Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 1, February 2026, Page. 730-742
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

Performance Comparison Of Xgboost Lightgbm And Lstm For E-Commerce Repeat Buyer Prediction

Lustiyono Prasetyo Nugroho*[1], Rujianto Eko Saputro[2], Fandy Setyo Utomo[3]

[1,2,3]Computer Science, Universitas Amikom Purwokerto, Indonesia

Email: [1]lustiyonoprasetyo2010@gmail.com

Received : Feb 26, 2026; Revised : Mar 3, 2026; Accepted : Mar 3, 2026; Published : Mar 8, 2025

Abstract

Repeat buyer behavior is a critical indicator of customer retention success in e-commerce platforms. However, accurately predicting repeat buyers remains a challenging problem due to the complexity of user behavior patterns and the temporal characteristics embedded in interaction data. Existing studies often focus on single modeling approaches or limited sequence exploration, resulting in insufficient comparative insight between ensemble-based machine learning and sequence-based deep learning models. Therefore, this study aims to systematically compare the performance of tree-based ensemble models (XGBoost and LightGBM) and a sequence-based deep learning model (LSTM) in predicting repeat buyers using user behavior data. To ensure fair evaluation, data preprocessing and feature engineering were carefully designed to prevent data leakage by utilizing user behavior prior to the first purchase. Model performance was evaluated using Accuracy, F1-score, and ROC–AUC metrics. Experimental results show that XGBoost and LightGBM achieve stable classification performance with accuracy values of 86.11% and 85.84%, respectively, while the LSTM model attains the highest ROC–AUC value of 0.937, indicating superior capability in capturing temporal behavioral patterns. This study provides valuable insights for e-commerce platforms seeking to optimize predictive models for repeat buyers, contributing to more effective customer retention strategies.

Keywords : *E-commerce, LightGBM, LSTM, Repeat Buyer Prediction, XGBoost.*

1. INTRODUCTION

The global digital economy has grown rapidly, positioning e-commerce as a key driver of national economic growth. E-commerce has reshaped market structures and consumer behavior through expanding internet penetration in both urban and rural areas, highlighting its role not only as a commercial platform but also as an instrument of digital and financial inclusion [1], [2]. Repeat buyers therefore represent a strategic asset for sustaining platform growth and profitability. Moreover, e-commerce has a positive impact on the development of micro, small, and medium enterprises (MSMEs). Previous study demonstrates that e-commerce adoption significantly increases employment, revenue, and profits among Indonesian MSMEs, strengthening local economic foundations [3]. In this context, maintaining and increasing repeat buyer rates is increasingly important, as repeat buyers contribute not only to revenue but also to customer loyalty that is difficult to replace through new customer acquisition. Machine learning as a branch of artificial intelligence that enables systems to learn from data and automatically improve their performance through pattern recognition, allowing them to generate accurate predictions for previously unseen data [4], [5]. This approach provides a critical foundation for data-driven.

Repeat buyer behavior is a key indicator of the effectiveness of customer retention strategies, as repeat customers contribute substantially to long-term revenue in e-commerce. Despite the continued growth of e-commerce, retaining repeat buyers remains a major challenge. Highlights that basic user interactions play a significant role in predicting repeat purchases, yet such factors are rarely integrated

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

comprehensively into prediction models [6]. Meanwhile, previous study demonstrates the effectiveness of session-based variables in improving predictive performance [7]. These findings indicate a research gap in comparative studies between traditional models and sequence-based models for repeat buyer prediction, underscoring the importance of this study in evaluating their respective performances.

In the context of repeat buyer prediction in e-commerce, algorithmic machine learning and deep learning approaches stand out due to their respective strengths. XGBoost has proven effective in predicting customer behavior and churn by improving accuracy and AUC [8], [9]. LightGBM is capable of robustly modeling relationships between user characteristics and purchasing behavior, despite challenges in interpretability [10]-[13]. Meanwhile, LSTM excels in leveraging sequential user behavior data to capture long-term dependencies in purchasing processes, achieving strong predictive performance with enhanced interpretability through SHAP-based analysis [14]-[16].

Although several studies have demonstrated the superiority of certain algorithms in the domain of customer behavior prediction, a research gap remains that needs to be addressed. The study previously shows that sequence-based approaches to consumer behavior still yield relatively limited performance, with AUC values ranging from 0.59 to 0.76 despite experiments conducted with various sequence lengths [17]. This indicates that exploring sequence length alone is insufficient to capture the complexity of repeat purchase behavior.

On the other hand, classical machine learning based studies such confirm that the XGBoost algorithm is capable of delivering stable predictive performance; however, such studies typically rely on a single model without comparison to other ensemble algorithms or deep learning models that account for temporal aspects [8], [18]. In line with this, emphasizes the importance of comparative modeling approaches, demonstrating that ensemble combinations such as Voting Classifier, XGBoost, and LightGBM can achieve high performance, with accuracy, F1-score, and an AUC above of 80% on global transaction data [19]-[21]. Previous research has shown that comparing the performance between different machine learning classifiers provides important insights in selecting the best model for a particular classification task [22], [23], . Meanwhile, sequence-based deep learning approaches such as LSTM have shown strong capability in modeling user navigation behavior as temporal sequences [24], [25], [26].

Based on this gap, the present study is designed to bridge the research gap by comparing the performance of tree-based ensemble models (XGBoost and LightGBM) and a sequence-based deep learning model (LSTM) in predicting repeat buyers using e-commerce user behavior data. Unlike prior studies that focus on optimizing sequence length, this study emphasizes consistent user behavior representation through feature engineering that is free from data leakage, as well as a comparative evaluation of models to assess the extent to which sequence-based approaches provide predictive advantages over non-sequential ensemble models in the context of repeat buyer behavior.

2. METHOD

This study adopts a quantitative research design with a comparative–experimental approach, aiming to compare the performance of three algorithms XGBoost, LightGBM, and LSTM in predicting repeat buyers on e-commerce platform [19].

2.1. Dataset

This study utilizes secondary data in the form of e-commerce user behavior logs obtained from the research of [17], specifically the public *Taobao UserBehavior.csv* dataset provided by Alibaba Tianchihttps://tianchi.aliyun.com/dataset/649. The dataset includes various types of user interactions, such as page views, add-to-cart actions, favorite markings, and purchase transactions.

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

Table 1. Feature and Type Data

Feature	Type Data
User_id	Numeric
Item_id	Numeric
Category_id	Numeric
Behavior	Category
Timestamp	Category

Table 1 shows that the dataset contains four features consisting of two data types (Numeric and Category). Numerical data are represented in the form of numbers, whereas categorical data are represented by character-based values.

2.2. Research Flow

The research workflow is designed based on best practices in international studies on user behavior analysis in e-commerce. This study follows a series of systematic stages adapted from previous research, starting from problem identification and concluding with the discussion of the experimental results of three models. The research workflow is illustrated using a flowchart diagram adapted from [17].

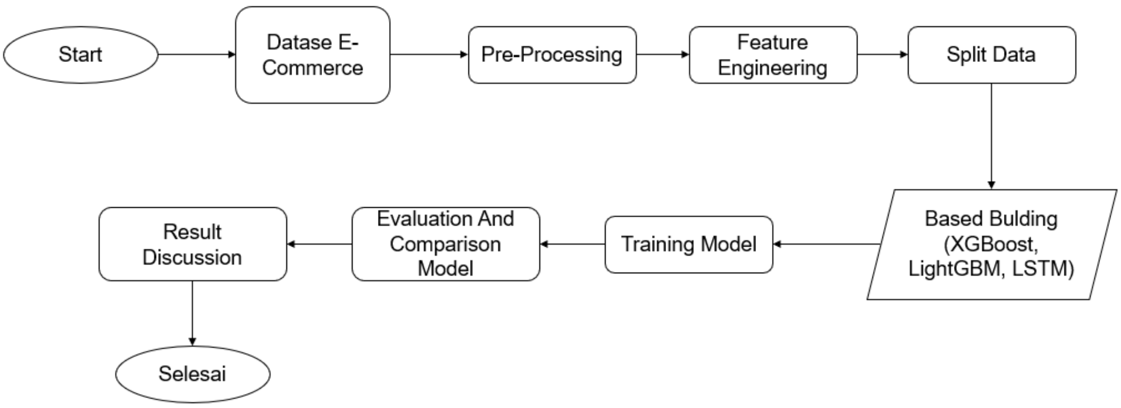


Figure 1. Research Flow

2.2.1. Pre-Processing

The data preprocessing stage aims to prepare raw data so that it is clean, structured, and consistent before feature construction and modeling are performed. This process focuses on the technical processing of user behavior data to ensure that the data are ready for the feature engineering and model development stages. Next, the data are sorted based on timestamps so that the sequence of user interactions accurately reflects the actual chronological order of events. The timestamps are then converted into an appropriate time format to facilitate time-based analysis in subsequent stages [17].

The final stage of preprocessing is the determination of the repeat buyer label, which involves classifying users based on their purchasing patterns. This label is used as the target variable in the training and evaluation of the predictive models in the subsequent stage.

2.2.2. Feature Engineering

Feature engineering is a crucial stage in this study, aimed at transforming the preprocessing data into informative and relevant feature representations for the modelling process. Through feature engineering, user behaviour patterns that are initially embedded in raw data can be extracted into

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

variables that more effectively reflect user characteristics, product attributes, and the dynamics of interactions [17].

Specifically for the LSTM model, user behaviour data are organized into sequences based on the chronological order of interactions. Each sequence represents a series of user activities over time, enabling the model to learn temporal dependencies and sequential patterns in user behaviour. These sequences are then formatted into inputs compatible with the LSTM architecture [25].

2.2.3. Split Dataset

After the feature engineering process is completed, the dataset is divided into training and testing sets 80:20 [9]. This partitioning aim to ensure that the developed models can be objectively evaluated using data that were not used during the training process. The training set is used to train the XGBoost, LightGBM, and LSTM models, while the testing set is employed to assess the models' generalization ability in predicting repeat buyer behaviour.

2.2.4. XGBoost Model Design

At this stage, the XGBoost model is developed by utilizing the features generated through the feature engineering process to predict repeat buyer behaviour. The model is configured based on the principles of gradient boosting, with the objective of minimizing prediction errors through iterative learning. XGBoost is employed due to its ability to handle large-scale data and effectively capture non-linear relationships within user behaviour data [11], [16], [17]. In calculating the XGBoost classifier in this study, three equations from previous research are employed [17]. The first is the loss function, which is presented in Equation (1) – (3).

$$L(y, f(x)) = (y - f(x))^2 \tag{1}$$

$$bj(\Theta) = \Sigma[L(yi, y^\wedge i) + RT(\Theta t)] \tag{2}$$

$$gi = \frac{\partial L(yi,y^\wedge i)}{\partial y^\wedge i}, h_i \frac{\partial 2L(yi,y^\wedge i)}{\partial y^\wedge{}_i^2} \tag{3}$$

In Equation (1), the objective is to measure the difference between the actual values (y) and the model predictions (f(x)). The smaller the value of this function, the better the model's performance. Equation (2), the regularization term $RT(\Theta_t)$, serves as a penalty to prevent the model from becoming overly complex (overfitting). Meanwhile, Equation (3) is used to compute the direction and magnitude of error changes at each iteration.

2.2.5. LightGBM Model Design

The LightGBM model is developed using a histogram-based gradient boosting approach designed to improve computational efficiency [12]. In this study, LightGBM utilizes the same set of features as XGBoost to ensure a fair comparison. This algorithm is selected due to its ability to process data quickly and efficiently, particularly on large-scale datasets, as well as its capability to determine optimal node splits during the training process. According to previous study in processing data, LightGBM employs three fundamental equations as follows [13].

$$f(x) = \sum_{t=0}^{T} ht(x) \tag{4}$$

$$f^\wedge = \frac{argminEx,y[L(y,f(x))]}{f} \tag{5}$$

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

$$Vj(d) = \frac{1}{n}\frac{(\sum xi\in Algi+\beta 1-\alpha\sum xi\in Blgi)^2}{n_{jl}\,(d)} + \frac{(\sum xi\in Argi+\beta 1-\alpha\sum xi\in Brgi)^2}{n_{jr}\,(d)} \quad (6)$$

In Equation (4), the prediction model is expressed as the sum of the outputs of Tdecision trees. Equation (5), known as the objective function, is used to minimize the loss function between the actual values and the predicted results. Meanwhile, Equation (6) describes the calculation of variance gain to determine the optimal split in the LightGBM tree using the Gradient-based One-Side Sampling (GOSS) method.

2.2.6. LSTM Model Design

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) that is capable of learning long-term dependencies in sequential data [16], [20]. LSTM employs three main mechanisms namely the forget gate, input gate, and output gate to control the flow of information, allowing the model to retain or discard important information from previous data sequences. The fundamental LSTM equations adopted from [16] are presented as follows.

$$\begin{aligned} ft &= \sigma g(Wfxt + Ufht - 1 + bf) \\ it &= \sigma g(Wixt + Uiht - 1 + bi) \\ ot &= \sigma g(Woxt + Uoht - 1 + bo) \\ ct = ft \odot ct - 1 + it &\odot \sigma c(Wcxt + Ucht - 1 + bc)ht = ot \odot \sigma h(ct) \end{aligned} \quad (7)$$

This set of equations explains how LSTM regulates the flow of information within its neural network. The forget gate (f_t) determines which information is removed from the previous memory state, the input gate (i_t) incorporates new information, and the output gate (o_t) controls the final output. Through this mechanism, LSTM excels at recognizing sequential patterns such as click, add-to-cart, and purchase sequences in e-commerce which are critical for predicting repeat buyer behaviour.

2.2.7. Training Model

At this stage, the XGBoost, LightGBM, and LSTM models are trained using the prepared training data to learn the relationships between features and the repeat buyer labels [5]. During the training process, the hyperparameters of each model are carefully tuned to optimize performance and prevent overfitting, ensuring that the resulting models exhibit stable performance and strong generalization capability [18].

2.2.8. Evaluation and Comparison Model

The evaluation stage is conducted to compare the performance of XGBoost, LightGBM, and LSTM in predicting repeat buyer behavior using the testing data. The models are consistently assessed using accuracy, precision, recall, F1-score, and AUC metrics, with the results presented in comparison tables and figures to identify the model with the best overall performance.

Accuracy is used to measure the overall proportion of correctly classified instances, providing a general indication of model correctness. However, since repeat buyer prediction may involve class imbalance, the F1-score is included to balance precision and recall, offering a more reliable measure of classification performance for both classes. Furthermore, ROC–AUC is utilized to evaluate the model's discriminative ability across different classification thresholds. In binary classification tasks such as repeat buyer prediction, ROC–AUC is particularly important because it reflects how well the model

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

distinguishes between repeat and non-repeat buyers independently of a specific decision threshold, making it a robust metric for comparing predictive models.

3. RESULT

3.1. Data Analysis

The dataset used consists of 1,018,018 rows of user interaction data reflecting user activities toward products, including page views (pv), add-to-cart actions (cart), favorites (fav), and purchases (buy), which is similar to the dataset used in the study by [17]. After applying data preprocessing and user-level aggregation, a group of users was obtained and analyzed based on their interaction history prior to the first purchase in order to avoid data leakage. Figure 2 is an example of raw data that has been loaded using Jupyter Notebook.

	user_id	item_id	category_id	behavior	timestamp
0	1	2268318	2520377	pv	1511544070
1	1	2333346	2520771	pv	1511561733
2	1	2576651	149192	pv	1511572885
3	1	3830808	4181361	pv	1511593493
4	1	4365585	2520377	pv	1511596146

Figure 2. Load Dataset

```
Raw rows: 1018010
Users with at least one buy: 6801
Sample aggregated features (safe):
```

	user_id	pv_count	cart_count	fav_count	unique_items	unique_cats	first_event_ts	last_event_ts
0	1	55	0	0	46	20	2017-11-24 17:21:10	2017-12-03 06:17:27
1	100	13	0	3	11	6	2017-11-24 19:08:36	2017-11-25 02:22:13
2	115	227	3	11	204	44	2017-11-25 07:43:44	2017-12-03 12:04:01
3	118	90	0	0	74	19	2017-11-24 23:01:32	2017-12-03 11:05:49
4	119	44	0	2	44	18	2017-11-24 20:44:03	2017-11-27 14:59:58

```
Label distribution (agg):
label
1    0.674704
0    0.325296
Name: proportion, dtype: float64
```

Figure 3. Pre-Processing and Feature Engineering

Based on figure 3, the analysis results indicate that there are 6,801 users who have made at least one purchase transaction. The labeling process classifies users as repeat buyers if they have made a purchase, and as non-repeat buyers if they have never made a purchase. The label distribution shows that approximately 67.47% of users are categorized as repeat buyers, while 32.53% are classified as non-repeat buyers. This distribution reflects realistic conditions on e-commerce platforms, where the majority of active users tend to make repeat purchases. The timestamps were converted into an appropriate datetime format to support time-based analysis in subsequent stages.

Overall, the results of the preliminary data analysis indicate that the dataset exhibits complex and informative user behavior characteristics in terms of interaction intensity, product diversity, and temporal dynamics. Therefore, this dataset is considered suitable for conducting repeat buyer prediction modeling experiments using machine learning and deep learning approaches.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

3.2. Experimental Design (Split Dataset)

The dataset that has undergone preprocessing and feature engineering is subsequently divided into training and testing sets using the hold-out validation method. The data are split using an 80% training and 20% testing ratio, with stratified sampling applied based on the target variable (repeat buyer) to preserve class proportions in both subsets. The data splitting is performed at the user level, ensuring that no users appear simultaneously in both the training and testing sets. This approach aims to prevent data leakage and to ensure that the models are evaluated on users who were not previously seen during the training process. This validation is confirmed by the absence of any overlap between user indices in the training and testing datasets.

3.3. Model Building Configuration

This study employs three primary models XGBoost, LightGBM, and LSTM each configured according to the characteristics of its respective approach and use training model. The models were implemented using standard configurations commonly adopted in prior studies. XGBoost and LightGBM were trained using default tree-based settings.

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	?	0 (unbuilt)
lstm_1 (LSTM)	?	0 (unbuilt)
dropout_1 (Dropout)	?	0
dense_1 (Dense)	?	0 (unbuilt)

Total params: 0 (0.00 B)
Trainable params: 0 (0.00 B)
Non-trainable params: 0 (0.00 B)

Figure 4. Architecture Sequential User Behaviour

Meanwhile, the LSTM Architecture model was designed to capture sequential user behavior through an embedding layer followed by a single LSTM layer and a dropout mechanism to mitigate overfitting, as illustrated in figure 4. No extensive hyperparameter tuning was conducted to ensure a fair comparison across models.

3.4. Result Model

Based on figure 5 specifically, the XGBoost model is able to identify the majority of repeat buyers with a high recall rate, indicating that most users who truly make repeat purchases are correctly recognized by the model. At the same time, the misclassification rate for non-repeat buyers remains within an acceptable range, indicating that the model is not overly aggressive in predicting all users as repeat buyers.

```
XGBoost metrics
Accuracy: 0.8611111111111112
F1      : 0.8953809711804185
AUC     : 0.9303877371510222
              precision    recall  f1-score   support

           0       0.77      0.82      0.79       621
           1       0.91      0.88      0.90      1287

    accuracy                           0.86      1908
   macro avg       0.84      0.85      0.84      1908
weighted avg       0.86      0.86      0.86      1908
```

Figure 5. Result of XGBoost Model

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

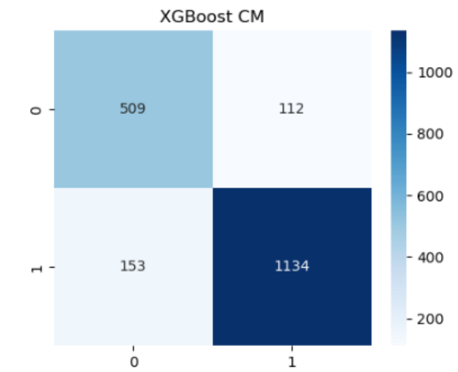


Figure 6. Confusion Matrix XGBoost

These values suggest that the model demonstrates strong capability in distinguishing between repeat buyers and non-repeat buyers. Based on the classification report, the model exhibits stable performance across both classes, with relatively balanced precision and recall levels. Based on figure 6, the XGBoost confusion matrix shows that out of 621 non-repeat buyers (class 0), 509 users are correctly classified, while 112 users are incorrectly predicted as repeat buyers. On the other hand, out of 1,287 repeat buyers (class 1), the model correctly identifies 1,134 users, whereas 153 users are misclassified as non-repeat buyers. These results confirm that the XGBoost model has a strong capability in identifying repeat buyers, as reflected by the high number of true positives

The LightGBM model is applied to the same dataset as the previous model, using identical training and testing data splits to ensure consistency in the evaluation process. Model performance is evaluated using Accuracy, F1-score, and AUC metrics, and is further analyzed through a confusion matrix to examine the distribution of predicted repeat buyer and non-repeat buyer classes.

```
LightGBM metrics
Accuracy: 0.8584905660377359
F1      : 0.8932806324110671
AUC     : 0.9323133477722849

Classification Report LightGBM:
              precision    recall  f1-score   support

           0       0.76      0.82      0.79       621
           1       0.91      0.88      0.89      1287

    accuracy                           0.86      1908
   macro avg       0.84      0.85      0.84      1908
weighted avg       0.86      0.86      0.86      1908
```

Figure 7. Result of LightGBM

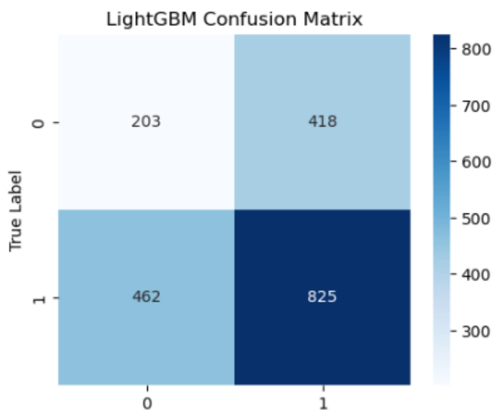
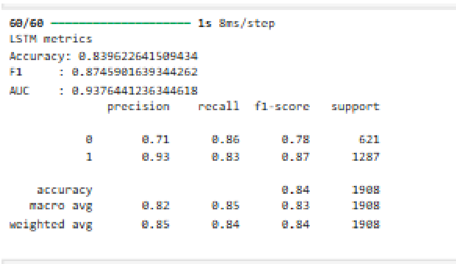


Figure 8. Confusion Matrix LightGBM

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

As in figure 7, the relatively high AUC value indicates that the LightGBM model demonstrates a strong capability in distinguishing between repeat buyers and non-repeat buyers. This performance suggests that LightGBM is well-suited for modelling user behaviour patterns relevant to repeat buyer prediction. Figure 8, LightGBM confusion matrix, the model successfully classified 825 repeat buyers correctly (true positives), however 462 repeat buyers were still misclassified as non-repeat buyers (false negatives). For the non-repeat buyer class, the model correctly identified 203 users (true negatives), while 418 non-repeat buyers were incorrectly predicted as repeat buyers (false positives). This result indicates that LightGBM exhibits a higher false negative rate, suggesting potential limitations in identifying all repeat buyers. In the context of repeat buyer prediction, a high number of false negatives suggests that a portion of potential repeat customers may go undetected, which could negatively affect the effectiveness of retention strategies based on this model.

The Long Short-Term Memory (LSTM) model is applied to model user behavior as a sequence of interactions, enabling the analysis of temporal patterns prior to purchase events. The evaluation is conducted on the same testing data as the previous models to ensure consistency in result comparison.



```
60/60 ━━━━━━━━━━━━━━━━━━━━ 1s 8ms/step
LSTM metrics
Accuracy: 0.839622641509434
F1      : 0.8745901639344262
AUC     : 0.9376441236344618
              precision    recall  f1-score   support

           0       0.71      0.86      0.78       621
           1       0.93      0.83      0.87      1287

    accuracy                           0.84      1908
   macro avg       0.82      0.85      0.83      1908
weighted avg       0.85      0.84      0.84      1908
```

Figure 9. Result of LSTM

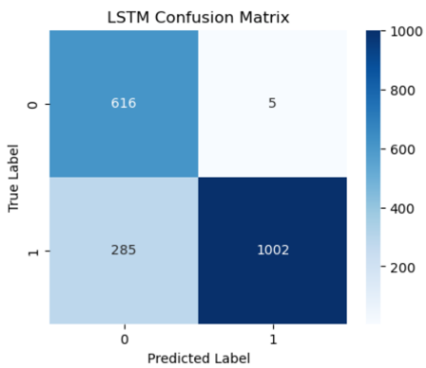


Figure 10. Confusion Matrix LSTM

Based on figure 9, LSTM model with a maximum sequence length of 50 demonstrated relatively strong performance in predicting repeat buyers, achieving an accuracy of 83.96%, an F1-score of 0.875, and a ROC–AUC value of 0.938. The high AUC value indicates the model's strong capability in distinguishing repeat buyers from non-repeat buyers based on sequential behavioral patterns.

As in figure 10 the LSTM confusion matrix, the model correctly classified 616 non-repeat buyers (true negatives) and 1,002 repeat buyers (true positives). Misclassification errors in the non-repeat buyer class were relatively low, with only 5 users incorrectly predicted as repeat buyers (false positives). However, there were still 285 repeat buyers that were misclassified as non-repeat buyers (false negatives). These results indicate that the LSTM model is highly effective in avoiding false repeat buyer predictions, but it continues to face challenges in capturing all behavioral patterns of users who actually engage in repeat purchases. Overall, the distribution of errors suggests that the sequence-based approach

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

of LSTM is effective in modeling users' temporal behavior, although it may still fail to identify a portion of true repeat buyers.

4. DISCUSSIONS

4.1. Evaluation Discussions

Table 2. Performance Model Comparison

Model	Accuracy	F1-Score	ROC-AUC
XGBoost	0.8611	0.8953	0.9303
LightGBM	0.8584	0.8932	0.9323
LSTM	0.8396	0.8745	0.9376

Based on the table 2, all models demonstrate strong performance in identifying repeat buyers, with ROC–AUC values exceeding 0.90. The XGBoost and LightGBM models produce relatively balanced results, achieving the highest accuracy and F1-scores, which indicate a strong ability to leverage aggregated user behavior features prior to the first purchase. This suggests that tree-based ensemble approaches are effective for stable and efficient repeat buyer classification.

On the other hand, the LSTM model yields the highest ROC–AUC value, indicating superior discriminative capability in distinguishing between repeat buyers and non-repeat buyers. This advantage suggests that sequence-based modeling of user interactions is able to capture temporal dynamics in user behavior that are not fully represented by static features. However, the slightly lower accuracy and F1-score of the LSTM model compared to the other models indicate a trade-off between discriminative power and classification stability.

4.2. Comparative Discussion

This study provides a comparative evaluation of tree-based ensemble models (XGBoost and LightGBM) and a sequence-based deep learning model (LSTM) for repeat buyer prediction, and the findings are largely consistent with, while also extending, prior research in this domain.

Previous studies have demonstrated the effectiveness of XGBoost in modeling customer behavior and retention. XGBoost achieved stable predictive performance in e-commerce customer loss prediction, emphasizing its strength in handling non-linear relationships within user behavior data [8]. Similarly studies, showed that XGBoost can leverage consumer interaction features, however, their study reported relatively limited performance, with AUC values ranging from 0.59 to 0.76, even after experimenting with different sequence lengths [17]. In contrast, the results of this study show that XGBoost achieves a substantially higher ROC–AUC value (0.9303), suggesting that careful feature engineering and leakage-free aggregation of pre-purchase behavior can significantly enhance predictive performance.

LightGBM has also been reported as an efficient and robust alternative to XGBoost in previous research [21]. Highlighted LightGBM's capability to model large-scale consumer data efficiently, though potential challenges remain in capturing complex temporal behavior [10], [13], [26]. In this study, LightGBM achieves performance comparable to XGBoost, with an ROC–AUC of 0.9323, confirming findings from previous study that ensemble-based approaches are highly competitive for repeat purchase prediction tasks [19]. However, the higher false negative rate observed in the LightGBM confusion matrix indicates that, while efficient, the model may overlook a portion of potential repeat buyers, which is a critical consideration in customer retention strategies.

Sequence-based modeling using LSTM has been explored in several prior studies to capture temporal dependencies in user behavior. LSTM-based models can effectively learn sequential purchase patterns, often achieving strong discriminative performance [14], [25]. Consistent with these findings,

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

the LSTM model in this study achieves the highest ROC–AUC value (0.9376), indicating superior ability in distinguishing repeat buyers from non-repeat buyers. Compared to [17], who focused primarily on optimizing sequence length, this study demonstrates that maintaining consistent sequence representation combined with comparative evaluation against ensemble models provides clearer insights into the relative strengths of sequence-based approaches.

Overall, the comparative results confirm that tree-based ensemble models offer strong and stable classification performance using aggregated behavioral features, while LSTM provides enhanced discriminative capability by modeling temporal user behavior. These findings align with recent comparative studies [19], [22], [16], reinforcing the importance of evaluating multiple modeling paradigms rather than relying on a single algorithm. By integrating ensemble and sequence-based perspectives within a unified experimental framework, this study contributes a more comprehensive understanding of repeat buyer prediction in e-commerce contexts.

Beyond methodological comparison, the findings of this study carry important practical implications for e-commerce platforms. Accurate repeat buyer prediction directly supports customer retention strategies, which are significantly more cost-efficient than acquiring new customers. By understanding the trade-off between classification stability (as offered by ensemble models) and temporal discriminative power (as demonstrated by LSTM), platform managers can align predictive modeling choices with specific business objectives. For example, models with higher recall for repeat buyers can be prioritized for proactive retention campaigns, personalized promotions, or loyalty program targeting. Meanwhile, models with strong ROC–AUC performance enable more flexible threshold adjustment based on marketing budgets and risk tolerance. Consequently, this research provides not only theoretical insight into modeling paradigms but also actionable guidance for improving long-term customer engagement, optimizing retention interventions, and enhancing revenue sustainability in competitive e-commerce environments.

5. CONCLUSION

This study aims to evaluate and compare the performance of different machine learning approaches for repeat buyer prediction in an e-commerce context. Three representative models, namely XGBoost, LightGBM, and LSTM, were employed to reflect tree-based ensemble methods and sequence-based deep learning approaches. All models were developed using a consistent preprocessing and feature engineering pipeline designed to prevent data leakage, ensuring an objective assessment of predictive performance.

The experimental results demonstrate that all three models achieve strong predictive performance in identifying repeat buyers. XGBoost and LightGBM exhibit stable results with high accuracy and F1-score values, highlighting the effectiveness of ensemble-based approaches in leveraging aggregated behavioral features. Meanwhile, the LSTM model achieves the highest ROC–AUC value, indicating a superior capability in distinguishing repeat buyers from non-repeat buyers by modeling temporal user behavior sequences.

Overall, the findings emphasize the importance of a comparative modeling approach in repeat buyer prediction, as each model offers distinct advantages. Ensemble methods provide robustness and efficiency, while sequence-based models offer enhanced discriminative power for capturing temporal behavior patterns. This study is expected to contribute practical insights for developing customer prediction systems in e-commerce platforms and to serve as a reference for future research exploring more advanced behavioral representations and modeling techniques. including the enrichment of behavioral and contextual features, the application of feature selection or representation learning techniques, and the adoption of more advanced modeling approaches such as attention-based or transformer architectures. For future research, evaluating the models on more diverse or cross-platform e-commerce datasets is necessary to improve generalization and enhance the practical relevance of the findings. Add the Hyperparameter Tuning method to create a more optimal performance comparison.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Universitas Amikom Purwokerto, for providing academic support and facilities during the completion of this research. Appreciation is extended to all parties who contributed valuable insights and technical assistance during the research process.

REFERENCES

[1] Z. M. Jannah, Miftahul, "The Digital Economy Boom: How E-Commerce is Reshaping Indonesia's Market," 2025. doi: https://doi.org/10.15294/indi.v2i1.23034.
[2] S. F. Mauludiah, Y. M. Arif, M. Faisal, and D. D. Putra, "Struggling Models: An Analysis of Logistic Regression and Random Forest in Predicting Repeat Buyers with Imbalanced Performance Metrics," *Appl. Inf. Syst. Manag.*, vol. 7, no. 2, pp. 31–38, 2024, doi: 10.15408/aism.v7i2.39326.
[3] S. Silaban, N. Kusnadi, and F. Feryanto, "the Impact of E-Commerce Implementation on the Performance of Micro and Small Industries in Indonesia," *Agrisocionomics J. Sos. Ekon. Pertan.*, vol. 8, no. 1, pp. 126–133, 2024, doi: 10.14710/agrisocionomics.v8i1.17725.
[4] M. Izhan and M. Yusoff, "Machine Learning : An Overview," pp. 89–99, 2024, doi: 10.4236/ojmsi.2024.123006.
[5] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies*, vol. 12, no. 7, 2024, doi: 10.3390/technologies12070113.
[6] E. Kuric, A. Puskas, P. Demcak, and D. Mensatorisova, "Effect of Low-Level Interaction Data in Repeat Purchase Prediction Task," *Int. J. Hum. Comput. Interact.*, vol. 40, no. 10, pp. 2515–2533, 2024, doi: 10.1080/10447318.2023.2175973.
[7] R. Esmeli and A. Gokce, "An Analysis of Consumer Purchase Behavior Following Cart Addition in E-Commerce Utilizing Explainable Artificial Intelligence," *J. Theor. Appl. Electron. Commer. Res.* , vol. 20, no. 1, pp. 1–18, 2025, doi: 10.3390/jtaer20010028.
[8] L. Gan, "XGBoost-Based E-Commerce Customer Loss Prediction," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/1858300.
[9] S. N. Ruscikasani, R. Roro, N. Oktalivia, F. R. Putra, and A. J. Wahidin, "Prediksi Pembelian E-Commerce Menggunakan XGBoost Berbasis Perilaku Sesi Pengguna," vol. 4, no. 4, pp. 5666–5672, 2025.
[10] J. Du, "Predictions for Consumer Behaviour of E-Commerce Sales Data 2023-2024 Based on the LightGBM Model," no. Ecai 2024, pp. 109–116, 2025, doi: 10.5220/0013207400004568.
[11] F. Alharbi, "A comparative study of SMOTE and ADASYN for multiclass classification of IoT anomalies," *Int. J. Inf. Technol. Secur.*, vol. 17, no. 2, pp. 15–24, 2025, doi: 10.59035/qefu7977.
[12] R. Sinaga and S. Widianto, "Understanding Telecommunication Customer Churn: Insights from LightGBM Predictive Modelling and SHAP Feature Interpretation," *ASEAN Mark. J.*, vol. 15, no. 1, 2023, doi: 10.7454/amj.v15i1.1229.
[13] Y. Ling, "RF-LighGBM: A PROBABILISTIC ENSEMBLE WAY TO PREDICT CUSTOMER REPURCHASE BEHAVIOUR IN COMMUNITY E-COMMERCE A," pp. 1–15, 2021.
[14] B. Predić, M. Ćirić, and L. Stoimenov, "Business Purchase Prediction Based on XAI and LSTM Neural Networks," *Electron.*, vol. 12, no. 21, 2023, doi: 10.3390/electronics12214510.
[15] D. Tran and A. W. Tham, "Accuracy Comparison Between Feedforward Neural Network, Support Vector Machine and Boosting Ensembles for Financial Risk Evaluation," *J. Risk Financ.*

Vol. 7, No. 1, February 2026, Page. 730-742
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.1.5746

Manag., vol. 18, no. 4, 2025, doi: 10.3390/jrfm18040215.

[16] M. Nasseri, T. Falatouri, P. Brandtner, and F. Darbanian, "Applying Machine Learning in Retail Demand Prediction—A Comparison of Tree-Based Ensembles and Long Short-Term Memory-Based Deep Learning," *Appl. Sci.*, vol. 13, no. 19, 2023, doi: 10.3390/app131911112.

[17] K. Cai and M. R. Rodavia, "XGBoost Analysis based on Consumer Behavior," *Front. Comput. Intel. Syst.*, vol. 5, no. 2, pp. 85–89, 2023, [Online]. Available: doi: https://doi.org/10.54097/fcis.v5i2.12974

[18] Sutarman, R. Siringoringo, D. Arisandi, E. Kurniawan, and E. B. Nababan, "Model Klasifikasi Dengan Logistic Regression Dan Recursive Feature Elimination Pada Data Tidak Seimbang," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 735–742, 2024, doi: 10.25126/jtiik.1148198.

[19] M. Z. Alam and T. Roy, "Predicting Online Repeat Purchases: A Comparative Analysis of Machine Learning Algorithms," *2025 Int. Conf. Electr. Comput. Commun. Eng. ECCE 2025*, no. May, pp. 1–6, 2025, doi: 10.1109/ECCE64574.2025.11013423.

[20] S. J. Haddadi, A. Farshidvard, F. dos S. Silva, J. C. dos Reis, and M. da Silva Reis, "Customer churn prediction in imbalanced datasets with resampling methods: A comparative study," *Expert Syst. Appl.*, vol. 246, no. September 2023, p. 123086, 2024, doi: 10.1016/j.eswa.2023.123086.

[21] W. Junhai, W. Yunfeng, O. Ibrahim, and S. Member, "Explainable E-Commerce Transaction Prediction Using LightGBM and Permutation Importance," *IEEE Access*, vol. 14, no. January, pp. 10153–10169, 2026, doi: 10.1109/ACCESS.2026.3654903.

[22] J. Rixen, N. Blass, S. Lyra, and S. Leonhardt, "Comparison of Machine Learning Classifiers for the Detection of Breast Cancer in an Electrical Impedance Tomography Setup," 2023.

[23] Z. Shao and M. N. Ahmad, "Comparison of Random Forest and XGBoost Classifiers Using Integrated Optical and SAR Features for Mapping Urban Impervious Surface," 2024.

[24] K. Diamantaras, M. Salampasis, A. Katsalis, and K. Christantonis, "Predicting shopping intent of e-commerce users using LSTM recurrent neural networks," *Proc. 10th Int. Conf. Data Sci. Technol. Appl. DATA 2021*, no. Data, pp. 252–259, 2021, doi: 10.5220/0010554102520259.

[25] R. Kasemrat and T. Kraiwanit, "Attention-enhanced LSTM for high-value customer behavior prediction: Insights from Thailand's E-commerce sector," *Intell. Syst. with Appl.*, vol. 26, no. December 2024, p. 200523, 2025, doi: 10.1016/j.iswa.2025.200523.

[26] I. Z. A. Illah, W. S. Jauharis Sapu, and A. T. Damaliana, "Implementasi Metode Klasifikasi LightGBM dan Analisis Survival dalam Memprediksi Pelanggan Churn," *J. Komtika (Komputasi dan Inform.*, vol. 8, no. 1, pp. 43–53, 2024, doi: 10.31603/komtika.v8i1.11194.