dan x_{12}. The third experiment used all variables (13 variables). Based on experiments using varying numbers of variables and iterations at T = 40, the results have various accuracy. A comparison of the accuracy results from all experiments is shown in Figure 5.

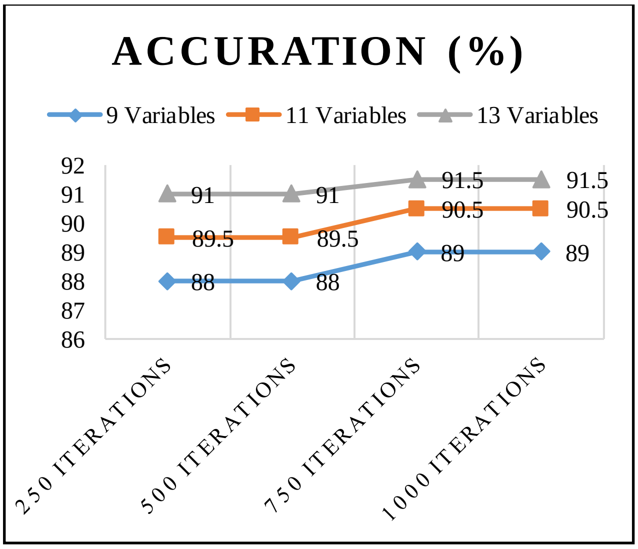

Figure 5. Comparison of accuracy in the T = 40 experiment

The accuracy comparison results from the T = 40 experiment are illustrated in Figure 5. The experiment's accuracy was 88% with 250 and 500 iterations and 9 variables. With 750 and 1000 iterations, the accuracy was enhanced to 89%. With 250 and 500 iterations, the accuracy was 89.5% when 11 variables were employed in the experiment. With 750 and 1000 iterations, the accuracy was enhanced to 90.5%. The accuracy was 91% when the experiment utilized 13 variables and was conducted with 250 and 500 iterations. With 750 and 1000 iterations, the accuracy was enhanced to 91.5%. The scenarios for the T = 20 and T = 40 experiments were also replicated for the T = 60 experiment. The primary distinction is that the initial experiment employed nine variables: x_1, x_3, x_6, x_4, x_8, x_2, x_{13}, x_7, and x_{10}. On the other hand, the second experiment employed eleven variables: x_1, x_3, x_6, x_2, x_8, x_4, x_{13}, x_7, x_{10}, x_{11}, and x_5. All 13 variables were implemented in the third experiment.

Based on experiments using varying numbers of variables and iterations at T = 60, the results have various accuracy. A comparison of the accuracy results from all experiments is shown in Figure 6.

The accuracy comparison results from the T = 60 experiment are shown in Figure 6. The experiment's accuracy was 88.5% with 250 and 500 iterations and 9 variables. With 750 and 1000 iterations, the accuracy was enhanced to 89.5%. The accuracy was 90% with 250 and 500 iterations when 11 variables were employed in the experiment. With 750 and 1000 iterations, the accuracy was enhanced to 91%. The accuracy was 91% when the experiment utilized 13 variables and was conducted with 250 and 500 iterations. With 750 and 1000 iterations, the accuracy was enhanced to 91.5%. This study achieved exceptional results, with a maximum accuracy of 91.5%, which was achieved through a series of experiments that varied the number of variables, iterations, and T values. The model's accuracy is significantly influenced by the selection of the T value. Accuracy is enhanced by an elevated T value. Nevertheless, the T = 40 and T = 60 experiments appear more robust, particularly the one that employs 13 variables.

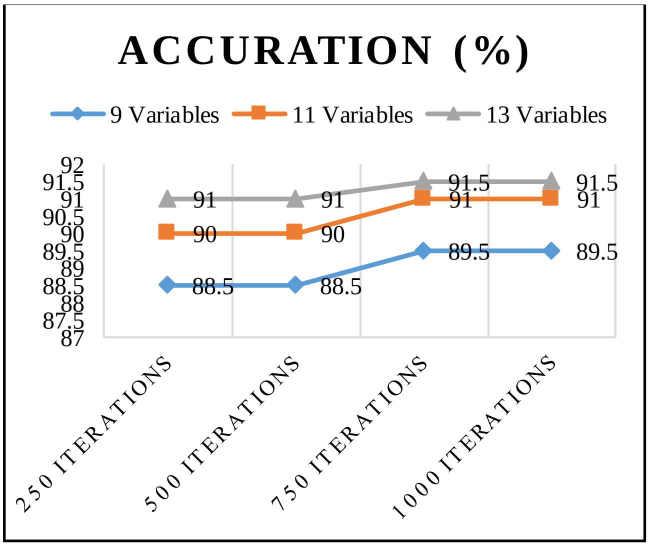

Figure 6. Comparison of accuracy in the T = 60 experiment

The model's accuracy was also influenced by experiments with different numbers of variables. The model's accuracy improved with the inclusion of variables, as indicated by the experimental results. The accuracy of experiments with 11 variables was higher than that of experiments with 9 variables. The accuracy of results from experiments with 13 variables was also higher than that of those with 9 and 11 variables.

The model's optimization was also influenced by the number of iterations. Model accuracy improved across all experiments, particularly in those that used over 750 iterations. The accuracy results remained consistent for over 750 iterations. The optimization of the intelligent decision-making system model based on the green economy was influenced by variations in the number of variables, T values, and iterations across all experiments.

The experiments were also evaluated in terms of computation time, in addition to accuracy. A personal computer with a Core i5 processor and 8 GB of RAM was employed in this investigation. Table 11 compares computational time.

The experimental variations and the computational time of each procedure are compared in Table 11. The K-Means clustering process operates at an exceptionally rapid pace in all experiments. In experiments T = 20 and T = 40, the RF procedure consumed 1 second of time. RF utilised 2 seconds of time for T = 60. The computational time is significantly influenced by the number of variables and iterations, particularly in the context of FLVQ classification. The time required for FLVQ classification increases as the number of variables and iterations increases. Memory requirements increase as the number of variables increases.

4.2 Discussion

This study achieved the highest accuracy at 91.5%. This demonstrates that the research model has excellent precision. The results were also compared to another studies, as illustrated in Table 12.

In Table 11, the research presented a comparison of the accuracy of its method with other studies. The KM-RF-FLVQ method demonstrated the highest accuracy among comparison methods.

Research in the health sector investigated several machine learning algorithms for predicting disease. The study concentrated on variable analysis. It also employed between 8 and 10 variables. The RF approach had the best accuracy, with a score of 82.26% [72]. Another study examined 18 different factors to help people with heart disease make decisions. The Gradient Boosting approach had the best accuracy, at 91.26% [73]. A study in the government sector employed a hybrid machine learning technique, K-Means FLVQ.

Table 11. Comparison of computing time

T	Var	Computing Time			
		Iterations	**K-Means Clustering**	**Random Forest Optimization**	**FLVQ Classification**
20	9	250	1 s	1 s	5 s
		500	1 s	1 s	6 s
		750	1 s	1 s	7 s
		1000	1 s	1 s	9 s
	11	250	1 s	1 s	7 s
		500	1 s	1 s	8 s
		750	1 s	1 s	9 s
		1000	1 s	1 s	10 s
	13	250	1 s	1 s	8 s
		500	1 s	1 s	9 s
		750	1 s	1 s	10 s
		1000	1 s	1 s	11 s
40	9	250	1 s	1 s	5 s
		500	1 s	1 s	6 s
		750	1 s	1 s	7 s
		1000	1 s	1 s	9 s
	11	250	1 s	1 s	7 s
		500	1 s	1 s	8 s
		750	1 s	1 s	9 s
		1000	1 s	1 s	10 s
	13	250	1 s	1 s	8 s
		500	1 s	1 s	9 s
		750	1 s	1 s	10 s
		1000	1 s	1 s	11 s
60	9	250	1 s	2 s	5 s
		500	1 s	2 s	6 s
		750	1 s	2 s	7 s
		1000	1 s	2 s	9 s
	11	250	1 s	2 s	7 s
		500	1 s	2 s	8 s
		750	1 s	2 s	9 s
		1000	1 s	2 s	10 s
	13	250	1 s	2 s	8 s
		500	1 s	2 s	9 s
		750	1 s	2 s	10 s
		1000	1 s	2 s	11 s

Table 12. Comparison of accuracy with other studies

Methods	Types of Variables	Number of Variables	Analysis of the Number of Variables	Accuracy (%)
Random Forest [72]	health variables	8, 9, 10	Yes	82.26
K-Nearest Neighbor [72]	health variables	8, 9, 10	Yes	80.55
Logistic Regression [72]	health variables	8, 9, 10	Yes	72.64
Naïve Bayesian [72]	health variables	8, 9, 10	Yes	70.56
Decision Tree [72]	health variables	8, 9, 10	Yes	81.02
Gradient Boosting [73]	health variables	18	Yes	91.26
XGBoost [73]	health variables	18	Yes	91.23
K-Means FLVQ [36]	survey variables	7	No	88.00
KM-RF-FLVQ	survey variables	9, 11, 13	Yes	91.50

The research utilized survey variable data. But it didn't talk about variable analysis. The best accuracy was 88.00%. This research presented the KM-RF-FLVQ methodology. The KM-RF-FLVQ approach serves as a foundation for an intelligent decision-making system grounded in a green economy. This framework or paradigm is helpful for starting businesses that do social forestry. The dataset used in this study consisted of survey findings. This model had the best accuracy, at 91.50%. This study also demonstrated that having more variables improves accuracy.

5.CONCLUSIONS

This research has effectively developed a model for a system capable of making smart decisions. The model is based on an environmentally friendly economy. Also, the concept helps social forestry enterprises grow in the best way possible. The KM-RF-FLVQ approach is used in the model. The KM-RF-FLVQ method is a machine learning method that combines multiple learning techniques. It brings together K-Means, RF, and FLVQ. The model's best accuracy, based on the tests, was 91.50%. In tests with 13 variables, this level of accuracy was reached. This level of accuracy is better than that of the models with 9 and 11 variables. This shows that the number of variables significantly affects model performance, especially for the hybrid machine learning method.

There are certain limits to this research. In this study, the optimization technique was equivalent to the effect of an increasing number of factors on the survey dataset. Researchers in the future could use different datasets to evaluate how well this hybrid machine learning method performs. Also, a more in-depth look at how well the model performs may be conducted with computation memory in mind.

ACKNOWLEDGMENT

We would like to extend our appreciation to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for their financial support of this research. Additionally, we would like to express our gratitude to Universitas Merdeka Madiun for providing research facilities. In addition, we would like to express our gratitude to the residents of the forest border in Ngawi Regency, East Java, Indonesia, for their collaboration in providing the assessment data for this study.

REFERENCES

[1] Supriyanto, B., Juniar Puspita, I., Nuryanto, I., Hasnawir. (2024). Integrated area development: A new social forestry landscape approach in Indonesia. IOP Conference Series: Earth and Environmental Science, 1299(1): 012006. https://doi.org/10.1088/1755-1315/1299/1/012006

[2] Raodah, P., Mulyana, S.P., Fathoni, L.A. (2024). IAD (integrated area development) agroforestry mandalika sebagai pendukung pengembangan usaha kelompok usaha perhutanan sosial (KUPS) di lombok tengah. Journal Kompilasi Hukum, 9(2): 156-168. https://doi.org/10.29303/jkh.v9i2.182

[3] Kwilinski, A., Dacko-Pikiewicz, Z., Szczepanska-Woszczyna, K., Lyulyov, O., Pimonenko, T. (2025). The role of innovation in the transition to a green economy: A path to sustainable growth. Journal of Open Innovation: Technology, Market, and Complexity, 11(2): 100530. https://doi.org/10.1016/j.joitmc.2025.100530

[4] Yakovleva, E.A., Subhonberdiev, A.S. (2019). Implementation of “green” economy principles in the forest sector. IOP Conference Series: Earth and Environmental Science, 392(1): 012016. https://doi.org/10.1088/1755-1315/392/1/012016

[5] Thathongkham, T., Sonsuphap, R. (2023). The social green economy: A perspective on a forest community project in a developing country. Corporate and Business Strategy Review, 4(4): 62-68. https://doi.org/10.22495/cbsrv4i4art7

[6] Thu, T.P., Moeliono, M., Wong, G.Y., Brockhaus, M., Dung, L.N. (2020). The politics of swidden: A case study from Nghe An and Son La in Vietnam. Land Use Policy, 99: 103050. https://doi.org/10.1016/j.landusepol.2017.10.057

[7] Thaler, G.M. (2015). Shifting cultivation, contentious land change, and forest governance in eastern Borneo. In An International Academic Conference, Chiang Mai University, pp. 1-14. https://www.iss.nl/sites/corporate/files/CMCP_33-_Thaler.pdf.

[8] Moeliono, M., Thuy, P.T., Bong, I.W., Wong, G.Y., Brockhaus, M. (2017). Social forestry-why and for whom? A comparison of policies in Vietnam and Indonesia. Forest and Society, 1(2): 78-97. https://doi.org/10.24259/fs.v1i2.2484

[9] Shavazipour, B., Sundström, L.E. (2024). Decision support for sustainable forest harvest planning using multi-scenario multiobjective robust optimization. arXiv preprint arXiv:2405.16612. https://doi.org/10.48550/arXiv.2405.16612

[10] Eberhard, B., Trailovic, Z., Magagnotti, N., Spinelli, R. (2025). A GIS-based decision support model (DSM) for harvesting system selection on steep terrain: Integrating operational and silvicultural criteria. Forests, 16(5): 854. https://doi.org/10.3390/f16050854

[11] Jauhar, S.K., Sethi, S., Kamble, S.S., Mathew, S., Belhadi, A. (2024). Artificial intelligence and machine learning-based decision support system for forecasting electric vehicles' power requirement. Technological Forecasting and Social Change, 204: 123396. https://doi.org/10.1016/j.techfore.2024.123396

[12] Li, J.C., Namvar, M., Im, G.P., Akhlaghpour, S. (2024). Machine learning based decision-making: A sensemaking perspective. Australasian Journal of Information Systems, 28: 1-22. https://doi.org/10.3127/ajis.v28.4781

[13] Yang, T.L., Lin, C.H., Chen, W.L., Lin, H.Y., Su, C.S., Liang, C.K. (2019). Hash transformation and machine learning-based decision-making classifier improved the accuracy rate of automated Parkinson’s disease screening. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(1): 72-82. https://doi.org/10.1109/TNSRE.2019.2950143

[14] Bondu, A., Achenchabe, Y., Bifet, A., Clérot, F., Cornuéjols, A., Gama, J., Hébrail, G., Lemaire, V., Marteau, P.F. (2022). Open challenges for machine learning based early decision-making research. ACM SIGKDD Explorations Newsletter, 24(2): 12-31. https://doi.org/10.1145/3575637.3575643

[15] Guha, S., Khan, F.A., Stoyanovich, J., Schelter, S. (2024). Automated data cleaning can hurt fairness in machine learning-based decision making. IEEE Transactions on Knowledge and Data Engineering, 36(12): 7368-7379. https://doi.org/10.1109/TKDE.2024.3365524

[16] Chen, Y., Zhou, Y. (2020). Machine learning based decision making for time varying systems: Parameter estimation and performance optimization. Knowledge-Based Systems, 190: 105479. https://doi.org/10.1016/j.knosys.2020.105479

[17] Prasad, B., McGeough, C., Eakin, A., Ahmed, T., Small, D., Gardiner, P., Pendleton, A., Wright, G., Bjourson, A.J., Gibson, D.S., Shukla, P. (2022). ATRPred: A machine learning based tool for clinical decision making

of anti-TNF treatment in rheumatoid arthritis patients. PLoS Computational Biology, 18(7): e1010204. https://doi.org/10.1371/journal.pcbi.1010204

[18] Barzizza, E., Biasetton, N., Ceccato, R., Molena, A. (2024). Machine learning-based decision-making approach for predicting defects detection: A case study. Machine Learning, 13(3): 3052-3060. https://doi.org/10.11591/ijai.v13.i3.pp3052-3060

[19] Pathak, T., Bhakta, S.S., Das, A., Sadhukhan, B. (2023). Development of an ensemble machine learning-based decision support system for patient admission in healthcare unit. In 2023 International Conference on Electrical, Electronics, Communication and Computers (ELEXCOM), Roorkee, India, pp. 1-6. https://doi.org/10.1109/ELEXCOM58812.2023.10370105

[20] Yang, C., Gu, W., Ito, T., Yang, X. (2021). Machine learning-based consensus decision-making support for crowd-scale deliberation: Machine learning-based consensus decision-making support for crowd-scale deliberation. Applied Intelligence, 51(7): 4762-4773. https://doi.org/10.1007/s10489-020-02118-z

[21] Lee, K.H., Choi, G.H., Yun, J., Choi, J., Goh, M.J., Sinn, D.H., Jin, Y.J., Kim, M.A., Yu, S.J., Jang, S., Lee, S.K., Jang, J.W., Lee, J.S., Kim, D.Y., Cho, Y.Y., Kim, H.J., Kim, S., Kim, J.H., Kim, N., Kim, K.M. (2024). Machine learning-based clinical decision support system for treatment recommendation and overall survival prediction of hepatocellular carcinoma: A multi-center study. NPJ Digital Medicine, 7(1): 2. https://doi.org/10.1038/s41746-023-00976-8

[22] Bougteb, Y., Akachar, E., Ouhbi, B., Frikh, B., Zemmouri, E. (2023). Deep learning-driven multi-criteria decision-making for effective recommender systems. International Journal of Intelligent Engineering & Systems, 16(5): 612-626. https://doi.org/10.22266/ijies2023.1031.52

[23] Kaliyan, K., Kothandaraman, R. (2018). Secure decision-making approach to improve knowledge management based on online samples. International Journal of Intelligent Engineering & Systems, 11(1): 50-61. https://doi.org/10.22266/ijies2018.0228.06

[24] Wibowo, R.S., Uman Putra, D.F., Puteri, I.M., Ardiyanto, A.D., Kurniawan, M., Fabian, O., Pratama, F.C. (2025). Machine learning-based dynamic reconfiguration of distribution network to improve system reliability. International Journal of Intelligent Engineering & Systems, 18(6): 522-533. https://doi.org/10.22266/ijies2025.0731.33

[25] Marco, R., Ahmad, S.S.S., Ahmad, S. (2023). An improving long short term memory-grid search based deep learning neural network for software effort estimation. International Journal of Intelligent Engineering & Systems, 16(4): 164-180. https://doi.org/10.22266/ijies2023.0831.14

[26] Irzoqe, F.N., Raheem, F.A., Nasser, A.R. (2024). Path planning improvement using a modified Q-learning algorithm based on artificial potential field. International Journal of Intelligent Engineering & Systems, 17(4): 411-423. https://doi.org/10.22266/IJIES2024.0831.32

[27] Al-Khamees, H.A., Smadi, A.A., Alsmadi, M.K., Alkannad, A.A., Abugabah, A., Almusfar, L.A., Althani, B. (2026). Advancing decision-making: A comprehensive review of intelligent systems, applications, and challenges. Intelligent Systems with Applications, 29: 200631. https://doi.org/10.1016/j.iswa.2026.200631

[28] Prawira, N., Natalia, F. (2025). Exploring the role of machine learning and big data analytics in enhancing decision-making processes: A systematic literature review. International Journal on Informatics Visualization, 9(4): 1783-1791. https://doi.org/10.62527/joiv.9.4.3244

[29] Salem, A.M., Eyupoglu, S.Z., Ma'aitah, M.K. (2024). The influence of machine learning on enhancing rational decision-making and trust levels in e-government. Systems, 12(9): 373. https://doi.org/10.3390/systems12090373

[30] Mishra, R.K., Mishra, D., Agarwal, R. (2024). Artificial intelligence and machine learning applications in forestry. Journal of Science Research International, 10(1): 43-55. https://doi.org/10.1002/9781119865728.ch11

[31] Wang, T., Zuo, Y., Manda, T., Hwarari, D., Yang, L. (2025). Harnessing artificial intelligence, machine learning and deep learning for sustainable forestry management and conservation: Transformative potential and future perspectives. Plants, 14(7): 998. https://doi.org/10.3390/plants14070998

[32] Prodhan, F.A., Hoque, M.Z., Nafee, K.M., Fahad, M.S.A., Rahman Sakib, M.N. (2025). Machine learning-based analysis of community perceptions on coastal forest ecosystem services, restoration willingness and their determinants in Bangladesh. Wild, 2(3): 26. https://doi.org/10.3390/wild2030026

[33] Zogaan, W.A., Nilashi, M., Ahmadi, H., Abumalloh, R.A., Alrizq, M., Abosaq, H., Alghamdi, A. (2024). A combined method of optimized learning vector quantization and neuro-fuzzy techniques for predicting unified Parkinson's disease rating scale using vocal features. MethodsX, 12: 102553. https://doi.org/10.1016/j.mex.2024.102553

[34] Duarte, M.L., da Silva, T.A., de Sousa, J.A.P., de Castro, A.L., Lourenço, R.W. (2024). Fuzzy inference system for mapping forest fire susceptibility in Northern Rondônia, Brazil. Geography, Environment, Sustainability, 17(1): 83-94. https://doi.org/10.24057/2071-9388-2023-2910

[35] Amezcua, J., Melin, P. (2019). A new fuzzy learning vector quantization method for classification problems based on a granular approach. Granular Computing, 4(2): 197-209. https://doi.org/10.1007/s41066-018-0120-7

[36] Utomo, P., Adi, K., Nurhayati, O.D. (2023). Success factors for citizen-based government decision making using K-means fuzzy learning vector quantization. Indonesian Journal of Electrical Engineering and Computer Science, 32(1): 506-516. https://doi.org/10.11591/ijeecs.v32.i1.pp506-516

[37] Yassin, S.S., Pooja. (2020). Road accident prediction and model interpretation using a hybrid K-means and random forest algorithm approach. SN Applied Sciences, 2(9): 1576. https://doi.org/10.1007/s42452-020-3125-1

[38] Al-Abadi, A.A.J., Mohamed, M.B., Fakhfakh, A. (2023). Enhanced random forest classifier with k-means clustering (ERF-KMC) for detecting and preventing distributed-denial-of-service and man-in-the-middle attacks in internet-of-medical-things networks. Computers, 12(12): 262. https://doi.org/10.3390/computers12120262

[39] Falfushynska, H. (2025). A multidimensional assessment

of CO_2-intensive economies through the green economy index framework. Environments, 12(6): 195. https://doi.org/10.3390/environments12060195

[40] Mohapatra, N., Shreya, K., Chinmay, A. (2019). Optimization of the random forest algorithm. Advances in Data Science and Management: Proceedings of ICDSM, 37: 201-208. https://doi.org/10.1007/978-981-15-0978-0_5

[41] Yan, S., Su, Q., Gong, Z., Zeng, X., Herrera-Viedma, E. (2023). Online public opinion prediction based on rolling fractional grey model with new information priority. Information Fusion, 91: 277-298. https://doi.org/10.1016/j.inffus.2022.10.012

[42] Pardede, J., Dwianto, R. (2025). The effect of feature selection on machine learning classification. JOIV: International Journal on Informatics Visualization, 9(4): 1419-1429. https://doi.org/10.62527/joiv.9.4.2926

[43] Ju, K.B., Jang, D.W. (2025). Variable selection and model comparison for optimizing machine learning-based TOC prediction. Water, 17(23): 3367. https://doi.org/10.3390/w17233367

[44] Godinho, N., Silva, H., Curado, M., Paquete, L. (2022). A reconfigurable resource management framework for fog environments. Future Generation Computer Systems, 133: 124-140. https://doi.org/10.1016/j.future.2022.03.015

[45] Coppola, A., Lui, D.G., Petrillo, A., Santini, S. (2023). Cooperative driving of heterogeneous uncertain nonlinear connected and autonomous vehicles via distributed switching robust PID-like control. Information Sciences, 625: 277-298. https://doi.org/10.1016/j.ins.2023.01.045

[46] Malah, M., Hemam, M., Abbas, F. (2023). 3D face reconstruction from single image with generative adversarial networks. Journal of King Saud University - Computer and Information Sciences, 35(1): 250-256. https://doi.org/10.1016/j.jksuci.2022.11.014

[47] Loucif, H. (2024). A hybrid deep learning approach for spam detection in Twitter. Ingénierie des Systèmes d'Information, 29(1): 117-123. https://doi.org/10.18280/isi.290113

[48] Deshmukh, P.V., Shahade, A.K., Shahade, M.R., Wankhede, D.S., Gohatre, P.H. (2025). Breast cancer detection using a novel hybrid machine learning approach. Ingenierie des Systemes d'Information, 30(3): 565-576. https://doi.org/10.18280/isi.300301

[49] Saravi, B., Hassel, F., Ülkümen, S., Zink, A., et al. (2022). Artificial intelligence-driven prediction modeling and decision making in spine surgery using hybrid machine learning models. Journal of Personalized Medicine, 12(4): 509. https://doi.org/10.3390/jpm12040509

[50] Parihar, H., M, S. (2024). Decision making in monopoly using a hybrid deep reinforcement learning approach. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 12(2): 58-67. https://doi.org/10.22214/ijraset.2024.58270.

[51] Okfalisa, O., Hafsari, R., Nawanir, G., Toto, S., Yanti, N. (2021). Optimizing placement of field experience program: An integration of MOORA and rule-based decision making. Pertanika Journal of Science & Technology, 29(2): 895-918. https://doi.org/10.47836/pjst.29.2.11

[52] Lewis, H.G., Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. International Journal of Remote Sensing, 22(16): 3223-3235. https://doi.org/10.1080/01431160152558332

[53] Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. Behavioural Processes, 148: 56-62. https://doi.org/10.1016/j.beproc.2018.01.004

[54] AlShammari, A.F. (2024). Implementation of model evaluation using confusion matrix in Python. International Journal of Computer Applications, 186(50): 42-48. https://doi.org/10.5120/ijca2024924236

[55] Guan, L., Merigó, J.M., Beydoun, G. (2025). 40 years of decision support systems: A bibliometric analysis. Decision Support Systems, 194: 114469. https://doi.org/10.1016/j.dss.2025.114469

[56] Hall, S.J., Smith, S., Baskent, C., De Raffaele, C. (2024). A taxonomy of knowledge management systems in the micro-enterprise. International Journal of Operations Management, 4(1): 7-20. https://doi.org/10.18775/ijom.2757-0509.2020.41.4001

[57] Shokare, C. (2025). Enhancing decision support systems with hybrid machine learning and operations research models. Asian Journal of Science, Technology, Engineering, and Art, 3(2): 240-253. https://doi.org/10.58578/AJSTEA.v3i2.4933

[58] Shulajkovska, M., Smerkol, M., Noveski, G., Bohanec, M., Gams, M. (2024). Artificial intelligence-based decision support system for sustainable urban mobility. Electronics, 13(18): 3655. https://doi.org/10.3390/electronics13183655

[59] Barkah, A.S., Selamat, S.R., Abidin, Z.Z., Wahyudi, R. (2023). Impact of data balancing and feature selection on machine learning-based network intrusion detection. JOIV: International Journal on Informatics Visualization, 7(1): 241-248. http://doi.org/10.30630/joiv.7.1.1041

[60] Sureja, N., Vasant, A., Chaudhari, N. (2021). A hybrid shuffled frog-simulated annealing algorithm for clustering. International Journal of Intelligent Engineering and Systems, 14(4): 490-499. https://doi.org/10.22266/ijies2021.0831.43

[61] Mukti, F.S., Junikhah, A., Putra, P.M.A., Soetedjo, A., Krismanto, A.U. (2022). A clustering optimization for energy consumption problems in wireless sensor networks using modified k-means++ algorithm. International Journal of Intelligent Engineering & Systems, 15(3): 355-365. https://doi.org/10.22266/ijies2022.0630.30

[62] Periyasamy, S., Khara, S., Thangavelu, S. (2016). Balanced cluster head selection based on modified k-means in a distributed wireless sensor network. International Journal of Distributed Sensor Networks, 12(3): 5040475. https://doi.org/10.1155/2016/5040475

[63] Liu, J., Zhou, Z., Kong, S., Ma, Z. (2022). Application of random forest based on semi-automatic parameter adjustment for optimization of anti-breast cancer drugs. Frontiers in Oncology, 12: 956705. https://doi.org/10.3389/fonc.2022.956705

[64] Zhu, N., Zhu, C., Zhou, L., Zhu, Y., Zhang, X. (2022). Optimization of the random forest hyperparameters for power industrial control systems intrusion detection using an improved grid search algorithm. Applied Sciences, 12(20): 10456.

https://doi.org/10.3390/app122010456

[65] Karayiannis, N.B., Bezdek, J.C. (2002). An integrated approach to fuzzy learning vector quantization and fuzzy c-means clustering. IEEE Transactions on Fuzzy Systems, 5(4): 622-628. https://doi.org/10.1109/91.649915

[66] Kusumadewi, S., Purnomo, H. (2010). Neuro-Fuzzy; Integrasi Sistem Fuzzy & Jaringan Syaraf Edisi 2. Yogyakarta, Indonesia: Graha Ilmu.

[67] Alamri, S.S.A., Bin-Sama, A.S.A., Bin-Habtoor, A.S.Y. (2016). Satellite image classification by using distance metric. International Journal of Computer Science And Information Security, 4(3): 65-68. https://doi.org/10.6084/m9.figshare.3153877

[68] Suwanda, R., Syahputra, Z., Zamzami, E.M. (2020). Analysis of euclidean distance and manhattan distance in the K-means algorithm for variations number of centroid K. Journal of Physics: Conference Series, 1566(1): 012058. https://doi.org/10.1088/1742-6596/1566/1/012058

[69] Heydarian, M., Doyle, T.E., Samavi, R. (2022). MLCM: Multi-label confusion matrix. IEEE Access, 10: 19083-19095. https://doi.org/10.1109/ACCESS.2022.3151048

[70] Hossain, M.F., Adnan, Z.H., Hasin, M.A.A. (2014). Improvement in weighting assignment process in analytic hierarchy process by introducing suggestion matrix and Likert scale. International Journal of Supply Chain Management, 3(4): 91-95. https://doi.org/10.59160/ijscm.v3i4.1000

[71] Utomo, P., Budiman, A., Amadi, D.N. (2023). Optimasi metode simple additive weighting menggunakan skala likert dalam aplikasi pemilihan peserta rapat bersih dusun. Journal of Computer and Information Systems Ampera, 4(3): 141-156. https://doi.org/10.51519/journalcisa.v4i3.411

[72] Chang, V., Ganatra, M.A., Hall, K., Golightly, L., Xu, Q.A. (2022). An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. Healthcare Analytics, 2: 100118. https://doi.org/10.1016/j.health.2022.100118

[73] Alqarni, A. (2025). Analysis of decision support systems (DSS) integration with AI using mobile health (m-health) and wearables. Intelligent Information Management, 17(6): 235-262. https://doi.org/10.4236/iim.2025.176013