Journal of Theoretical and Applied Information Technology
15th April 2024. Vol.102. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MACHINE LEARNING MODEL OF CUSTOMER BEHAVIOR ON E-BANKING TRANSACTION USING CLASSIFICATION TECHNIQUE

SEPTIAN EKA ADY BUANANTA[1] , LAY CHRISTIAN [2] , PARADISE [3]

[1,2]Information Systems Department, School of Information Systems,

Bina Nusantara University, Jakarta, Indonesia 11480

[3]Department of Informatics, Institut Teknologi Telkom Purwokerto,

Jl. D. I. Panjaitan No. 128, Purwokerto 53147, Indonesia

E-mail: [1]septian.buananta@binus.ac.id, [2]lchristian@binus.edu, [3]paradise@ittelkom-pwt.ac.id

ABSTRACT

E-banking services provide easy access for users to make purchase and payment transactions anytime and anywhere via devices connected to the internet. However, e-banking adoption rates may vary between groups and evolve as awareness and trust in the service increases. Customers perceive products and services differently based on their experiences, beliefs, and values. Their attitudes toward a brand or product influence their decision to buy or continue using it. Therefore, Machine Learning Development can be used by businesses to analyze customer behavior through market research, surveys, and data analytics to gain insights that can inform product development, marketing strategies, and customer relationship management efforts. Knowing customer behavior allows banks to provide a more personalized user experience. This may include presenting product recommendations or providing notifications that match customer preferences. The classification results of the best-selected technique, with an accuracy rate of 98.61% and an execution time of 10 seconds, can be used as a reference to determine customer behavior in important e-banking transactions because they significantly impact business strategy and development of digital banking services. Customer behavior data is a valuable source of information that can be used to make strategic decisions. For example, some services are often used, namely bill payment and QRIS. So, the Bank can use data analysis to design more effective marketing campaigns, optimize operational processes, and design new products that suit customer needs.

Keywords: *Machine Learning, Customer Behavior, E-banking Transaction, Classification Technique, Decision Tree*

1. INTRODUCTION

The e-banking transaction process in Indonesia involves several stages and involves various parties, including customers, banks, and electronic payment system providers [1][2]. It is important to note that each bank may have slightly different procedures and provide additional features that can be tailored to customer needs. Besides, the regulations and security standards enforced by Bank Indonesia also influence the e-banking transaction process in Indonesia [3]. The transaction process for purchasing and paying for products through e-banking services in Indonesia can vary depending on the service type and features each bank provides. E-banking services for purchases and payments in Indonesia are needed because they provide convenience and efficiency for users. This allows access to a wide range of financial services without requiring a physical bank, saving time and effort. In addition, e-banking can also improve financial control, enable real-time transactions, and provide instant access to financial information [4].

E-banking services increase user convenience, speed, and control of daily financial transactions. E-banking services are also essential in purchasing and payment transactions in banking because they provide easy access for users to carry out purchase and payment transactions anytime and anywhere via devices connected to the internet. Additionally, it reduces dependence on physical banks and provides a faster and more efficient transaction solution, saving users time [5]. Users can easily monitor financial activities, view transaction history, and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

check balances through the e-banking platform, increasing awareness and control over their finances. E-banking services provide high transparency by providing real-time transaction details, helping users clearly understand their finances. Besides that, it allows users to carry out various types of transactions, such as bill payments, online purchases, interbank transfers, investments, and others, on one platform. E-banking services often have strong security measures, including two-factor authentication and data encryption, to protect users' financial information. Overall, e-banking can help reduce operational and administrative costs for users and financial institutions, as it reduces the need for manual processes. E-banking services continue to evolve with innovations and integration with other technologies, such as mobile banking, providing a more sophisticated banking experience. Thus, e-banking services significantly benefit users and banks regarding efficiency, convenience, and better financial management [6].

E-banking services in Indonesia have a legal basis for purchase and payment transactions involving several rules and regulations. Law of the Republic of Indonesia Number 7 of 2011 concerning currency, which controls the fundamental aspects of money and payment systems in Indonesia. In addition, Bank Indonesia Regulation Number 18/40/PBI/2016 concerning the Implementation of Electronic Payment Systems which establishes the regulatory framework for the implementation of payment systems, including e-banking; Bank Indonesia Regulation Number 19/12/PBI/2017 concerning Electronic Money Services which establishes provisions related to the performance of electronic money services, which include payment transactions; Bank Indonesia Regulation Number 20/6/PBI/2018 concerning Digital Financial Innovation in the Financial Services Sector which regulates digital financial innovation, including e-banking services, to encourage the development of the financial sector in the digital era. On the other hand, Financial Services Authority (OJK) Regulation Number 12/POJK.03/2016 concerning Information Technology-Based Money Lending and Borrowing Services which stipulates regulations regarding electronic money lending and borrowing, which can also involve e-banking services; OJK Regulation Number 13/POJK.03/2018 concerning Digital Financial Innovation in the Financial Services Sector which establishes a framework for digital financial innovation in the financial services sector. With this legal basis, e-banking services in Indonesia have a lawful basis that regulates various aspects, including security, privacy, and the obligations of service providers [7].

E-banking services in Indonesia involve various groups, including the General Public or Individuals from multiple age groups and backgrounds who can use e-banking services for different financial transactions, bill payments, and other banking needs. In addition, entrepreneurs and entrepreneurs often use e-banking services to manage business transactions, fund transfers, and corporate banking needs [8]. Young Professionals tend to actively use e-banking services because of their familiarity with technology and high mobility.

Corporate customers and large companies also use e-banking services to manage financial transactions, employee salaries, and other corporate banking needs [9] online business actors who are individuals. Small online businesses often use e-banking services to pay and receive online transactions [10]. Therefore, knowing customer behavior in e-banking transactions is necessary with various customer groups. This is important because it significantly impacts business strategy and the development of digital banking services [11]. Customer behavior refers to individuals' actions, decisions, and activities when searching for, purchasing, using, evaluating, and disposing products and services. Understanding customer behavior is critical for businesses and marketers as it helps them tailor products, services, and marketing strategies to meet the needs and preferences of their target audience. By leveraging artificial intelligence, banks can improve operational efficiency, provide better service to customers, and remain competitive in an ever-changing marketplace. Machine learning models can analyze user behavior and offer product recommendations that better suit individual needs and preferences, increasing the personalization of the user experience. Additionally, by being able to provide relevant offers, this model can help improve customer retention. It should be noted that users tend to continue using e-banking services if they feel they are getting added value from the products and services offered [12].

2. BACKGROUND AND RELATED WORK

2.1 Machine Learning

Machine learning is a branch of artificial intelligence that focuses on developing algorithms and computer models that enable systems to learn from data, identify patterns, and make decisions without explicit programming. In machine learning, computers can “learn” from experience and data, improving the performance of specific tasks over time. Several machine learning approaches include supervised, unsupervised, and reinforcement learning. The model is given labeled data in supervised learning to learn the relationship between input and output. In unsupervised learning, the model identifies patterns or structures in unlabeled data [13].

Meanwhile, in reinforcement learning, the model learns through interaction with its environment and getting feedback based on its actions. Examples of machine learning applications include facial recognition, automatic translation, recommendation systems, and more. Machine learning is becoming an active and widely applied area of research due to its ability to process and understand complex data and make predictions or decisions without detailed manual programming. Machine learning is needed to handle various challenges and tasks that are difficult to solve with conventional approaches [14].

By leveraging machine learning models' adaptability and generalization capabilities, we can better explore and understand data, make smarter decisions, and provide more effective solutions to complex problems and tasks. In the Indonesian banking world, the application of machine learning has become a growing trend. The use of machine learning in assessing credit can help banks optimize the loan approval process by analyzing several variables and predicting credit risk more accurately; machine learning models are used to detect suspicious transaction patterns or fraudulent acts in the banking system, helping banks protect customers and their assets, optimizing banking services, machine learning can provide more precise product recommendations according to customer needs and preferences, using machine learning-based sentiment analysis to understand customer views and reviews which allows them to respond better to customer needs and satisfaction [15]. In addition, machine learning in risk management helps banks identify and manage risks, including credit, market, and operational risks. Machine learning can also understand customer behavior and preferences, enabling banks to provide more personalized service and improve user experience. To enhance operational efficiency, banks can use machine learning to automate internal processes such as supply chain management and worker scheduling.

On the other hand, machine learning to detect suspicious patterns related to money laundering and other illegal activities to strengthen system security. In investment banking services, machine learning can be used to analyze the feasibility of investments and provide more accurate advice to investors. Banks leverage machine learning in chatbot development to improve customer service, providing faster responses and more personalized customer solutions [16].

Applying machine learning in the Indonesian banking sector helps improve operational efficiency, provide more adaptive financial solutions, and create a better customer experience. With these technological advances, the banking sector can continue to innovate to meet growing market and regulatory demands [17].

2.2 Classification Techniques

Classification techniques are methods or approaches used in machine learning to separate data into categories or classes based on specific attributes [18]. The main goal of classification techniques is to create a model that can learn patterns from given training data and then classify new data into appropriate categories. Classification techniques are necessary for machine learning because they systematically group or organize data or classes based on certain features or attributes. Classification techniques enable automatic decision-making based on patterns found in training data. This helps automate the grouping or categorization process. Using classification techniques can analyze and understand complex patterns in data, which may be difficult or even impossible for humans to identify. Classification makes it possible to make predictions or forecasts based on previously observed data. It is helpful in various contexts, such as finance, business, and science. In recommendation systems, classification techniques enable building models that provide users with more personalized and relevant recommendations based on their preferences [19].

Additionally, Classification can detect anomalies or abnormal behavior in data, such as fraudulent acts or rare events. In banking and finance, classification techniques help manage risk by predicting potential credit, market, or operational risks. On the other hand, classification techniques can identify and group customers into specific market segments based on their preferences or behavior. Companies can optimize their operations using classification techniques by grouping data or elements with similar characteristics or behavior [19]. Using classification techniques, we can produce models to make decisions and solve various problems involving data analysis and clustering. It helps improve efficiency, accuracy, and precision in multiple aspects of life and industry.

2.3 Customer Behaviour for E-Banking Transaction

Understanding customer behavior provides valuable information for creating novel products and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

services. Financial institutions can introduce innovative features that align with customer expectations by grasping the elements customers appreciate and require in their digital banking interactions, ensuring engagement and satisfaction [1][2]. Additionally, insights into how customers engage with e-banking platforms enhance the overall user experience. Banks can refine digital interfaces, streamline processes, and incorporate features that match customer preferences, resulting in a more user-friendly and effective e-banking environment.

Comprehending customer behavior in e-banking transactions also facilitates the optimization of operational processes [1][2]. By identifying peak transaction times, popular services, and shared user preferences, banks can allocate resources more judiciously, improve service delivery, and reduce operational costs. Thus, analyzing customer behavior is a strategic approach benefiting financial institutions and customers. This approach bolsters security measures and personalizes services, guides decision-making, and ultimately contributes to a more positive and secure e-banking experience.

The novel aspects of customer behavior in e-banking transactions include evolving patterns, emerging preferences, and changing trends that shape how customers interact with electronic banking services [10]. Furthermore, the objectives related to customer behavior in e-banking transactions include understanding, analyzing, and responding to customers' preferences, needs, and expectations that will be implemented in this research. The aim is to enhance user experience, optimize service delivery, and tailor electronic banking services to meet customer requirements effectively.

The research on customer behavior in e-banking transactions utilizing classification focuses on identifying and understanding distinct patterns or categories within customer actions [10]. By employing classification methods, researchers aim to differentiate various behaviors exhibited by customers during electronic banking transactions. This approach allows for categorizing customers based on specific criteria, enabling a deeper analysis of preferences, trends, and tendencies in their interactions with e-banking platforms. The ultimate goal is to uncover meaningful insights that can inform personalized strategies, improve user experiences, and enhance the overall effectiveness of e-banking services.

2.3.1 Decision Tree

A decision tree in classification is a predictive modeling tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. In classification, decision trees are often employed as a supervised machine learning algorithm to predict the class or category of an item or instance based on its features. Decision trees are popular for their interpretability and ease of visualization. They are versatile and can be used for both classification and regression tasks. However, they can be sensitive to slight variations in the data, and sometimes, they may overfit the training data, which is why techniques like pruning are employed to enhance their generalization capabilities [20]. Creating a decision tree classifier for e-banking recommendations involves designing a model to predict the most suitable e-banking services or products based on user characteristics and preferences. This decision tree is a basic structure and should be adapted to incorporate more sophisticated criteria and features for a real-world application. Additionally, the model's performance can be improved by utilizing a larger dataset, refining feature selection, and employing techniques such as pruning to avoid overfitting.

2.3.2 Random Forest

Random Forest is an ensemble learning method that can be used for both classification and regression tasks. In the context of classification, Random Forest combines the predictions from multiple decision trees to improve overall accuracy and generalization. Random Forest is widely used in practice due to its effectiveness, versatility, and ability to handle complex datasets. It is a choice for many machine learning practitioners for classification tasks [21]. Implementing a Random Forest classifier for e-banking recommendation involves using an ensemble of decision trees to improve predictive accuracy and robustness. It is important to note that the success of a Random Forest classifier depends on the quality and representativeness of the dataset, appropriate feature engineering, and careful tuning of hyperparameters. Additionally, interpreting the results, understanding the importance of features, and addressing any biases in the data are crucial aspects of building an effective e-banking recommendation system.

2.3.3 Naïve Bayes

In classification, naive Bayes is a probabilistic machine learning algorithm based on Bayes' theorem. It is particularly popular for text classification tasks, such as spam detection and sentiment analysis, but it can be applied to various classification problems. Naive Bayes relies on Bayes' theorem, which describes the probability of an event based on prior knowledge of conditions that might be related to the event. In the context of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classification, Bayes' theorem calculates the probability of a particular class given a set of features. However, the "naive" assumption of independence may not hold in all real-world scenarios, and as a result, Naive Bayes might not be the best choice for all types of data. Despite its simplicity, naive Bayes is often a strong baseline model in classification tasks [22]. A Naive Bayes classifier is a probabilistic model that makes predictions based on Bayes' theorem, assuming independence among features. It is often used for text classification tasks but can also be applied to recommendation systems, including e-banking recommendations. Naive Bayes is known for its simplicity and efficiency, but it may need to capture more complex relationships in the data and more advanced models. Its performance heavily relies on the independence assumption among features, which may only sometimes hold in real-world scenarios. Nevertheless, it can be a good starting point for recommendation systems, especially when interpretability and computational efficiency are essential.

2.3.4 Logistic Regression

Logistic Regression in classification is a statistical method used for binary and multiclass classification problems. Despite its name, logistic regression is used for classification rather than regression. It predicts the probability of an instance belonging to a particular class. Logistic Regression is widely used for simplicity, interpretability, and efficiency. It is a linear model, meaning it assumes a linear relationship between the input features and the log odds of the response variable. While initially designed for binary classification, logistic regression can be extended to handle multiclass classification problems using techniques like one-vs-rest or one-vs-one [23]. Logistic Regression is a linear model used for binary classification tasks, and it can be applied to e-banking recommendation systems. Logistic Regression is a simple and interpretable model, but it may not capture complex relationships in the data as effectively as more sophisticated models. Its performance depends on the underlying relationships' linearity and the feature set's appropriateness. It is generally well-suited for scenarios where interpretability is crucial, and the relationships are approximately linear.

3. MATERIALS AND METHODS

The material and methodology are formulated for the proposed work to handle large-scale transaction data, low computing time, and automation in determining customer behavior in carrying out e-banking transactions. The section of a research paper typically aims to describe the procedures and materials used in the study.

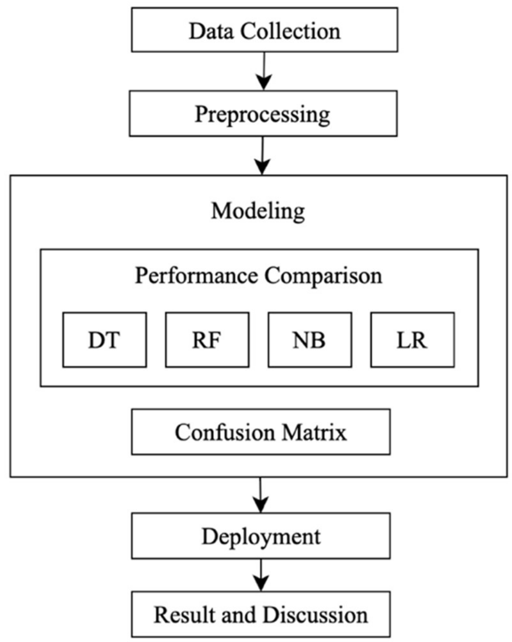

Figure 1: Machine Learning Model Using Classification Techniques

Figure 1 illustrates the proposed machine learning model for payment or purchase transactions in e-banking to support product promotion decision-making according to customer behavior. Therefore, the primary goals of this section are to ensure transparency, reproducibility, and understanding of how the research was conducted. Modeling using machine learning is presented to provide valid and actual monitoring information that can assist company management in making promotional decisions following the depiction of customer behavior.

3.1 Data Collection

The data comprises unprocessed financial information of e-banking (internet and mobile banking) sourced from Bank PT XXX spanning the years 2016 to 2020. This data encompasses customer-related details, transactional records, and product information. Transactional data captures various aspects of the transaction process, including sales, sales orders, purchase orders, delivery orders, and quotations. Customer data includes all customer information and products they have purchased. An exemplification of e-banking data is provided in Figure 2. There are 4595287 rows of data and 8 columns.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

	ServiceName	TglTransaksi	NamaPemilikRekening	NamaInstansi	BiayaAdministrasi	TotalJumlah	KodeRegional	Nominal
0	Bill Payment	2016-01-01 00:00:00.0000000	RIZKI PRXTXMX	Kartu Kredit Mega	NaN	681716.55	804.0	681716.55
1	Bill Payment	2016-01-01 00:00:00.0000000	RXTIH CHXNDRX DXWI	Kartu Kredit Mega	NaN	83350.00	807.0	83350.00
2	Bill Payment	2016-01-01 00:03:21.0000000	RXTIH CHXNDRX DXWI	Kartu Kredit Mega	NaN	5125000.00	807.0	5125000.00
3	Purchase	2016-01-01 00:39:11.0000000	RXTIH CHXNDRX DXWI	Listrik Prabayar (PLN)	3500.0	203500.00	807.0	200000.00
4	Bill Payment	2016-01-01 07:49:50.0000000	XGXS BXSRI L	Kartu Kredit Mega	NaN	1668750.00	811.0	1668750.00

Figure 2: Combined Data Sources on E-Banking Transactions (Internet and Mobile Banking)

3.2 Modeling

In machine learning, modeling refers to creating and using mathematical or statistical models that can learn patterns from data [21]. This model can then be used to make predictions, classify data, or make decisions without explicit programming of rules or instructions. The modeling process in machine learning involves several key steps:

3.2.1 Model Selection

Several classification technique models were chosen by several related studies, including Decision Tree, Random Forest, Naive Bayes, and Logistic Regression [21]. Each model type has its characteristics and uses depending on the kind of problem and data at hand. Selecting an appropriate model for e-banking recommendation is a crucial step that can significantly impact the overall performance of your recommendation system. The choice of the model depends on various factors, including the nature of your data, the task's complexity, and the e-banking domain's specific requirements. Model selection and good modeling are key to getting accurate and relevant results from data.

3.2.2 Model Training

It provided the model input and corresponding output training data [20]. The model learns from this data to adjust parameters and generate internal representations that can be used to perform specific tasks. Model training performance is critical to developing an effective e-banking recommendation system. Efficient and effective training ensures that the model can learn patterns from the data, make accurate predictions, and scale to handle the demands of a production environment. By paying attention to these considerations, you can enhance the efficiency and effectiveness of the model training process for e-banking recommendations. Regular evaluation and refinement based on performance metrics will contribute to developing a robust recommendation system.

3.2.3 Tuning Hyperparameters

It is optimizing model parameters (hyperparameters) to improve model performance. This can involve trying different combinations of hyperparameters and measuring their impact on model performance [20]. Tuning hyperparameters is a crucial step in optimizing the performance of your e-banking recommendation model. It involves adjusting the settings not learned from the data set before the training process. Proper hyperparameter tuning can significantly impact the model's ability to generalize and make accurate predictions.

3.2.4 Validation and Evaluation

Validation and evaluation are crucial steps in building and assessing the performance of an e-banking recommendation system. These processes help ensure that the model generalizes well to new, unseen data and meets the desired criteria for effectiveness—measure model performance using previously unseen data (validation or test data). Evaluation metrics, such as accuracy, precision, recall, and so on, are used to assess how much the model can generalize from training data to new data f[20]. The model can be updated, adjusted, or replaced based on the evaluation to improve its performance. This process can involve multiple iterations to develop a better model.

3.3 Deployment

Deployment in machine learning refers to running a model that has been trained and evaluated in a production environment so that it can be used to make predictions or provide solutions to a system or application [20]. This involves moving from the development and testing stage to a stage where the model can effectively process real-world data. This process uses a trained model to make predictions or classify new data that has never been seen before. Successful deployment is a critical stage in the machine learning lifecycle that ensures that the value derived from a model can be used effectively in the business context or environment in which it is deployed. Deploying an e-banking recommendation system requires careful planning and execution to ensure the model performs well in a real-world environment. By addressing these considerations, you can deploy an e-banking recommendation system that is technically sound and performs reliably and ethically in a real-world environment. Regular monitoring, maintenance, and iteration are essential for keeping the system up-to-date and aligned with evolving user needs and industry standards.

Journal of Theoretical and Applied Information Technology
15th April 2024. Vol.102. No 7
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. RESULT AND DISCUSSION

The data collection results are then subjected to data pre-processing to clean, normalize, and convert the data into a format suitable for model training. This involves understanding the input data format the model expects and ensuring that the input data meets the model's requirements. It is also necessary to ensure that regulations protect privacy data. This section paper aims to present and analyze the findings obtained from your experiments, evaluations, and tests.

4.1 Preprocessing

Data preprocessing is a crucial step in machine learning, which aims to convert and clean raw data into a format that machine learning models can understand and process well. Besides, data preprocessing is used to clean, transform, and organize the raw data into a suitable format for model training and analysis. These steps help improve model performance and accuracy. This process varies depending on the type of data and the machine-learning problem at hand. Through good data preprocessing, models can be fed clean, well-structured data, improving their ability to understand patterns and make accurate predictions. Data cleaning is carried out for missing AI or deleting incomplete or damaged data entries in this process. Then, convert categorical or qualitative variables into a form that can be processed by the model, such as with "one-hot encoding" or label coding techniques. The data normalization process is also needed to change the scale of numerical variables showing several uniform ranges of values. This helps prevent large-scale variables from dominating the model. The correlation between features is depicted in Figure 3. Therefore, the relevant parts are extracted from the data to improve model representation and performance. This involves selecting the right components and eliminating irrelevant or redundant features. Correlation between features in machine learning refers to the statistical measure of how strongly pairing elements are related. It helps to understand how changes in one variable correspond to changes in another. The correlation between features is crucial in feature selection, model interpretation, and avoiding issues like multicollinearity. Understanding the correlation between features is an essential step in the data preprocessing phase of machine learning, as it helps optimize the performance and interpretability of models.

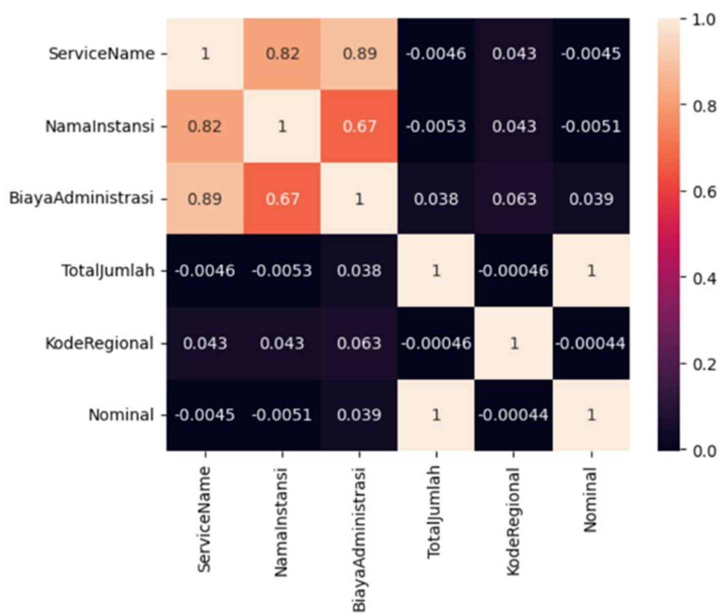

Figure 3: Correlation between Features

Additionally, anomaly cleaning is used to detect and address anomalies or outliers in the data that may impact model performance. This may involve deleting or customizing suspicious data. A robust scaler is used for normalization and dealing with outliers, fitting the training and testing data transform to avoid data leakage. Figure 4 illustrates the normalized dataset. By achieving these aims in data preprocessing, you set the foundation for a robust and effective e-banking recommendation system, ensuring that the input data is of high quality and suitable for training machine learning models.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

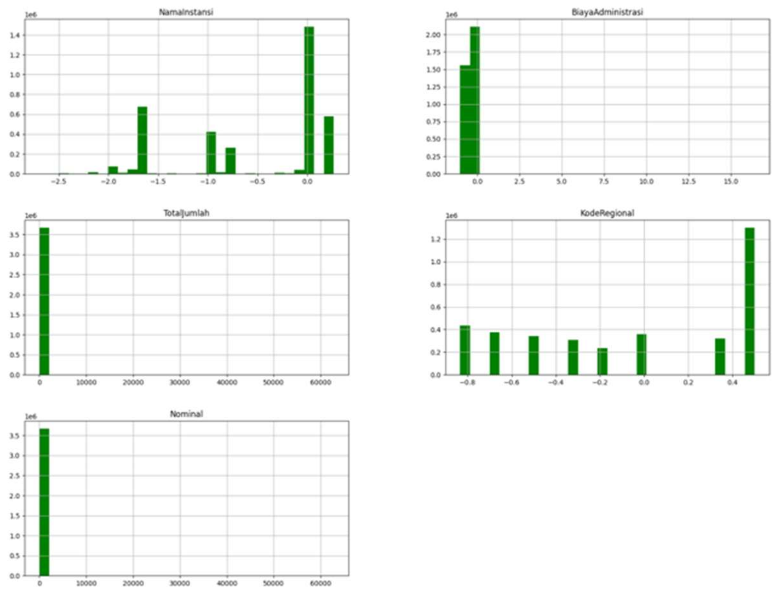

Figure 4: Normalized Data

4.2 Performance Comparison of Classification Techniques

Performance comparisons between classification techniques in machine learning often depend on the nature of the data used, the complexity of the problem, and the evaluation metrics used. Several previous researchers proposed classification techniques considering their performance, namely Decision Tree, Random Forest, Naive Bayes, and Logistic Regression. The aim is to systematically evaluate and compare the effectiveness of different classification algorithms. The performance comparison between several classification techniques is presented in Table 1.

Table 1: Performance Comparison of Classification Techniques of E-Banking Transaction.

Algorithm	Time (s)	Accuracy (%)
Decision Tree	10.34	98.61
Random Forest	442.96	98.61
Naïve Bayes	1.32	64.24
Logistic Regression	781.23	85.79

The aim of performance comparison of classification techniques of e-banking transactions is to compare various algorithms to determine their effectiveness in accurately classifying transactions. Table 1 shows the highest accuracy is between Decision Tree and Random Forest with 98.61%. However, the fastest time execution is the Decision Tree algorithm. Figures 5 and 6 illustrate the visual of the performance comparison by time execution and the algorithm chosen. Moreover, the confusion matrix results can be seen in Figure 7.

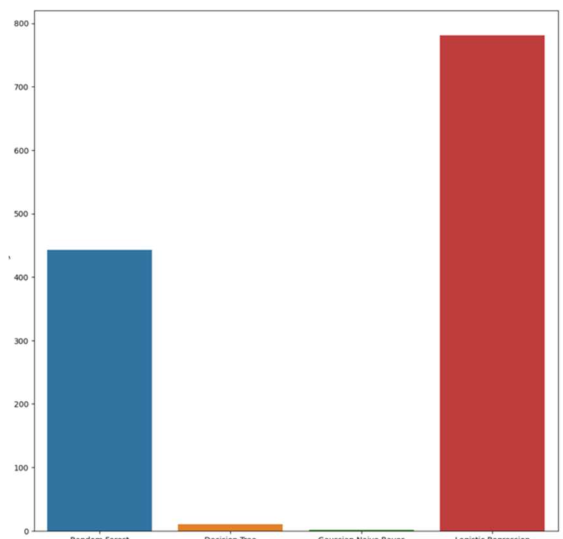

Figure 5: The Performance Comparison of Classification Techniques based on Time Execution

Accuracy is the ratio of correctly predicted instances to the total cases. However, optimizing for accuracy in scenarios where the classes are imbalanced may lead to many false negatives (fraudulent transactions misclassified as non-fraudulent), a critical issue in e-banking security. Besides, Precision and recall are alternative metrics that provide a more nuanced understanding of the model's performance. Precision is the ratio of correctly predicted positive instances to predicted positive cases. At the same time, recall (sensitivity) is the percentage of correctly predicted positive

Journal of Theoretical and Applied Information Technology
15th April 2024. Vol.102. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

models to actual positive models. Therefore, achieving a high precision may reduce false positives (non-fraudulent transactions misclassified as fraudulent), but it might result in a lower recall, meaning some fraudulent transactions are not detected. In e-banking transactions for customer behavior, the consequences of missing a fraudulent transaction (false negative) could be severe. On the other hand, marking a legitimate transaction as fraudulent (false positive) may inconvenience the user but might be more acceptable from a security standpoint. The choice between optimizing for precision or recall depends on the business priorities and the proper trade-offs between false positives and false negatives.

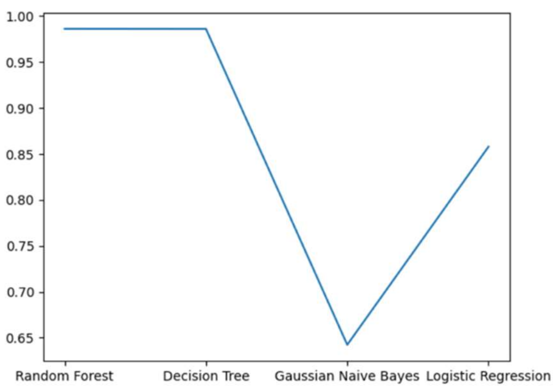

Figure 6: The Performance Comparison of Classification Techniques based on Accuracy Value

Confusion Matrix is a table used in classification to measure model performance. It presents the correct and incorrect prediction results in each category or class. The confusion matrix allows us to calculate various model performance evaluation metrics in Figures 7 a and b. The confusion matrix helps provide a deeper understanding of the model's performance in multiple aspects. It can be used to determine whether the model is better at handling certain classes or whether there is an imbalance between the classes. Based on this, the best classification technique is the Decision Tree. This technique will implement the machine learning model to show customer behavior based on e-banking data transactions. By addressing these aims, the confusion matrix performance provides a comprehensive evaluation of the customer behavior of the e-banking transaction model's classification performance, offering insights into strengths and weaknesses. This analysis aids in understanding the model's behavior and guides potential improvements for enhanced recommendation accuracy in e-banking transactions for customer behavior scenarios.

4.3 Implementation of The Best Classification Technique for E-Banking Transactions

The implementation process involves selecting the best algorithm for fast processing, namely the Decision Tree. Classification results using this selected algorithm can show customer behavior. The classification results in Figure 8 show that the E-Banking (Internet and Mobile) services that are frequently used are Internet Bill Payments for Credit Cards and Mobile Phone Balance Purchases by using a T-SEL provider. This can also be seen because of the low administration costs. Therefore, it can be interpreted that many customers often use these product services.

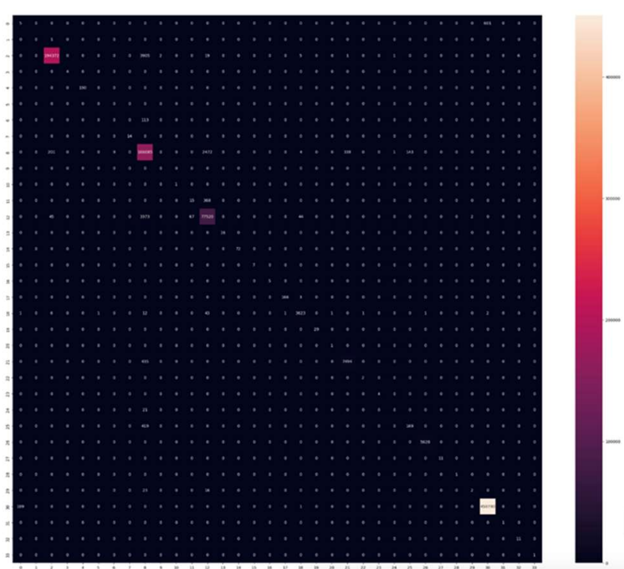

Figure 7 a: The Confusion Matrix

	0	1	2	3	4	5	6	7	8	9	...	25	26	27	28	29	30	31	32	33	34
0	52333	5	0	41	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	0	7632	0	0	0	0	0	0	0	440	...	0	0	0	0	0	0	0	0	0	0
2	0	0	5	0	0	0	0	0	0	0	...	0	0	0	0	0	0	575	0	0	0
3	27	0	0	142168	0	0	0	0	0	3935	...	0	0	0	0	0	0	0	0	4	0
4	0	0	0	0	4	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	199	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	114	...	0	0	0	0	0	0	0	0	0	0

Figure 7 b: The Confusion Matrix

	NamaInstansi	BiayaAdministrasi	TotalJumlah	KodeRegional	Nominal	ServiceName
0	Kartu Kredit [redacted]	50000.0	535000.0	407.0	5000000.0	INTERNET Bill Payment
1	PULSA REGULAR-TSEL	2000.0	153500.0	707.0	900000.0	MOBILE Isi Ulang HP

Figure 8: The Sample of Classification Result using Decision Tree

4.4 Discussion

Mobile and internet banking as an e-banking platform has risen in Indonesia. With the increasing penetration of smartphones and improved internet connectivity, more customers are using mobile apps to access banking services conveniently. Currently, Customers in Indonesia access mobile and internet banking through their smartphones. With the increasing convenience of using e-banking platforms for various transactions, including fund transfers, bill payments, online purchases, etc., The convenience and accessibility of these services contribute to their popularity. It's important to note that the landscape of e-banking can change rapidly with technological advancements and shifts in consumer behavior. For the most up-to-date information, consider checking recent reports, surveys, or official statements from relevant financial institutions and authorities in Indonesia. Customer behavior refers to individuals' actions, decisions, and activities when searching for, purchasing, using, evaluating, and disposing products and services. Understanding customer behavior is crucial for businesses and marketers as it helps them tailor their products, services, and marketing strategies to meet the needs and preferences of their target audience. Companies analyze customer behavior through market research, surveys, and data analytics to gain insights about product development, marketing strategies, and customer relationship management efforts.

Based on the experiment result, the Decision Tree is the best classification algorithm with the highest accuracy and fastest execution. By analyzing past transaction data, machine learning models can predict how customers will likely respond to specific marketing campaigns. This enables banks to optimize their marketing strategies for better engagement and conversion rates.

Understanding customer behavior in e-banking transactions is crucial for several reasons, and it plays a significant role in improving services, fraud detection, and overall customer satisfaction. Besides, analyzing customer behavior helps identify unusual patterns or activities that may indicate fraudulent transactions. By establishing a baseline of normal behavior for each customer, anomalies such as large transactions, irregular spending patterns, or transactions from unusual locations can be flagged for further investigation. On the other hand, analyzing e-banking transaction behavior helps segment customers based on their usage patterns, preferences, and financial needs. This segmentation allows banks to develop targeted marketing strategies, offer relevant products, and provide personalized communication to different customer segments. Analysis of transaction behavior aids in predicting customer churn. By identifying signs of dissatisfaction or reduced engagement, banks can take proactive measures to retain customers. This may involve offering incentives, personalized promotions, or addressing specific pain points to enhance customer loyalty.

Monitoring customer behavior enhances the security of e-banking platforms. Financial institutions can implement multi-factor authentication or additional security measures when deviations from normal behavior are detected. This proactive approach helps in preventing unauthorized access and protecting customer accounts. Therefore, understanding customer behavior enables banks to offer personalized services and recommendations. By analyzing transaction history, spending habits, and preferences, financial institutions can tailor product offerings and promotions to meet individual customer needs, ultimately enhancing the customer experience. Furthermore, Customer behavior analysis contributes to effective risk management. Financial institutions can implement risk mitigation strategies by identifying high-risk transactions or customers, such as adjusting credit limits, enhancing fraud monitoring, or conducting additional verification checks.

Customer behavior insights inform the development of new products and services. By understanding what customers value and need in their digital banking experience, financial institutions can innovate and introduce features that resonate with customer expectations, keeping them engaged and satisfied. Besides, understanding how customers interact with e-banking platforms helps improve the overall user experience. Banks can optimize their digital interfaces, streamline processes, and introduce features that align with customer preferences, leading to a more user-friendly and efficient e-banking experience. Insights into customer behavior enable banks to optimize operational processes. Banks can allocate resources more efficiently, enhance service delivery, and reduce operational costs by understanding peak transaction times, popular services, and shared user preferences. Therefore, analyzing customer behavior on e-banking transactions is a strategic approach that benefits financial institutions and customers. It enhances security, personalizes services, informs decision-making, and ultimately contributes to a more positive and secure e-banking experience.

5. CONCLUSION

The machine learning model has been successfully implemented through the experiment by comparing several classification techniques such as Decision Tree, Random Forest, Naïve Bayes, and Logistic Regression. The Decision Tree is implemented as the best classifier to understand the e-banking transaction data. A deep understanding of customer behavior by comparing banking transactions provides banks significant strategic and operational benefits. This allows them to provide better services, offer promotions, improve security, and stay relevant amidst the tight competition in the current digital banking landscape. The customer mostly used Internet Bill Payments for Credit Cards and Mobile Phone Balance Purchases by using a T-SEL provider. This is also evident due to the minimal expenses related to administration because many customers frequently utilize this product's services. However, this study's effectiveness can be elevated by incorporating alternative approaches to categorizing comparable items or entities according to specific attributes or features. Beyond its utilization in e-banking services, other classification techniques can be implemented across diverse domains for many objectives.

REFERENCES:

[1] C. J. Anwar, I. Suhendra, E. Purwanda, A. Salim, N. A. Rakhmawati, and F. Jie, “Investigating the relationship between monetary policy, macro-prudential policy and credit risk in Indonesia banking industry,” *Heliyon*, vol. 9, no. 7, p. e18229, 2023, doi: 10.1016/j.heliyon.2023.e18229.

[2] A. Zulkifli, A. Yani, and E. Nilda, “CUSTOMER PERCEPTION OF E-BANKING SERVICES AT PT. BANK RAKYAT INDONESIA,” *AICOS: Asian Journal Of Islamic Economic Studies*, vol. 01, pp. 87–93, 2024.

[3] I. A. Firmansyah, R. Yasirandi, and R. G. Utomo, “The influence of efficacy, credibility, and normative pressure to M-banking adoption level in Indonesia,” *Procedia Comput Sci*, vol. 197, no. 2021, pp. 51–60, 2021, doi: 10.1016/j.procs.2021.12.117.

[4] R. A. Kasri, B. S. Indrastomo, N. D. Hendranastiti, and M. B. Prasetyo, “Digital payment and banking stability in emerging economy with dual banking system,” *Heliyon*, vol. 8, no. 11, p. e11198, 2022, doi: 10.1016/j.heliyon.2022.e11198.

[5] H. Ben Abdennebi, “M-banking adoption from the developing countries perspective: A mediated model,” *Digital Business*, vol. 3, no. 2, p. 100065, 2023, doi: 10.1016/j.digbus.2023.100065.

[6] M. H. Rahman, J. Rahman, T. Tanchangya, and M. A. Esquivias, “Green banking initiatives and sustainability: A comparative analysis between Bangladesh and India,” *Research in Globalization*, vol. 7, no. December, p. 100184, 2023, doi: 10.1016/j.resglo.2023.100184.

[7] M. Wijayaningtyas, K. Lukiyanto, E. Nursanti, and D. I. Laksmana, “The effect of economical phenomenon on informal construction workers earnings within Covid-19 pandemic: A mixed method analysis,” *Heliyon*, vol. 8, no. 8, p. e10321, 2022, doi: 10.1016/j.heliyon.2022.e10321.

[8] M. Rahman, I. Ismail, S. Bahri, and M. K. Rahman, "An Empirical Analysis of Cashless Payment Systems for Business Transactions," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 8, no. 4, p. 213, 2022, doi: 10.3390/joitmc8040213.

[9] X. Chen, M. H. Miraz, M. A. I. Gazi, M. A. Rahaman, M. M. Habib, and A. I. Hossain, "Factors affecting cryptocurrency adoption in digital business transactions: The mediating role of customer satisfaction," *Technol Soc*, vol. 70, no. July, p. 102059, 2022, doi: 10.1016/j.techsoc.2022.102059.

[10] R. Schmuck, "The use of online business models," *Procedia Manuf*, vol. 54, pp. 45–51, 2020, doi: 10.1016/j.promfg.2021.07.008.

[11] M. Hosseini, N. Abdolvand, and S. R. Harandi, "Two-dimensional analysis of customer behavior in traditional and electronic banking," *Digital Business*, vol. 2, no. 2, p. 100030, 2022, doi: 10.1016/j.digbus.2022.100030.

[12] C. R. Lages, R. Perez-Vega, S. Kadić-Maglajlić, and N. Borghei-Razavi, "A systematic review and bibliometric analysis of the dark side of customer behavior: An integrative customer incivility framework," *J Bus Res*, vol. 161, no. February, 2023, doi: 10.1016/j.jbusres.2023.113779.

[13] B. Brkovi, M. Zarkovi, and T. Miljkovi, "International Journal of Electrical Power and Energy Systems Machine learning for power transformer SFRA based fault detection ˇ," vol. 156, no. July 2023, 2024, doi: 10.1016/j.ijepes.2023.109779.

[14] M. H. Mobarak *et al.*, "Scope of machine learning in materials research—A review," *Applied Surface Science Advances*, vol. 18, no. November, p. 100523, 2023, doi: 10.1016/j.apsadv.2023.100523.

[15] S. Z. El Mestari, G. Lenzini, and H. Demirci, "Preserving data privacy in machine learning systems," *Comput Secur*, vol. 137, no. March 2023, p. 103605, 2024, doi: 10.1016/j.cose.2023.103605.

[16] R. Gardas and S. Narwane, "An analysis of critical factors for adopting machine learning in manufacturing supply chains," *Decision Analytics Journal*, vol. 10, no. April 2023, p. 100377, 2024, doi: 10.1016/j.dajour.2023.100377.

[17] M. M. Chowdhury, R. S. Ayon, and Md. S. Hossain, "An investigation of machine learning algorithms and data augmentation techniques for diabetes diagnosis using class imbalanced BRFSS dataset," *Healthcare Analytics*, vol. 5, no. December 2023, p. 100297, 2023, doi: 10.1016/j.health.2023.100297.

[18] P. Talari *et al.*, "Hybrid feature selection and classification technique for early prediction and severity of diabetes type 2," *PLoS One*, vol. 19, no. 1 January, Jan. 2024, doi: 10.1371/journal.pone.0292100.

[19] M. Salauddin Khan *et al.*, "Comparison of multiclass classification techniques using dry bean dataset," *International Journal of Cognitive Computing in Engineering*, vol. 4, no. March 2022, pp. 6–20, 2023, doi: 10.1016/j.ijcce.2023.01.002.

[20] L. J. Shangguan and Z. Sun, "Enhancing data quality and real-time sharing performance in water informatics through decision tree mining algorithm," *Measurement: Sensors*, vol. 31, no. November 2023, p. 101024, 2024, doi: 10.1016/j.measen.2023.101024.

[21] B. K. Meher, M. Singh, R. Birau, and A. Anand, "Forecasting stock prices of fintech companies of India using random forest with high-frequency data," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 10, no. 1, p. 100180, 2024, doi: 10.1016/j.joitmc.2023.100180.

[22] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, "Variable selection for Naïve Bayes classification," *Comput Oper Res*, vol. 135, p. 105456, 2021, doi: 10.1016/j.cor.2021.105456.

[23] J. J. Wimhurst, J. S. Greene, and J. Koch, "Predicting commercial wind farm site suitability in the conterminous United States using a logistic regression model," *Appl Energy*, vol. 352, no. February, p. 121880, 2023, doi: 10.1016/j.apenergy.2023.121880.