EAI Endorsed Transactions
on Intelligent Systems and Machine Learning Applications Research Article EAI.EU

Development of Churn Prediction System for a Marketing Company Using Machine Learning Techniques

I. Yusuf[1*], S.O. Abdulsalam[2], T.O. Jimoh[3] and I.I. Akuji[4]

[1]Department of Business and Entrepreneurship, Faculty of Management and Social Sciences, Kwara State University, Malete, Nigeria
[2]Department of Computer Science, Faculty of Information and Communication Technology, Kwara State University, Malete, Nigeria
[3]Department of Business and Entrepreneurship, Faculty of Management and Social Sciences, Kwara State University, Malete, Nigeria
[4]Department of Computer Science, Faculty of Information and Communication Technology, Kwara State University, Malete, Nigeria

Abstract

INTRODUCTION: In today's competitive marketing landscape, customer churn prediction is vital for marketing organizations to identify patterns, factors, and indicators contributing to customer attrition. This paper focuses on developing a customer churn prediction system using machine learning algorithms.
OBJECTIVES: This paper employed an e-commerce dataset, obtained from the Kaggle repository, and was preprocessed. Important features were selected from the preproccessed dataset before models' development. The parameters of AdaBoost, Gradient Boosting (GB), and Extreme Gradient Boosting (XGB) were optimized to improve their performance.
METHODS: Techniques such as label encoder, mean imputation, and synthetic minority over-sampling technique (SMOTE) were applied during data preprocessing stage. Ensemble learning algorithms, namely AdaBoost, GB, and XGB were used to develop the model while random search was employed for parameter optimization. Accuracy, precision, recall, and F1-score metrics were used to evaluate the models' performance.
RESULTS: The results of the models with 15 important selected features before parameter tuning yielded the following scores: AdaBoost attained 87% accuracy, 77% precision, 81% recall, and an 79% F1-score. Gradient boosting outperformed AdaBoost with 89% accuracy, 80% precision, 82% recall, and an 81% F1-score. XGB outperformed the two algorithms (AdaBoost and GB), achieving 97% accuracy, 96% precision, 94% recall, and 95% F1-score. Notably, the Random Search significantly improves Gradient boosting's performance, increasing accuracy from 89% to 97%, precision from 80% to 97%, recall from 82% to 93%, and F1-score from 81% to 95%., making it comparable to XGB results. SHAP analysis reveals that the "Complain" feature was a consistent and key positive driver of predicted churn across all models, implying that customers who register complaints are significantly more likely to churn than those who do not. Additionally, the 'tenure' feature which has a strong negative impact on the prediction across the three models implies that longer tenure logically means less likely to churn, which makes the model's behaviour intuitive and trustworthy.
CONCLUSION: The results demonstrate the effectiveness of the system in identifying at-risk customers, enabling businesses to proactively retain customers and reduce churn rates. The findings of this paper showcase the critical importance of effective complaint management and rapid response strategies in customer retention efforts, and suggest that fostering long-term relationships and increasing customer loyalty can be highly effective in reducing churn.

Keywords: Churn Prediction, Machine Learning, Ensemble Method, Adaboost, Gradient Boosting, Extreme Gradient Boost.

Received on 19 June 2025, accepted on 29 October 2025, published on 30 October 2025

doi: 10.4108/eetismla.9577

*Corresponding author. Email: ismaila.yusuf@kwasu.edu.ng

1. Introduction

In the context of a marketing organization, a churn prediction system could significantly enhance business outcomes. Marketing companies invest substantial resources in acquiring customers, and each churned customer represents both a lost investment and a potential gain for competitors. Studies have shown that retaining an existing customer is often more cost-effective than acquiring a new one, emphasizing the need for effective churn management strategies [1]. In a highly competitive market, failure to address customer churn can lead to substantial losses, eroding the customer base and threatening business viability. Hence, developing an accurate, data-driven churn prediction system can provide a competitive edge in the market. Traditional methods for managing churn, such as manual analysis or heuristic-based approaches, often fail to capture the complexity and scale of modern customer behaviour. Furthermore, marketing companies often deal with heterogeneous data, including demographic information, transaction histories, and engagement metrics, which complicates the prediction process.

[2] assert that the growing accessibility of customer data, combined with affordable computational infrastructure and data storage, has driven the adoption of machine learning for customer churn prediction. Machine learning (ML) techniques offer robust tools for analyzing customer data and detecting patterns that can lead to improved prediction accuracy. ML models can process vast amounts of historical data, including demographics, purchasing behaviours, and engagement patterns, to predict the likelihood of churn. Notably, machine learning algorithms have been applied by previous researchers, namely, logistic regression, decision trees, and neural networks among others, to churn prediction tasks, each demonstrating the capability to identify high-risk customers with a reasonable degree of accuracy [3]. These techniques also allow for continuous refinement, ensuring the model adapts to new data trends over time.

Essentially, ML models offer a powerful alternative by automating the analysis of customer data and generating precise churn predictions. Whereas, despite its potential, many marketing companies have yet to adopt or fully leverage ML-based systems, often due to a lack of tailored solutions or expertise in implementing such technologies. Existing churn prediction models were built using various ensemble methods, particularly boosting method such as XGBoost [4] [5], Gradient Boosting [6] [7], and Adaboost [8], however, there is a lack of comprehensive comparison of the efficacy of these boosting methods in customer churn prediction. Moreover, existing studies relied on automated feature selection techniques, namely analysis of variance recursive feature elimination, feature importance, and Pearson correlation among others, which are considered worthwhile, nevertheless, there is a need to explore alternative methods, such as Information Gain. Information Gain, a filter method, is employed for its efficiency and lower computational cost compared to wrapper methods. Notably, parameter optimization is slightly considered by existing studies. Optimizing model parameter, using various techniques such as random search, can significantly impact model performance, leading to improved reliance and adoption.

Consequently, this study addresses the critical need for a robust and scalable churn prediction system tailored specifically for marketing companies through the development of a system that accurately identifies high-risk customers and provides actionable insights for retention strategies using a promising ensemble method, feature selection technique and hyperparameter tuning strategy. Precisely, this study is guided by the following specific objectives: (i) acquiring a benchmark customer churn dataset from Kaggle repository. (ii) employing automated feature selection using Information Gain (iii) developing customers' churn prediction system using an optimized boosting techniques and random search. (iv) evaluating and comparing the effectiveness of the optimized ensemble boosting methods (Adaboost, Gradient Boosting, and XGBoost) using accuracy, precision, recall and F1-score.

The proposed system enables decreased customers' attrition by accurately forecasting customers at high risk of churn. Having a predictive model to identify churn gives the marketing company an edge over competitors who rely on reactive approaches to customer retention. Most important, the proposed system contributes to technology advancement and the existing body of knowledge in the areas of machine learning and churn prediction.

2. Literature Review

2.1 Related Works

Studies of customer churn prediction cut across traditional algorithms, ensemble methods, neural networks, and deep learning techniques. Each of these machine learning categories contributing distinct strengths to the field. In this study, existing studies are synthesised to enable comparative performance and justification.

2.1.1 Traditional Algorithms

Traditional methods such as Logistic Regression, and Decision Trees have consistently demonstrated reliable baseline performance in customer churn prediction across various domains. Logistic Regression frequently yields competitive accuracy and recall scores, making it a robust for churn identification, as evidenced in [6][7][9][10]. Decision Trees remain useful for scenarios requiring transparency and explainability as evidenced in [11]. Although, traditional models show promise in churn prediction but are often being outperformed by ensemble and deep learning methods especially in complex tasks.

2.1.2 Neural Networks & Deep Learning

Neural network architectures such as multilayer perceptrons (MLP), artificial neural networks (ANN), convolutional neural networks (CNNs), and domain-specific deep learning models offer better performance especially when applied to rich, high-dimensional customer data [12]. CNNs in particular have recently shown breakthrough results with accuracy rates as high as 97.62% for churn prediction in retail contexts [13]. Deep learning models are noted for effectively capturing complex nonlinear patterns in customer behaviours, and are increasingly favoured for large-scale, multi-categorical datasets such as those found in e-commerce, despite challenges with imbalanced data [14]. However, they require significant computational resources and expertise.

2.1.3 Ensemble Methods

Ensemble methods such Random Forests, Gradient Boosting (GBM), XGBoost, AdaBoost, CatBoost, and stacking had consistently outperformed traditional machine learning algorithms in churn prediction research. Studies [2][6][4][8][15][16][17] found that ensemble models such as XGBoost, CatBoost, and Random Forest achieve superior metrics in accuracy, precision, recall, and F1-score across domains like telecom, retail, and e-commerce. XGBoost regularly emerges as the leading single algorithm, attaining accuracy rates from 73% to 96% depending on dataset and feature engineering. Ensemble-fusion models integrating multiple algorithms have pushed accuracy even higher, surpassing 95% in some benchmarks, and demonstrate strong generalizability and robustness for business applications [18].

Notably, ensemble methods offer consistent highest predictive accuracy and robustness, making them highly effective for practical deployment in churn prediction. While previous research establishes ensemble methods and boosted algorithms as gold standards for churn prediction, few studies have specifically leveraged automated feature selection and lack of hyperparameter optimization to build scalable, interpretable systems tailored for churn prediction. The ability to turn raw predictions into real business value by not only identifying potential churners but also identifying the contributing features remains inadequately addressed. Therefore, this study tackles the pressing need to develop a robust, scalable, and business-centric customer churn prediction system

2.2 Related Concepts

2.2.1 Overview of Customer Churn

Customer churn is defined as the loss customers by an organization [19]. This concept is crucial in today's fiercely competitive market, where customer retention is key to an organization's survival. According to [20], customer churn prediction has become a vital marketing strategy, particularly in the financial services industry. While customer churn prediction is a complex issue, it also presents opportunities for organizations to leverage innovative technologies such as machine learning [21].

Overall, customer churn prediction is a critical marketing strategy that marketing organizations cannot afford to ignore. Through understanding the causes and consequences of customer churn, organizations can develop effective retention strategies to maintain customer loyalty and stay competitive in the market. According to [17] there are two main types of customer churn, namely, voluntary and involuntary churn. According to [7], the former (voluntary churn) is a category of churn whereby customers intentionally terminate their relationship with a company, often as a result of factors such as dissatisfaction with the service or product, attractive offers from competing organizations, evolving consumer needs, or financial limitations, whereas the latter (involuntary churn) occurs when clients discontinue their relationship with an organization due to circumstances beyond their control. Examples of such incidents include the death of a client, relocation to a location where the service is not accessible, or facing technical challenges that remain unresolved.

Companies need to understand the varying factors that drive customer churn to effectively retain customers. Key contributors include customer dissatisfaction, which often prompts customers to seek better alternatives [8]. Regular monitoring and quality improvements are essential to address this. Additionally, a lack of customer engagement can lead to disinterest and defection to competitors [22]. To mitigate this, a focus on creating engaging experiences through loyalty programs, active communication, and personalization becomes imperative.

2.2.2 Machine Learning

Machine learning, a subset of artificial intelligence, involves creating algorithms that enable systems to learn from data and improve their performance without explicit programming. [23]. However, machine learning is categorized in terms of the learning approach into three categories, namely, supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, a model is trained on a labelled dataset while unlabelled dataset is used in unsupervised learning [24]. Reinforcement Learning (RL) enables agents to learn decision-making strategies in complex, dynamic environments. Ensemble method is another powerful machine learning approach that combines the predictions of multiple base models to improve overall performance [25].

Ensemble learning method consists of bagging, boosting, and stacking. In bagging, random forest is the common algorithm [26], that aggregates multiple decision trees to enhance the overall model performance [27]. The popular categories of boosting method namely, Adaptive Boosting (AdaBoost), Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost). [27] describes AdaBoost as an ensemble learning technique that iteratively trains basic classifiers on subsets of data, focusing on incorrectly classified observations. Gradient Boosting creates sequential models that correct previous inaccuracies, minimizing mean squared errors [5]. XGBoost is a highly scalable and enhanced algorithm that builds upon Gradient Boosting Machines (GBM), offering faster computational speed and superior predictive performance in machine learning applications [29].

3. Methodology

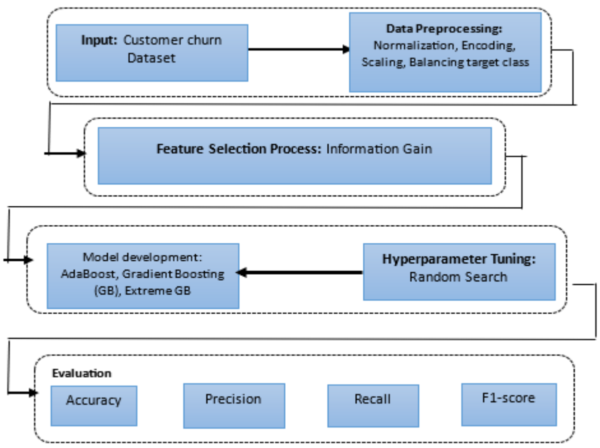

Figure 1: System Approach

Figure 1 shows that the first stage of the system development process is dataset import, follow by data preprocessing using normalization, encoding, scaling and class balancing. The next step is feature selection using information gain, followed by model development and hyper parameter tuning using AdaBoost, Gradient boosting and Extreme GB and random search, respectively. Lastly, the system is tested using four evaluation metrics, namely accuracy, precision, recall and F1-score.

3.1 Data Acquisition

The dataset acquired from the Kaggle repository comprises 5,630 instances and 20 features. Table 1 provides a brief description of the datasett

Table 1: Dataset Description

S/N	Variable	Type	Number of missing values
1	CustomerID	Integer	Non
2	Churn	Integer	Non
3	Tenure	Float	264
4	PreferredLoginDevice	Object	Non
5	CityTier	Integer	Non
6	WarehouseToHome	Float	251
7	PreferredPaymentMode	Object	Non
8	Gender	Object	Non
9	HourSpendOnApp	Float	255
10	NumberOfDeviceRegistered	Integer	Non
11	PreferedOrderCat	Object	Non
12	SatisfactionScore	Integer	Non
13	MaritalStatus	Object	Non
14	NumberOfAddress	Integer	Non
15	Complain	Integer	Non
16	OrderAmountHikeFromlastYear	Float	265
17	CouponUsed	Float	256
18	OrderCount	Float	258
19	DaySinceLastOrder	Float	307
20	CashbackAmount	Float	Non

3.2 Data Preprocessing

The categorical features, labelled 'object' in Table 1, in the dataset were converted to numerical data using **Label Encoder.** Label Encoder assigns a unique integer value to each category in a dataset, enabling machine learning algorithms to process categorical data. Furthermore, the mean imputation technique was employed to fill in the missing values as they appeared in the features shown in Table 1.

Class imbalance poses a significant challenge in machine learning model development. It causes the training model to become biased toward the majority class, ultimately reducing its overall performance [30] An exploratory analysis conducted on the dataset also reveals that the dataset is not balanced across the target classes; thus, SMOTE is applied to balance the class distribution.

3.3 Feature Selection

Information gain, a feature selection method used in machine learning to evaluate the relevance of a feature in predicting the target variable, was employed to select important features to train and test the model.

3.4 Model Development

In developing the proposed robust customer churn prediction models, AdaBoost, GB, and XGB were configured using suitable parameters. Table 2 presents the key parameters used in each algorithm and the functional role of the parameter in each model construction

Table 2: Key Parameters Configuration for the Proposed Models

Algorithm	Parameters
Adaboost	n_estimators
	learning_rate
	algorithm
Gradient Boosting	n_estimators
	min_samples_split
	min_samples_leaf
	max_depth learning_rate
XGBoost	n_estimators max_depth
	learning_rate
	gamma
	subsample colsample_bytree
	reg_alpha reg_lambda

3.5 Parameter Optimization

Random Search is a hyperparameter tuning algorithm that randomly samples the hyperparameter space to find the optimal combination of hyperparameters. It's a simple and efficient method that can be used when the hyperparameter space is large or complex. Random search pseudocode is presented in figure 6.

1. Define the hyperparameter space
2. Set a budget for the number of iterations
3. For each iteration:
   i. Randomly sample a combination of hyperparameters
   ii. Train a model using the sampled hyperparameters
   iii. Evaluate the performance of the model
   iv. Store results
4. After the budget is exhausted
5. Train developed Boosting models

Figure 6: Random search pseudocode

Table 3: Hyperparameter tuning values

Algorithm	Parameter	Value
Adaboost+RandomSearch	estimator	AdaBoostClassifier()
	n_estimators	[50, 100, 200, 500]
	learning_rate	[0.1, 0.5, 1, 2]
	algorithm	['SAMME', 'SAMME.R']
	cv	5
	n_iter	10
	random_state	42
Gradient Boosting + RandomSearch	Estimator	GradientBoostingClassifier()
	n_estimators	[50, 100, 200, 500]
	learning_rate	[0.1, 0.5, 1, 2]
	max_depth	[3, 5, 7, 10]
	min_samples_split	[2, 5, 10]
	min_samples_leaf	[1, 5, 10]
	cv	5
	n_iter	10
	random_state	42
Extreme GB + Random Search	estimator	XGBClassifier()
	max_depth	[3, 5, 7, 10]
	learning_rate	[0.1, 0.5, 1, 2]
	n_estimators	[50, 100, 200, 500]
	gamma	[0, 0.1, 0.5, 1]
	subsample	[0.5, 0.8, 1]
	colsample_bytree	[0.5, 0.8, 1]
	reg_alpha	[0, 0.1, 0.5, 1]
	reg_lambda	[0, 0.1, 0.5, 1]
	cv	5
	n_iter	10
	random_state	42

3.6 Performance Evaluation

Accuracy determines the proportion of correctly predicted classes to all the samples that were analysed (Ahmad et al., 2023). Eq. (1) represents accuracy formula:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad \ldots\ldots\ldots (1)$$

Precision is the positive patterns that every predicted pattern in a positive class has correctly predicted (Ahmad et al., 2023). It can be determined using Eq (2)

$$Pre = \frac{TP}{TP + FP} \quad \ldots\ldots\ldots (2)$$

Recall is the percentage of positive patterns that are accurately identified is determined by the sensitivity or recall (Ahmad et al., 2023). The Eq (3) can be used to calculate recall:

$$Recall = \frac{TP}{TP + FN} \quad \ldots\ldots\ldots (3)$$

F1-score is the harmonic average of the recall and precision rates (Ahmad et al., 2023). It is determined by the formula in Eq (4).

$$F1 - score = 2 \times \frac{Pre \times Recall}{Pre + Recall} \quad \ldots\ldots\ldots (4)$$

where:

- TP denotes True Positive
- TN denotes True Negative
- FP denotes False Positive
- FN denotes False Negatives

4. Results and Discussion

4.1 Class balancing

The acquired dataset initially exhibited a significant class imbalance, with 4682 instances belonging to class 0 (churn) and only 948 instances belonging to class 1 (not churn). This disparity can lead to biased models that favor the majority class; therefore, it is addressed using SMOTE. Figures 7 and 8 reveal the class distribution before and after SMOTE application, respectively.

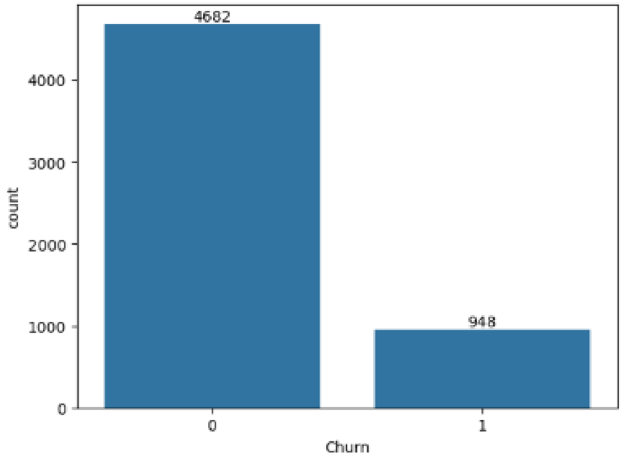

Figure 7: Class distribution before SMOTE

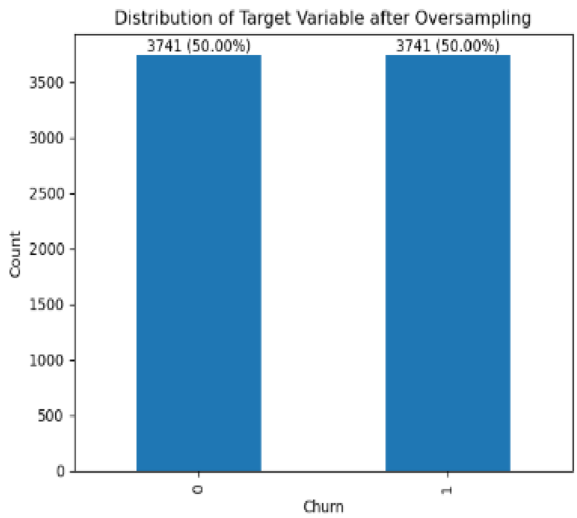

Figure 8: Class distribution after SMOTE

4.2 Important features

Figure 9 and 10 depicts the top 9 and 15 features of the dataset.

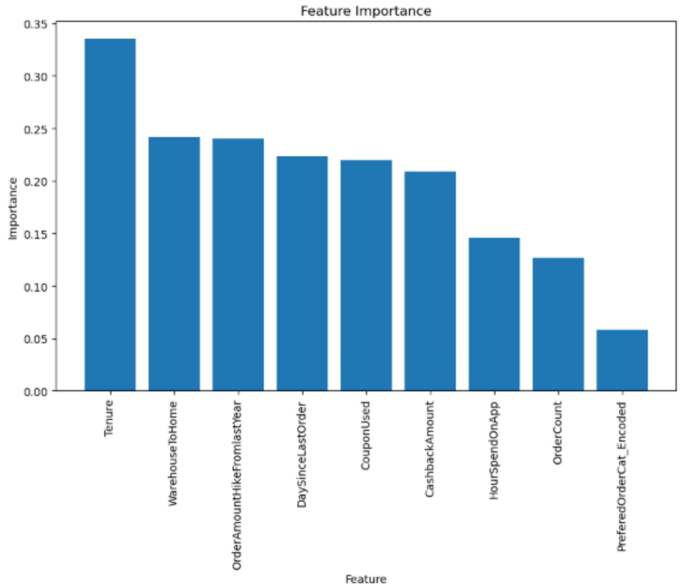

Figure 9: Top 9 features

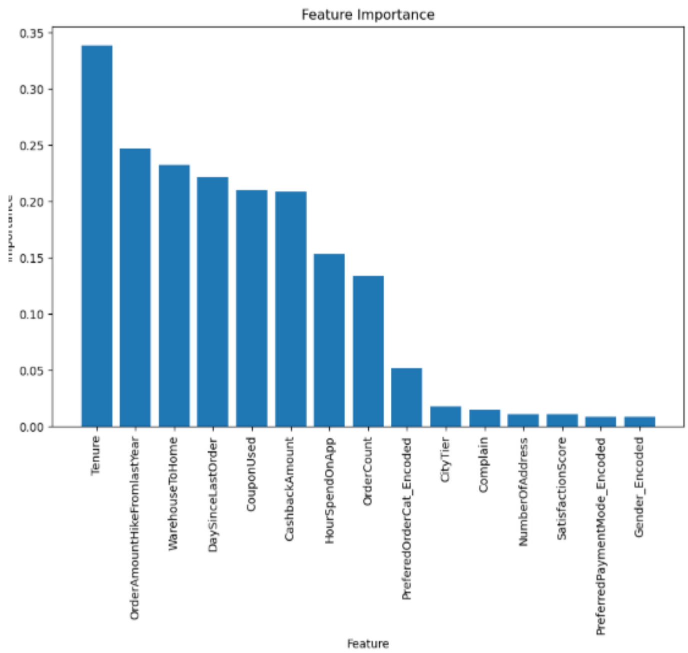

Figure 10: Top 15 features

Table 4: Feature Importance Rates

Feature Importance	Percentage	Importance
Tenure	0.338505	18.138722
OrderAmountHikeFromlastYear	0.246472	13.207145
WarehouseToHome	0.232475	12.457166
DaySinceLastOrder	0.221814	11.885872
CouponUsed	0.210096	11.257982
CashbackAmount	0.208305	11.161996
HourSpendOnApp	0.152761	8.185692
OrderCount	0.133744	7.166662
PreferedOrderCat_Encoded	0.051543	2.761905
CityTier	0.017918	0.960126
Complain	0.014237	0.762914
NumberOfAddress	0.010926	0.585482
SatisfactionScore	0.010621	0.569140
PreferredPaymentMode_Encoded	0.008646	0.463279
Gender_Encoded	0.008135	0.435919

Table 4 reveals that 'tenure' is most important feature while Gender is the least important features, considering top 15 features.

4.3 Performance evaluation

The models' results are presented in Table 5 based on the use of all features, top 9 features, and top 15 features. This study retains 15 top features because it maximizes the explanatory variance and minimizing the risk of discarding useful information, allowing models to capture more subtle churn patterns. Furthermore, this study's leverage on 15 features instead of all available features maintains computational efficiency, reducing complexity and overfitting risk.

Table 5: Evaluation result of AdaBoost, gradient boosting, and extreme gradient boosting (XGB) without hyperparameter tuning.

Algorithms/metrics	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Before feature selection				
Adaboost	87	77	80	78
Gradient Boosting	90	81	83	82
XG Boost	97	97	93	95
Using top 9 features				
Adaboost	84	73	77	74
Gradient Boosting	85	74	78	76
XG Boost	93	89	85	87
Using top 15 features				
Adaboost	87	77	81	79
Gradient Boosting	89	80	82	81
XG Boost	97	96	94	95

Figure 11-14 presents the summary of the Table 5.

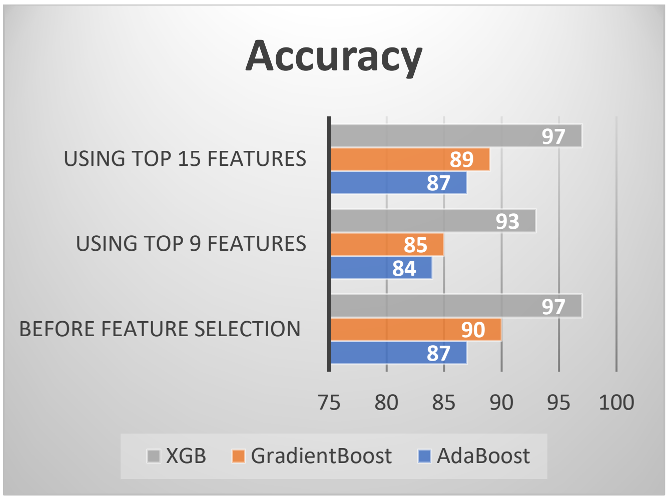

Figure 11: Accuracy scores of XGB, GradientBoost, and Adaboost

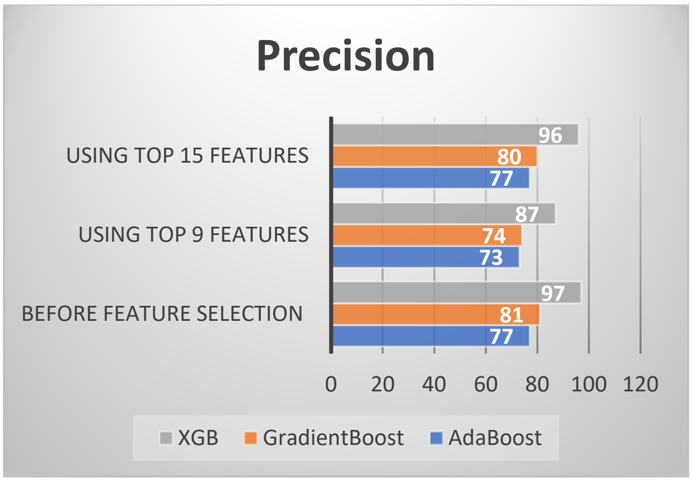

Figure 12: Precision scores of XGB, GradientBoost, and Adaboost

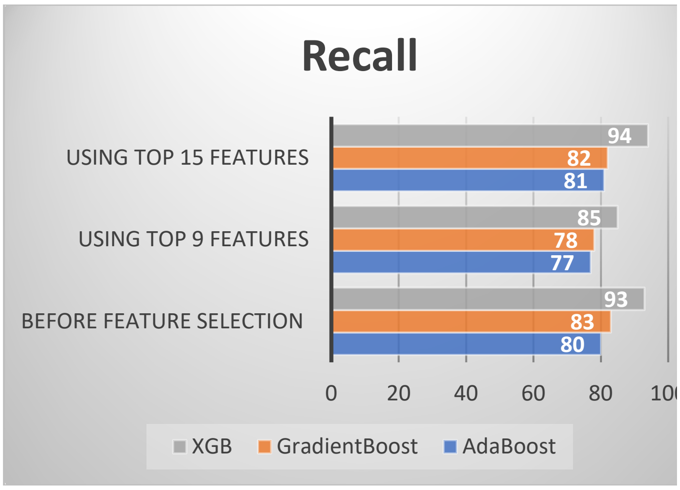

Figure 13: Recall scores of XGB, GradientBoost, and Adaboost

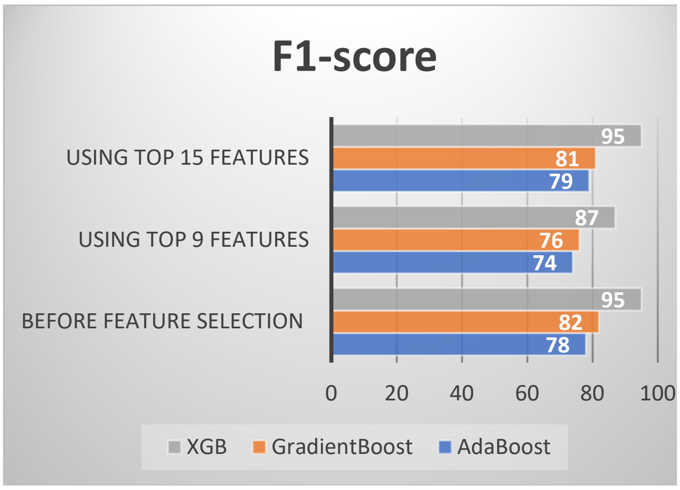

Figure 14: F1-score of XGB, GradientBoost, and Adaboost

When using the top 15 features, XGBoost significantly outperforms both Adaboost and Gradient Boosting across all evaluated metrics with the highest accuracy (97%), precision (96%), recall (94%), and F1-score (95%). Gradient Boosting follows with an accuracy of 89%, precision of 80%, recall of 82%, and an F1-score of 81%, while Adaboost records slightly lower but still competitive scores with 87% accuracy, 77% precision, 81% recall, and a 79% F1-score. These results underscore XGBoost's superior predictive performance and robustness, making it the most reliable choice for customer churn prediction when leveraging the optimal set of 15 features.

Table 6: Best Parameters for Each Algorithm

Algorithm	Parameter	Values
Adaboost	n_estimators	100
	learning_rate	1
	algorithm	SAMME.R
Gradient Boosting	n_estimators	500
	min_samples_split	5
	min_samples_leaf	1
	max_depth	10
	learning_rate	0.5
XGBoost	n_estimators	500
	max_depth	10
	learning_rate	0.5
	gamma	0.1
	subsample	1
	colsample_bytree	0.8
	reg_alpha	0
	reg_lambda	0

The results for the top 15 features have been further optimized using Random Search hyperparameter tuning. The post-optimization performance metrics and results are presented in Table 7.

Table 7: Evaluation result of AdaBoost, gradient boosting and extreme gradient boosting (XGBoost) with hyperparameter tuning (Random Search) and 15 features

Algorithms/metrics	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Adaboost	87	77	81	79
Gradient Boosting	97	97	93	95
XG Boost	97	94	93	94

Table 8: Comparison of the models' result before and after tuning using 15 important features

Algorithms / metrics	**Accuracy (%)**		**Precision (%)**		**Recall (%)**		**F1-score (%)**	
	Before	After	Before	After	Before	After	Before	After
Adaboost	87	87	77	77	81	81	79	79
GB	89	97	80	97	82	93	81	95
XGB	97	97	96	94	94	93	95	94

The comparison of results before and after tuning using 15 important features reveals notable improvements for gradient boosting. After tuning, Gradient Boosting's accuracy increases from 89% to 97%, precision from 80% to 97%, recall from 82% to 93%, and F1-score from 81% to 95%. AdaBoost shows no change, with identical scores before and after tuning. XGBoost maintains high performance, with slight decreases in precision (96% to 94%), recall (94% to 93%), and F1-score (95% to 94%), while accuracy remains consistent. These results indicate that tuning significantly enhances gradient boosting's performance, making it comparable to XGBoost. Therefore, hyperparameter tuning can significantly improve the performance of machine learning models, as evident in the substantial boost in gradient boosting's accuracy, precision, recall, and F1-score after tuning.

Figure 15 describes how the AdaBoost model made predictions using SHapley Additive exPlanations (SHAP) analysis. The figure reveals that the predicted number of ‘churn’ is 0, while the average prediction is 0.5. Moreover, the figure reveals that only the ‘complain’ feature contributes to increasing the predicted churn by 0.04, denoting a positive impact on the model's prediction, while other features contribute to decreasing the predicted churn, implying that the features have a negative impact on the model's prediction.

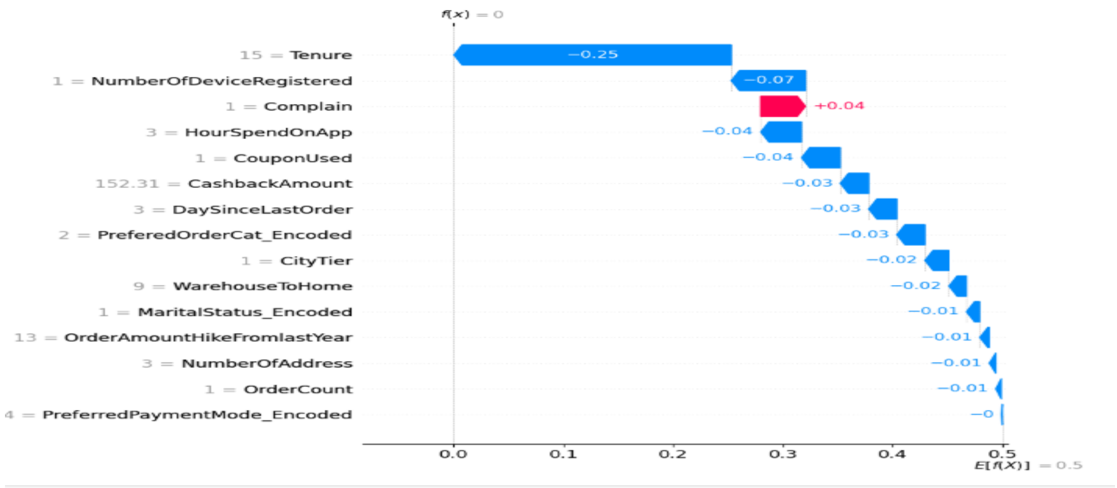

Figure 15: SHAP Analysis for Adaboost model

Figure 16 describes how Gradient Boosting model made prediction using SHapley Additive exPlanations (SHAP) analysis. The figure reveals that the predicted number of ‘churn’ is 0 while the average prediction is 0.51. Furthermore, the figure shows that only ‘complain’ feature contributes to increasing the predicted churn by 0.03, denoting a positive impact on the model's prediction, while other features contributes to decreasing the predicted churn, implying that the features has a negative impact on the model's prediction.

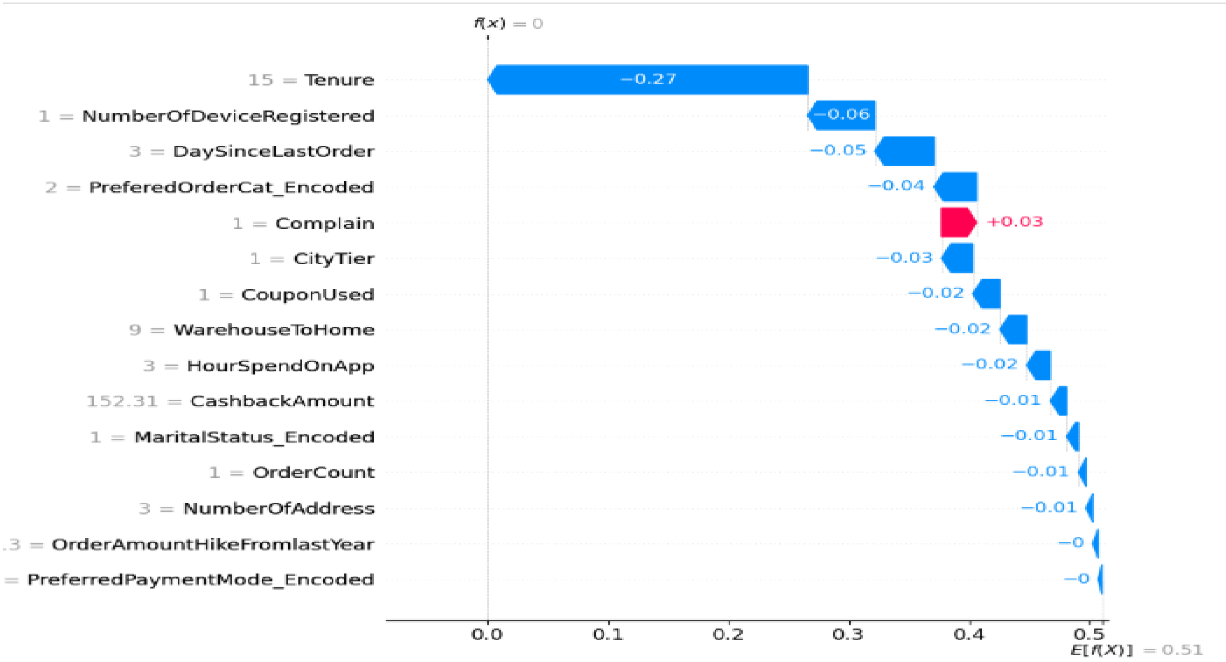

Figure 16: SHAP Analysis for Gradient Boosting model

Figure 17 describes how the XGB model made predictions using SHapley Additive exPlanations (SHAP) analysis. The figure reveals that the predicted number of 'churn' is -0 while the average prediction is 0.49. The figure, however, shows that only the 'complain' feature contributes to increasing the predicted churn by 0.02, denoting a positive impact on the model's prediction, while other features contribute to decreasing the predicted churn, implying that the features have a negative impact on the model's prediction.

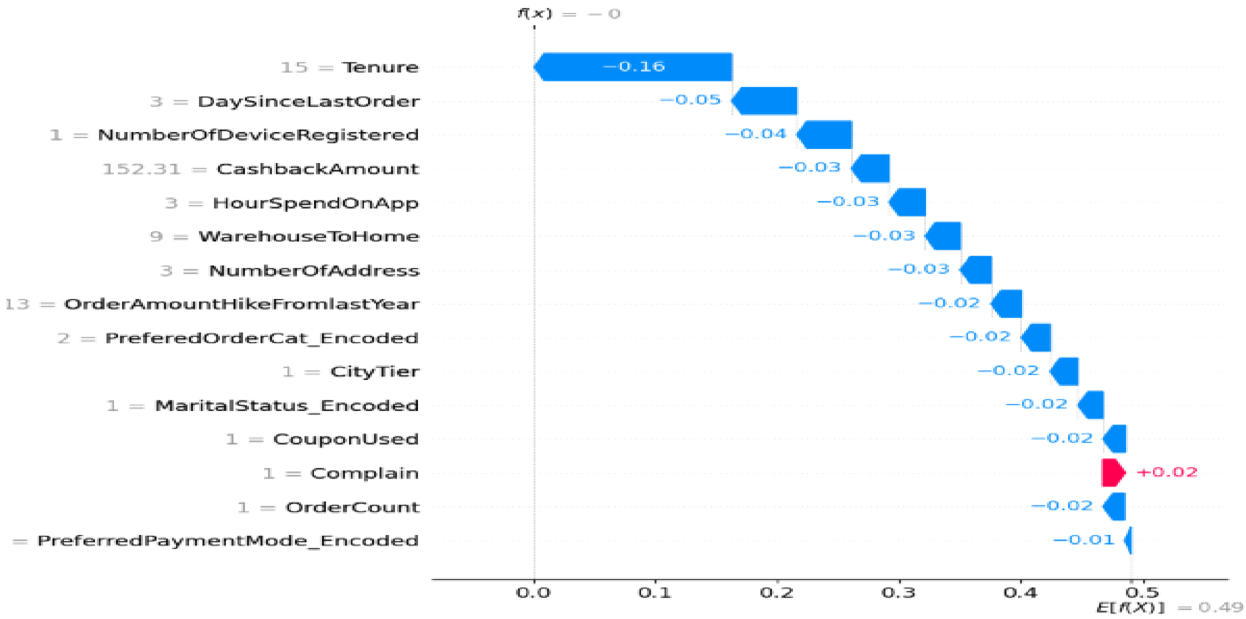

Figure 17: SHAP Analysis for Gradient Boosting model

Overall, the SHPA analysis highlights the importance of the "complain" feature in predicting customer churn. Additionally, the 'tenure' feature which has a strong negative impact on the prediction across the three models implies that longer tenure logically means less likely to churn, which makes the model's behaviour intuitive and trustworthy.

Table 9: Comparison of the models' results with existing studies

Author(s) & Year	Methodology	Accuracy	Precision	Recall	F1-score
[6] (2024)	Gradient Boosting Recursive Feature Elimination	79.86%	-	-	-
[26] (2024)	XGBoost & Random Forest	96%	96%	96%	96%
[5] (2024)	XGBoost	72.88%	-	-	-
[8] (2023)	XGBoost	87%	-	-	-
[7] (2023)	Stochastic Gradient Booster	83.9%	-	-	-
Current Study	AdaBoost + Random Search + Information Gain	87%	77%	81%	79%
	Gradient Boosting + Random Search + Information Gain	97%	97%	93%	95%
	XGBoost + Random Search + Information Gain	97%	94%	93%	94%

A comparison with existing studies reveals that the current study's Gradient Boosting and XGBoost models outperform most previous studies. [30] reported similar performance metrics for XGBoost and Random Forest, with an accuracy of 96%, precision of 96%, recall of 96%, and F1-score of 96%. The current study's optimized XGBoost model achieves comparable results, with an accuracy of 97%, precision of 96%, recall of 94%, and F1-score of 95%. This shows that the current model (XGBoost) outperformed existing studies ([30]; [5]; [8] in terms of accuracy (97%). The optimized Gradient Boosting model in the current study also achieves high performance metrics, with an accuracy of 89%, precision of 80%, recall of 82%, and F1-score of 81%, outperforming existing studies [6]; [7].

Notably, the use of Random Search and Information Gain in the current study appears to have a positive impact on the performance of the Gradient Boosting and XGBoost models. These findings suggest that the combination of ensemble methods with feature selection and hyperparameter tuning can lead to improved performance in customer churn prediction tasks.

5. Conclusion and Recommendations

This study demonstrates the effectiveness of ensemble methods, specifically AdaBoost, Gradient Boosting, and XGBoost, in customer churn prediction. The use of random search and information gain for feature selection and hyperparameter tuning significantly improves the performance of these models. The results show that Gradient Boosting and XGBoost achieve high accuracy, precision, recall, and F1-score, outperforming AdaBoost. These findings suggest that ensemble methods with proper feature selection and hyperparameter tuning can be a valuable tool for businesses to predict customer churn and develop targeted retention strategies.

Consequently, future studies can investigate whether Random Forest (a bagging algorithm) can achieve comparable performance to XGBoost with less hyperparameter sensitivity. Additionally, investigating the impact of different datasets and industries on the performance of these models can provide valuable insights. Researchers can also consider using LIME to generate local, instance-level explanations and comparing them with SHAP's more game-theory-grounded approach to see if they provide different insights for actionable retention strategies.

Data Availability: The dataset used for this paper are accessible through the link below:

https://www.kaggle.com/code/wonderdavid/e-commerce-customer-churn-prediction

References

[1] Huang, X., Wu, Y., Zhang, L., & Lyu, J. A survey on churn analysis: How and why customers leave. *Knowledge-Based Systems.* 2020; 188.

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

[3] Nasiri, A., & Gharebagh, F. H. A hybrid machine learning approach for customer churn prediction in telecommunication sector. *Telecommunication Systems*. 2019; 71(3): 399-411.

[4] Karaarslan, H., Baştuğ, M., Şen, C. G., & Işik, E. E. A Comparative Study on Customer Churn Analysis Using Machine Learning and Data Enrichment Techniques. *Journal of Soft Computing and Decision Analytics*; 2024; *2*(1): 225-235.

[5] Sikri, A., Jameel, R., Idrees, S. M., & Kaur, H. Enhancing customer retention in telecom industry with machine learning driven churn prediction. *Scientific Reports*. 2024; *14*(1): 13097.

[6] Kusnawi, K., Ipmawati, J., Asadulloh, B. P., Aminuddin, A., Abdulloh, F. F., & Rahardi, M. Leveraging various feature selection methods for churn prediction using various machine learning algorithms. *JOIV: International Journal on Informatics Visualization*. 2024; *8*(2): 897-905.

[7] Prabadevi, B., Shalini, R., & Kavitha, B. R. Customer churning analysis using machine learning algorithms. *International Journal of Intelligent Networks.* 2023; *4:* 145-154.

[8] Lukita, C., Bakti, L. D., Rusilowati, U., Sutarman, A., & Rahardja, U. Predictive and analytics using data mining and machine learning for customer churn prediction. *Journal of Applied Data Sciences.* 2023; *4*(4): 454-465.

[9] Yamini, B., Ramana, K. V., Nalini, M., Devi, D. C., Maheswari, B., & Siva, S. R. Customer churn prediction model in enterprises using machine learning. *International Journal of Advanced Technology and Engineering Exploration.* 2024; *11*(110): 94.

[10] Matuszelański, K., & Kopczewska, K. Customer churn in retail e-commerce business: Spatial and machine learning approach. *Journal of Theoretical and Applied Electronic Commerce Research.* 2022; *17*(1): 165-198.

[11] Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, J. R., Ambadekar, S. P., & Gawande, S. H. Customer churn prediction in telecom sector using machine learning techniques. *Results in Control and Optimization.* 2024; *14*.

[12] Khare, P., & Arora, S. Predicting customer churn in saas products using machine learning. *International Research Journal of Engineering and Technology (IRJET).* 2024; *11*(5): 754-765.

[13] Seymen, O. F., Ölmez, E., Doğan, O., Er, O., & Hiziroğlu, K. Customer churn prediction using ordinary artificial neural network and convolutional neural network algorithms: A comparative performance assessment. *Gazi University Journal of Science*. 2023; 36(2): 720-733.

[14] Pondel, M., Wuczyński, M., Gryncewicz, W., Łysik, Ł., Hernes, M., Rot, A., & Kozina, A. Deep learning for customer churn prediction in e-commerce decision support. In *Business Information Systems.* 2021; 3-12.

[15] Sam, G., Asuquo, P., & Stephen, B. Customer churn prediction using machine learning models. *J Eng Res Rep.* 2024; *26*(2): 181-193.

[16] Li, A., Yang, T., Zhan, X., Shi, Y., & Li, H. Utilizing Data Science and AI for Customer Churn Prediction in Marketing. *Journal of Theory and Practice of Engineering Science.*2024; *4*(05): 72-79.

[17] Gurung, N., Hasan, M. R., Gazi, M. S., & Chowdhury, F. R. AI-Based Customer Churn Prediction Model for Business Markets in the USA: Exploring the Use of AI and Machine Learning Technologies in Preventing Customer Churn. *Journal of Computer Science and Technology Studies.* 2024; *6*(2): 19-2.

[18] He, C., & Ding, C. H. A novel classification algorithm for customer churn prediction based on hybrid Ensemble-Fusion model. *Scientific Reports.* 2024; *14*(1): 20179.

[19] Maduna, M., Telukdarie, A., Munien, I., Onkonkwo, U., & Vermeulen, A. Smart Customer Churn Management System Using Machine Learning. *Procedia Computer Science.* 2024; *237*: 552-558.

[20] Vo, N. N., Liu, S., Li, X., & Xu, G. Leveraging unstructured call log data for customer churn prediction. *Knowledge-Based Systems.* 2021; *212:* 106586.

[21] Hoppner, S. Stripling, E. Baesens, B. Broucke, S. & Verdonck, T. Profit driven decison trees for churn prediction. European Journal of Operational Research. 2020; 284(3): 920 – 933.

[22] Sook-Ling, L., Mustafa, N., & Abdul-Razak, S.F. Customer churn prediction for telecommunication industry: A Malaysian Case Study. F1000Research. 2021; 10(1): 1-13, doi: 10.12688/f1000research.73597.1.

[23] Tayarani, M. Applications of artificial intelligence in battling against covid-19: A literature review, Chaos, Solitons & Fractals. 2021; 142(1): 110338-110345, doi: https://doi.org/10.1016/j.chaos.2020.110338.

[24] Khan, S., Iqbal, K., Faizullah, S., Fahad, M., Ali, J., & Ahmed, W. Clustering based privacy preserving of big data using fuzzification and anonymization operation. *arXiv preprint arXiv.* 2020; *2001.01491*.

[25] Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., & Faruki, P. Network Intrusion Detection for IoT Security Based on Learning Techniques. IEEE Communications Surveys Tutorials. 2019; 21(3): 2671–2701

[26] Al-Najjar, D., Al-Rousan, N., & Al-Najjar, H. Machine learning to develop credit card customer churn prediction. *Journal of Theoretical and applied electronic commerce research.* 2022; *17*(4): 1529-1542.

[27] Alotaibi, M. Z., & Haq, M. A. Customer Churn Prediction for Telecommunication Companies using Machine Learning and Ensemble Methods. *Engineering, Technology & Applied Science Research.* 2024; *14*(3): 14572-14578.

[28] Ahmad, N., Awan, M. J., Nobanee, H., Zain, A. M., Naseem, A., & Mahmoud, A. Customer personality analysis for churn prediction using hybrid ensemble models and class balancing techniques. *IEEE Access.* 2023; *12*: 1865-1879.

[29] Dhand, G., Rao, M., Chaudhary, P., & Sheoran, K. (2025). A secure routing and malicious node detection in mobile Ad hoc network using trust value evaluation with improved XGBoost mechanism. *Journal of Network and Computer Applications*, *235*, 104093.

[30] Sadar, U., Agarwal, P., Parveen, S., Dhand, G., & Sheoran, K. Heart disease prediction using machine learning classifiers with various balancing techniques. *Proceedings on Engineering Sciences*, 2024; 6(4), 1871–1878. ISSN 2620-2832.