Indonesian Journal of Sustainability
Vol. 4 No. 2, June : 15-24
P-ISSN 2827-8526
E-ISSN 2827-8917

INDONESIAN JOURNAL of SUSTAINABILITY

CLASSIFICATION MODEL FOR ONLINE PAYMENT TRANSACTION FRAUD USING A DECISION TREE BASED ON THE GINI INDEX

Sabir Rosidin[1], Allina Kurniasari[2]

[1]*Diponegoro University, Semarang, Indonesia*
[2]*Sultan Agung Islamic University, Semarang, Indonesia*

sabirrosidin@gmail.com

ABSTRACT

The development of digital technology has led to a significant increase in the use of online payment systems. However, this rise in transaction volume has also resulted in heightened risks of fraud, which threaten the security of digital transactions. Early detection of fraudulent activities is essential to prevent financial losses and maintain user trust in these payment systems. This study aims to build a classification model for detecting fraudulent online payment transactions using the Decision Tree algorithm based on the Gini Index. The dataset used consists of digital financial transactions, which include both numeric and categorical features. The model's performance is evaluated using confusion matrix metrics such as accuracy, precision, and recall. The results indicate that employing the Gini Index for feature selection enhances the model's performance, achieving an accuracy of 93.50% and a notable increase in recall for minority classes, such as DEBIT transactions. The Gini Index-based Decision Tree has proven effective for the interpretive and efficient detection of fraudulent transactions. This study contributes to the development of a more accurate digital fraud detection system that can be implemented in real-world online payment systems.

Keywords : *Fraud Detection; Decision Tree; Gini Index; Online Transaction; Classification*

INTRODUCTION

The rapid advancement of digital technology has led to a significant rise in the use of online payment systems. However, this progress has also brought about increased security threats to digital transactions, particularly in the form of fraud. Fraudulent activities in online payment transactions can result in substantial financial losses for both individuals and financial institutions, as well as erode user trust in digital payment systems.

The decision tree algorithm is one of the most popular methods for fraud detection due to its ability to create models that are both easy to interpret and efficient to train. One of the advantages of decision trees is their capability to handle both categorical and numerical features without requiring extensive pre-processing. There are various types of decision tree algorithms, including

ID3, C4.5, C5.0, CART, conditional inference trees, and CHAID. Additionally, there are tree-based ensemble methods such as random forests, rotation forests, and gradient-boosting decision trees (Mienye & Jere, 2024).

In building a Decision Tree model, selecting relevant features is very important to improve the accuracy and efficiency of the model. One of the feature selection methods used is the Gini Index, the reason is that the Gini Index can measure the level of purity of a node in a decision tree. The Gini Index helps in determining which attributes are most effective in separating classes in the data, thus improving the performance of the model in detecting fraudulent transactions (Park & Kwon, 2011).

Several studies have shown the effectiveness of using a Decision Tree with the Gini Index in detecting Fraud, research conducted by (S. Shaankari et al., 2025) shows that the use of a Decision Tree with Gini index separation criteria can increase accuracy. In a recent study by (Xu et al., 2023) a new method for detecting fraud called DBDT was introduced, which improves accuracy.

Meanwhile, research by (Appavu, 2025) presents a comparative study between the Logistic Regression and Decision Tree machine learning techniques in detecting credit card fraud, with a unique approach based on shopping behavior, the utilization of historical transaction data, and application of SMOTE to overcome class imbalance, thereby improving accuracy and MCC.

Another study by (Editya et al., 2025) focuses on optimizing the classification of fraudulent transactions in online payment systems using supervised machine learning algorithms. The research evaluates several algorithms, including Naïve Bayes, Decision Tree, Random Forest, Gradient Boosting Tree, and SVM, and finds that ensemble-based methods particularly Gradient Boosting Tree achieve the highest accuracy and robustness. This highlights the importance of selecting the appropriate algorithm and fine-tuning parameters to enhance fraud detection in digital payment platforms.

A study conducted by (Mishra et al., 2024) highlights the necessity of improving fraud detection methods due to the limited accuracy of current techniques. The researchers advocate for the use of the Decision Tree algorithm to detect credit card fraud in real-time, suggesting the combination of approaches such as random forests and gradient boosting to enhance accuracy.

A study conducted by (B. Palad et al., 2020) evaluated five decision tree algorithms: J48, Hoeffding Tree, Decision Stump, REPTree, and Random Forest. The researchers used these algorithms to classify computer fraud data derived from police reports and victim narratives. The results indicated that the J48 algorithm achieved the highest accuracy and the lowest error rate, suggesting that it could be a valuable tool in cybercrime investigations.

Another study (Alraddadi, 2023) proposed a theoretical model for detecting and preventing credit card fraud using the Decision Tree algorithm. This study also included a survey to understand students' perceptions of credit card fraud incidents, providing additional insights into the development of a fraud detection system.

Various previous studies have shown that the Decision Tree algorithm is effective in detecting fraud, but still, the classification model used is not effective enough in detecting fraudulent transactions effectively, especially in data with unbalanced characteristics, where the number of fraudulent transactions is much less than normal transactions. So the selection of features used in the classification process is not optimal.

Therefore, this study aims to develop an effective classification model for detecting fraud in online payment transactions using the Decision Tree algorithm based on the Gini Index, in order to enhance accuracy and the model.

RESEARCH METHOD

This research has stages that must be passed, including data collection, data processing, and presentation of results.

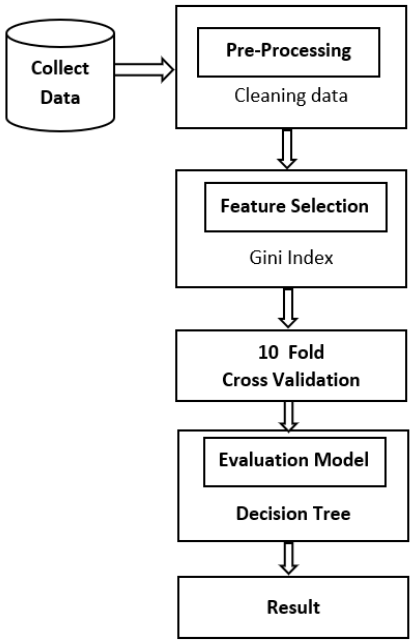

Fig. 1 Methodology

Collect Data

Data retrieval was conducted on Kaggle.com (Shah, 2022), The dataset used is public and focuses on online payment fraud detection. The features of the dataset include the transaction date and time, the transaction amount, the type of transaction (payment, transfer, cash out), information about the sender and recipient, and the fraud status (1 for fraud, 0 for normal).

Preprocessing Data

At this stage, handle empty data/missing values.

Feature Selection

Using the Gini Index to select the most relevant features, this index measures the impurity level of each attribute for optimal data division.

Split Data

The model is trained using a Decision Tree Classifier and the k-Fold Cross Validation Technique with training and test data.

Decision Tree Classifier

A Decision Tree is a machine-learning algorithm used for classification and regression. It works by dividing a dataset into subsets based on certain features, forming a decision tree structure that facilitates interpretation and decision-making (Rokach & Maimon, 2005).

In the context of online transaction fraud detection, Decision Tree can identify patterns that distinguish legitimate transactions from fraudulent ones. This algorithm is effective in handling large and complex datasets, and is able to provide clear interpretations of the decisions taken by the model.

Gini Index

The Gini Index is a statistical measure used to determine how often a randomly selected element from a set would be misclassified if classified based on the distribution of labels in that subset. The Gini Index is used to determine the best split in Decision Tree algorithms such as CART (Classification and Regression Tree) (Breiman, 1984).

There is a dataset with K classes, so the Gini Index formula for the nodes is (Raileanu et al., 2004):

$$\text{Gini}(t) = 1 - \sum_{i=1}^{K} p_i^2$$

Where:

t: is a node

P_i: s the proportion (probability) of the i-th class in the node

k: total number of classes in the node

Tabel 1. Example of dataset subset:

ID	amount	type	is Fraud
1	1000	TRANFER	1
2	850	CASH_OUT	1
3	1000	TRANSFER	0
4	950	CASH_OUT	0
5	1050	TRANSFER	0

From the 5 transactions above:

- Number of classes isFraud=1 (fraud): 2 transactions
- Number of classes isFraud=0 (not fraud): 3 transactions

Formula :

$$p_1 = \frac{2}{5} = 0.4 \quad ; \quad p_0 = \frac{3}{5} = 0.6$$

Calculate Gini Index:

Gini (t) = 1 – $(p_0)^2$ – $(p_1)^2$=1-$(0.6)^2$-$(0.4)^2$=1 - 0.36 – 0.16 = 0.48

Evaluation

Evaluation model or conducting testing to determine the performance of the Decision Tree algorithm using the Gini Index selection feature. Evaluation results with the accuracy of each method, namely before using the Gini Index and after using the Gini Index. The evaluation results will be seen using the confusion matrix by looking at accuracy, precision, and recall with the following formula:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i}$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

RESULT AND DISCUSSION

The research results are divided into two sections. The first section presents the test results of the Decision Tree algorithm, while the second section shows the results after incorporating the Gini Index selection feature.

Decision Tree Testing

The dataset labels are classified into 5 parts, including PAYMENT, TRANSFER, CASH_OUT, DEBIT AND CASH_IN.

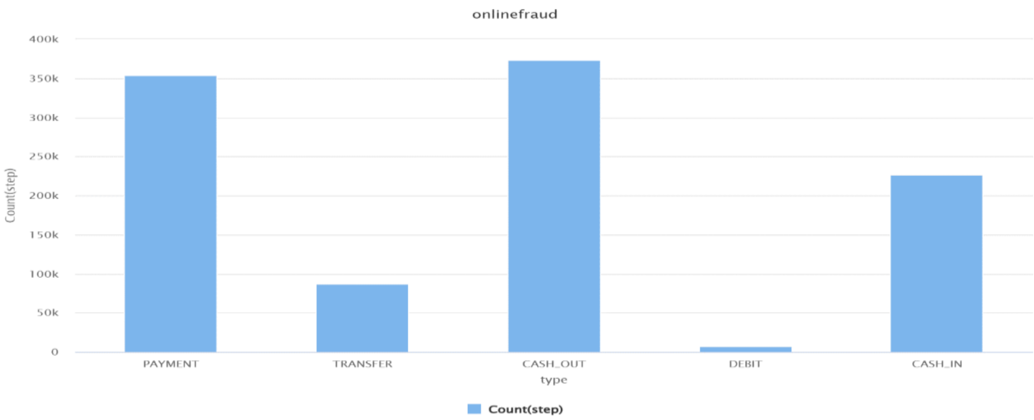

Fig. 2 Distribution of Transaction Types

Testing Decision Tree Without Using Gini Index

Table 2. Testing Decision Tree Classifier without using Selection Feature

Algorithm	Transaction Class	Precision TP / (TP + FP)	Recall TP / (TP + FN)	Accuracy
Decision Tree	PAYMENT	98.75%	99.93%	90.98%
	TRANSFER	88.12%	42.25%	
	CASH_OUT	85.16%	93.10%	
	DEBIT	58.06%	2.51%	
	CASH_IN	89.82%	94.95%	

The best performing classes are PAYMENT and CASH IN with very high precision and recall. Then in testing without using the Gini Index, it produces an accuracy of 90.98%.

Table 3. Confusion Matrix

	true PAYMENT	true TRANSFER	true CASH_OUT	true DEBIT	true CASH_IN	class precision
pred. PAYMENT	70727	3	94	0	799	98.75%
pred. TRANSFER	0	7331	987	0	1	88.12%
pred. CASH_OUT	2	9583	69569	1052	1485	85.16%
pred. DEBIT	0	2	15	36	9	58.06%
pred. CASH_IN	46	432	4063	348	43132	89.82%

Table 3, Confusion matrix is used to evaluate the performance of the classification model against five financial transaction categories, namely PAYMENT, TRANSFER, CASH_OUT, DEBIT, and CASH_IN. Based on the evaluation results, the model produces an overall accuracy of 90.98%, indicating that around 91% of the total predictions generated match the actual labels.

The table presents the distribution of predictions against the actual labels, where:

- Each row shows the amount of data predicted by the model for each class.
- Each column represents the amount of actual data (actual labels) from each class.

Analysis of the precision and recall of each class shows the following:

- PAYMENT has a precision value of 98.75% and a recall of 99.93%. This indicates that the model can accurately predict PAYMENT transactions, both in terms of accuracy and completeness.
- TRANSFER has a precision of 88.12%, but its recall value is low, which is 42.25%. This means that the model is quite accurate when predicting TRANSFER, but there are still many TRANSFER data that it fails to recognize correctly.
- CASH_OUT is recorded with a precision of 85.16% and a recall of 93.10%, indicating that the model is quite effective in detecting this type of transaction.
- DEBIT shows very low performance with a precision of 58.06% and a recall of only 2.51%. This indicates that the model almost completely fails to recognize DEBIT transactions, which may be due to an imbalance in the amount of data or a lack of

features that support the identification of these transactions.

- CASH_IN has a precision of 89.82% and a recall of 94.95%, indicating that the model can recognize these transactions consistently and accurately.

Testing Decision Tree Using Gini Index

Table 4. Decision Tree Classifier Testing using Gini Index Selection Feature

Algorithm	Transaction Class	Precision TP / (TP + FP)	Recall TP / (TP + FN)	Accuracy
Decision Tree	PAYMENT	99.33%	99.96%	93.50%
	TRANSFER	86.96%	43.98%	
	CASH_OUT	87.36%	96.81%	
	DEBIT	53.67%	59.05%	
	CASH_IN	98.22%	97.99%	

Table 4, overall class types have improved performance, all lifts increased after using the Gini Index. The accuracy reached 93.50%. Accuracy increased by 2.52% after using the Gini Index. This shows that the Gini Index contributes positively to separating classes more effectively.

Table 5. Confusion Matrix

	true PAYMENT	true TRANSFER	true CASH_OUT	true DEBIT	true CASH_IN	class precision
pred. PAYMENT	70749	2	88	0	390	99.33%
pred. TRANSFER	0	7631	1143	1	0	86.96%
pred. CASH_OUT	15	9621	72345	505	330	87.36%
pred. DEBIT	0	32	505	848	195	53.67%
pred. CASH_IN	11	65	647	82	44511	98.22%
class recall	99.96%	43.98%	96.81%	59.05%	97.99%	

PAYMENT has a precision of 99.33% and a recall of 99.96%. This shows that the model is almost perfect in predicting PAYMENT transactions.

TRANSFER is achieved with a precision of 86.96% and a recall of 43.98%. Although the precision is quite high, the recall is still low, indicating that many TRANSFER transactions are not detected.

CASH_OUT has a precision of 87.36% and a recall of 96.81%, which is a very good performance. The model is very good at recognizing and classifying this transaction.

DEBIT only has a precision of 53.67%, but its recall increases to 59.05% compared to after using the Gini Index. Although it has increased, its performance is still relatively low compared to other classes.

CASH_IN has a high precision of 98.22% and a recall of 97.99%, which shows that the model is very accurate and consistent in predicting this transaction.

CONCLUSION

After testing the classification model both before and after applying the Gini Index, it can be concluded that the Gini Index significantly enhances overall model performance. This improvement is reflected in the increase in accuracy from 90.98% to 93.50% following the application of the Gini Index.

Not only has there been a rise in overall accuracy, but there has also been an improvement in recall values for several classes that previously showed low performance. For instance, the DEBIT class, which initially had a recall of just 2.51%, soared to 59.05%. Similarly, the recall for the TRANSFER class increased from 42.25% to 43.98%.

Previously, the model exhibited a bias towards majority classes such as PAYMENT and CASH_IN. However, it is now much better at recognizing minority classes without significantly compromising performance in the majority class. Precision for the majority class remains high, with PAYMENT achieving 99.33% and CASH_IN at 98.22% after applying the Gini Index.

In conclusion, the use of the Gini Index effectively increases the accuracy of the model while improving the distribution of predictions across different classes. This makes it a valuable method for addressing data imbalance in transaction classification scenarios.

REFERENCES

Alraddadi, A. S. (2023). A Survey and a Credit Card Fraud Detection and Prevention Model using the Decision Tree Algorithm. *Engineering, Technology & Applied Science Research*, *13*(4), 11505–11510. https://doi.org/10.48084/etasr.6128

Appavu, N. (2025). AI and ML Approaches for Credit Card Fraud Detection: A Comparative Study of Logistic Regression and Decision Tree Techniques. *2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 2068–2074. https://doi.org/10.1109/IDCIOT64235.2025.10915092

B. Palad, E. B., Burden, M. J. F., Torre, C. R. D., & C. Uy, R. B. (2020). Performance evaluation of decision tree classification algorithms using fraud datasets. *Bulletin of Electrical Engineering and Informatics*, *9*(6), 2518–2525. https://doi.org/10.11591/eei.v9i6.2630

Breiman, L. (1984). *Classification And Regression Trees*.

Editya, A. S., Alamin, Moch. M., Pramana, A. L., & Kurniati, N. (2025). Fraud Classification in Online Payments Using Supervised Machine Learning Algorithms. *Green Intelligent Systems and Applications*, *5*(1), 40–50. https://doi.org/10.53623/gisa.v5i1.552

Mienye, I. D., & Jere, N. (2024). A Survey of Decision Trees: Concepts, Algorithms, and Applications. *IEEE Access*, *12*, 86716–86727. https://doi.org/10.1109/access.2024.3416838

Mishra, P., Dewangan, P., Dewangan, S., & Ansari, S. (2024). *Credit Card Fraud Detection Using Decision Tree Algorithm*.

Park, H., & Kwon, H.-C. (2011). Improved Gini-Index Algorithm to Correct *Feature*-Selection Bias in Text Classification. *IEICE Transactions on Information and Systems*, *E94-D*(4), 855–865. https://doi.org/10.1587/transinf.E94.D.855

Raileanu, L. E., Stoffel, K., Raileanu, L., & Stoffelg, K. (2004). *Theoretical Comparison between the Gini Index and Information Gain Criteria*.

Rokach, L., & Maimon, O. (2005). Decision Trees. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 165–192). Springer US. https://doi.org/10.1007/0-387-25465-X_9

S. Shaankari, M. Aishwarya, B. Vaishnavi, & Syeda Hifsa Naaz. (2025). *Online Fraud Detection: A Decision Tree Approach*. https://doi.org/10.5281/ZENODO.15055119

Shah, J. (2022). *Online Payment Fraud Detection*. Online Payment Fraud *Detection. https://www.kaggle.com/datasets/jainilcoder/online-payment-fraud*-detection

Xu, B., Wang, Y., Liao, X., & Wang, K. (2023). Efficient fraud detection using deep boosting decision trees. *Decision Support Systems*, *175*, 114037. https://doi.org/10.1016/j.dss.2023.114037