Indonesian Journal of Data and Science

Volume 6 Issue 2
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v6i2.244

Research Article

Comparative Analysis of Random Forest and LSTM Models for Customer Churn Prediction Based on Customer Satisfaction and Retention

Babajide Gegeleso [1,*], Oluwaseun Ebiesuwa [2]

[1] Babcock University, Ilisan, Ogun State, Nigeria, gegeleso0172@pg.babcock.edu.ng
[2] Babcock University, Ilisan, Ogun State, Nigeria, ebiesuwao@babcock.edu.ng

Correspondence should be addressed to Babajide Gegeleso; gegeleso0172@pg.babcock.edu.ng

Received 21 January 2025; Accepted 22 May 2025; Published 31 July 2025

Abstract:

Forecasting of Customer churn and prediction is important for sustaining long-term customer relationships and enhancing profitability in competitive markets. This study outlines the comparison of the performance of Random Forest (RF) and Long Short-Term Memory (LSTM) models in predicting customer churn using a dataset of 2,850 customers. The dataset comprises of behavioral, transactional, and satisfaction metrics. Key evaluation metrics include accuracy, precision, recall, F1-score, and AUC-ROC. The result clearly shows that while Random Forest offers strong baseline performance with interpretable results, LSTM captures temporal patterns very effectively and performs better in identifying subtle churn indicators, especially in sequential customer satisfaction data. The result of metrics evaluated shows LSTM has an Accuracy of 88.6%, Precision of 85.3%, Recall of 82.5%, F1-score of 83.9% and AUC-ROC of 0.92 while Random Forest has Accuracy of 85.2%, Precision of 81.5%, Recall of 77.0%, F1- Score of 79.2% and AUC-ROC of 0.89. This shows the preference of LSTM for rapidly changing and large volume dataset over RF excellence in less complicated and sparse dataset.

Keywords: Churn; Customer; LSTM; Recommendation Systems; Retention.

1. Introduction

Exchange of goods and services has metamorphosed from physical line queues to punching of few buttons on our gadget from our comfort zones [1]. This shift has caused revolution from traditional brick-and-mortar stores to either adapt to an online model or incorporate a hybrid approach. The early 2000s and the Covid 19 pandemic has catalyzed a significant surge in e-commerce activity. The introduction of secure online payment methods and the proliferation of user-friendly websites have made online shopping increasingly appealing [2]. The competition for customers has reached a dimension that needed to be taken with utmost importance. Customer churn refers to the likelihood of loss of clients or customers and is a critical concern for businesses, especially in subscription-based models or service industries. Predicting or forecasting churn allows companies to proactively engage at-risk customers by identifying them through personalized retention strategies. Traditional machine learning methods like decision trees, logistic regression, and ensemble models (e.g., Random Forest) have been extensively used in churn prediction. Recently, deep learning models such as LSTM, capable of modeling time-series and sequential behavior, have gained traction due to their ability to learn complex patterns over time [3].

The dataset consists of 79,850 customers data obtained from Jumia website for fifty-four week. The features of the dataset are customer demographics which include (age, location, and tenure), transactional history (purchase frequencies, average spend), customer satisfaction scores (CSAT, NPS over time), interaction logs (support tickets, feedback), churn label (binary: churned or retained). This dataset was later streamlined to 2,850 which was used for the research [4].

Data preprocessing was conducted on the dataset to clear all the anomalies discovered during the process of collection and prepare it for the task ahead. Handling missing values was carried out by finding the mean and uses it to interpolate the missing values. Normalizing of satisfaction metrics was also carried out. Time-series structuring was implemented for LSTM input and one-hot encoding of categorical variables was carried out for Random Forest. This research focuses on merits and demerits of RF and LSTM on the dynamic and rapidly changing E-Commerce environment [5]. This research aims to:

- Evaluate the effectiveness of Random Forest and LSTM in churn prediction.
- Determine which model better utilizes satisfaction and retention features.
- Provide actionable insights into customer behavior over time.

2. Method:

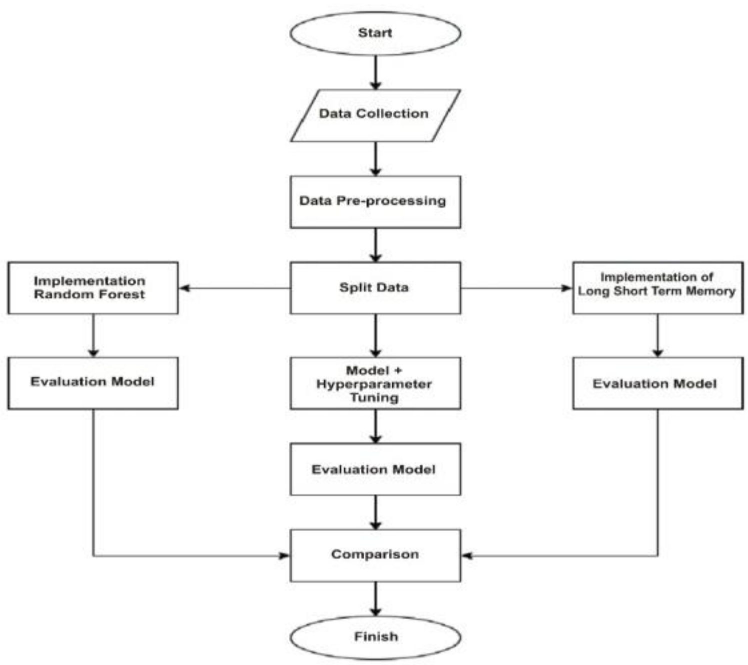

Figure 1. Framework for the Design

Figure 1 illustrates the framework for the design employed in this study, beginning with data collection from relevant sources. The first step is data pre-processing, it is carried out, concurrently to involve handling of missing values, encoding categorical variables, and scaling numerical features to ensure the data is suitable for modeling. The dataset is then split into training and testing subsets, after which two main approaches—Random Forest and Long Short-Term Memory—are implemented independently. An initial evaluation of each model's performance is conducted before proceeding to hyperparameter tuning, where key parameters are optimized to enhance predictive accuracy [6], [7]. The tuned models undergo a final evaluation, and their results are compared to identify differences in performance and interpretability, culminating in conclusions and recommendations regarding the most appropriate classification approach for the dataset adopted for the comparison [8].

Comparison of Machine Learning techniques and Deep Learning techniques on consumers review helps businesses and organization to continually improve their marketing strategies and obtain a comprehensive analysis of consumer feedback concerning their product and brand. This research adopts a way to find out the applicability of Random Forest and Long Short-Term Memory approach [9]. This is to clarify why there is a shift from the traditional machine learning technique. Data preprocessing steps like lower case processing, stop word removal and punctuation removal as well as tokenization will be used for data cleaning. Systematic and integrated approach to enhance data processing in e-commerce, specifically tailored for analysis and improvement in customer churn will be investigated. This comprehensive design seamlessly combines various components to form a cohesive and efficient model [10].

Data Collection

The Data Collection System forms the foundation of this research. It is tasked with aggregating a wide range of data, including customer interactions, transactions and feedback from Jumia Nigeria's e-commerce platform. The richness and variety of this collected data are crucial as they provide the raw materials for subsequent analysis and model development. Focusing on these goals and implementing strategies and initiatives to achieve them, e-commerce companies like Jumia can build strong relationships with their customers, drive loyalty and ultimately drive business growth and reduce the rate of customer churning. Highlighted below are the selected items whose customer behavior while purchasing them will be analysed for this research [11].

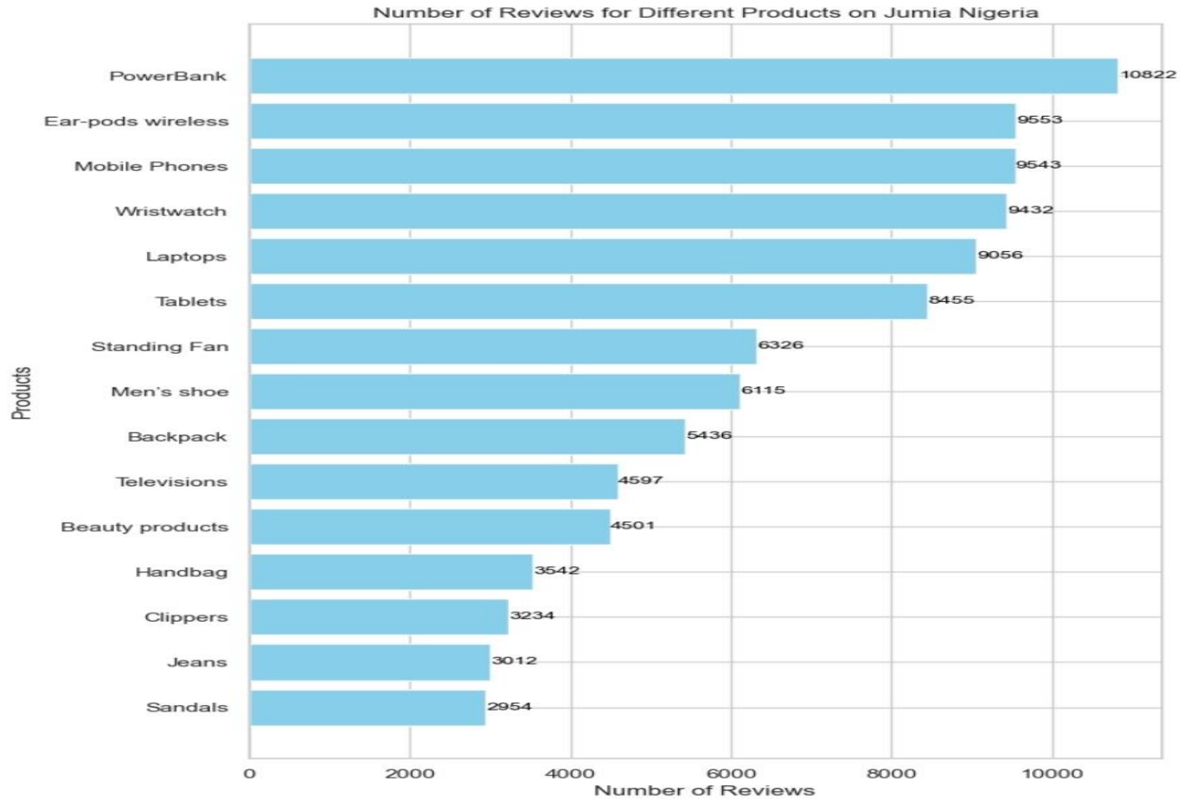

Figure 2. Products with the highest review on Jumia Nigeria

Table 1. Evaluation of customer experience reviews per product category considered

Product Name	Review Count
Power Bank	10822
Ear-pods wireless	9553
Mobile Phones	9543
Wristwatch	9432
Laptops	9056
Tablets	8455
Standing Fan	6326
Mens shoe	6515
Backpack	5436
Televisions	4597
Total	79,850

Data Description

Jumia Nigeria has a record unit that collects feedback from customers through surveys, reviews, or feedback forms available on the Jumia website or mobile app. Analyzing historical data and customer interactions to assess these metrics by using customer reviews, ratings and complaints to evaluate product quality, delivery speed, customer service, website usability, pricing, order accuracy and available payment options are done in reports. This report for the top reviewed products was gathered for this research. The dataset can be categorized to customer retention variable and customer metric. After sorting of the dataset, which was done manually, the initial dataset of 79,850 was left with 2,850 data entries for the research [12].

Customer Retention Variables: These variables are specifically related to customer retention and loyalty. They measure customers' likelihood to continue purchasing from the platform, their overall satisfaction with the service, and their long-term value to the business. Further analysis such as NPS, retention rate, churn rate, AOV, and CLV are crucial for understanding customer loyalty and developing strategies to enhance retention and maximize lifetime value will still be conducted to establish churning rate [13], [14].

Customer Metrics: These variables focus on assessing various aspects of the customer experience, including the quality of products and services, the efficiency of delivery, the effectiveness of customer service and the usability of the website. They provide insights into different facets of customer interactions with the e-commerce platform [15].

Table 2 represents the variables we had access to and their description. Some of the variables as explained in the previous section were structured and some were semi-structured, **Table 3** details the different type of data structure.

Table 2. The description of customer feedback dataset

Variable	Definition
Customer Feedback	Feedback provided by customers regarding their shopping experience, including comments, reviews, and ratings.
Delivery Speed	The time taken for products to be delivered to customers after placing an order.
Product Quality	Assessment of the quality of products offered by Jumia Nigeria based on customer perceptions.
Customer Service Quality	Evaluation of the quality of customer service provided by Jumia Nigeria, including responsiveness, helpfulness, and professionalism.
Website Usability	Measurement of the ease of use, navigation, and overall user experience of the Jumia Nigeria website.
Pricing	Assessment of the competitiveness and affordability of product prices offered by Jumia Nigeria compared to competitors.
Order Accuracy	Accuracy and correctness of orders fulfilled by Jumia Nigeria, including the correct items, quantities, and specifications.
Payment Options	Availability and variety of payment methods accepted by Jumia Nigeria for customer transactions
Promotions and Discounts	Frequency and attractiveness of promotional offers, discounts, and deals provided to customers.
Return and Refund Policy	Clarity, flexibility, and effectiveness of the return and refund policy implemented by Jumia Nigeria
Net Promoter Score (NPS)	A metric used to gauge customer loyalty and satisfaction based on the likelihood of customers recommending the company to others.
Customer Satisfaction	Overall satisfaction level of customers with their shopping experience on Jumia Nigeria.
Customer Retention Rate	Percentage of customers who continue to make purchases from Jumia Nigeria over a specific period.
Repeat Purchase Rate	The proportion of customers who make multiple purchases from Jumia Nigeria within a defined timeframe
Customer Churn Rate	Percentage of customers who stop purchasing from Jumia Nigeria over a specific period.
Average Order Value (AOV)	The average monetary value of orders placed by customers on Jumia

Table 3. The description of the dataset variables

Structured Data Variables	Semi-Structured Data Variables

Delivery Speed	Customer Feedback
Product Quality	Promotions and Discounts
Customer Service Quality	Return and Refund Policy
Pricing	Net Promoter Score (NPS)
Order Accuracy	Customer Satisfaction

Data Pre-processing

Cleaning, normalizing, and transforming the data using NumPy and Pandas. These are fundamental libraries for data manipulation in Python. NumPy provides support for numerical operations, while Pandas offers powerful data structures and tools for data analysis. They will be used to load data, manipulate it and prepare it for training of LSTM models, into a suitable format for analysis to be carried out. This includes handling missing values, encoding categorical variables, and tokenizing textual data [16].

Data pre-processing plays a pivotal role in ensuring the quality and reliability of the dataset for subsequent analysis. In this chapter, we detail the steps involved in pre-processing the data collected from Jumia Nigeria's e-commerce platform. These steps include data cleaning, transformation and feature engineering to prepare the dataset for analysis using Long Short-Term Memory (LSTM) models [17].

Software Requirements

Python machine learning libraries have developed greatly over the years and are now the most popular language for using machine learning techniques. The following are some fundamental Python libraries for machine learning and model training:

1. Matplotlib: This data visualization toolkit is used to produce 2D graphics and common picture plots in several different formats.
2. NumPy: This is a well-known all-purpose toolkit for array processing that provides a wide range of sophisticated mathematical functions. It is particularly helpful for modelling multidimensional arrays and matrices.
3. Scikit-learn: With the help of a Python interface, you may use a variety of supervised and unsupervised learning techniques thanks to the Scikit-learn module. Numerous machine learning functionalities, such as pre-processing, model selection, clustering, regression, classification, and dimensionality reduction, are available with Scikit-learn.
4. Pandas: This is a widely used data analysis package that is renowned for its flexibility, quickness, and potent data manipulation powers. It functions flawlessly with different data formats and enables effective data combining and transformation.

Using the software and hardware requirements, exploratory data analysis of the dataset was carried out, which is an important step in carrying out the algorithm learning process [18].

Train-test Split

In this study, the train-test split is a crucial step in ensuring the reliability and validity of the RF and LSTM model for predicting customer churn. From sklearn.model_selection import train_test_split the entire dataset which is divided into two subsets 80:20 ratio: the training set and the testing set. The training set, typically comprising 80% of the data, is used to train the RF and LSTM model. This involves adjusting the model's parameters to minimize the prediction error on the training data. The remaining 20% of the data forms the testing set, which is used to evaluate the model's performance on unseen data. This split ensures that the model is not just memorizing the training data but is capable of generalizing its predictions to new, unseen customer interactions. By assessing the model's accuracy, precision, recall, and F1 score on the testing set, we can determine how well the model is likely to perform in real-world scenarios, thereby validating its effectiveness for enhancing e-commerce customer satisfaction and retention.

Hyperparameter Tuning

Hyperparameter tuning is a critical process in the development of the LSTM model for this study, aimed at optimizing its performance for predicting customer satisfaction and retention in e-commerce. Hyperparameters, which include the number of LSTM units, the number of layers, learning rate, batch size and dropout rate, significantly influence the model's ability to learn and generalize from the data. Through systematic experimentation, such as grid search or randomized search, we explore different combinations of these hyper parameters to identify the configuration that yields the best performance metrics. This involves training multiple models with varying hyperparameter values and evaluating their performance on a validation set [19]. By fine-tuning these hyperparameters, we ensure that the LSTM model achieves a balance between underfitting and overfitting, thereby enhancing its predictive accuracy and robustness. Effective hyperparameter tuning ultimately leads to a more reliable model that can better capture the intricate patterns in customer behavior, contributing to improved customer satisfaction and retention strategies in e-commerce.

Wrap the Model for Use with Scikit-Learn: model = KerasRegressor(build_fn=create_model, verbose=0)

Implementation Algorithm

The classification algorithm for Random Forest is implemented in this study.

For the Random Forest classifier, an ensemble of decision trees is constructed. Each tree Ti n the ensemble makes its own prediction $h(x)$ for an input x and the final prediction $\hat{y}$ is obtained via majority voting:

$\hat{y} = \{h1\ (x), h2\ (x), \ldots , hN(x)\}$

This approach leverages the diversity of multiple decision trees to improve robustness and generalization.

Random Forest Classifier is a powerful and widely used machine learning algorithm, especially for classification tasks like predicting customer churn. A Random Forest is an ensemble of decision trees. It builds many decision trees (hence a "forest") and lets them vote on the final output. It’s robust, handles both numerical and categorical data, and reduces the risk of overfitting that a single decision tree might have [20].

RF has the advantage of handling complex relationships between features. It can rank the importance of features (like which factors most influence churn). It works well even with imbalanced datasets if tuned properly. High interpretability, robustness to overfitting, feature importance [21].

LSTM Neural Network

LSTM networks were introduced to overcome the vanishing gradient problem in RNNs, enabling the modeling of long-term dependencies in sequential data [10]. With their unique memory cell architecture, LSTMs have become the preferred method for applications involving time-series data, such as stock market analysis, speech recognition, and customer behavior prediction [22].

LSTM Architecture

The LSTM model designed for enhancing e-commerce customer satisfaction and retention consists of several key layers that work together to process the input data and generate predictions, the representation is shown in **Figure 4**. The architecture begins with the input layer, which receives the pre-processed customer data sequences. This data is then passed to the LSTM layers, which are the core of the model. The LSTM layers are responsible for capturing the sequential dependencies and patterns in the customer interaction data over time, with each layer comprising multiple LSTM units that contain forget, input, and output gates to manage the flow of information through the network. Following the LSTM layers, the data moves through dense layers (fully connected layers), which perform further transformations and extract high-level features from the LSTM outputs. These dense layers use activation functions, such as ReLU, to introduce non-linearity and improve the model's ability to learn complex relationships in the data. Finally, the output layer produces the final predictions, which could be a single value indicating customer satisfaction or a probability score indicating the likelihood of customer retention [23]. This comprehensive architecture allows the model to effectively learn from historical customer data and make accurate predictions, thereby enhancing e-commerce strategies for improving customer satisfaction and retention. Activation functions are mathematical equations that determine the output of a neural network node. They introduce non-linear properties to the network, enabling it to learn from complex data patterns and make accurate predictions. Without activation functions, a neural network would only be able to model linear relationships. ReLU (Rectified Linear Unit) activation functions are widely used in hidden layers of neural networks due to their simplicity and effectiveness. It is used in dense layers, while the sigmoid activation function is used in the output layer to predict binary outcomes (customer retention) [24], [25]. It introduces non-linearity and efficient computation. ReLU (Rectified Linear Unit) Activation Function: Equation: $f(x) = max(0, x)$. Equation: $f(x) \ = 1/\ 1 + e - x$[26].

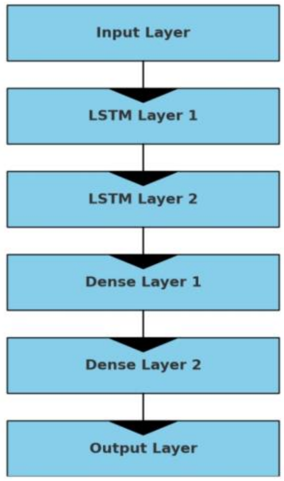

Figure 4. LSTM Architecture

3. Results and Discussion

Results

The result of the experimental evaluation revealed that both the Random Forest and Long Short-Term Memory achieved comparable performance on the grade classification task. Initially, the Random Forest model reached an overall accuracy of approximately 85.2%, while the Long Short-Term Memory attained around 88.6%. This indicates LSTM's stronger ability to correctly classify churners and non-churners. However, accuracy alone can be misleading, especially in imbalanced datasets.

LSTM also achieved higher precision of 85.3% compared to Random Forest of 81.5%. This means when LSTM predicts a customer will churn; it is more likely to be a correct prediction. High precision is crucial for avoiding false alarms, which saves resources by not targeting customers who aren't at risk.

LSTM dominance also continues in recall. It scored higher on recall 82.5% compared to Random Forest of 77.0%, indicating it catches more of the actual churners. This is particularly valuable in churn prediction, where missing a churner can be costlier than mistakenly flagging a loyal customer.

LSTM's F1-score of 83.9% shows it has a better balance between precision and recall than Random Forest (79.2%). F1 is especially helpful when determining the harmonic average of precision and recall.

LSTM achieves an AUC-ROC of 0.92, compared to 0.89 from Random Forest. This indicates better ranking ability, meaning LSTM is more effective at distinguishing between churners and non-churners across thresholds.

Figure 5 and **6** below clearly shows the comparison of the metrics of the two models considered for the analysis carried out in the research.

Table 4. Comparison of Performance Metrics Scores for the RF and LSTM Models

Metric	Random Forest	LSTM
Accuracy	85.20%	88.60%
Precision	81.50%	85.30%
Recall	77.00%	82.50%
F1-score	79.20%	83.90%

Metric	Random Forest	LSTM
AUC-ROC	0.89	0.92

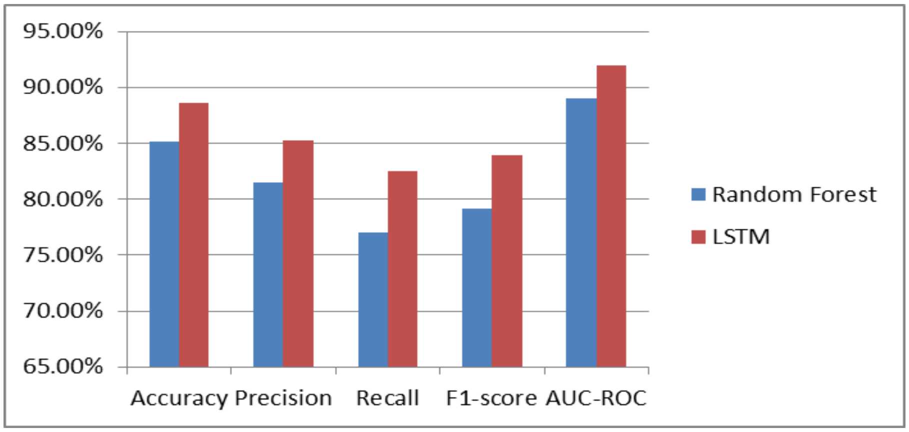

Figure 5. Graph of Comparison of Metric Values of Random Forest and LSTM

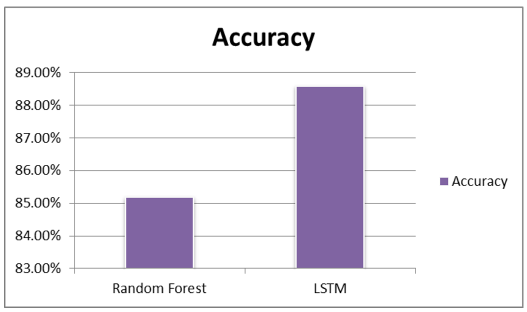

Figure 6. Graph of Comparison of Accuracy of RF and LSTM

Random Forest identified key features like tenure, average spends and recent support interactions as most predictive features revealed. However, LSTM outperformed RF due to its ability to track changes in satisfaction over time. RF is easier to interpret and deploy in real-time dashboards. LSTM requires more resources but captures churn signals that evolve gradually.

Confusion Matrix: The confusion matrix is a crucial tool for understanding how well the model distinguishes between different classes, which could represent various customer satisfaction levels or retention outcomes in this study. In **Figure 7** the classes labeled as 0, 1, and 2, the confusion matrix is a 3×3 matrix where each element (i,j) represents the number of instances of class i that were predicted as class j. The diagonal elements of the matrix show the counts of correctly classified instances for each class, while the off-diagonal elements indicate the misclassifications. For instance, the model correctly classified 3 instances of class 0 as class 0 but misclassified one instance of class 0 as class 2. Similarly, it correctly identified 1 instance of class 1 but misclassified others as class 0 and class 2.

The heatmap visualization helps to quickly identify these patterns, with darker cells along the diagonal indicating higher counts of correct predictions, and lighter cells off the diagonal indicating fewer misclassifications. This visual representation is invaluable for diagnosing specific strengths and weaknesses of the model, providing insights into how well it performs in predicting different outcomes related to customer satisfaction and retention. This can guide further tuning and improvement of the model to enhance its accuracy and reliability in real-world e-commerce applications.

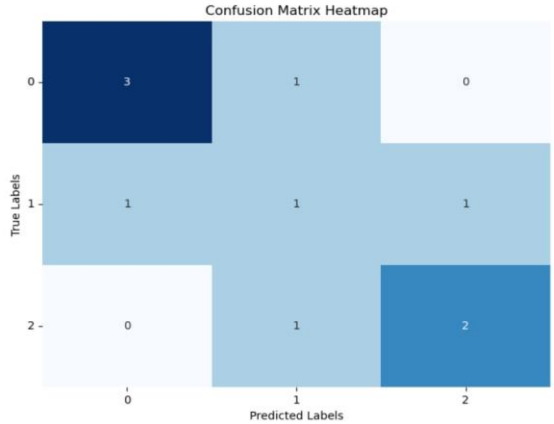

Figure 7. LSTM Network Confusion Matrix Heatmap

Discussion

The results suggest that while Random Forest offers robust performance on static features, LSTM excels when longitudinal customer data is available. Customer satisfaction and retention patterns are inherently temporal, and LSTM is better suited to model these. Random Forest is useful for fast, interpretable insights compared to Long Short-Term Memory. LSTM is best suited for richer behavioral modeling and proactive churn intervention. Deploying LSTM can improve early identification of potential churners, allowing for timely interventions (e.g., discounts, personalized outreach). Random Forest remains a strong choice where interpretability and quick deployment are priorities, especially in low-resource environments.

Future Work:

- Incorporation of more granular time-series data (e.g., monthly satisfaction trends) can be researched on to enhance the outcome.
- Combination of both models in an ensemble or hybrid architecture and application of explainable AI (e.g., SHAP) to LSTM outputs for interpretability are another area that will be of immense value to the research.

4. Conclusion

Both Random Forest and LSTM are effective in churn prediction, but their utility depends on the nature of the data and business constraints. LSTM is better when historical customer interaction and satisfaction trends are key to decision-making.This review highlights the transformative potential of LSTM networks in enhancing customer satisfaction and retention in e-commerce. By addressing limitations in traditional predictive models, LSTMs enable more accurate and context-aware predictions. However, challenges related to scalability, interpretability, and data requirements must be addressed to maximize their impact. Future research should explore hybrid models and real-time applications to further advance the role of LSTMs in e-commerce.

References:

[1] A. Mittal and P. Gupta, “An Empirical Study on Enhancing Product Quality and Customer Satisfaction Using Quality Assurance Approach in an Indian Manufacturing Industry,” Int. J. Math. Eng. Manag. Sci., vol. 6, no. 3, pp. 878–893, 2021, doi: 10.33889/ijmems.2021.6.3.052.

[2] G. Gunawardane, “Enhancing customer satisfaction and experience in financial services: a survey of recent research in financial services journals,” J. Financ. Serv. Mark., vol. 28, no. 2, pp. 255–269, 2023, doi: 10.1057/s41264-022-00148-x.

[3] K. Lepistö, M. Saunila, and J. Ukko, “Enhancing customer satisfaction, personnel satisfaction and company reputation with total quality management: combining traditional and new views,” Benchmarking, vol. 31, no. 1, pp. 75–97, 2024, doi: 10.1108/BIJ-12-2021-0749.

[4] X. Yin and X. Tao, “Prediction of Merchandise Sales on E-Commerce Platforms Based on Data Mining and Deep Learning,” Sci. Program., vol. 2021, 2021, doi: 10.1155/2021/2179692.

[5] M. S. Hossen, A. H. Jony, T. Tabassum, M. T. Islam, M. M. Rahman, and T. Khatun, “Hotel review analysis for the prediction of business using deep learning approach,” Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021, pp. 1489–1494, 2021, doi: 10.1109/ICAIS50930.2021.9395757.

[6] J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu, and M. Arifuzzaman, “Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models,” 2022 Int. Conf. Adv. Electr. Electron. Eng. ICAEEE 2022, 2022, doi: 10.1109/ICAEEE54957.2022.9836420.

[7] S. Beltozar-Clemente, O. Iparraguirre-Villanueva, F. Pucuhuayla-Revatta, J. Zapata-Paulini, and M. Cabanillas-Carbonell, “Predicting customer abandonment in recurrent neural networks using short-term memory,” J. Open Innov. Technol. Mark. Complex., vol. 10, no. 1, 2024, doi: 10.1016/j.joitmc.2024.100237.

[8] L. Zhang and Q. Wei, “Personalized and Contextualized Data Analysis for E-Commerce Customer Retention Improvement With Bi-LSTM Churn Prediction,” IEEE Trans. Consum. Electron., 2024, doi: 10.1109/TCE.2024.3376672.

[9] W. eds. Marcus, A. and Wang, “Design, User Experience, and Usability. User Experience in Advanced Technological Environments. Lecture Notes in Computer Science. Cham: Springer,” no. International Publishing. doi: 10.1007/978-3-030-23541-3., 2019.

[10] L. Marozzo, F., and Belcastro, “Cloud Computing for Big Data Analysis. In: Zomaya, A., Taheri, J., Sakr, S. (eds) Encyclopedia of Big Data Technologies. Springer, Cham.,” doi. 10.1007/978-3-319-63962-8_136-2, 2022.

[11] P. S. Navarro, V.G., Gomez, H.G., Badenes, R.O. and Acosta, “Customer relationship management and its impact on entrepreneurial marketing: a literature review. International Entrepreneurship and Management Journal, [online],” no. 1(1). doi: 10.1007/s11365-022-00800-x., 2022.

[12] R. Munné, “Big Data in the Public Sector. In: Cavanillas, J., Curry, E., Wahlster, W. (eds) New Horizons for a Data-Driven Economy.,” no. Springer, Cham. 10.1007/978-3-319-21569-3_11, 2016.

[13] V. S. Nirala, K.K., Singh, N.K. and Purani, “A survey on providing customer and public administration based services using AI:,” no. chatbot. Multimedia Tools and Applications. doi: 10.1007/s11042-021-11458-y., 2022.

[14] A. H. Okolie, U.C. and Ojomo, “E-Commerce in Nigeria: Benefits and Challenges. Humanities and Social Sciences:” vol. Latvia, 28, p. pp.69–90.

[15] J. Ologunebi, “An analysis of customer retention strategies in e-commerce fashion business in the UK: A case study of Primark. [online] Social Science Research Network. d,” no. doi: 10.2139/ssrn.4628521., 2023.

[16] T. Alves Gomes, M., and Meisen, “A review on customer segmentation methods for personalized customer targeting in e-commerce use cases.,” vol. Inf Syst E, 2023.

[17] J. J. Ashiru, F., Nakpodia, F. and You, “Adapting emerging digital communication technologies for resilience: evidence from Nigerian SMEs. Annals of Operations Research,” vol.

[18] M. . Ayawei, “Assessing the Factors Underlying the Adoption of E-Commerce Among B2B SMEs.,” J. Electron. Commer. Organ. 21(1), pp.1–26. doi: 10.4018/jeco.333612., 2023.

[19] S. W. Chang, H.H. and Chen, “The impact of customer interface quality, satisfaction and switching costs on e-loyalty: Internet experience as a moderator.,” Comput. Hum. Behav. 24(6), pp.2927–2944. Doi: 10.1016/j.chb.2008.04.014., 2008.

[20] K. Chenthara, S., Wang, H., and Ahmed, “Security and Privacy in Big Data Environment. In: Sakr, S., Zomaya, A.Y. (eds) Encyclopedia of Big Data Technologies.,” vol. Springer, 2019.

[21] F. Perifanis, N.-A. and Kitsios, “Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. Information, [online] 14(2), 2023.

[22] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science, 2(3),” no. doi: 10.1007/s42979-021-00592-x., p. pp.1–21., 2021.

[23] C. van Nes, A., and Yamu, “Empirical Data Collection and Analysis, and Connecting Data with Space Syntax. In: Introduction to Space Syntax in Urban Studies. Springer, Cham.,” no. 10.1007/978-3-030-59140-3_5, 2021.

[24] L. . Huawei Technologies Co., “A General Introduction to Artificial Intelligence. In: Artificial Intelligence Technology.,” no. Springer, Singapore. 10.1007/978-981-19-2879-6_1, 2023.

[25] S. Wassouf, W.N., Alkhatib, R., Salloum, K. and Balloul, “Predictive analytics using big data for increased customer loyalty: Syriatel Telecom Company case study. Journal of Big Data, [online] 7(1).,” no. doi: 10.1186/s40537-020-00290-0., 2020.

[26] I. Alsolbi, F. H. Shavaki, R. Agarwal, G. K. Bharathy, S. Prakash, and M. Prasad, “Big data optimisation and management in supply chain management: a systematic literature review,” Artif. Intell. Rev., vol. 56, no. a systematic literature review. pp. 253–284, 2023, doi: 10.1007/s10462-023-10505-4