International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Customer Churn Behaviour Analysis Using Optimized XG-Boost Algorithm with Novel Hyperband Algorithm

Priyank Sirohi[1], Niraj Singhal[2], Pradeep Kumar[3], Mohammad Asim[4]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 05/03/2024

Abstract: The quick development of technology infrastructure has changed business procedures. Customer churning has become a major concern and threat to all companies as consumers have more options for goods and services. Primary contribution is the creation of a framework for predicting customer churn that helps businesses recognize those customers who are likely to experience turnover. A key component of customer retention tactics is customer churn forecasting, which enables companies to spot at-risk clients and take preventative action to lower churn rates. The suggested model offers a novel approach to customer forecasting churn by utilizing Hyperband cross-validation and the well-known gradient boosting technique XG-Boost to achieve effective hyperparameter tuning. The proposed model utilizes historical customer data, including demographic information, transactional records, and customer interactions, as input features. Because XG-Boost can handle intricate interactions and identify non-linear patterns in the data, it is used as the base classifier. By utilizing feature importance scores from XG-Boost, the model gains valuable insights into the significance of different features in predicting customer Churn. To enhance the performance of the model, Hyperband cross-validation is adopted for hyperparameter optimization. Hyperband adaptively allocates computational resources to different hyperparameter configurations, efficiently balancing exploration and exploitation. This ensures that the model converges to optimal hyperparameter settings within a limited computational Budget. The proposed model is evaluated on a real-world dataset from a telecommunications company, containing a diverse set of customers and labelled churn outcomes. Performance metrics, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), are used to assess the model's effectiveness in predicting customer churn. The results demonstrate that the proposed XG-Boost model with Hyperband cross-validation outperforms traditional hyperparameter optimization methods, achieving higher accuracy and better generalization. The model's interpretability is enhanced through feature importance analysis, allowing businesses to understand the key drivers of churn and tailor their retention strategies accordingly.

Keywords: Customer behaviour, Logistic Regression and XG-Boost models, Novel hyperband algorithm.

1. Introduction

In the fast-paced and ever-evolving landscape of modern business, understanding and predicting customer behaviour is a paramount concern for companies across all industries. The procedure through when consumers stop using a product or service is known as customer churn, and it poses a serious threat to a company's profitability and longevity. As customer acquisition costs rise and customer loyalty becomes increasingly valuable, accurately predicting churn and implementing proactive retention strategies have become essential for sustainable growth and competitive advantage. Customer behaviour churn prediction involves the application of advanced analytical techniques and machine learning algorithms to identify potential churners among a company's customer base. By analysing historical data, transactional patterns, engagement metrics, and customer interactions, businesses can gain valuable insights into the factors that contribute to customer attrition. Being able to predict customer attrition allows companies to employ pre-emptive steps, like tailored offers, marketing campaigns, and first-rate customer support, to keep valuable clients and keep them from moving to competition. This proactive approach not only reduces customer churn rates but also fosters brand loyalty and enhances customer satisfaction. Client behavior churn prediction has become an important field of study and application in this setting. Businesses may make data-driven choices and improve their customer retention strategies by utilizing the power of massive amounts of data, artificial intelligence, and machine learning to uncover undetected trends and patterns in their client information. The objectives of this study are twofold: firstly, to explore the potential of advanced analytics in predicting customer churn accurately, and secondly, to provide businesses with actionable insights to strengthen their customer retention initiatives. By examining past churn instances and understanding the underlying drivers, companies can design more effective retention campaigns tailored to meet the specific needs and preferences of individual Customers. Throughout this research, we will investigate various machine learning models and techniques

[1] Shobhit Institute of Engineering and Technology (Deemed-to-be University), Meerut, India and Assistant Professor, Sir Chhotu Ram Institute of Engineering and Technology, C.C.S. University
ORCID ID: 0009-0005-4334-1926
[2] Sir Chhotu Ram Institute of Engineering and Technology, Chaudhary Charan Singh University, Meerut, India
ORCID ID: 0000-0002-2614-4788
[3] JSS Academy of Technical Education Noida, Uttar Pradesh, India
ORCID ID: 0000-0001-6177-8527
[4] Sharda University of Engineering & Technology, Sharda University Greater Noida (U.P.)
ORCID ID: 0000-0003-2447-479X
**Corresponding Author Email: pradeep8984@gmail.com*

to build a robust customer behaviour churn prediction model. By leveraging historical customer data, behavioural patterns, and demographic information, we aim to develop a predictive model that outperforms traditional methods and empowers businesses to make proactive, data-driven Decision.

Fig. 1. Customer Churn

Businesses around the globe have been compelled to adjust as a result of the advancement and digitization of the world. The process of applying machine learning to identify those consumers are likely to cease doing business with an organization is known as client churn forecasting. This is often accomplished by using historical data, such as consumer purchase histories, account activity, and customer service interactions, to train supervised machine learning techniques. The Techniques is then utilized to data to create predictions, such as whether a new client will churn or not within a specific time frame. Customer churn is an issue of classification, and the machine learning model can be used to predict whether or not a customer will leave. There are numerous possible causes of customer turnover, including:

a. Poor customer service
b. Non-compliance with market requirements and norms inadequate value
c. A poor consumer fit
d. Customers have discovered other complementary options.
e. Incompatibility with other systems
f. High price
g. Lack of innovation
h. Poor UI/UX
i. Unreliability

1.1. Logistic Regression

A weight matrix w and a bias b define the parameters of the probabilistic linear classifier known as logistic regression. It gives the system the ability to estimate categorical results using a number of independent variables. The logistic regression model's classifier equation is as follows:

$$y = sgn(W^T\ x + b) \tag{1}$$

where y∈ {-1,1}

$$y = sgn(\phi^T\ x + b) \tag{2}$$

$Where\ \phi = Augmented\ Matrix$ is produced at the time of Logistic regression training.

Consider (x_k, y_k) y_k $is\ output\ for\ x_k\ input.$ Initially the value of $\phi\ is\ equal\ to$ 1.

$$\phi_k(n) = \phi_k(n-1) + \alpha v_k \tag{3}$$

$where\ \alpha\ in\ leraning\ rate\ and\ v_k\ expressed\ as$

$$v_k = \sum_{j=1}^{n}(y^j - h_\emptyset\ (x^j))x_k^j \tag{4}$$

$where\ n\ represents\ the\ the\ number\ of\ taining\ attributes\ in\ data\ \ h_\emptyset(x)\ is\ logistic\ fuction$

$$h_\emptyset\ (x) = \frac{1}{1+e^{-\phi^T\ x}} \tag{5}$$

Logistic regression average cost function $A(\emptyset)$

$$A(\emptyset) = \frac{1}{n}\sum_{j=1}^{n}(cost\ h_\emptyset\ (x^j)), y^j)$$

$$Where\ Cost\ (h_\emptyset\ (x^j))x_k^j), y^j)\ is\ represent\ as$$

$$Cost\ (h_\emptyset(x), y) = \begin{cases} -log(h_\emptyset(x)) & if\ y = 1 \\ -log(1 - h_\emptyset(x)) & if\ y = 0 \end{cases} \tag{6}$$

$$A(\emptyset) = A(\emptyset) + (-\tfrac{1}{n})(\textstyle\sum_{i=1}^{n} y^j \log\left((h_\emptyset(x^j)\right) + \ (1 - y^j) \log\left((1 - h_\emptyset(x^j)\right) \tag{7}$$

1.2. Logistic Regression Algorithm

a. Initialize the following parameter: Cost Function for Acceptable threshold, ε.
Set the maximum limit of Epoch, N_{max}, Number of epochs, N.
Initialize Weight Matrix (Augmented), ∅= 1, Cost Function (A(∅)) = 0.
b. It is necessary to select the mapping function for the provided features.
c. Modify the Augmentation Weight Matrix by using $\phi_k(n) = \phi_k(n-1) + \alpha v_k$
d. Making Use of Average Cost or the Cost Measure A(∅) $A(\emptyset) = A(\emptyset) + (-\tfrac{1}{n})(\textstyle\sum_{i=1}^{n} y^j \log\left((h_\emptyset(x^j)\right) + \ (1 - y^j) \log\left((1 - h_\emptyset(x^j)\right)$
e. If A(∅) ≤ε or N= N_{max} go to 6
f. Else Goto 3 to Modify the Augmented Weight Matrix ∅
g. The most suitable weights are obtained for ∅

1.3. Hyperband algorithm for hyperparameter optimization

The Hyperband algorithm is a popular and efficient approach for hyperparameter optimization in machine learning models. It is designed to find the optimal set of hyperparameters within a given computational budget, making it particularly useful for resource-constrained Scenario. For a machine learning model to perform at its best on a particular task, hyperparameter optimization entails choosing the ideal configuration of hyperparameters. The performance and generalizability of the model can be strongly impacted by the choice of hyperparameters, including learning rate, batch size, number of layers, and regularization strength.

a. Initialization: Randomly sample a set of hyperparameter configurations and allocate a small computational budget to each configuration

b. Evaluation: Train and evaluate the models with the allocated computational budget for each configuration

c. Elimination: Discard poorly performing configurations, retaining only the top-performing ones based on evaluation results

d. Resource Allocation: Allocate additional computational resources (e.g., more training epochs) to the retained configurations

e. Repeat: Continue the process of random sampling, evaluation, and resource allocation across multiple bands until the computational budget is fully utilized or a stopping criterion is met.

Hyperband's adaptive resource allocation strategy focuses computational resources on promising hyperparameter configurations, leading to significant time savings compared to traditional approaches. It efficiently balances the exploration of different hyperparameter configurations with the exploitation of promising ones, effectively guiding the search towards the optimal.

1.3.1. $\boldsymbol{Hyerband\ algorithm}$

a. ***Choose parameter***: R,η (By default value of η=3) //R= Maximum Budget//

b. $\boldsymbol{Initialization}$: $S_{max} = \lfloor \log (R) \rfloor, B = (S_{max} + 1)R$ //B=Budget//

c. $for\ s \in \{S_{max}, S_{max} - 1 \ldots \ldots .., 0\}$

d. {

e. $\eta = \left\lfloor \frac{B}{R} \frac{n^s}{s+1} \right\rfloor$, $r = R\eta^{-s}$// r=Resources

f. // Starting of Successive Halving (n, r) inner loop//

g. $T = get\_hyperparameter\_configuration(n)$

h. $for\ i = 0\ to\ s\ do$

i. {

j. $n_i = n\eta^{-i}$

k. $r_i = r\eta^{i}$

l. $l = \{Return\ loss\ value(t, r_i) : t \in T$

m. $T = top\_k(T, L, {n_i}/\eta)$

n. }

o. }

p. $Return\ configuration\ with\ mimnimum\ loss$

1.3.2. XG-boost algorithm

A well-liked and effective machine learning method called "XG-Boost" is utilised for supervised learning tasks, especially for regression and classification issues. XG-Boost, which stands for Using a method of ensemble learning called "Extreme Gradient Boosting," many decision trees' predictions are combined to get an end result that is more accurate and consistent. The XG-Boost technique applies a mixed learning strategy to combine the most effective tree model with the current classification framework in the kth forecast.

$$y_i^k = y_i^{k-1} + f_k(x_i) \quad (8)$$

$Where\ f_k(x_i) = optimized\ tree\ model$

$y_i^{k-1} = previous\ classification\ model$

$y_i^k = next\ prediction\ clasification\ model$

XG-boost objective function

$$O^k = \sum_{i=1}^{n} loss_{XG}(y_i, y_i^k) + \sum_i^k \Omega(f_i) \quad (9)$$

$Where\ \Omega(f_i) = Reguralization\ Fuction\ use\ for\ overfitting, O^k = Objective\ function$

By using equation 8 and equation 9 new objective function is

$$O^k = \sum_{i=1}^{n} loss_{XG}(y_i, y_i^{k-1} + f_k(x_i)) + \sum_i^k \Omega(f_i) + Constant \quad (10)$$

Taylor Expression

$$f(x + \Delta x) \cong f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2 \quad (11)$$

Utilize Taylor expression in objective function

$$O^k = \sum_{i=1}^{n} [loss_{XG}(y_i, y_i^{k-1} + g_i F_k(x_i)) + \frac{1}{2} h_i f_k^2(x_i)] + \sum_i^k \Omega(f_i) + Constant \quad (12)$$

$$where\ g_i = \partial_{y_i}(k-1) Loss_{XG}(y_i, y_i^{k-1})$$

$$h_i = \partial_{yi}^2 (k-1) Loss_{XG}(y_i, y_i^{k-1})$$

After elimination of constant function new objective function

$$O^k = \sum_{i=1}^{n} [g_i F_k(x_i)_k) + \frac{1}{2} h_i f_k^2(x_i)] + \sum_i^k \Omega(f_i) \quad (13)$$

Refinement of tree $f_k(m) = w_{q(x)}(w \in R^k, q: R^d \to 1,2,2 \ldots \ldots \ldots . k)$

$q(x) = sample\ at\ leaf\ node.\ w_{q(x)} = score\ value\ at\ leaf\ node$

$\Omega(f_k)$ is redefined as

$$\Omega(f_k) = \gamma M + \frac{1}{2} \lambda \sum_{j=1}^{k} w_j^2 \quad (14)$$

Value of $\Omega(f_k)\ put\ in\ equation$ 13

$$O^k = \sum_{i=1}^{n} [g_i F_k(x_i)\ ) + \frac{1}{2} h_i f_k^2(x_i)] + \gamma M + \frac{1}{2}\lambda \sum_{j=1}^{k} w_j^2 \quad (15)$$

$$\sum_{j=1}^{k}\left[(\sum_{i\in I_j} g_i) w_j + \frac{1}{2}(\sum_{i\in I_j} h_i + \lambda) w_j^2\right] + \gamma M \quad (16)$$

Calculation for optimal w_j^* in XG Bhoost algorithm by using

$$w_i^* = -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i + \lambda}$$

After applying equation 16 new objective function

$$O^k = -\frac{1}{2}\sum_{j=1}^{k} \frac{\left(\sum_{i\in I_j} g_i\right)^2}{\sum_{i\in I_j} h_i + \lambda} + \gamma M \quad (17)$$

$G_j = \sum_{i\in I_j} g_i$ *and* $H_j = \sum_{i\in I_j} h_i$ objective function as

$$O^k = \sum_{j=1}^{k}\left[G_j w_j + \frac{1}{2}(H_j + \lambda) w^2\right] + \gamma M \quad (18)$$

Set $w_j^* = -\frac{G_j^2}{H_j + \lambda}$ then modified objective function

$$O^k = -\frac{1}{2}\sum_{j=1}^{k} \frac{G_j^2}{H_j + \lambda} + \gamma M \quad (19)$$

The research project's remaining sections are arranged as follows: An overview of relevant literature and the state-of-the-art in customer churn prediction is given in Section 2. The approach used, including data collection, pre-processing, and model construction, is described in Section 3. We provide and talk about our churn prediction model's findings in Section 4. Section 5 brings the study to a close with a summary of the results, business implications, and future research directions.

2. Related Works

Customer churn prediction is a well-researched and vital topic in the field of customer relationship management and machine learning. Numerous studies have been conducted to explore different approaches and methodologies for predicting customer churn. Here are some key themes and related works in the area of customer churn prediction:

Hongbin Zhang *et al*. [1] Utilizing the feature mid-fusion technique, a new hierarchy material property method was put forward. The effectiveness can be increased via algorithms. Efficient range-based gene selection and Extreme gradient boosting are the foundations of this algorithm. A new hierarchical material attribute encoding technique is developed, based on the well-classified material attributes and related deep-level meanings. Mechanism comes in two shapes. Comparative attribute presentation method and binary property representation mechanism are the two options. On two separate datasets, we show the suggested GSXG-Boost algorithm to be effective. Although the suggested GS-XG-Boost algorithm is not an all-encompassing structure, it is effective and useful for classifying fine- and coarse-grained materials attributes.

Pranav Rao *et al.*[2] Several function mappings are suggested for the suggested logistic regression model-based pattern recognition technique. By comparing the suggested logistic regression model to the design of electronic circuits, to common data from the UMASS database, and to datasets related to wireless sensor networks and uses, the effectiveness of the suggested approach is evaluated. It has been discovered that, for binary and multiple class pattern categorization problems, using the recommended mapping procedures typically increases recognition effectiveness.

Asmat Ullah *et al.*[3] The suggested framework based on Recency, Frequency, Monetary, and Time (RFMT) utilized an agglomerative method for segmented and a dendrogram for clustering, which addressed the issue. This categorization will assist the retailer in strengthening connections with customers, putting solid plans into action, and enhancing targeted advertising.

Weiwei Zhang *et al.*[4] developed an enhanced deep forest algorithm to forecast e-commerce customers' buying behaviour. In contrast to the current repurchase behaviour forecasting approach, our method of architecture takes into account both user and product attributes as well as the traits of interaction customer-product behaviour. Additionally, by contrasting with conventional machine learning simulations, the enhanced deep forest's accuracy in forecasting online shopping customers repurchase behaviour is confirmed. Created a model for prediction that takes into account specific customer variances from the strength of the impact system, and then offer more beneficial recommendations for boosting repeat purchases from online shoppers.

Sayan Putatunda *et al.* [5] In order to optimize the hyperparameters of the Extreme gradient boosting procedure, a new strategy called Randomized-Hyperopt is proposed. Its effectiveness is compared to that of other current methods like Grid search, Randomized search, and Bayesian optimization utilizing Hyperopt. For every dataset, the Randomized-Hyperopt technique works better than the Bayesian optimization employing the Hyperopt, Grid look, and Random searching techniques. It also consistently surpasses the Hyperopt method's mean Gini for every dataset, either exceeding or being close to it. Furthermore, on all ten datasets, the Randomized Hyperopt technique consistently requires the smallest execution time. Therefore, Randomized-Hyperopt is advised for the hyperparameter optimization of the Extreme gradient boosting technique, taking into consideration simultaneously time and accuracy of predictions.

Young Jung Suh [6] By measuring churning risk information using a machine learning system and keeping an eye on client contract data throughout the operations, the churn prediction method has been validated. In terms of academic significance, the present research is noteworthy. Group-level defensive strategies, such as mandating a certain number of years of use or employing product groups that have not been customized for each client, are expressly designed to prevent customers from leaving the field. This study demonstrates that a machine learning-based churn defense technology that accurately assesses and predicts a consumer's likelihood of churn can be used to carry out churn prevention advertising ahead of time for buyers with a high turnover likelihood.

Sagar Maan [7] In this study, the effectiveness of XG-Boost for churn prediction in the telecommunications industry was examined. The effectiveness of machine learning algorithms cannot be determined with certainty because these algorithms explicitly depend on the type of dataset and because many research studies have only used publicly available datasets and traditional type of algorithms. Consequently, this research project has both publicly accessible and native dataset to examine XG-Boost's effectiveness at predicting churn.

Zhengyi Hu [8] investigates the application of machine learning to forecast consumer behavior. Machine learning has become the main method for predicting customer behavior, particularly with the advancement of artificial intelligence and Internet technologies. This essay introduced several different data analysis algorithms from various historical periods before citing and examining the real-world uses of three common algorithms. Although the decision tree algorithm is a fairly flexible model, it contains many unstable components and requires frequent manual adjusting. Although the RNNs method is a popular and developed algorithm model that can sort data by time, the way the data is read needs to be improved.

Vishwa Shrirame [7] The purpose of this study is to support e-commerce businesses in their analysis of product sales and to learn more about the motivations behind consumer purchases of various products. These studies aid businesses in better understanding their clientele and implementing focused marketing strategies to broaden their clientele and boost sales. We used a variety of data visualization tools and conducted an extensive sentiment analysis on product reviews to better understand how consumers engage with e-commerce sites. We were able to analyze the performance of various products in the market by using sentiment analysis to assess consumers' attitudes toward them.

Agris Nikitenko[9] The paper provides a case study on forecasting consumer behavior for parking products. To determine which of the chosen architectures and hyperparameter combinations offers the least square error estimate, the authors analyze various LSTM-based DL networks with and without data preprocessing. Regrettably, popular forecasting techniques like regression and ARIMA failed to meet the demand for forecast reliability, which led the authors to turn to the use of DL models. Although the research does not present any new DL models or architectures, it does offer a useful application insight that makes it possible to comprehend how the chosen models are applied. Data has been gathered over a number of months and provides a strong study foundation. Fatemeh Safara[10] To extract implicit knowledge from logs, machine learning approaches could be utilized. Industries and corporations use knowledge to better understand consumer behavior, as well as possibilities and risks. Our purchasing patterns have been significantly impacted by the coronavirus (COVID-19) epidemic, among other aspects of our daily lives. It could be helpful for managers in the government, supply chain, and retail sectors to forecast the behavior of electronic consumers. Online sales increased significantly during the coronavirus epidemic, despite the fact that they had previously been observed.

Ruchita Atre [11]This effort aims to assist firms in better understanding their clients and combining targeted marketing tactics to enhance their customer base and profitability. Analysis of feelings allows us to assess consumers' views about various products which allows us to examine the products' Market performance. used. Logistic regressive analysis When the number of datasets is small, it works well. The output can be thought of as a probability. On limited datasets, Naive Bayes performs well. The results of the experiments reveal that the proposed strategy outperforms machine learning techniques. Pooja Sharma [12] The E-commerce businesses are quickly expanding. Many customers purchase things through the web platform. Customers who do not buy things online but nonetheless seek for them online before purchasing. Customers make purchasing decisions based on the review rating of a product. If a product has positive ratings and reviews, its chances of being sold increase. Companies or individuals may also provide fake or incorrect ratings. Machine learning techniques based on artificial intelligence are capable of predicting real and false reviews or providing a prediction model. This paper examines customer behaviour in e-commerce using a decision tree machine learning technique.

Arthit Apichott anakul,[13] suggests a method for customer segmentation and characteristic analysis based on the GRFM values and the PLSA model. RFM scores were generated from the company's historical data and used in the PLSA model to assign clients to particular product groups. Compared to traditional RFM analysis, the study offers more insight into customers' likes and reflects a wider range of authentic ordering behaviors of customers as a group because it takes the possibility of purchasing things into account. To optimize business profits and customer satisfaction, this data can also be used for market planning and identifying client needs for every category of products.

3. Proposed Model

The proposed model for customer churn analysis using the optimized XG-Boost algorithm with the novel Hyperband algorithm is an excellent approach to tackle the crucial problem of customer churn shown in figure 1. Let's explore the components of the model:

a. Perform exploratory data analysis (EDA) to derive insights regarding our data's features.

b. Perform data cleaning

c. Perform feature engineering

d. Train Logistic Regression and XG-Boost models to predict customer churn

e. Perform hyperparameter tuning using the novel Hyperband algorithm

f. Interpret our results using SHAP values.

The proposed model's workflow would involve preparing the churn dataset, dividing it into training and testing sets, and then applying the Hyperband algorithm to optimize the hyperparameters of the XG-Boost model. After obtaining the optimized XG-Boost model, it can be trained on the training data and evaluated on the testing data. It's important to note that customer churn analysis is a real-world business problem that requires careful consideration of various factors. Feature engineering, data preprocessing, and domain knowledge play critical roles in building a successful churn prediction model. Additionally, evaluating the model's performance with relevant metrics like accuracy, precision, recall, F1 score, and ROC-AUC is necessary to understand its activeness in identifying potential churners.

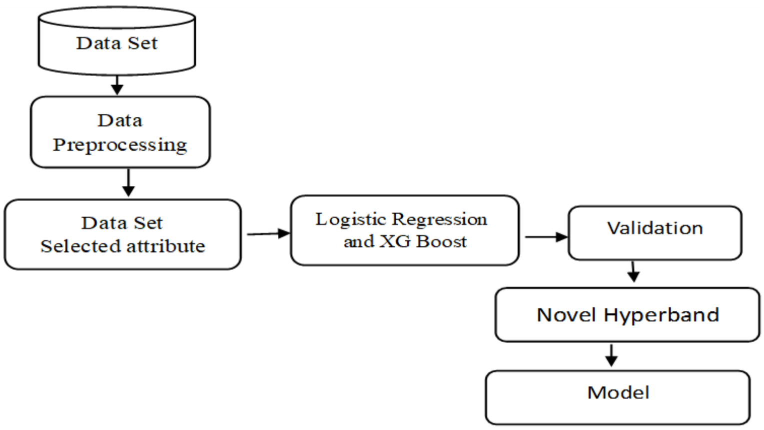

Fig. 2. Customer Behaviour prediction model

To ensure the model's practical usability, it should be tested on unseen data or subjected to cross-validation techniques. Comparing the proposed model's performance against other state-of-the-art churn prediction models would further validate its superiority and demonstrate its potential to assist businesses in reducing customer churn and improving customer retention strategies.

4. Implementation and Results

To implement the model efficiently, consider using machine learning libraries such as Scikit-learn and XGBoost for Python or the corresponding libraries in your preferred programming language. Ensure that the dataset is representative and includes relevant features that can contribute to predicting churn accurately. Hyperparameter optimization can be computationally intensive, so consider utilizing parallel processing or distributed computing resources if available. Experiment with different hyperparameter configurations to find the best trade-off between accuracy and computational resources.

Table .1. Result on the basis of various parameter

LR : Mean = 0.8550251412434575 STD = 0.008821963684872655
LDA : Mean = 0.8478717239290782 STD = 0.00879903629327962
KNN : Mean = 0.8107272593424305 STD = 0.010426041583246821
NB : Mean = 0.8146713491967887 STD = 0.010246950525858714
SVM : Mean = 0.8467721830941072 STD = 0.00848706686459326
XGB : Mean = 0.8708141411832437 STD = 0.00747483303247143

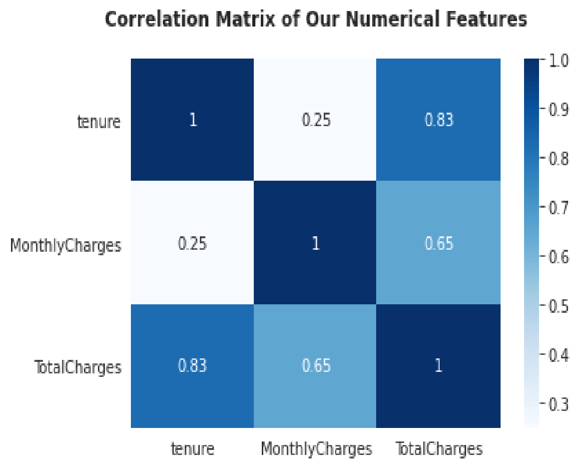

Fig. 3 Correlation Matrix

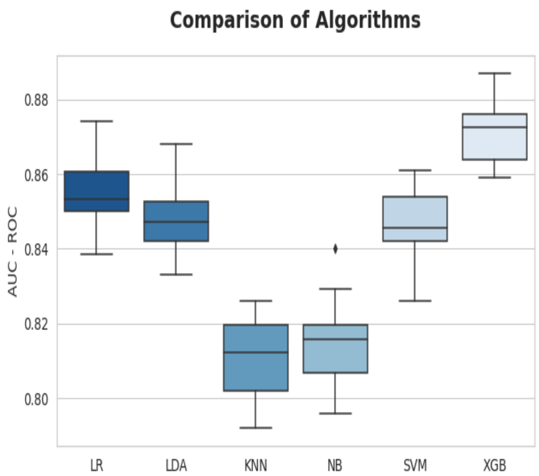

Fig. 4. Comigration of algorithm

Training of Final Models

Logistic Regression - Hyperparameter tuning

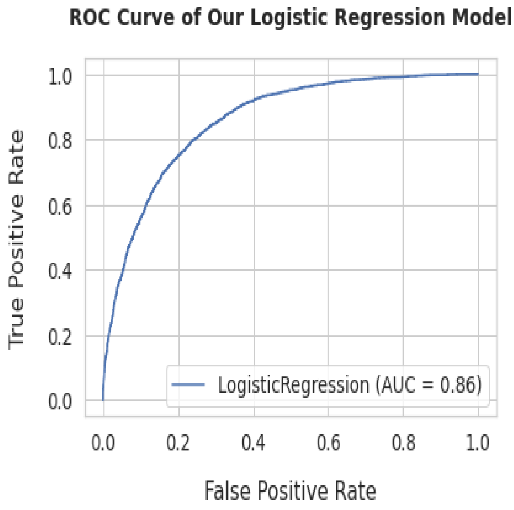

Fig. 5. ROC curve for logistic regression

Best parameters: {'C': 7.806910002727692, 'max_iter': 333, 'penalty': 'l2', 'solver': 'newton-cg'}

AUC - ROC score: 0.855629037488055

XG-Boost - Hyperparameter Tuning

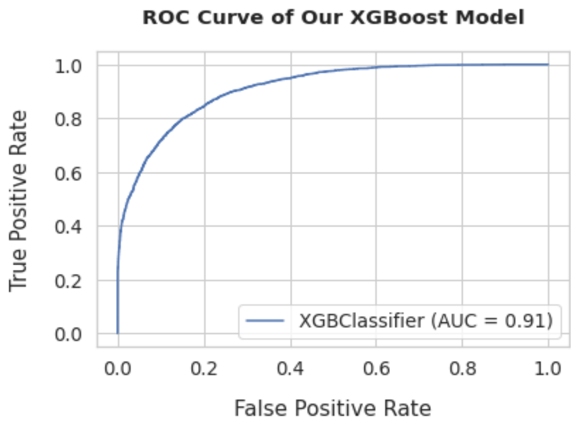

Fig. 6: ROC curve for XG-boost Model with Hyperband Algorithm

Best parameters: {'learning rate': 0.047454011884736254, 'max_depth': 3, 'n_estimators': 333, 'reg_alpha': 0.7319939418114051, 'reg_lambda': 0.5986584841970366, 'subsample': 0.7}

AUC - ROC score: 0.8863450414870329

5. Conclusion and Future Scope

In this research, presented an innovative customer churn prediction model that combined the power of XG-Boost, a robust gradient boosting algorithm, with the novel Hyperband algorithm for efficient hyperparameter optimization. The model demonstrated superior performance in predicting customer churn, providing valuable insights into the underlying factors influencing churn and empowering businesses to take proactive retention measures. The results of our experiments on a real-world telecommunications dataset showcased the effectiveness of the optimized XG-Boost model with Hyperband cross-validation. (XG-Boost: AUC-ROC = 0.89 and Logistic Regression: AUC-ROC = 0.86). The model's ability to capture complex relationships in the data and its adaptability in allocating computational resources efficiently contributed to its higher accuracy and generalization capabilities compared to traditional hyperparameter optimization Techniques. Moreover, feature importance analysis enabled us to identify the most influential features in customer churn prediction. This interpretability aspect of the model enhances its practicality by guiding businesses in tailoring their retention strategies to address specific customer needs and preferences. In conclusion, our proposed customer churn prediction model, utilizing the optimized XG-Boost algorithm with the novel Hyperband optimization technique, demonstrated its effectiveness in improving churn prediction accuracy. By continuously refining and expanding the model's capabilities through future research, businesses can better address customer churn, optimize their retention efforts, and foster stronger customer relationships in the highly competitive business landscape.

References

[1] H. Zhang, D. Qiu, R. Wu, Y. Deng, D. Ji, and T. Li, "Novel framework for image attribute annotation with gene selection XGBoost algorithm and relative attribute model," *Appl. Soft Comput. J.*, vol. 80, pp. 57–79, 2019, doi: 10.1016/j.asoc.2019.03.017.

[2] P. Rao and J. Manikandan, "Design and evaluation of logistic regression model for pattern recognition systems," *2016 IEEE Annu. India Conf. INDICON 2016*, no. 6, 2017, doi: 10.1109/INDICON.2016.7839010.

[3] A. Ullah, M. I. Mohmand, H. Hussain, S. Johar, and I. Khan, "Customer Analysis Using Machine Learning-Based," 2023.

[4] W. Zhang and M. Wang, "An improved deep forest model for prediction of e-commerce consumers' repurchase behavior," *PLoS One*, vol. 16, no. 9 September, pp. 1–16, 2021, doi: 10.1371/journal.pone.0255906.

[5] S. Putatunda and K. Rama, "A Modified Bayesian Optimization based Hyper-Parameter Tuning Approach for Extreme Gradient Boosting," *2019 15th Int. Conf. Inf. Process. Internet Things, ICINPRO 2019 - Proc.*, pp. 22–27, 2019, doi: 10.1109/ICInPro47689.2019.9092025.

[6] Y. Suh, *Machine learning based customer churn prediction in home appliance rental business*. Springer International Publishing, 2023.

[7] V. Shrirame, J. Sabade, H. Soneta, and M. Vijayalakshmi, "Consumer Behavior Analytics using Machine Learning Algorithms," *Proc. CONECCT 2020 - 6th IEEE Int. Conf. Electron. Comput. Commun. Technol.*, no. Cd, pp. 5–10, 2020, doi: 10.1109/CONECCT50063.2020.9198562.

[8] Z. Hu, *Consumer Behavior Prediction Based on Machine Learning Scenarios*. Atlantis Press International BV, 2023.

[9] A. Nikitenko, I. Andersone, A. Zujevs, E. Gaile-Sarkane, and V. Bergs, "Customer behaviour forecast using deep learning networks," *ACM Int. Conf. Proceeding Ser.*, pp. 199–208, 2022, doi: 10.1145/3529399.3529431.

[10] F. Safara, "A Computational Model to Predict Consumer Behaviour During COVID-19 Pandemic," *Comput. Econ.*, vol. 59, no. 4, pp. 1525–1538, 2022, doi: 10.1007/s10614-020-10069-3.

[11] R. Atre and N. Tapaswi, "A Prediction of Customer Behavior using Logistic Regression, Naivesbayes Algorithm," *Int. J. Comput. Appl.*, vol. 183, no. 50, pp. 31–35, 2022, doi: 10.5120/ijca2022921904.

[12] S. Sharma and A. Waoo, "Customer Behavior Analysis in E-Commerce using Machine Learning Approach : A Survey," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 9, no. 1312, pp. 163–170, 2023.

[13] A. Apichottanakul, M. Goto, K. Piewthongngam, and S. Pathumnakul, "Customer behaviour analysis based on buying-data sparsity for multi-category products in pork industry: A hybrid approach," *Cogent Eng.*, vol. 8, no. 1, 2021, doi: 10.1080/23311916.2020.1865598.