Volume XI Issue 3 Year 2026 | Page 1053-1063 | ISSN: 2527-9866
Received: 05-07-2026 | *Revised: 30-07-2026* | *Accepted: 20-08-2026*

Customer Lifetime Value Prediction Using Long Short-Term Memory with RFM-T Features Based on E-Commerce Customer Data

Sabilla Laili Ramadhani Putri[1], Anik Vega Vitianingsih[2], Anastasia Lidya Maukar[3], Achmad Muzakki[4], Hewa Majeed Zangana[5]
[1,2]Informatics Department, Universitas Dr. Soetomo, Surabaya, Indonesia
[3]Industrial Engineering Department, President University, Bekasi, Indonesia
[4]Information Systems Department, Telkom University, Surabaya, Indonesia
[5]IT Department, Duhok Technical College, Duhok Polytechnic University, Duhok, Iraq
e-mail: sabillalaili2@gmail.com[1], vega@unitomo.ac.id[2], almaukar@president.ac.id[3], achmad.muzakki@telkomuniversity.ac.id[4], hewa.zangana@dpu.edu.krd[5]
**Correspondence*: vega@unitomo.ac.id

Abstract: *CLV prediction plays an important role in e-commerce by helping companies identify valuable customers and develop effective retention strategies. However, predicting CLV remains challenging due to the sequential nature of customer purchasing behavior. This study proposes a CLV prediction model using the LSTM algorithm with Recency, Frequency, Monetary, and Tenure features extracted from the Online Retail II dataset. The CLV target is calculated as the accumulated monetary value generated during the three months following the historical observation period. The proposed approach consists of data preprocessing, monthly RFM-T feature engineering, feature normalization using Min-Max Scaling, LSTM model training, denormalization, and performance evaluation using MAE and RMSE. The LSTM model incorporates stacked LSTM layers, Batch Normalization, Dropout, and L2 regularization to improve learning stability and generalization. Experimental results indicate that the model was able to capture customer purchasing patterns based on sequential RFM-T features, with training and validation loss trends showing stable convergence. The proposed model achieved an MAE of 43.16 and an RMSE of 125.39 on the original monetary scale, reflecting the prediction performance obtained on the evaluated dataset. These findings suggest that the proposed LSTM model with RFM-T features provides an approach for CLV prediction in e-commerce.*

Keywords: *Customer Lifetime Value, Long Short-Term Memory, LSTM, RFM-T, E-commerce.*

1. Introduction

Advances in digital technology have driven the growth of electronic commerce (e-commerce) as one of the primary channels for modern commercial activity. Ease of access, a wide variety of product choices, and increasingly convenient transaction systems have led to a continuous increase in the number of e-commerce users and transaction volume [1], [2]. This increase in activity generates large amounts of historical customer data that companies can use to gain insights into customer behavior as a basis for business decision-making [3]. One particularly important metric is Customer Lifetime Value (CLV), which is an estimate of the economic value a customer can provide over the course of their relationship with a company. Information regarding CLV can help companies develop marketing strategies, retain potential customers, and improve the efficiency of marketing resource allocation [4]. Although large volumes of customer transaction data are available, many companies still struggle to identify customers who make a long-term contribution to the company's revenue. Most marketing strategies remain focused on short-term sales growth without taking into account the long-term value of customers. Consequently, companies risk losing high-value customers and may fail to optimise their customer retention strategies. Therefore, a predictive model is required that can estimate CLV based on historical customer transaction patterns, enabling companies to segment their customers more effectively and support data-driven decision-making.

Recent advances in CLV prediction have increasingly relied on deep learning models instead of conventional machine learning techniques. Among these models, LSTM has shown consistent performance across different application domains. For example, the integration of an Attention Mechanism with LSTM achieved 99.83% accuracy, an AUC of 0.985, an RMSE of 0.0391, and an MAE of 0.0039 on Thai e-commerce transaction data [5]. In a different study, using Monte Carlo Dropout with an LSTM model led to a root mean square error of 0.19 when forecasting the customer lifetime value of global online shoppers [6]. Similarly, an LSTM model used in a Software as a Service setting achieved a Root Mean Square Error of 2.271 and an adjusted Symmetric Mean Absolute Percentage Error of 22.9% [7]. These results consistently show that LSTM-based deep learning techniques can effectively provide precise customer lifetime value predictions in different business situations [8].

Among the developed approaches, the LSTM method has been widely adopted for modelling sequential and time-series customer transaction data [9], [10], [11]. In contrast to standard RNNs, LSTM has a memory system that is managed by input, forget, and output gates. This feature helps the network keep important information across lengthy sequences and reduces the issue of vanishing gradients [12], [13]. Therefore, the LSTM method was selected in this study as the primary model for predicting customer CLV based on e-commerce transaction history [14].

Nevertheless, previous studies still have several limitations. In previous CLV prediction studies, the Recency, Frequency, and Monetary (RFM) framework has been the primary set of features used for model development [15] , while the integration of customer relationship duration (Tenure) into CLV prediction models has not been widely explored compared with traditional RFM features [16], [17]. Furthermore, most studies have focused primarily on improving model performance without evaluating the contribution of additional features capable of representing the length of time a customer has remained active with the company [18], [19]. These limitations indicate that there is still scope to develop CLV prediction models that utilise customer information more comprehensively. To overcome these challenges, this research creates a customer lifetime value prediction model that relies on LSTM and utilizes RFM-T characteristics. Unlike previous studies that primarily rely on the conventional RFM framework, the proposed model incorporates the Tenure feature to provide a more comprehensive representation of customer behaviour. By accounting for the length of a customer's active relationship with the company, the additional feature is expected to improve the identification of high-value customers and enhance CLV prediction performance [20].

The objective of this research is to construct an LSTM-based Customer CLV prediction model using the RFM-T framework. Model development is carried out using the Online Retail II dataset obtained from the UC Irvine Machine Learning Repository. The resulting model is expected to assist companies in identifying high-value customers, supporting the development of customer retention strategies, and enhancing the effectiveness of data-driven marketing decision-making. This research is an original study developed on the basis of a review of previous research, with an approach and model implementation designed independently.

2.Methods

This section explains the methodology employed in the study in sufficient detail to enable the experiments to be replicated. All established methods are supported by appropriate references, while any modifications or adaptations introduced in this research are clearly described. This study employs an experimental research approach to develop and evaluate a CLV prediction model using the LSTM algorithm with RFM-T features as representations of customer transaction behavior. The experimental approach was selected because it enables the model development process to be conducted systematically, allowing the obtained results to be objectively evaluated and reproduced. This research uses LSTM because its design helps it understand time-based trends from ordered data and keeps important information for extended periods. Consequently, LSTM has become a widely used model for prediction tasks based on historical data [21].

The model for predicting Customer Lifetime Value is created with the Online Retail II dataset, which can be accessed by anyone through the UC Irvine Machine Learning Repository. The overall research methodology was designed systematically to ensure that the proposed model effectively represents customer transaction behavior and provides reliable prediction results. Furthermore, the methodology was structured to facilitate the reproducibility of the experimental process and to enable an objective evaluation of the effectiveness of integrating RFM-T features with the LSTM algorithm for predicting Customer Lifetime Value [22].

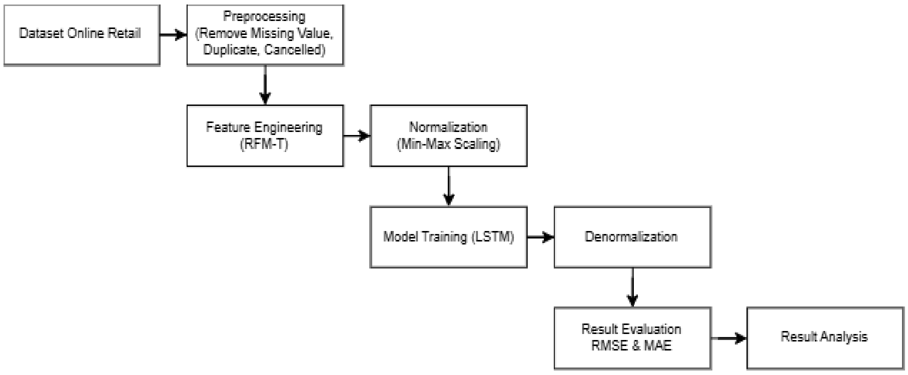

Figure 1. Research Flow

Figure 1 shows the complete process used in this research to create the suggested LSTM-based CLV prediction model by employing the RFM-T framework. The process begins with the acquisition of the Online Retail II dataset, after which data preprocessing is carried out to ensure data quality. Customer behaviour features are subsequently derived using the RFM-T approach, and the resulting features are normalized through Min-Max Scaling prior to model training. After the LSTM model generates the predicted CLV values, the outputs are transformed back to their original scale using a denormalization process. The effectiveness of the model is evaluated using RMSE and MAE, and then a review of the experimental findings is conducted.

A. Data Description

The proposed model is developed using transaction data from the Online Retail II dataset, which is publicly accessible through the UCI Machine Learning Repository. The information shows the buying habits of a UK online shop from December 1, 2010, to December 9, 2011. Although the dataset contains multiple attributes, this study focuses exclusively on Customer ID, InvoiceDate, Quantity, and Price, as these variables are sufficient for generating customer behaviour features under the RFM-T framework. Before being used in the modeling process, the data underwent preprocessing to improve data quality. These steps included removing records with empty Customer IDs, removing canceled transactions, and removing duplicate records. After preprocessing is complete, the data is transformed into a monthly dataset, with the RFM-T features used as input to the LSTM model.

B. Data Preprocessing

During the preprocessing stage, the dataset was cleaned to improve data quality before RFM-T feature engineering was performed. The process involved standardizing column names, converting the InvoiceDate attribute to a datetime format, converting the Price attribute to a numeric value, removing canceled transactions, removing records without a Customer ID, and removing duplicate records. Next, the data is sorted by InvoiceDate so that the customer's transaction history is arranged chronologically and ready for use in the feature engineering stage.

C. Feature Engineering Using RFM-T

Feature construction involved converting raw transaction records into four behavioural indicators: Recency, Frequency, Monetary, and Tenure. Adding the Tenure factor expands the usual RFM model by taking into account how long the customer has been actively engaged with the business. To reflect changes in purchasing behaviour over time, the RFM-T indicators were calculated at monthly intervals and then used as sequential inputs for the LSTM model.

Table 1. Definition of RFM-T Features

	Feature	Description
1	Recency	Calculates how long ago a customer bought something by figuring out the days from their most recent purchase to the end of the month being reviewed.
2	Frequency	Represents customer purchasing activity as the total number of transactions recorded within the observation month.
3	Monetary	Total monetary value of customer transactions during the observation month, measured in Pound Sterling (£).
4	Tenure	Demonstrates the duration of the customer's activity by calculating the number of days from the first purchase to the final day of the month in question.

In this study, CLV is defined as the cumulative monetary value generated by a customer during the future prediction period. The model utilizes historical monthly RFM-T sequences as input features, while the CLV target is calculated from future Monetary values after the observation period. The historical observation window consists of six consecutive months, and the prediction horizon is defined as the following three months. Therefore, the Monetary feature used as input does not contain information from the target period, preventing target leakage during the prediction process.

$$CLV_i = \sum_{t=T+1}^{T+h} Monetary_{i,t} \qquad (1)$$

Where T represents the final month in the historical observation window and h represents the prediction horizon. In this study, h is set to three months.

Table 2. Input and Target Construction

	Component	Description
1	Historical observation window	6 consecutive months
2	Input features	Recency, Frequency, Monetary, Tenure
3	Prediction horizon	3 future months
4	Target variable	Accumulated future Monetary value
5	Input Dimensions	(samples, 6, 4)
6	Time steps	6

To construct sequential input data for the LSTM model, the monthly RFM-T observations were arranged chronologically for each customer. Each input sequence consists of six consecutive monthly observations, where each time step contains four behavioural features: Recency, Frequency, Monetary, and Tenure. Therefore, the input dimension of the LSTM model is represented as (samples, 6, 4), where 6 indicates the number of historical time steps and 4 represents the number of input features.

For each six-month input sequence, the corresponding CLV target was generated from the accumulated Monetary value of the following three months. This sliding-window approach allows the model to learn temporal purchasing patterns from historical customer behaviour and predict future customer value. Months without customer transactions were retained in the sequence to preserve the temporal structure. For these inactive months, Frequency and

Monetary values were assigned as zero, while Recency and Tenure were updated according to the monthly observation period. The resulting monthly RFM-T observations were then arranged chronologically for each customer to construct sequential input data for the LSTM model. An example of the resulting Monthly RFM-T dataset is presented in Table 3.

Table 3. Example of Monthly RFM-T Features

	Customer ID	Year Month	Recency	Frequency	Monetary (£)	Tenure
1	13	2011-05	23	2	10.35	25
2	13	2011-06	53	0	0.00	55
3	13	2011-07	6	1	0.83	86
4	13	2011-08	13	1	4.13	117
5	13	2011-09	19	1	1.65	147
6	13	2011-10	1	1	6.13	178
7	13	2011-11	0	2	2.07	208
8	13	2011-12	31	0	0.00	239

D. Data Normalization

Data normalization was performed to standardize the scale of each feature, thereby making the LSTM model training process more stable. Prior to normalization, the RFM-T features were first clipped using the 1st and 99th percentiles of the training data to reduce the influence of outliers. To ensure consistent feature values, Min-Max Scaling was employed to normalize all features within the interval of 0–1. The normalization process was fitted using only the training data, while the resulting parameters were reused for the validation and test sets, thereby preventing data leakage.

In contrast to the input features, the target variable, CLV, did not undergo Min-Max Scaling for normalization. The target value is transformed using the Signed Logarithm (Signed Log1p) function to reduce the influence of a skewed distribution caused by customers with very high transaction values. The formula for Min-Max Scaling utilized in this research is displayed in Equation (2).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (2)$$

E. Dataset Splitting Strategy

The dataset was divided into training, validation, and testing sets based on customer identity. Customer-level splitting was applied to ensure that transaction records from the same customer were not distributed across different subsets, thereby preventing customer information leakage. This strategy enables the model to evaluate its generalization capability on unseen customers. The dataset was split into 70% training data, 10% validation data, and 20% testing data. Although the transaction data contain temporal information, customer-based splitting was selected because this study focuses on evaluating CLV prediction performance across different customer profiles while maintaining customer independence between datasets. Within each customer subset, transaction records were maintained in chronological order before sequence generation for LSTM input construction.

Table 4. Dataset Splitting Strategy

Dataset	Proportion	Purpose
Training	70%	Model training
Validation	10%	Validation monitoring and early stopping
Testing	20%	Final performance evaluation

F. LSTM

LSTM overcomes one of the main limitations of conventional RNNs by employing an internal memory architecture that preserves information across long sequences. As a result, the model can effectively capture temporal relationships in sequential data, which is essential for accurate CLV prediction. The flow of information in an LSTM cell is managed through a set of three gates: forget, input, and output. The initial phase includes the forget gate, which looks at details from the last hidden state and carefully eliminates unneeded information before adding the current input into the memory cell. The formulation of this operation is shown in Equation (3).

$$f_t = \sigma\,(W_f[W_{t-1}, x_t] + b_f) \qquad (3)$$

The input gate controls how new data is added to the memory cell by determining which pieces of information should be kept at this moment. This operation is mathematically defined in Equation (4).

$$i_t = \sigma\,(W_i[h_{t-1}, x_t] + b_i) \qquad (4)$$

The candidate cell state is then generated using the hyperbolic tangent activation function, as expressed in Equation (5).

$$\hat{C}_t = \tanh\,(W_c[h_{t-1}, x_t] + b_c) \qquad (5)$$

Following this, the memory cell gets refreshed by merging the old cell state with the suggested cell state using the input gate, as indicated in Equation (6).

$$c_t = f_t * c_{t-1} + i_t \times \hat{C}_t \qquad (6)$$

In the end, the output gate decides which part of the hidden state moves forward to the next time step and helps make the prediction, as shown in Equation (7).

$$o_t = \sigma\,(W_o[h_{t-1}, x_t] + b_o) \qquad (7)$$

The output gate controls the information flow from the cell state to the hidden state. The hidden state is then generated by combining the output gate with the activated cell state, as expressed in Equation (8).

$$h_t = o_t \times \tanh(c_t) \qquad (8)$$

A stacked LSTM network is constructed using two recurrent layers with 64 and 32 hidden units, respectively. To improve the stability and generalization capability of the model, Batch Normalization, Dropout, and L2 regularization are incorporated during training. Parameter optimization is carried out using the Adam algorithm by minimizing the Mean Squared Error (MSE) loss. Moreover, the training strategy includes Early Stopping, Model Checkpoint, and Reduce Learning Rate on Plateau, enabling efficient convergence while preserving the optimal model based on validation performance.

F. Denormalization

The prediction results produced by the LSTM model are still represented in the transformed target space. Therefore, a denormalization process is performed to convert the predicted values back to their original monetary scale (£), making the prediction results easier to interpret and suitable for performance evaluation. In this study, the target values were transformed using the Signed Log1p function during the normalization stage. Consequently, the prediction results were restored to their original scale by applying the inverse transformation before calculating the evaluation metrics. This process ensures that the predicted CLV can be interpreted as actual monetary values.

G. Evaluation Metrics

Performance evaluation was conducted after converting the normalized predictions back to their original monetary scale. Model performance was evaluated using MAE and RMSE. Equation (9) defines Mean Absolute Error (MAE) as the average absolute difference between the observed and predicted CLV values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (9)$$

Observed values are denoted by y_i, predicted values by $\hat{y}_i$, and $n$indicates the total samples considered. RMSE is also employed to evaluate prediction accuracy by measuring the square root of the mean squared error between actual and estimated values. Because larger errors receive greater weight during computation, RMSE penalizes substantial prediction deviations more heavily than MAE. Its formulation is shown in Equation (10).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \qquad (10)$$

Lower values of MAE and RMSE indicate better prediction performance because they represent smaller differences between the predicted and actual CLV.

3. Results and Discussion

The following section reports the outcomes of the LSTM model developed for CLV prediction based on RFM-T features. Evaluation focuses on RMSE and MAE values, together with visual representations that describe the model training dynamics and predictive capability.

A. Comparison of RFM and RFM-T Models

To evaluate the contribution of the Tenure feature, an additional experiment was conducted by comparing two LSTM models: LSTM-RFM using Recency, Frequency, and Monetary features, and LSTM-RFM-T using Recency, Frequency, Monetary, and Tenure features. Both models were trained and evaluated using the same dataset splitting strategy and experimental settings to ensure a fair comparison.

Table 5. Performance Comparison between RFM and RFM-T Models

Model	Features	MAE	RMSE
LSTM-RFM	Recency, Frequency, Monetary	43.3025	124.8661
LSTM-RFM-T	Recency, Frequency, Monetary, Tenure	43.2911	125.6761

The comparison results show that the LSTM-RFM-T model achieved a slightly lower MAE value than the LSTM-RFM model, decreasing from 43.3025 to 43.2911. However, the RMSE value increased from 124.8661 to 125.6761. These results indicate that the Tenure feature provides additional information regarding customer relationship duration, but its contribution to overall CLV prediction performance is limited in this dataset. Table 5 presents the comparative experiment results, whereas Table 6 reports the final evaluation of the selected LSTM-RFM-T model.

B. Model Training Performance

Figure 2 presents the progression of the loss values obtained during model training. The evolution of training and validation losses was analyzed to examine convergence stability and the model's capacity to generalize. The training loss indicates the error computed from the training samples, whereas the validation loss reflects the model's performance on a separate validation dataset. Similar trends between the two curves suggest that the model learns representative patterns while preserving stable predictive performance on unseen data.

To improve training efficiency and prevent unnecessary iterations, the Early Stopping technique was employed during the optimization process. This strategy automatically terminated the training when the validation loss showed no further improvement for a predefined number of epochs while restoring the model weights that achieved the lowest validation loss. Consequently, the selected model represents the best-performing configuration obtained during training and is expected to provide reliable prediction performance on unseen customer transaction data.

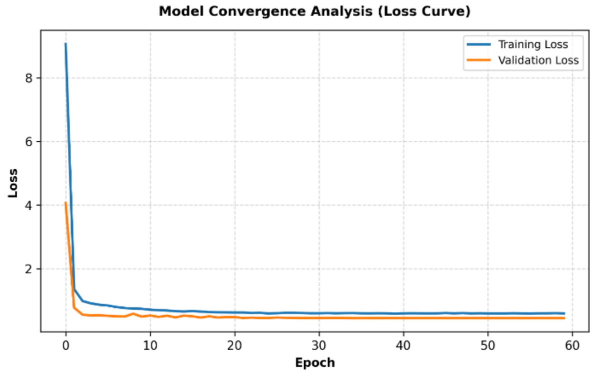

Figure 2. Training and Validation Loss

Figure 2 provides a visualization of the training and validation loss trends observed during optimization. Both metrics decreased substantially in the initial epochs and then stabilized as the model approached convergence. Moreover, the validation loss consistently followed the same pattern as the training loss without experiencing a noticeable increase, suggesting that the learning process remained stable and that the model generalized effectively to unseen data. The small gap between the two curves suggests that the proposed LSTM model achieved good generalization performance while avoiding significant overfitting. Overall, these results demonstrate that the selected model converged successfully and was capable of learning the temporal purchasing patterns represented by the RFM-T features.

C. Model Training Performance

Following the completion of model training, the proposed LSTM architecture was evaluated on an independent testing dataset to examine its ability to estimate CLV. Figure 3 presents the prediction results for the first 100 testing samples. The close correspondence between predicted and actual CLV values reflects the model's ability to capture purchasing behavior.

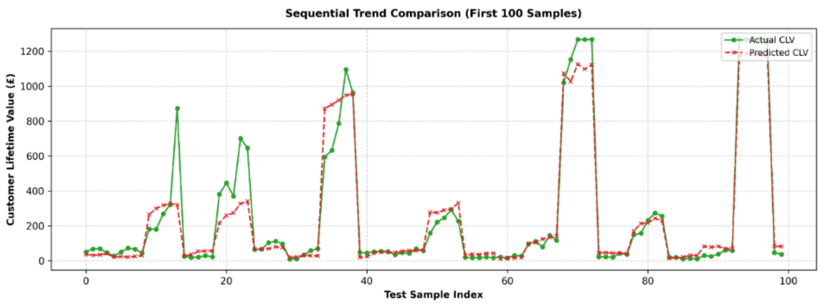

Figure 3. CLV Prediction Results on the Testing Dataset

Figure 3 shows that the predicted CLV values generally follow the same trend as the actual CLV values throughout the observation period. The proposed model is capable of capturing both low-value and high-value customer patterns, including several noticeable fluctuations in customer lifetime value. Although prediction errors are still observed at several peak values, the overall prediction trend remains consistent with the actual data.

The consistency between the two curves reflects the model's capability to extract temporal patterns from the sequential RFM-T representation. This result demonstrates that the model can effectively represent changes in customer purchasing behavior over time and produce predictions that closely approximate the actual Customer Lifetime Value. Nevertheless, the deviations observed at several extreme values suggest that customers with exceptionally high transaction values remain more difficult to predict accurately than customers with moderate purchasing behavior.

D. Prediction Accuracy Evaluation

To further evaluate the prediction performance, a scatter plot was generated to compare the actual CLVwith the predicted values for all samples in the testing dataset, as shown in figure 4.

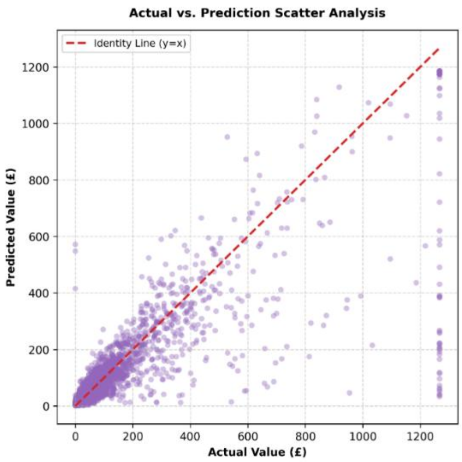

Figure 4. Scatter Visualization of CLV Predictions

The scatter plot in Figure 4 reveals that most prediction points lie close to the reference line ($y = x$), demonstrating that the model maintains accurate CLV estimates for the majority of customers. The close alignment also confirms that the sequential RFM-T representation effectively supports the LSTM model in learning customer purchasing behaviour associated with CLV. A wider dispersion can be observed for customers with relatively high CLV values, indicating that customers with exceptionally large transaction values remain more difficult to predict than those with moderate or low purchasing behavior. Such deviations are expected because high-value customers generally exhibit greater variability in their purchasing patterns, making their future lifetime value more challenging to estimate accurately. Nevertheless, the overall distribution confirms that the proposed LSTM model provides reliable predictions across different customer segments while maintaining a strong correlation between the predicted and actual values.

E. Model Evaluation

Prediction quality was examined using MAE and RMSE, which indicate the deviation of the model outputs from the reference CLV values. These metrics were used to determine the reliability of the developed model, where smaller error values reflect higher predictive accuracy.

Table 6. Model Evaluation

	Metric	Value
1	MAE	43.16
2	RMSE	125.39

As presented in Table 6, the proposed LSTM model achieved an MAE of 43.16 and an RMSE of 125.39 on the original CLV scale. The MAE indicates that the predicted CLV differs from the actual value by approximately £43.16 per customer on average. Meanwhile, the RMSE is higher than the MAE because RMSE places greater emphasis on larger prediction errors, making it more sensitive to customers with exceptionally high CLV values.

The discrepancy between MAE and RMSE implies that, although the proposed model predicts most customer CLV values with relatively low error, several high-value customers generate larger deviations that have a stronger influence on RMSE. A similar pattern can be observed in Figure 4, where the spread of prediction points increases at higher CLV values. Even so, the experimental results confirm that the model provides consistent and dependable CLV estimates for the overall customer population.

4. Conclusions

Using the Online Retail II e-commerce dataset, this research developed an LSTM-based approach for CLV prediction using the RFM-T feature framework. The experimental results demonstrate that the proposed model was able to learn customer purchasing patterns from sequential RFM-T features and generate CLV predictions based on the evaluated metrics. The training and validation loss trends showed stable convergence, while the prediction analysis indicated that the model was able to capture the general pattern of customer lifetime value.

The quantitative evaluation produced an MAE of 43.16 and an RMSE of 125.39 on the original monetary scale. The comparison between LSTM-RFM and LSTM-RFM-T models shows that the Tenure feature provides additional information regarding customer relationship duration; however, its contribution to prediction performance was limited in this dataset. The proposed model may assist companies in analyzing customer value patterns and supporting customer retention strategies. Future research may explore additional behavioural features, alternative deep learning architectures, or attention mechanisms to improve CLV prediction performance.

References

[1] G. Yılmaz Benk, B. Badur, and S. Mardikyan, “A New 360° Framework to Predict Customer Lifetime Value for Multi-Category E-Commerce Companies Using a Multi-Output Deep Neural Network and Explainable Artificial Intelligence,” *Information (Switzerland)*, vol. 13, no. 8, Aug. 2022, doi: https://doi.org/10.3390/info13080373.

[2] Samidi, R. Y. Suladi, and D. Kusumaningsih, “Comparison of the RFM Model’s Actual Value and Score Value for Clustering,” *Jurnal RESTI*, vol. 7, no. 6, pp. 1430–1438, Dec. 2023, doi: https://doi.org/10.29207/resti.v7i6.5416.

[3] F. Alzami *et al.*, “Implementation of RFM Method and K-Means Algorithm for Customer Segmentation in E-Commerce with Streamlit,” *ILKOM Jurnal Ilmiah*, vol. 15, no. 1, pp. 32–44, Apr. 2023, doi: https://doi.org/10.33096/ilkom.v15i1.1524.32-44.

[4] V. Norouzi, “Predicting e-commerce CLV with neural networks: The role of NPS, ATV, and CES,” *Journal of Economy and Technology*, vol. 2, pp. 174–189, Nov. 2024, doi: https://doi.org/10.1016/j.ject.2024.04.004.

[5] R. Kasemrat and T. Kraiwanit, “Attention-enhanced LSTM for high-value customer behavior prediction: Insights from Thailand’s E-commerce sector,” *Intelligent Systems with Applications*, vol. 26, no. April, p. 200523, 2025, doi: https://doi.org/10.1016/j.iswa.2025.200523.

[6] X. Cao, Y. Xu, and X. Yang, “Customer Lifetime Value Prediction with Uncertainty Estimation Using Monte Carlo Dropout,” vol. 4, p. 9, 2024, doi: https://doi.org/10.48550/arXiv.2411.15944.

[7] H. Chen, E. Ng, S. Smyl, and G. Steininger, “Predicting Customer Lifetime Value Using Recurrent Neural Net,” *stat.AP*, vol. 15, pp. 1–4, 2025, doi: https://doi.org/10.48550/arXiv.2412.20295.

[8] K. A. Rijal, A. Vega Vitianingsih, Y. Kristyawan, A. Lidya Maukar, S. Fitri, and A. Wati, “Jurnal Teknologi dan Manajemen Informatika Forecasting Model of Indonesia’s Oil & Gas and Non-Oil & Gas Export Value using Var and LSTM Methods,” *Jurnal Teknologi dan Manajemen*

Informatika (JTMI), vol. 10, pp. 59–69, 2024, [Online]. Available: http://http://jurnal.unmer.ac.id/index.php/jtmi

[9] A. Gasparin, S. Lukovic, and C. Alippi, "Deep learning for time series forecasting: The electric load case," Mar. 01, 2022, *John Wiley and Sons Inc*. doi: https://doi.org/10.1049/cit2.12060.

[10] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," Sep. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: https://doi.org/10.3390/info15090517.

[11] L. Aritonang, B. Aryowindo, R. Syarif, and E. Sabarita Barus, "Forecasting Red Chilli Plant Growth Using Time Series Method With Long Short-Term Memory Model Forecasting Pertumbuhan Tanaman Cabai Merah Menggunakan Metode Time Series Dengan Model Long Short-Term Memory," *Jurnal Inovtek Polbeng - Seri Informatika*, vol. 10, no. 2, p. 2025.

[12] X. Zhang, F. Guo, T. Chen, L. Pan, G. Beliakov, and J. Wu, "A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research," Dec. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: https://doi.org/10.3390/jtaer18040110.

[13] E. Damayanti, A. V. Vitianingsih, S. Kacung, H. Suhartoyo, and A. Lidya Maukar, "Sentiment Analysis of Alfagift Application User Reviews Using Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) Methods," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 2, pp. 509–521, Jun. 2024, doi: 10.51454/decode.v4i2.478.

[14] K. Li, G. Shao, N. Yang, X. Fang, and Y. Song, "Billion-user Customer Lifetime Value Prediction: An Industrial-scale Solution from Kuaishou," in *International Conference on Information and Knowledge Management, Proceedings*, Association for Computing Machinery, Oct. 2022, pp. 3243–3251. doi: https://doi.org/10.1145/3511808.3557152.

[15] Y. Sun, H. Liu, and Y. Gao, "Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model," *Heliyon*, vol. 9, no. 2, Feb. 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e13384.

[16] G. Mena, K. Coussement, K. W. De Bock, A. De Caigny, and S. Lessmann, "Exploiting time-varying RFM measures for customer churn prediction with deep neural networks," *Ann. Oper. Res.*, vol. 339, no. 1–2, pp. 765–787, Aug. 2024, doi: https://doi.org/10.1007/s10479-023-05259-9.

[17] E. B. Firmansyah, M. R. Machado, and J. L. R. Moreira, "How can Artificial Intelligence (AI) be used to manage Customer Lifetime Value (CLV)—A systematic literature review," Nov. 01, 2024, *Elsevier B.V.* doi: https://doi.org/10.1016/j.jjimei.2024.100279.

[18] A. Solichin, "Prediksi Nilai Pengadaan Barang Dan Jasa Pada Sebuah Perusahaan Pariwisata Menggunakan Metode Arima dan Fuzzy Time Series," *JURNAL INOVTEK POLBENG - SERI INFORMATIKA*, vol. 9, no. 1, p. 2024.

[19] A. B. N. Sinurat, W. A. Arifin, and W. A. Larasati, "Sea Level Prediction Using Gated Recurrent Unit And Bidirectional Long Short-Term Memory Methods Prediksi Tinggi Muka Air Laut Menggunakan Metode Gated Recurrent Unit Dan Bidirectional Long Short-Term Memory," *Jurnal Inovtek Polbeng - Seri Informatika*, vol. 9, no. 2, p. 2024.

[20] P. A. Duran, A. V. Vitianingsih, Moch. S. Riza, A. L. Maukar, and S. F. A. Wati, "Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression," *Teknika*, vol. 13, no. 1, pp. 27–34, Jan. 2024, doi: 10.34148/teknika.v13i1.712.

[21] S. M. Al-Selwi *et al.*, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," Jun. 01, 2024, *King Saud bin Abdulaziz University*. doi: https://doi.org/10.1016/j.jksuci.2024.102068.

[22] X. Liu and W. Wang, "Deep Time Series Forecasting Models: A Comprehensive Survey," May 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: https://doi.org/10.3390/math12101504.