Application of Machine Learning Techniques to Classify Intention to Pay for Forest Ecosystem Services

Pham Thu Thuy[1,2](✉), Nguyen Thanh Tung[3], and Luu Quoc Dat[4]

[1] VNU School of Interdisciplinary Studies, Vietnam National University, 144 Xuan Thuy Road, Hanoi 100000, Vietnam
phamthuthuy@vnu.edu.vn
[2] Science and Technology Department, Vietnam National University, 144 Xuan Thuy Road, Hanoi 100000, Vietnam
[3] International School, Vietnam National University, 144 Xuan Thuy Road, Hanoi 100000, Vietnam
tung_nt@vnu.edu.vn
[4] VNU University of Economics and Business, Vietnam National University, 144 Xuan Thuy Road, Hanoi 100000, Vietnam

Abstract. Capturing the ability to take part in the payment of forest ecosystem services by beneficiaries is the result that policy-making agencies are always concerned. This research selects several machine learning techniques, including single classifiers (Multilayer Perceptron, Naive Bayes, SMO) and ensemble classifiers (LogitBoost, Random Forest, Bagging) to evaluate and classify willingness-to-pay intention for mangrove ecosystem services of people in PhuLong commune, Vietnam. Research data is inherited from a previous contingent valuation survey, with a sample size of 235. The results show that the machine learning algorithms are workt with small sample-size data sets with feasibility prediction results in behavioral intent classification. The LogitBoost model achieves the best classification performance compared to the remaining models. Besides, socio-psychological factors are ranked as important factors in classifying behavioral intentions related to payment for forest ecosystem services.

Keywords: Intention to pay · forest ecosystem services · classification · machine learning

1 Introduction

Forest ecosystems provide public goods characterized by non-excluded and non-competitive nature (Tóth et al., 2010). The value of ecosystem services is often determined through hypothetical markets (established in the Contingent valuation method) rather than traditional markets (Asmare et al., 2022; Obeng Aguilar, 2021; Getachew, 2018). There is ample evidence to show that climate change and economic activities

P. Cong Vinh and N. Thanh Tung (Eds.): ICCASA 2023, LNICST 579, pp. 54–69, 2024.
https://doi.org/10.1007/978-3-031-58878-5_4

to maximize forest-related benefits such as logging, land use conversion, and forest product exploitation is a major causes of the decline in the quality and quantity of forest ecosystem services worldwide (Roesch Rabotyagov, 2016; Taylor et al., 2015; Van, 2012; Raudsepp-Hearne et al., 2010).

In climate change, strengthening forest ecosystem services and the potential exploitation of forests has become a global concern in recent decades (IPCC, 2022; Pagiola et al., 2004; Edward, 2004). According to Obeng Aguilar (2018), when implementing ecosystem service payment programs, it is necessary to classify and evaluate the public's readiness (willingness to pay) for these programs to ensure the financial viability of payment programs.

Many statistical analysis methods have been used to identify the critical factors that affect respondents' decisions about payment intent related to the environment in a multivariate regression model (State et al., 2016; Mojiol et al., 2022), Logit regression model (Pham et al., 2018; Pouta Rekola, 2001), Tobit model (Asmare et al., 2022; Tran et al., 2002; Pouta Rekola, 2001), Structured models (Li et al., 2015; Lin Syrgabayeva, 2016; Liu et al., 2020). According to Sokratous et al. (2023), probabilistic models come with a wide range of problems stemming from reality, (i) many types of behavior can arise from stochastic processes, (ii) many theoretical models have never been explored because they do not have convenient probability density functions for different combinations of parameters, (iii) it is challenging to combine all the variables from many questions to determine conditional probability.

Machine learning models began to be approached as an alternative in several studies related to pay preferences (Sokratous et al., 2023). Phan et al. (2021) have adopted a logistic regression model and Bayes network to identify determinants that affect willingness to pay (willingness to pay) for reservoir construction, increasing water prices. Subhan et al. (2023) also used machine learning and quantum regression approaches to research willingness to pay to enhance road safety. Ayansola et al. (2022) used K-Nearest Neighbors, Random Forest, Support Vector Machine Decision Tree, and Boosting to model consumers' willingness to pay for electricity.

Machine learning methods can use data from a variety of sources or integrate survey questions into models without transforming or combining survey data (Shen et al., 2021; Thuy et al., 2021). Besides, when applied to a new situation that has never appeared, the trained model will make predictions using the trained generalization samples (Crook et al., 2007; Sokratous et al., 2023). Many studies have shown that machine learning models can provide information with higher accuracy and reliability than traditional regression models (Bashar et al., 2023; Zeng et al., 2019; Selz, 2020).

Machine learning algorithms often prioritize large-sample-size data sets due to the growing capacity of computing systems; the more significant the data size, the greater the statistical capacity. In addition, recent data collection methods are becoming cheaper and more accessible (Vabalas et al., 2019; Raudys Jain, 1991). On the other hand, studies based on traditional methods, such as contingent valuation (CVM), often require face-to-face interviews. The US National Oceanic and Atmospheric Administration (NOAA) recommends that face-to-face interviews minimize bias when implementing CVM. However this way of implementation often requires significant expense, so large data samples are often not preferred. Therefore, developing machine learning methods to solve problems with small datasets needs further exploration.

Although machine learning methods are increasingly applied to solve environmental management issues, the literature review shows that it needs to be of more use in classifying human preferences for ecosystem services. The objective of this research are, (i) to apply machine learning models to classify people's willingness to pay for mangrove ecosystem services, (ii) discover whether machine learning models give positive results on a small-sample-sizes dataset that collected earlier through CVM, (iii) evaluates the performance results of models with data sets with and without important feature options.

The rest of this study is organized as follows, Sect. 2 proposes the research design. Section 3 presents the research methodology. Section 4 overviews the research data. Section 5 presents about research results. Finally Discussion and Conclusion are drawn in Sect. 6 (Fig. 1).

2 Research Design

2.1 Research Design Diagram

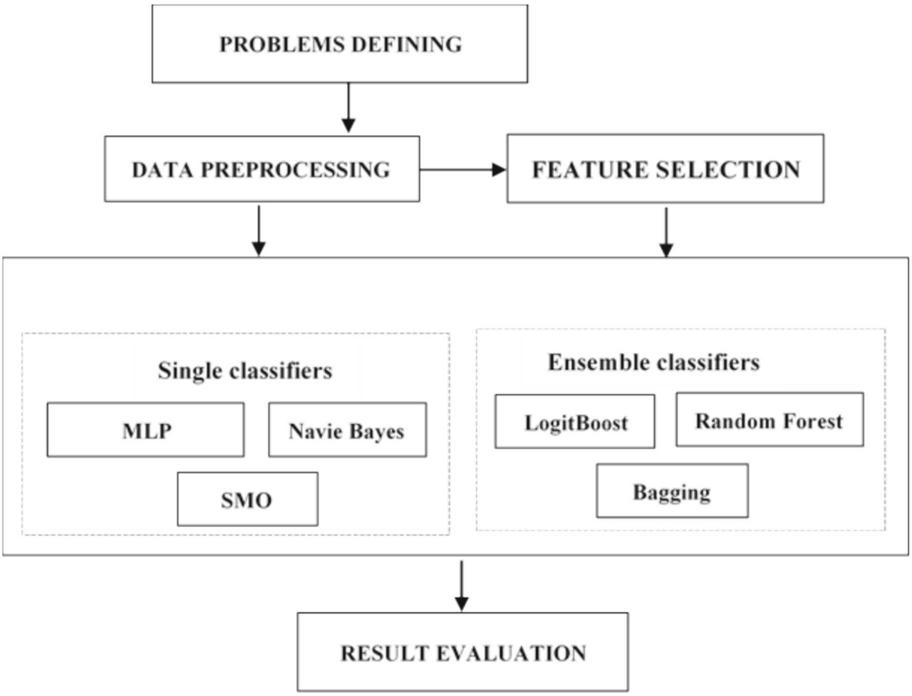

Fig. 1. Research design model (author)

2.2 Implementation Steps

Step 1, Problems defining.
The research problem can be stated as follows:
For dataset X, $X = \{x_1, x_2, x_3, \ldots, x_n\}$, n is the number of sample, xn is the vector including the demographic, socio-economic, and psychological characteristics of the nth object. Set of classification labels W, $W = \{ w_1, w_2, w_3, \ldots.. w_n\}; w_n \in \{1; 0\}$. The problem is giving correct prediction of a label for each sample of dataset.

Step 2, Data Preprocessing.

Data preprocessing involves transforming raw data into a format the machine can comprehend. Some specific fields of information in the dataset might lack meaning or fail to meet the conditions for execution, necessitating removal or normalization.

After preprocessing, the dataset will be cleaned, standardized, and encoded into categories and sets, as the machine learning model requires.

Step 3.1, Features Selection.

Machine learning models often come with diverse datasets in terms of features. Feature selection is necessary to ensure accurate predictive results. This work helps reduce the number of parameters or training time and overfitting issues. It also aims to achieve a balanced classification and establish the best model with high prediction capability and minimal error.

Step 3.2. Intention to pay classification.

In this study, single classifiers and ensemble classifiers were chosen for experimentation. The selected single classifiers include Multilayer Perceptron, Naive Bayes, and SMO (SVO). The selected ensemble classifiers include LogitBoost, Random Forest, and Bagging. The aim was to evaluate and compare the performance of these classifiers in the context of the research.

Step 4, Evaluate and compare attribute classification results and model performance.

2.3 Evaluation Metrics

The evaluation metrics Accuracy, Precision, Recall, and F1-score were chosen to classify and evaluate the performance of the machine learning methods, Accuracy, Precision, Recall, and F1-score. These metrics are widely used and commonly employed (Goutte Gaussier, 2005). However, the relative importance of each metric depends on the specific problem and the associated costs related to classification outcomes for each task. These criteria are measured on a scale from 0 (terrible classification) to 1 (perfect classification) using the following equations,

$$\text{Accuracy} = \frac{\text{Number of true predictions}}{\text{Total number of predictions}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F1 - \text{score} = \frac{2 \text{ x } (\text{Precision x Recall})}{\text{Precision} + \text{Recall}}$$

In there, TP, True Positive; TN, True Negative; FP, False Positive; FN, False Negative.

2.4 Experimental Setups

The k-fold cross-validation technique is employed to assess the predictive performance. This approach divides the dataset into k subsets of equal size. One subset is used as validation data, while the remaining (k-1) subsets are used as training data to parameterize the models. The training and testing process of the models is repeated until all k subsets have been used. For this study, the chosen value of k for the models' execution is 10.

2.5 Tools for Experiments

In this study, the Weka software version 3.8.6 has been chosen as the tool to conduct machine learning algorithms. Weka also provides tools to preprocess data and select features in this research.

3 Research Methods

3.1 Single Classifiers Techniques

3.1.1 Sequential Minimal Optimization (SMO)

Sequential Minimal Optimization (or SMO) is an algorithm used for training Support Vector Machine (SVM), a type of machine learning model commonly used for classification and regression tasks (Liang et al., 2017; Naik Desai, 2017). SMO is designed to efficiently solve the optimization problem of training an SVM, which involves finding a solution that maximizes the margin between different classes while minimizing classification errors (Platt, 1998). SMO breaks this problem into a series of smallest possible sub-problems, which are then solved analytically. Because of the linear equality constraint involving the Lagrange multipliers αi, the smallest possible problem involves two such multipliers. Then, for any two multipliers α_1 and α_2, the constraints are reduced to,

$$0 \leq \alpha_1, \alpha_2 \leq C$$

$$y_1 a_1 + y_2 a_2 = k$$

Moreover, this reduced problem can be solved analytically, one must find a minimum of a one-dimensional quadratic function. k is the negative of the sum over the rest of the terms in the equality constraint, which is fixed in each iteration. The algorithm proceeds as follows,

- Find a Lagrange multiplier $\alpha 1$ that violates the Karush–Kuhn–Tucker (KKT) conditions for the optimization problem.
- Pick a second multiplier $\alpha 1$ and optimize the pair (α_1, α_2).

- Repeat steps 1 and 2 until convergence.

When all the Lagrange multipliers satisfy the KKT conditions (within a user-defined tolerance), the problem has been solved. Although this algorithm is guaranteed to converge, heuristics are used to choose the pair of multipliers to accelerate the convergence rate, and this is critical for large data sets since a number p represents the possible choices for α_i and α_j. This number p is calculated by this formula,

$$p = \frac{n(n-1)}{2}$$

Although SMO was designed for SVM training, it is worth mentioning that there are other optimization techniques and libraries available that can train SVMs effectively as well. Nonetheless, SMO remains a foundational algorithm in the SVM literature and has contributed to the widespread use of SVMs in various machine-learning applications (Noronha et al., 2019).

3.1.2 Multilayer Perceptron (MLP)

A Multilayer Perceptron is an artificial neural network consisting of multiple layers of interconnected nodes, or "neurons". It is a fundamental architecture used in machine learning and deep learning for various tasks such as classification, regression, and pattern recognition (Botalb et al., 2018). Neurons within an MLP are organized feedforward, meaning the data flows from the input layer through the hidden layers to the output layer without forming any cycles. Each connection between neurons has an associated weight that the network adjusts during training to optimize its performance on the given task.

Training an MLP involves feeding it with labeled data (input-output pairs) and using a process called backpropagation to adjust the weights based on the difference between predicted and target outputs. This iterative optimization process aims to minimize a chosen loss function, which quantifies the difference between predictions and ground truth. The two historically common activation functions are both sigmoid functions which are described by two formulas,

$$y(v_i) = \tanh(v_i)$$

$$y(v_i) = \left(1 + e^{-v_i}\right)^{-1}$$

MLPs are considered shallow neural networks compared to more complex architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs). However, they can still model various functions and have been used successfully in various applications.

3.1.3 Naive Bayes

Naive Bayes is a probabilistic machine learning algorithm for classification and sometimes regression tasks. It is based on Bayes' theorem and the assumption of conditional independence among features given the class label (Chaudhary et al., 2016). Despite the simplifying assumption of independence, Naive Bayes can be surprisingly effective for many real-world problems. At the core of Naive Bayes is Bayes' theorem, which calculates the probability of a hypothesis (class label) given the observed evidence (features). It is expressed as the below mathematical formula,

$$\mathrm{p}\ (\mathrm{C}_k |\ \mathrm{x}) = \frac{\mathrm{p}\ (\mathrm{C}_k) * \mathrm{p}\ (\mathrm{x} \mid \mathrm{C}_k)}{\mathrm{p}\ (\mathrm{x})}$$

The fraction's numerator plays a vital role in the equation because of the independence between C and the value of feature $\mathrm{x_i}$. It equals to model p $(\mathrm{C_k}, \mathrm{x_1}, \ldots, \mathrm{x_n})$. Thus, this equation was rewritten using the chain rule for repeated applications of the definition of conditional probability.

$$\begin{aligned}
\mathrm{p}(\mathrm{C_k}, \mathrm{X_1}, \ldots, \mathrm{X_n}) &= \mathrm{p}(\mathrm{X_1}, \mathrm{X_2}, \ldots, \mathrm{X_n}, \mathrm{C_k}) \\
&= \mathrm{p}(\mathrm{X_1}|\mathrm{X_2}, \ldots, \mathrm{X_n}, \mathrm{C_k}).\mathrm{p}(\mathrm{X_2}, \ldots, \mathrm{X_n}, \mathrm{C_k}) \\
&= \mathrm{p}(\mathrm{X_1}|\mathrm{X_2}, \ldots, \mathrm{X_n}, \mathrm{C_k}).\mathrm{p}(\mathrm{X_2}|\mathrm{X_3}, \ldots, \mathrm{X_n}, \mathrm{C_k}).\mathrm{p}(\mathrm{X_3}, \ldots, \mathrm{X_n}, \mathrm{C_k}) \\
&= \ldots \\
&= \mathrm{p}(\mathrm{X_1}|\mathrm{X_2}, \ldots \mathrm{X_n}, \mathrm{C_k}).\ \mathrm{p}(\mathrm{X_2}|\mathrm{X_3}, \ldots, \mathrm{X_n}, \mathrm{C_k}).....\mathrm{p}(\mathrm{X_n} - 1|\mathrm{X_n}, \mathrm{C_k}) \\
&\quad .\ \mathrm{p}(\mathrm{X_n}|\mathrm{C_k}).\ \mathrm{p}(\mathrm{Ck})
\end{aligned}$$

Naive Bayes has limitations, mainly when the independence assumption does not hold actual or features are highly correlated (Saritas Yasar, 2019). However, its simplicity, efficiency, and ability to perform well on specific tasks make it a valuable tool in many machine-learning applications.

3.2 Ensemble Classifiers Techniques

3.2.1 Logit Boost

Logit Boost is an ensemble machine-learning algorithm that belongs to the boosting family of techniques. Boosting is a method that combines multiple weak learners (typically simple models) into strong learners, iteratively focusing on the instances that the current ensemble is struggling to classify correctly (Tehrany et al., 2019). (Tehrany et al., 2019). Logit Boost is designed for binary classification tasks. Logit Boost can be interpreted as a convex optimization process. To elaborate, when we aim to obtain an additive model in the formula (Dettling Bühlmann, 2003),

$$f = \sum_t \alpha_t h_t$$

Thus, the LogitBoost algorithm reduces the logistic loss to the lowest value,

$$\sum_i \log\left(1 + e^{-y_i f(x_i)}\right)$$

Work Process of LogitBoost,

- Initialization, Initialize the training data weights. Initially, all data points are given equal weights.
- Iterative Process, LogitBoost performs a series of iterations
- Final Prediction, The final prediction is made by combining the predictions of all weak learners in the ensemble, weighted by their corresponding boosting coefficients.

LogitBoost is a robust algorithm that has demonstrated exemplary performance in various applications. However, tuning hyperparameters is essential to avoid overfitting and achieve optimal results like other boosting methods.

3.2.2 Random Forest

Random Forest is a powerful ensemble learning algorithm used for classification and regression tasks in machine learning (Chaudhary et al., 2016). It is particularly effective for handling complex datasets and producing accurate predictions. The algorithm belongs to the family of ensemble methods, which combine the predictions of multiple individual models to create a more robust overall model (Liu et al., 2012).

The initial step in assessing variable importance within dataset A is expressed by the formula below, First, a random forest is fitted to the dataset. During this fitting process, the out-of-bag error is calculated for each data point and averaged across the entire forest. If bagging is not employed during training, this can be replaced with errors from an independent test set.

$$D_n = \{X_i, Y_i\}_{i-1}^{n}$$

To gauge the importance of feature rank j post-training, the values corresponding to the jth feature rank are permuted across the training data. Subsequently, the out-of-bag error is recalculated on this perturbed dataset. The critical score for feature rank j is computed by taking the average of the differences in out-of-bag error before and after the permutation across all trees. This score is then normalized using the standard deviation of these computed differences.

Random Forests are widely used in practice and have found applications in various fields, including finance, medicine, marketing, and more. They are relatively easy to use and require fewer tunable hyperparameters than other complex algorithms. The only primary parameter that typically needs to be adjusted is the number of trees in the ensemble.

3.2.3 Bagging

Bagging, short for "Bootstrap Aggregating", is an ensemble learning technique used to improve the accuracy and robustness of machine learning models, especially decision trees (Breiman, 1996). It involves training multiple instances of the same model on different subsets of the training data and then combining their predictions to make a final prediction (Yaman Subasi, 2019). The goal of bagging is to reduce overfitting and increase the model's generalization performance. The training process for random forests employs the fundamental method of bootstrap aggregating, often called bagging, in the context of tree learners. In the presence of a training set $X = x_1, \ldots, x_n$ with corresponding responses $Y = y_1, \ldots, y_n$, the bagging procedure involves iteratively selecting a random sample from the training set, employing replacement, and then constructing trees based on these samples. This process is repeated B times, and for each iteration $b = 1, \ldots,$ B, the process begins by initially selecting a random sample, with replacement, of n training examples from the sets X and Y. This sampled subset is denoted as X_b and Y_b. Subsequently, a classification or regression tree f_b is trained using the data from X_b and Y_b. Following the training phase, predictions for previously unseen samples x' are generated by averaging the predictions derived from all individual regression trees on x',

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_B(x')$$

Bagging is not limited to decision trees but can also be applied to other base models (Yaman Subasi, 2019). As previously mentioned, Random Forest is a specific ensemble method that uses bagging with decision trees as its base models. Bagging can also be a foundation for more advanced ensemble techniques like boosting. Overall, bagging is a powerful tool to enhance the performance and reliability of machine learning models.

4 Research Data

The research data is taken from the survey results of households living in Phu Long commune, Hai Phong City, Vietnam, in 2022. The survey questionnaire is designed according to the Contingent valuation method with this dataset. Respondents in the survey are heads of households. In the interview scenario, respondents were first provided information about the status of mangroves. After that, they were asked to answer questions to explore the intention to pay (willingness to pay) for benefiting from mangrove ecosystem services through the forest payment project.

The database has a sample size of 235. Demographic, socio-psychological factors, and socio-economic characteristics classify intention to pay. The data structure is summarized in Table 1. The data fields are normalized and labeled to ensure the execution conditions of machine learning algorithms. One hundred eighty (180) samples are labeled willing to pay with the value "yes", and fifty-five (55) samples are labeled unwilling to pay with the value "no" which are encoded as 1 and 0, respectively.

Table 1. Data structure

Variables symbol	Variables name	Measurement
Socio-psychological factors		The mean value of observed variables
Att	Attitude	Real
Sjn	Subjective norms	Real
Pbc	Perceived behavioral control	Real
Knl	Knowledge	Real
Wtp (*predictor variable*)	Willingness to pay for forest environment service	{1; 0}
Socio-economic characteristics		
Age	Age of household head	Natural number
lnIncome	Average household income/month (million VietNam Dong)	Real
Edu	Education level of household head	1, Primary school 2, Secondary school 3, High school 4, College 5, University 6, Postgraduate
Gender	Gender of household head	1, Male 2, Female
Job	The main occupation of the household	1, Raising, planting, and exploiting seafood; 2, Other
Hhsize	Number of people in the household	Natural number

5 Result

Feature selection is crucial in determining relevant features from a dataset for classification and prediction purposes. The classification performance is expected to improve when the dataset contains predictive variables with high-value contributions. This research uses two technicals to feature selection, Information Gain and Correlation Attribute Ranking Filters. From the original dataset, five attributes that significantly influence the dependent variable are ranked in descending order of importance, as shown in Table 2. The results indicate that most of the socio-psychological variables play a dominant role in machine learning models for classifying willingness-to-pay intention. Only the "income" variable holds relative importance in the model among the demographic and socio-economic variables.

Table 2. Ranking values of Information Gain and Correlation Attributes Evaluator

Attributes	Information Gain ranking values	Correlation Attributes ranking values
Att	0.250	0.579
Knl	0.209	0.484
Sjn	0.098	0.366
Pbc	0.072	0.327
Income	0.053	0.274
Job	0.004	0.171
Gender	0.002	0.085
Hhsize	0	0.081
Edu	0	0.052
Age	0	0.048

Table 3 shows the result of using the full 11 parameters (including age, edu, gender, hhsize, job, pbc, att, knl, sjn, income, wtp) for machine learning models. In this case, all six selection algorithms obtained relatively good accuracy ($\geq$ 80%). In the Ensemble classifiers group, the LogitBoost model gave a significantly better accuracy than Bagging and Random Forest. The values achieved were 87.234%, 83.404%, and 87.234%, respectively.

Table 3. Experimental results of classification of willingness to pay using machine learning algorithms with full input parameters

Machine learning methods	Accuracy	Precision	Recall	F1-Score
Single classifiers				
Naive Bayes	84.255%	0.838	0.843	0.840
SMO	84.680%	0.840	0.840	0.830
Multilayer Perceptron	80.000%	0.792	0.800	0.795
Ensemble classifier				
LogitBoost	87.234%	0.867	0.872	0.865
Random Forest	84.255%	0.835	0.843	0.836
Bagging	83.404%	0.824	0.834	0.824

Table 4 shows the results of conducting machine learning algorithms with six (06) essential parameters selected by two technicals for feature selection before (including pbc, att, knl, sjn, income, and WTP). After reducing the input parameters, the algorithms in the single classifiers group tended to increase performance, while the ensemble classifiers group tended to differ from the three (3) algorithms.

After reducing the input parameters, the algorithms in the single classifiers group increased accuracy, while the ensemble classifiers group tended to differ from the three algorithms. Multilayer Perceptron has the most impressive change (from 80% to 85.1%). On the contrary, Random Forest is the only algorithm out of the six models with an accuracy reduction (84.255% to 82,128%). The LogitBoost model is almost unaffected regarding results when reducing the number of model input parameters.

Table 4. Experimental results of classification of willingness to pay using machine learning algorithms and important attribute selection

Machine learning methods	Accuracy	Precision	Recall	F1-Score
Single classifiers				
Naive Bayes	84.680%	0.845	0.847	0.846
SMO	85.957%	0.856	0.860	0.847
Multilayer Perceptron	85.106%	0.843	0.851	0.844
Ensemble classifier				
LogitBoost	87.234%	0.867	0.872	0.865
Random Forest	82.128%	0.819	0.821	0.820
Bagging	83.405%	0.824	0.834	0.824

In both cases, LogitBoost is a model for better results than other methods for all evaluation indicators, details shown in Table 5. The correct accuracy reached 87.234%, and the F1-Score values for classes "1" and "0" were 0.920 and 0.688 respectively. The Precision and Recall scores for the case "willing to pay" is relatively high, respectively 0.865 and 0.956. However, for the case of "not willing to pay", the Precision and Recall scores have relative differences. The values are 0.805 and 0.600, respectively.

Table 5. LogitBoost model experimental results (best)

Machine learning methods	Accuracy		Class	Precision	Recall	F1-Score
	Correctly	Incorectly				
LogitBoost	87.234%	12.766%	1	0.887	0.956	0.920
			0	0.805	0.600	0.688
			ave	0.867	0.872	0.865

6 Discussion and Conclusion

The machine learning algorithms tested in this study provided workable prediction results for classifying willingness-to-pay intentions for mangrove ecosystem services. This outcome enhances comparability and confidence in determining strategies for payment programs related to the mangrove forest. Experimental results show that LogitBoost achieved the best performance among the models regarding accuracy and precision when tested with the research dataset.

Changing the number of attributes has a different effect on performing the tested machine-learning models. Notably, most models improved performance, with Multilayer Perceptron exhibiting the most impressive classification enhancement. Conversely, Random Forest was the only model to experience a decrease in performance when keeping only the more critical variables. LogitBoost remained unaffected by changes in the number of input attributes.

Psychological factors were ranked higher in importance than demographic and socioeconomic factors in classifying willingness-to-pay intentions. Four out of five independent variables ranked as important and significant for the classification model belonged to the psychological group. This underscores that willingness to pay is a behavioral intention that needs to be examined based on socio-psychological contexts. However, in this small-sample dataset, the limited number of instances labeled as "not willing to pay" (55) likely contributed to the relative difference in Precision and Recall scores for this class. Therefore, combined and enhanced techniques are recommended to further explore to improve prediction performance of the models.

7 Declaration of Competing Interest.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Asmare, E., Bekele, K., Fentaw, S.: Households willingness to pay for the rehabilitation of wetlands, evidence from Gudera Wetland, Northwest Ethiopia. Heliyon **8** (2022). https://doi.org/10.1016/j.heliyon.2022.e08813
2. Ayansola, O., Ogundunmade, T., Adedamola, A.: Modelling willingness to pay of electricity supply using machine learning approach. Modern Econ. Manag. 0–6 (2022). https://doi.org/10.53964/mem.2022009
3. Bashar, M.A., Nayak, R., Astin-Walmsley, K., Heath, K.: Machine learning for predicting propensity-to-pay energy bills. Intell. Syst. Appl. **17**, 200176 (2023). https://doi.org/10.1016/j.iswa.2023.200176
4. Botalb, A., Moinuddin, M., Al-Saggaf, A.S.: Contrasting convolutional neural network (CNN) with multi-layer perceptron (MLP) for big data analysis. In: 2018 International Conference on Intelligent and Advanced System (ICIAS), Kuala Lumpur, Malaysia, pp. 1–5 (2018) https://doi.org/10.1109/ICIAS.2018.8540626
5. Breiman, L.: Bagging predictors. Mach. Learn. **24**(2), 123–140 (1996). https://doi.org/10.1007/BF00058655

6. Chaudhary, A., Kolhe, S., Kamal, R.: An improved random forest classifier for multi-class classification. Inform. Process. Agri. **3**(4), 215–222 (2016). https://doi.org/10.1016/j.inpa.2016.08.002
7. Crook, J., Edelman, D., Thomas, L.: Recent developments in Consumer Credit Risk assessment. Eur. J. Oper. Res. **183**, 1447–1465 (2007). https://doi.org/10.1016/j.ejor.2006.09.100
8. Dettling, M., Bühlmann, P.: Boosting for tumor classification with gene expression data. Bioinformatics **19**(9), 1061–1069 (2003). https://doi.org/10.1093/bioinformatics/btf867
9. Edward, B.B., Suthawan, S.: Shrimp Farming and Mangrove Loss in Thailand. In: Edward B.B (ed.), Elgar Publisher, Chulalongkorn University, Thailand (2004)
10. Getachew, T.: Estimating willingness to pay for forest ecosystem conservation the case of Wof-Washa forest, North Shewa Zone, Amhara National regional state, Ethiopia. J. Res. Develop. Manag. **46**, 46–61 (2018)
11. Goutte, C., Gaussier, E.: A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds.) Advances in Information Retrieval. ECIR 2005. Lecture Notes in Computer Science, vol. 3408, pp. 345–359. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-31865-1_25
12. IPCC: Climate change 2022, impacts, adaptation, and vulnerability. In: Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York. https://doi.org/10.1017/9781009325844
13. Li, Y., Deng, H., Dong, R.: Prioritizing protection measures through ecosystem services valuation for the Napahai Wetland, Shangri-La County, Yunnan Province, China. Int J Sust Dev World **22**(2), 142–150 (2015). https://doi.org/10.1080/13504509.2014.926298
14. Liang, X., Zhu, L., Huang, D.S.: Multi-task ranking SVM for image cosegmentation. Neurocomputing, **247**, 126–136 (2017). https://doi.org/10.1016/j.neucom.2017.03.060
15. Lin, C.Y., Syrgabayeva, D.: Mechanism of environmental concern on intention to pay more for renewable energy, application to a developing country. Asia Pac. Manag. Rev. **21**(3), 125–134 (2016). https://doi.org/10.1016/j.apmrv.2016.01.001
16. Liu, P., Teng, M., Han, C.: How does environmental knowledge translate into pro-environmental behaviors? The mediating role of environmental attitudes and behavioral intentions. Sci. Total. Environ. **728**, 138126 (2020). https://doi.org/10.1016/j.scitotenv.2020.138126
17. Liu, Y., Wang, Y., Zhang, J.: New machine learning algorithm: Random Forest. In: Liu, B., Ma, M., Chang, J. (eds.) Information Computing and Applications, ICICA 2012. LNCS, vol. 7473, pp. 246–252. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34062-8_32
18. Mojiol, A.R., Hong, K.Y., Saleh, E.: Willingness to pay for mangroves conservation by the local communities in Salut Mengkabong Lagoon Tuaran Sabah. Hutan Tropika, **14**(1), 1–9 (2022). https://doi.org/10.36873/jht.v14i1.310
19. Naik, V., Desai, A.: Online handwritten Gujarati character recognition using SVM, MLP, and K-NN. In: 8th International Conference on Computing, Communication and Networking Technologies, 1–6 (2017). https://doi.org/10.1109/ICCCNT.2017.8203926
20. Noronha, D.H., Torquato, M.F., Fernandes, M.A.C.: A parallel implementation of sequential minimal optimization on FPGA. Microprocess. Microsyst. **69**, 138–151 (2019). https://doi.org/10.1016/j.micpro.2019.06.007
21. Obeng, E.A., Aguilar, F.X.: Value orientation and payment for ecosystem services, Perceived detrimental consequences lead to willingness-to-pay for ecosystem services. J. Environ. Manage. **206**, 458–471 (2018). https://doi.org/10.1016/j.jenvman.2017.10.059

22. Obeng, E.A., Aguilar, F.X.: Willingness-to-pay for restoration of water quality services across geo-political boundaries. Curr. Res. Environ. Sustain. **3**, 100037 (2021). https://doi.org/10.1016/j.crsust.2021.100037
23. Pagiola, S., Ritter, K., Bishop, J.: How Much is an Ecosystem Worth? Assessing the Economic Value of Conservation (2004)
24. Pham, T.D., Kaida, N., Yoshino, K., Nguyen, X.H., Nguyen, H.T., Bui, D.T.: Willingness to pay for mangrove restoration in the context of climate change in the Cat Ba biosphere reserve Vietnam. Ocean Coast. Manag. **163**, 269–277 (2018). https://doi.org/10.1016/j.ocecoaman.2018.07.005
25. Phan, T.D., Bertone, E., Pham, T.D., Pham, T.V.: Perceptions and willingness to pay for water management on a highly developed tourism island under climate change, a Bayesian network approach. Environ. Challenges **5**, 100333 (2021). https://doi.org/10.1016/j.envc.2021.100333
26. Platt, J.: Fast training of support vector machines using sequential minimal optimization. Adv. kernel Methods - Support Vector Learn. (1998). https://www.microsoft.com/en-us/research/publication/fast-training-of-support-vector-machines-using-sequential-minimal-optimization/
27. Pouta, E., Rekola, M.: The theory of planned behavior in predicting willingness to pay for abatement of forest regeneration. Soc. Nat. Resour. **14**(2), 93–106 (2001). https://doi.org/10.1080/08941920130000517
28. Raudsepp-Hearne, C., Peterson, G., Bennett, E.: Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. Proc. Natl. Acad. Sci. U.S.A. **107**, 5242–5247 (2010). https://doi.org/10.1073/pnas.0907284107
29. Raudys, S.J., Jain, A.K.: Small sample size effects in statistical pattern recognition, recommendations for practitioners. IEEE Trans. Pattern Anal. Mach. Intell. **13**(3), 252–264 (1991). https://doi.org/10.1109/34.75512
30. Roesch-McNally, G.E., Rabotyagov, S.S.: Paying for forest ecosystem services, voluntary versus mandatory payments. Environ. Manage. **57**(3), 585–600 (2016). https://doi.org/10.1007/s00267-015-0641-7
31. Saritas, M.M., Yasar, A.: Performance analysis of ANN and Naive Bayes classification algorithm for data classification. Int. J. Intell. Syst. Appl. Eng. **7**(2), 88–91 (2019). https://doi.org/10.18201//ijisae.2019252786
32. Selz, D.: From electronic markets to data driven insights. Electron. Mark. **30**(1), 57–59 (2020). https://doi.org/10.1007/s12525-019-00393-4
33. Shen, Y., Hamm, J.A., Gao, F., Ryser, E.T., Zhang, W.: Assessing consumer buy and pay preferences for labeled food products with statistical and machine learning methods. J. Food Prot. **84**(9), 1560–1566 (2021). https://doi.org/10.4315/JFP-20-486
34. Sokratous, K., Fitch, A.K., Kvam, P.D.: How to ask twenty questions and win, Machine learning tools for assessing preferences from small samples of willingness-to-pay prices. J. Choice Model. **48**, 100418 (2023). https://doi.org/10.1016/j.jocm.2023.100418
35. State, D., Ogeh, K.T., Jimoh, S.O., Ajewole, O.I.: Willingness to pay for environmental service functions of mangrove forest. J. Res. Dev. Manag. **16**, 1–7 (2016)
36. Subhan, F., Ali, Y., Zhao, S.: Unraveling preference heterogeneity in willingness-to-pay for enhanced road safety, a hybrid approach of machine learning and quantile regression. Accid. Anal. Prev. **190**, 107176 (2023). https://doi.org/10.1016/j.aap.2023.107176
37. Taylor, P., Bhatta, L.D., Eric, B., Oort, H., Van Stork, N.E., Baral, H.: Ecosystem services and livelihoods in a changing climate, understanding local adaptations in the Upper Koshi, Nepal. Int. J. Biodivers. Sci. Ecosyst. Serv. Manag. 37–41 (2015). https://doi.org/10.1080/21513732.2015.1027793

38. Tehrany, M.S., Jones, S., Shabani, F., Martínez-Álvarez, F., Tien Bui, D.: A novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using logitboost machine learning classifier and multi-source geospatial data. Theoret. Appl. Climatol. **137**(1–2), 637–653 (2019). https://doi.org/10.1007/s00704-018-2628-9
39. Thuy, P.T., Thanh Tung, N., Hoang, C.D.: Predicting the level of hypertension using machine learning BT. In: Vinh, P.C., Rakib, A. (eds.) Context-Aware Systems and Applications, and Nature of Computation and Communication, pp. 113–122. Springer, Heidelberg (2021). https://doi.org/10.1007/978-3-319-29236-6
40. Tóth, S., Ettl, G., Rabotyagov, S.: ECOSEL, an auction mechanism for forest ecosystem services. MCFNS **2**, 99–116 (2010)
41. Tran, Y.L., Siry, J.P., Bowker, J.M., Poudyal, N.C.: Atlanta households' willingness to increase urban forests to mitigate climate change. Urban For. Urban Green. (2017). https://doi.org/10.1016/j.ufug.2017.02.003
42. Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J.: Machine learning algorithm validation with a limited sample size. PLoS ONE **14**(11), 1–20 (2019). https://doi.org/10.1371/journal.pone.0224365
43. Yaman, E., Subasi, A.: Comparison of bagging and boosting ensemble machine learning methods for automated EMG signal classification. Biomed. Res. Int. **2019**, 9152506 (2019). https://doi.org/10.1155/2019/9152506
44. Zeng, M., Cao, H., Chen, M., Li, Y.: User behaviour modeling, recommendations, and purchase prediction during shopping festivals. Electron. Mark. **29**(2), 263–274 (2019). https://doi.org/10.1007/s12525-018-0311-8