Design of Long Short Term Memory Based Deep Learning Model for Customer Churn Prediction in Business Intelligence

Zahraa Hasan [1], Dasha Stablichenkova [2]

[1.]University Of Applied Science, Faculty of Literature and Science, Manama, Bahrain
[2.]University of Debrecen, Department of Mathematical and Computational Science, Debrecen, Hungary

Emails: Zahraamathscience1243@gmail.com; dashaukr.latic1322@gmail.com

Abstract

Innovations in business intelligence are crucial in the digital era to staying popular and competitive across the increasing business trends. Businesses have started scrutinizing the next level of data analytics and business intelligence solutions. Customer Churn Prediction (CCP), on the other hand, a crucial for making business decisions, which correctly recognizes the churn customers and acts appropriately for customer retention. Customer churn is an unavoidable consequence when the user is not satisfied with the company's service for a longer period. Service unsubscription by the user does not emerge unexpectedly; instead, it comes from the customer as a vigorous act owing to its accumulation of long-term disappointment. Thus, there is a need for the service provider to find and address their challenges related to customer satisfaction and service for retaining irate customers. The possibilities to predict customer churn have dramatically increased with the advances in artificial intelligence (AI) and machine learning (ML) algorithms. Therefore, this study introduces an Optimal Long Short Term Memory Based Customer Churn Prediction for Business Intelligence (OLSTM-CCPBI) method. The proposed OLSTM-CCPBI method incorporates many innovative components, such as Min-Max scaling for normalization, LSTM networks for temporal sequence modelling, and Adam optimization for hyperparameter tuning. The OLSTM-CCPBI method effectively captures temporal dependency in sequential customer data by leveraging the dynamic nature of the LSTM network, which enables correct prediction of churn events. Through detailed investigations on real-time customer churn datasets, OLSTM-CCPBI achieves better predictive capabilities than classical approaches, showcasing its promising solution to aid businesses in preemptively addressing customer attrition and considerably enhancing churn prediction accuracy.

Keywords: Customer Churn Prediction; Long Short Term Memory; Business Intelligence; Adam Optimization; Data Normalization

1. Introduction

Recently, developments in business intelligence have become extremely significant to be effective and competitive over the emerging corporate trends [1]. Consequently, businesses of each size (small and medium-scale enterprises) started to invest in the next level of business intelligence and data analysis outcome [2]. The efficient deployment of business intelligence techniques is extracted from the analysis and visualizes the effectiveness indication from a huge quantity of big data enterprises. It reduces the cost and increases the quickness of problem-solving [3]. However, business organizations must make suitable decisions due to many complexities. The text analytics technique extremely employs robust business intelligence with improved outcomes [4].

In recent times, research has demonstrated that the primary aim is to recognize useful churn customers through a large quantity of data obtained from the telecommunication sector [5]. Almost there are several constraints to utilizing existing systems that can face several complexities and hurdles for the issue of churning in nowadays.

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

Although development of modeling, various information-rich features could be ignored in Feature selection (FS) techniques. Primarily statistical techniques are employed for different fields that tend to provide unwanted outcomes of the existing predictive models [6]. The FS is an alternative large issue with the present models. Each user can be an individual or group and various churning causes. Churn customer classification is called a churner, regardless of noting causes and parameters for churn [7]. Numerous patterns of behavior in the churn technique must not be treated all similarly. A few customers are not churn simpler than others.

Nowadays, the requirement is for a more accurate prediction model that will predict the churn customers in an advanced manner. It is significant support for offering powerful retention approaches for the various collections of churners that will be diverse promotions that depend upon the churn aspects for the distinct churner's group [8]. Inspired by the aforementioned information and observations in this research, a technique for CCP using various machine learning (ML) methods was developed. Predictive modelling is an alternative essential feature of an application. This can be combined with ML techniques or statistical techniques for customer churn prediction (CCP). With the help of produced churn prediction and analyzing the possibility of churn for the customer, the app supports banks in recognizing the higher-risk customers and obtaining active steps for avoiding churn [9]. Likewise, the app enables reports and communication by permitting the development of modified presentations and reports. This facet allows for simply sharing churn insights and marketing teams, recommendations with management, and other related events [10].

This study introduces an Optimal Long Short Term Memory Based Customer Churn Prediction for Business Intelligence (OLSTM-CCPBI) method. The proposed OLSTM-CCPBI method incorporates many innovative components, such as Min-Max scaling for normalization, LSTM networks for temporal sequence modelling, and Adam optimization for hyperparameter tuning. The OLSTM-CCPBI method effectively captures temporal dependency in sequential customer data by leveraging the dynamic nature of the LSTM network, which enables correct prediction of churn events. Through detailed investigations on real-time customer churn datasets, OLSTM-CCPBI achieves better predictive capabilities than classical approaches, showcasing its promising solution to aid businesses in preemptively addressing customer attrition and considerably enhancing churn prediction accuracy.

2. Related Works

Akhmetshin et al. [11] presented Intelligent Data Analytics employing a Hybrid Gradient Optimization Algorithm with ML (IDA-HGOAML) technique. This method first undertakes data preprocessing through Z-score normalization. The proposed methods create the usage of an equilibrium optimization algorithm (EOA) for the FS. Moreover, the convolutional autoencoder (CAE) algorithm applied the CCP technique. In conclusion, the HGOA was utilized in the optimal hyperparameter selection of the CAE technique. In [12], a churn prediction model was proposed. An improved ensemble learning method (ELM) has been developed for classification. Different alternative ensemble methods, the developed classification method was an enhanced weighted soft voting ensemble with a series of weights implemented to forecast base learning. Powell's optimization technique was implemented during the optimization to enhance the ensemble weights of effect based on the significance of base learners. In [13], an innovative two-level stacking-mode ELM was introduced that could be applied to the Whale Optimization Algorithm (WOA) technique for FS and hyperparameter optimization methods. The developed technique by integrating the K-member clustering and WOA for efficiently handling the class imbalance in churn databases.

Öztürk et al. [14] aimed to propose churn systems for the sellers in the e-commerce marketplace by employing ML techniques. For developing such methods, three techniques have been implemented to design the systems. Logistic Regression (LR) and Random Forest (RF) have been utilized to analyze the churn in every method. In the 1st technique, models have been devised without implementing preprocessing functions under the database. In the 2nd and 3rd techniques, oversampling and under-sampling techniques could be utilized correspondingly for balancing the data set. Fatima et al. [15] projected a new method by leveraging a combined customer analytics environment. The developed model comprises AutoML oversampling techniques. In addition, this research employs the power of oversampling techniques such as synthetic minority oversampling with encoded nominal and continuous features (SMOTE-ENC) and synthetic minority oversampling methods for nominal and continuous features (SMOTE-NC). Sun et al. [16] designed a customer segmentation system. This study initially performed the feature engineering, followed by executing the customer value segmentation dependent upon ML methods and customer correlation management analysis methods and diverse the customer values segmentation recognition system. In conclusion, experimentation analysis could be implemented with actual customer transaction information of the real online business environment.

3. The Proposed Model

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

In this work, we have established a novel OLSTM-CCPBI method. The proposed OLSTM-CCPBI method incorporates many innovative components, such as Min-Max scaling for normalization, LSTM networks for temporal sequence modelling, and Adam optimization for hyperparameter tuning. Fig. 1 represents the entire flow of the OLSTM-CCPBI method.

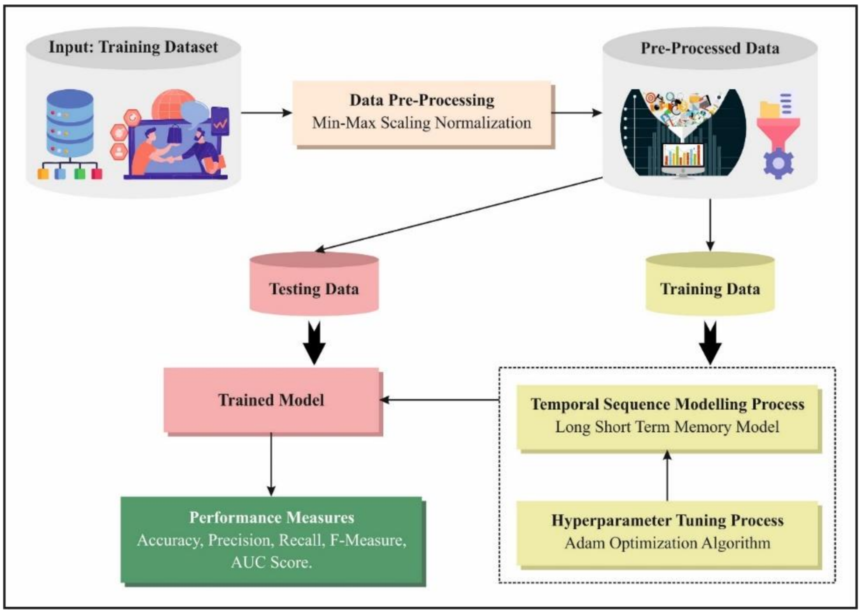

Figure 1: Overall flow of OLSTM-CCPBI method

A. Data Normalization

A Standard Scaler helps to get a uniform distribution, with zero mean and variance of one [17]. It normalizes features by subtracting the mean values from the feature and then dividing the result by the feature standard deviation.

$$z = \frac{(x - u)}{s} \tag{1}$$

Where s shows the standard deviation of the training samples, z refers to scaled data, x symbolizes the scaled data, and u specifies the mean of training samples.

B. LSTM-based Classification Process

A recurrent neural network (RNN) is one kind of NN that exactly intended to simplify the forecast of information with time-based relations [18]. In RNN, the output from an exact time step is approved as an input to the subsequent time step, which allows the method to execute forecasts dependent upon the mixture of the input and output of the existing and preceding time steps, respectively. LSTM is the most commonly used RNN. It integrates an interior memory unit which permits the method to efficiently collect the state through an input feature, creating them which is appropriate for time-sequence predicting where relationships of unidentified duration among the observations and significant actions may occur.

The LSTM structure was stimulated by examining the error movement in conventional RNNs, resulting in the end that long-time dependences might not be controlled by such kinds of methods. The reputation of LSTM originates precisely from their essential capability to overwhelm the important issues of general RNN structures such as exploding and vanishing gradients. RNNs commonly consume a restricted collection of data which can be procedure and included in their forecasts. The foremost cause is that the influence of the input on the system yield may destroy rapidly, producing the issue of vanishing gradient, or blast greatly, by initiating the problem of exploding gradient. To contract with the abovementioned problems, LSTMs contain cells that include an additive gradient framework, permitting a preferred performance from the error gradient at each stage of the training procedure

The conventional RNNs contain a single unit, whereas normal LSTM cells comprise 4 units. The most significant idea of the defined structure is the cell state portrayed by the straight line at the topmost. Each line can be measured as a vector to numerous point-wise processes are executed. The cell state captures the long-term memory ability of systems keeping and uploading data that are not essentially directly preceding events. The LSTM cell also

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

comprises the functioning memory ability which handovers data from directly preceding events, recognized as a hidden layer (HL). The HL of the LSTM is over-written at each step.

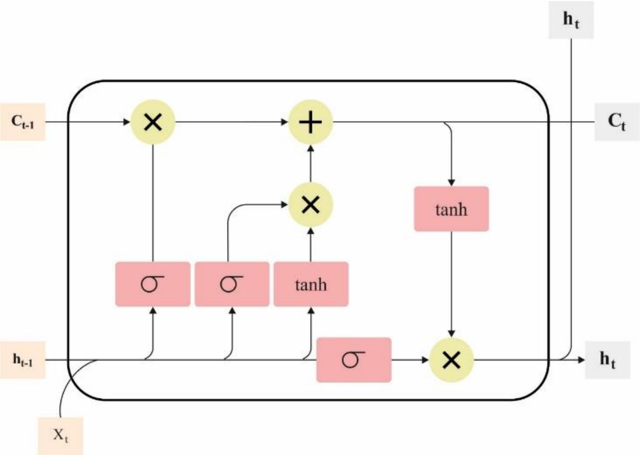

Figure 2: Architecture of LSTM

LSTM cells need 3 inputs and offer dual outputs. The inputs distributed in the vector method are the preceding HL h_{t-1}, the preceding cell state c_{t-1}, and the existing input x_t. The fundamental functions of an LSTM cell are executed over its 3 gates namely forget, the input, and the output. Fig. 2 defines the infrastructure of LSTM.

The forget gate is the initial block signified in the LSTM structure. The main part is to define which portion of the data should be kept or rejected. Then, the inputs preceding HL h_{t-1} and the existing input x_t. These inputs are delivered over the sigmoid function σ_g outcomes in output values among 0 and 1. The forget gate f_t activation vector is set below:

$$f_t = \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right). \tag{2}$$

The input gate aids as an input in order to upgrade the cell state. It generally works in dual portions. Initially, the preceding HL and the existing input were distributed into the 2nd sigmoid function σ_g. Next, similar inputs are delivered into the hyperbolic tangent function σ_c to adjust the system. Lastly, the cell state c_t is the outcome of the element-by-element invention of the cell input activation vector $\tilde{c}_t$ and the upgrade gate i_t. The activation functions of the input gate are set below:

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \tag{3}$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{4}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \tag{5}$$

Lastly, the output gate defines what the subsequent HL h_t delivered by the cell must be. The HL contains data on preceding inputs and is used for forecast. The preceding HL h_{t-1} and the existing input x_t were distributed into the 3rd sigmoid function σ_g. Next, the altered cell state is delivered to the hyperbolic tangent function σ_h. These outputs are increased element-by-element by permitting the system in order to define which data HL must transport.

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{6}$$

$$h_t = o_t \circ \sigma_h(c_t) \tag{7}$$

Whereas, the parameters $b \in \mathbb{R}^h$ and $W \in \mathbb{R}^{h\times d}, U \in \mathbb{R}^{h\times h}$ signify bias and weight matrices vector, correspondingly, which are acquired through the training procedure. It is noticeable that this explanation shows how one LSTM cell functions.

C. Hyperparameter Tuning using Adam optimization

The predictable values of the random parameter for the power of n are known as $the\ N\text{-}th$ moment [19]. This will be denoted as given below:

$$m_n = B[R^n] \tag{8}$$

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

Here, R represents the random variable, and m refers to the moment.

The technique can measure rates of adaptive learning for each feature described in Adaptive Moment Estimation (Adam). This computes exponential changing averages dependent upon the gradient of the existing minibatch for evaluating the moments. This conserves an exponential decaying average of previous squared gradients v_t along with an exponential decay average of the previous gradient m_t. This describes the first moment, and the uncentered variance will be the next one.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)gt \tag{9}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2 \tag{10}$$

Now, m_t and v_t refer to the gradient estimation of the initial moment and the next moment. The default values of β_1 and β_2 are 0.9 and 0.999, correspondingly.

Adam allowed DL experts to considerably increase the optimization of present methods over normal and stochastic gradient descent.

$$m_t = \varphi_1 m_{t-1} + (1-\varphi_1)G_t \tag{11}$$

$$V_t = \varphi_2 V_{t-1} + (1-\varphi_2)G_t^2 \tag{12}$$

Here, φ defines the new method hyperparameters presented, G means the gradient under the existing minibatch, and m and v refer to moving averages. The properties forms and v are required since they measure the 1st and 2nd moments:

$$E[m_t] = E[G_t] \tag{13}$$

$$E[V_t] = E[G_t^2] \tag{14}$$

The change for such biases will be acquired by calculating bias-corrected evaluates for the primary and secondary moments as expressed:

$$\widehat{m}_t = \frac{m_t}{1-\varphi_1^t} \tag{15}$$

$$\hat{V} = \frac{V_t}{1-\varphi_2^t} \tag{16}$$

These bias-corrected evaluations are further employed for upgrading the parameters, Same as Adadelta and RMSprop, which leads to the Adam upgrade rule to upgrade weight:

$$w_t = w_{t-1} - \beta \frac{\widehat{m}_t}{\sqrt{\hat{V}+\epsilon}} \tag{17}$$

where w refers to the model step size and weights.

The Adam optimization derives an FF to achieve better classifier outcomes. It determines a positive integer to represent the high efficiency of the candidate solutions. Here, the decline of the classifier error rate is assumed as the FF.

$$fitness(x_i) = ClassifierErrorRate(x_i)$$

$$= \frac{No. of\ misclassified\ samples}{Total\ No. of\ samples} * 100 \tag{18}$$

4. Result Analysis

The performance analysis of the OLSTM-CCPBI technique takes place using the dataset from the Kaggle dataset [20]. It includes data samples with two classes. Table 1 and Fig. 3 inspect the overall churn prediction outcomes of the OLSTM-CCPBI technique. The results highlighted that the OLSTM-CCPBI technique properly identified the existence of churn and non-churn samples. With 70%TRAS, the OLSTM-CCPBI technique provides an

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

average $accu_y$ of 94.37%, $prec_n$ of 94.69%, $reca_l$ of 94.37%, $F_{measure}$ of 94.28%, and AUC_{score} of 94.37%. Also, based on 30%TESS, the OLSTM-CCPBI method offered an average $accu_y$ of 93.75%, $prec_n$ of 94.83%, $reca_l$ of 93.75%, $F_{measure}$ of 93.94%, and AUC_{score} of 93.75%.

Table 1: Churn prediction outcome of OLSTM-CCPBI technique on 70:30 of TRAS/TESS

Classes	$Accu_y$	$Prec_n$	$Reca_l$	$F_{Measure}$	AUC_{Score}
TRAS (70%)					
Churn	99.42	90.00	99.42	94.48	94.37
Non-churn	89.33	99.38	89.33	94.08	94.37
Average	94.37	94.69	94.37	94.28	94.37
TESS (30%)					
Churn	100.00	89.66	100.00	94.55	93.75
Non-churn	87.50	100.00	87.50	93.33	93.75
Average	93.75	94.83	93.75	93.94	93.75

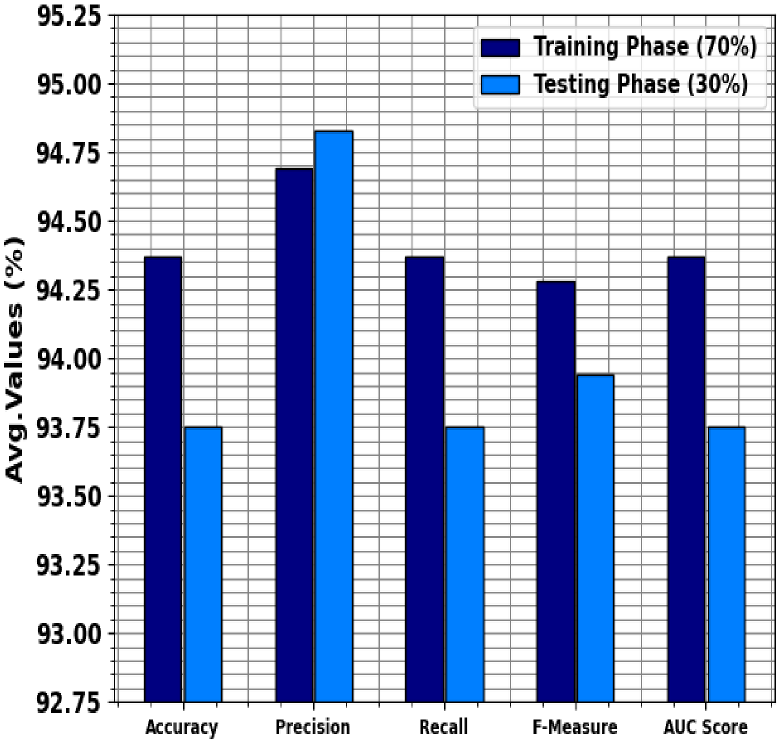

Figure 3: Average of OLSTM-CCPBI technique on 70:30 of TRAS/TESS

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

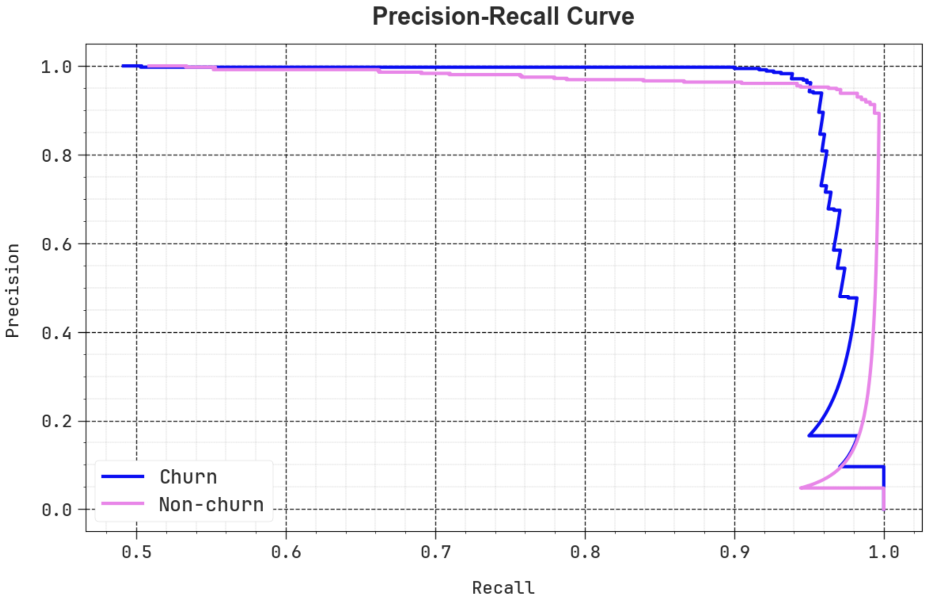

Figure 4: PR curve of the OLSTM-CCPBI model

Examining the precision-recall (PR) curve, as demonstrated in Fig. 4, the results confirmed that the OLSTM-CCPBI method gradually achieves boosted PR values in every class. It confirms the improved capabilities of the OLSTM-CCPBI technique for the identification of diverse classes, displaying the ability in the recognition classes.

Likewise, in Fig. 5, ROC curves acquired by the OLSTM-CCPBI technique outperformed in the classification of diverse labels. It offers a comprehensive understanding of the tradeoff among TPR and FRP over distinctive recognition threshold values and epoch counts. The figure emphasized the greater classifier outcomes of the OLSTM-CCPBI algorithm with every class, pointing out the efficiency of overcoming several classification complexities.

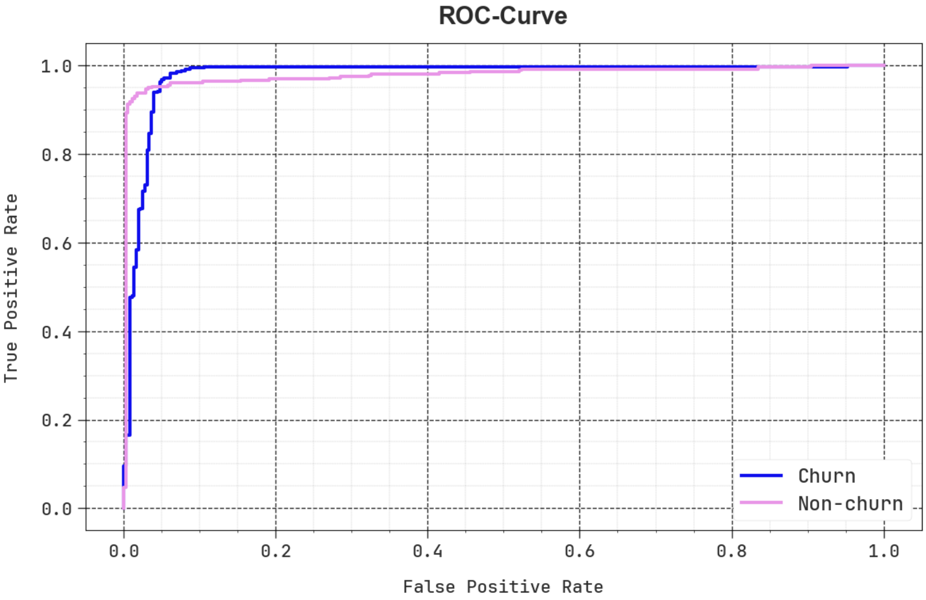

Figure 5: ROC curve of the OLSTM-CCPBI method

In Table 2 and Fig. 6, the results of the OLSTM-CCPBI technique have undergone comparison with recent approaches in terms of distinct measures [21, 22]. Based on $accu_y$, the OLSTM-CCPBI technique has resulted in boosted $accu_y$ of 94.37% while the LR, DT, KNN, RF, NB, SVM linear, and XGBoost models have obtained reduced $accu_y$ of 82.22%, 81.74%, 81.18%, 79.62%, 78.78%, 80.78%, and 82.32%, correspondingly. Along with that, based on $prec_n$, the OLSTM-CCPBI method offers an increased $prec_n$ of 94.69% whereas the LR, DT,

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023

KNN, RF, NB, SVM linear, and XGBoost techniques have acquired decreased $prec_n$ of 81.93%, 81.79%, 81.45%, 80.48%, 78.75%, 81.57%, and 82.26%. Meanwhile, based on $reca_l$, the OLSTM-CCPBI system is provided a higher $reca_l$ of 94.37% although the LR, DT, KNN, RF, NB, SVM linear, and XGBoost techniques get reduced $reca_l$ of 80.68%, 80.31%, 79.97%, 79.25%, 79.15%, 80.19%, and 81.85%, respectively. Moreover, with $F_{measure}$, the OLSTM-CCPBI method provides improved $F_{measure}$ of 94.28% however the LR, DT, KNN, RF, NB, SVM linear, and XGBoost algorithms have acquired minimized $F_{measure}$ of 80.61%, 80.41%, 78.76%, 79.64%, 78.87%, 80.50%, and 80.49%, correspondingly.

Table 2: Comparative outcomes of OLSTM-CCPBI technique with other algorithms

Model	$Accu_y$	$Prec_n$	$Reca_l$	$F_{Measure}$
OLSTM-CCPBI	94.37	94.69	94.37	94.28
Logistic Regression	82.22	81.93	80.68	80.61
Decision Tree	81.74	81.79	80.31	80.41
KNN Classifier	81.18	81.45	79.97	78.76
Random Forest	79.62	80.48	79.25	79.64
Naive Bayes (Gaussian)	78.78	78.75	79.15	78.87
SVM Classifier Linear	80.78	81.57	80.19	80.50
XGBoost	82.35	82.26	81.85	80.49

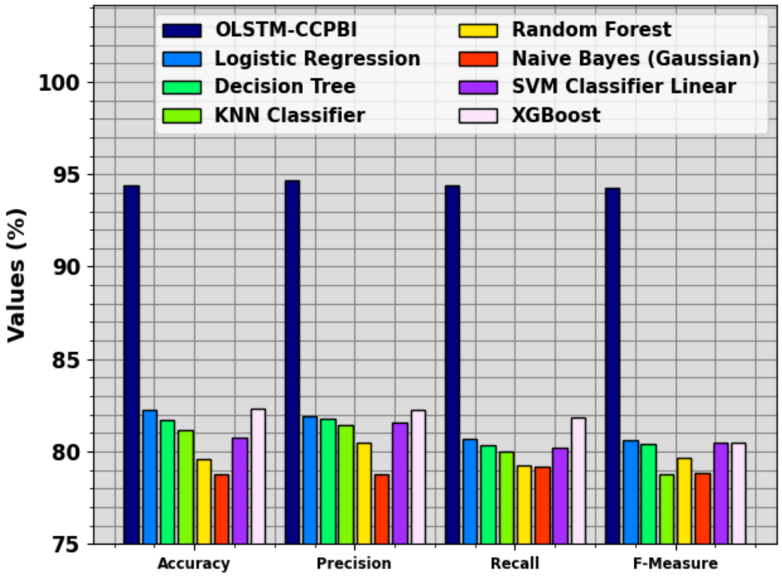

Figure 6: Comparative outcomes of OLSTM-CCPBI system with recent models

Thus, the OLSTM-CCPBI technique can be applied for enhanced prediction of customer churns.

5. Conclusion

In this study, we have introduced a novel OLSTM-CCPBI method. The proposed OLSTM-CCPBI method incorporates many innovative components, such as Min-Max scaling for normalization, LSTM networks for temporal sequence modelling, and Adam optimization for hyperparameter tuning. The OLSTM-CCPBI method effectively captures temporal dependency in sequential customer data by leveraging the dynamic nature of the LSTM network, which enables correct prediction of churn events. Through detailed investigations on real-time customer churn datasets, OLSTM-CCPBI achieves better predictive capabilities than classical approaches, displaying its promising solution to aid businesses in preemptively addressing customer attrition and considerably enhancing churn prediction accuracy.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Abbasimehr H, Setak M, Tarokh M (2011) A neuro-fuzzy classifier for customer churn prediction. International Journal of Computer Applications 19(8):35–41

[2] De Caigny, A., Coussement, K., De Bock, K.W., et al., 2020. Incorporating textual information in customer churn prediction models based on a convolutional neural network. Int. J. Forecast. 36 (4), 1563–1578.

[3] de Lima Lemos, R.A., Silva, T.C., et al., 2022. Propension to customer churn in a financial institution: a machine learning approach. Neural Comput. Appl. 34 (14), 11751–11768.

[4] Adwan O, Faris H, Jaradat K, Harfoushi O, Ghatasheh N (2014) Predicting customer churn in telecom industry using multilayer perceptron neural networks: Modeling and analysis. Life Science Journal 11(3):75–81

[5] Ahmad AK, Jafar A, Aljoumaa K (2019) Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data 6(1):28

[6] Dias, J., Godinho, P., Torres, P., 2020. Machine learning for customer churn prediction in retail banking. In: International Conference on Computational Science and its Applications. Springer, Berlin, pp. 576–589.

[7] Ali, M. (2023). The Decay of The Solutions of a Nonlinear Viscoelastic Hyperbolic Equation. *Pure Mathematics for Theoretical Computer Science*, 1(1), 47-55.

[8] Asthana P (2018) A comparison of machine learning techniques for customer churn prediction. International Journal of Pure and Applied Mathematics 119(10):1149–1169

[9] Geiler, L., Affeldt, S., Nadif, M., 2022. An effective strategy for churn prediction and customer profiling. Data Knowl. Eng. 142 (Nov.), 102100.

[10] Karvana, K.G.M., Yazid, S., Syalim, A., et al., 2019. Customer churn analysis and prediction using data mining models in banking industry. In: 2019 International Workshop on Big Data and Information Security. IEEE, pp. 33–38.

[11] Akhmetshin, E., Fayzullaev, N., Klochko, E., Shakhov, D. and Lobanova, V., 2024. Intelligent Data Analytics using Hybrid Gradient Optimization Algorithm with Machine Learning Model for Customer Churn Prediction. Fusion: Practice and Applications, 14(2), pp.159-59.

[12] Khoh, W.H., Pang, Y.H., Ooi, S.Y., Wang, L.Y.K. and Poh, Q.W., 2023. Predictive churn modeling for sustainable business in the telecommunication industry: optimized weighted ensemble machine learning. Sustainability, 15(11), p.8631.

[13] Moradi, B., Khalaj, M., Herat, A.T., Darigh, A. and Yamcholo, A.T., 2024. A swarm intelligence-based ensemble learning model for optimizing customer churn prediction in the telecommunications sector. AIMS Mathematics, 9(2), pp.2781-2807.

[14] Öztürk, M.E., Tunç, A.A. and Akay, M.F., 2023. Machine learning based churn analysis for sellers on the e-commerce marketplace. International Journal of Mathematics and Computer in Engineering, 1(2), pp.171-176.

[15] Fatima, G., Khan, S., Aadil, F., Kim, D.H., Atteia, G. and Alabdulhafith, M., 2024. An autonomous mixed data oversampling method for AIOT-based churn recognition and personalized recommendations using behavioral segmentation. PeerJ Computer Science, 9, p.e1756.

[16] Sun, Y., Liu, H. and Gao, Y., 2023. Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model. Heliyon, 9(2).

[17] Shantal, M., Othman, Z. and Bakar, A.A., 2023. A Novel Approach for Data Feature Weighting Using Correlation Coefficients and Min–Max Normalization. Symmetry, 15(12), p.2185.

[18] Sarmas, E., Spiliotis, E., Stamatopoulos, E., Marinakis, V. and Doukas, H., 2023. Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent Long Short-Term Memory models. Renewable Energy, 216, p.118997.

[19] Abouelmagd, L.M., Shams, M.Y., Marie, H.S. and Hassanien, A.E., 2024. An optimized capsule neural networks for tomato leaf disease classification. EURASIP Journal on Image and Video Processing, 2024(1), p.2.

[20] https://www.kaggle.com/code/kmalit/bank-customer-churn-prediction/data

[21] Lalwani, P., Mishra, M.K., Chadha, J.S. and Sethi, P., 2022. Customer churn prediction system: a machine learning approach. Computing, 104(2), pp.271-294.

[22] Singh, P.P., Anik, F.I., Senapati, R., Sinha, A., Sakib, N. and Hossain, E., 2024. Investigating customer churn in banking: A machine learning approach and visualization app for data science and management. Data Science and Management, 7(1), pp.7-16.

Doi: https://doi.org/10.54216/IJAACI.050105
Received: June 09, 2023 Revised: November 06, 2023 Accepted: December 06, 2023