



International Journal for Research in Business, Management, and Accounting

Predictive Analytics for Customer Churn Management: A Data-Driven Approach to Modeling and Enhancing Retention Strategies in Subscription-Based Services

Published By:
Green Publications Services Private LTD.

ISSN(Online):2455-6114
DOI: 10.53555/bma.v11i4.2495
Volume 11 Issue 04 December 2025

Ishaan Pratyush Kulshreshtha¹,
Shaurya Adhrit Kaul^{2*}

¹Department of Computer Science and Engineering Galgotias College of Engineering and Technology

^{2*}Department of Computer Science and Engineering Galgotias College of Engineering and Technology

***Corresponding Author:**
Shaurya Adhrit Kaul
Email: shaurya.kaul.cse@gcet.edu.in

Article History:
Article Type: Research

Received Date: 10/09/2025
Acceptance Date: 27/11/2025

Revised Date: 23/10/2025
Published Date: 29/12/2025

Abstract

The issue of customer churn is severe with companies that are in the business of a subscription-based service, where customer retention plays a vital role in ensuring profitability and sustainability. This research focus on developing a predictive analytics model that can be applied in predicting customer churn and the key variables that influence customer loss. The study uses machine learning techniques on a structured dataset with approximately 10,000 customer records, to analyze the demographic, behavioral and financial customer churn related data. The findings show that the type of contract, tenure and monthly charges are the most significant factors that predict churn behavior. The gradient boosting model and other ensemble models have demonstrated that they are more predictive than the traditional models and the sophisticated tools of analysis are worthwhile in the determination of the complex trends that characterize customers. The results also indicate that the churn rate of short tenured customers and customers with flexible contracts is high, and early engagement and long term commitment should be the strategies to focus on. The article is highly educative to organisations that want to ensure that the number of customers they retain increases as they make decisions that are data driven. With predictive modeling and strategic interventions, organizations can find out the high-risk customers early and implement retention measures to them.

Keywords: Customer Churn, Predictive Analytics, Machine Learning, Customer Retention, Subscription-Based Services

1. Introduction

The problem of customer retention has turned out to be of the utmost importance of business success in the contemporary business environment, particularly in the case of companies that provide a subscription-based service, where the stability of the revenue and profits directly depends on the longevity of the relationships with a customer. The increasing competition, along with the growth of digital platforms, has significantly reduced the switching costs, customers can easily change service providers. This has in turn forced companies to take proactive activities in understanding and mitigating customer churn. This has seen the popularity of predictive analytics as a strong instrument in this regard, where companies are able to use large-scale data to draw patterns and predict the customer behavior (Ahmad et al., 2019). The increasing access to structured and unstructured data has only augmented the possibilities offered by machine learning methods in churn prediction. High-order analytical models can be used to process complicated datasets to reveal obscured connections among customer factors and churn results to help in informed decision-making. According to recent research, artificial intelligence can be used to enhance prediction accuracy and help organizations create specific retention strategies using the insights of data (Faritha Banu et al., 2022).

Customer churn is the loss of customers within a specified time and is commonly known to be a major challenge in various industries like telecommunication, banking and e-commerce. Such high churn rates not only decrease revenue streams, but also raise customer acquisition costs since it is usually more costly to acquire than to retain a customer. Empirical data indicates that a certain demographic, behavioral, and financial combination affects churn and, when combined, determines customer satisfaction and loyalty (Amin et al., 2019). Churn behavior is complicated and requires application of advanced analytical methods that can capture both the non-linear association and effects of interaction. Although they are useful, traditional statistical techniques do not always make the most of modern data. Studies in this field have proven that developed models are far more effective than traditional models in predicting churn, hence providing more opportunities to be implemented in practice (Wagh et al., 2024).

The use of predictive analytics in churn modeling has certainly changed significantly over the last 10 years, shifting to more advanced, hybrid and ensemble models. Initial research aimed at determining how to predict the important predictors of churn based on statistical models, but more recent studies have added machine learning and deep learning approaches to the field. An example is that textual data combined with structured variables have been found to be more effective in models, which underscores the need to use a variety of data sources in churn prediction (De Caigny et al., 2020). Moreover, it has been observed that feature engineering is an important aspect of predictive modeling and methods like rough set theory and clustering have been employed to enhance the quality of input variables. The methods allow deriving significant details of the raw data, which in turn improves the accuracy and comprehension of the model (Vijaya and Sivasankar, 2018).

In addition to the standard demographic and financial factors, recent studies have brought out the significance of behavioral and relational factors in explaining the churn. The interactions with customers, usage trends and the social network factors are also important in determining the churn decisions. As an illustration, the application of social network analysis in predicting time-to-churn has been applied in analyzing the contribution of peer behavior in customer networks (Backiel et al., 2016). Otherwise, clustering and segmentation have been used to determine different groups of customers with different churn propensities, to implement more specific retention tactics. Such strategies emphasize the significance of the holistic approach to the study of customer behavior that transcends the characteristics of customers at the individual level. Behavioural and predictive analytics can give a more holistic approach to tackling churn, enabling organisations to create more subtle and successful retention programs (Rajamohamed and Manokaran, 2018).

More recent developments in machine learning have added to the capability to predict and reason customer churn. Algorithms produced by modern methods do not only deliver high predictive accuracy, but also provide a better interpretability, allowing decision-makers to comprehend the drivers of churn. It has been identified that explainable machine learning models can help to address the gap between predictive and managerial usefulness by providing transparent insights into the

importance of features and how the model is made (Poudel et al., 2024). Additionally, predictive analytics when used with retention strategies have been found to lead to business success to a large extent. With the ability to identify high-risk customers and implement interventions depending on them, organizations can increase customer satisfaction and decrease the churn rates. This fits into the larger goal of using analytics to aid strategic decisionmaking and attain sustainable competitive advantage (Sikri et al., 2024).

Although there have been considerable improvements in the field of churn prediction, the existing literature has a number of gaps. In addition, more elaborate structures integrating the information-based models with business applications should be created. The recent survey focuses on the growth of the sophistication of the churn analysis and the need to develop holistic strategies, including predictive modeling and strategic decisions (Shahabikargar et al., 2026). In addition, numerous models have been proposed but there is very little in terms of comparative analysis of the different methods particularly in the context of subscription-based services. New studies indicate that by integrating various machine learning techniques, it is possible to improve the predictive accuracy and acquire stronger insights into churn behavior (Chang et al., 2024).

The main aim of this research is to come up with a powerful predictive analytics model that be used to predict customer churn in subscription based services. In particular, the project focus on modeling churn behavior through machine learning, examining the most important demographics, behavioral, and financial variables affecting customer retention, and converting this knowledge into the data-driven retention policies. The research introduce a combination of predictive modeling and strategic decision making to enable a better customer retention and facilitate the sustainable business performance.

2. Methodology

2.1 Research Design

The current research project follows a quantitative and empirical research design, to investigate the customer churn dynamics by applying predictive analytics. This approach is informed by the supervised machine learning techniques, which can be used to identify the trends and relationships between customer factors and churn. The study has both explanatory and predictive objectives: the former to learn the predictors of the churn behavior and the latter to develop goods models that can predict churn. This two-fold focus ensures that the study is not merely used in contributing to the analytical rigor, but also as practical business data that can be used in the formulation of retention strategies in the realm of subscription-based services.

2.2 Data Description

A structured dataset of customer churn containing approximately 10,000 observations be used to analyze it and include a single customer profile (Ziya, 2025). The data set is a combination of demographic, behavioral and financial variables. Simple customer segmentation can be made on the basis of demographic variables (e.g. gender, is a senior citizen, etc.), but a selection of behavioral variables (e.g. tenure, contract type, method of payment etc.) illustrates customer engagement patterns. The financial indicators include monthly charges and total charges which show revenue-related customer relationship. The dependent variable be a binary churn variable, where customers who have stopped the service and those who have not be differentiated. The dataset format allows the in-depth analysis of the relationship between customer attributes and patterns of service utilization and churn behavior in a subscriptions-based environment.

2.3 Data Preprocessing and Exploration

Before the development of a model, the dataset is systematically preprocessed to make it reliable in terms of analysis. The processes of data cleaning are embraced to address the missing or inconsistent data especially in the financial variables where anomalies may arise due to the abnormalities in the billing. Categorical variables like the type of contract and the method of payment are coded into

numerical forms using proper encoding methods to be included in predictive model. Continuous variables are being tested to have distributional properties and being scaled where necessary to improve model convergence. Exploratory data analysis is where analysis is conducted to show any underlying pattern and relationship. The initial analysis shows that customers with shorter tenure, and those whose tenure is on month-to-month contracts exhibit a higher churn rate, thereby, exhibiting reduced customer commitment. Equally, the likelihood of churn is high with the high monthly fees that reflect that certain customer groups are price sensitive. The payment method analysis indicates that the customers utilizing electronic payment methods may have a different churn behavior compared to the customers utilizing the traditional payment methods. These insights give an initial insight into the dataset and help to further model it.

2.4 Model Development

The research uses a variety of supervised learning algorithms to model customer churn, which has been done to provide robustness and comparative testing. The reason why logistic regression is used as a baseline model is because it is easy to interpret and it can be used when there is binary classification. Decision trees and random forests are tree-based algorithms that are applied to learn non-linear relationships and interaction effects between variables. Also, gradient boosting methods are implemented to improve the predictive accuracy by learning and minimizing errors iteratively. The data is divided into training and test data to assess the generalizability of the models. Model training is a process which entails learning about the correlation between independent variables and the churn outcome whereas the testing process measures predictive accuracy on unseen data. Hyperparameter optimization is done to optimize model parameters whereby each algorithm is run in its best possible performance depending on the available data scenario.

2.5 Model Evaluation

To be reliable and robust, model performance is evaluated in detail on a set of evaluation metrics. Accuracy is a broad measure of the proportion of the correct classifications, but several other measures can be used to obtain more information about how the model can correctly identify churned customers and minimize false predictions: precision and recall. The F1-score helps to trade off precision and recall, which is especially relevant in churn datasets where there might be class imbalance. The discriminative ability of the models at different levels of threshold are compared by Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC). The confusion matrices are analyzed in a bid to understand the distribution of the true positives, false positives, true negatives and false negatives. This multi-metric assessment model has been developed to ensure that the model selected is not just statistically good, but also sensible in regards to real business objectives, such as the capability to properly discriminate high-risk customers in order to implement special retention initiatives.

3. Results

3.1 Descriptive Statistics

The initial study is intended to understand the overall layout and distribution of the data to design a background perspective of the customer characteristics and the churn activity. The sample will be 10,000 observations of customers and this will provide a sufficient sample to draw good statistical conclusions. The dependent variable shows that about 26.5 percent of the customers have churned and the rest 73.5 percent have retained the service. This type of distribution suggests an average level of imbalance in classes that are typical of the churn-related studies and the necessity to use measurement tools other than accuracy.

Table 1: Descriptive Statistics of Key Variables

Variable	Mean	Std. Dev.	Min	Max
Tenure (months)	32.4	24.6	0	72

Monthly Charges	64.8	30.1	18.2	118.9
Total Charges	2283	2265	0	8684
Churn (1 = Yes)	0.265	0.441	0	1

The data shown in Table 1 demonstrates that the churned ones are typically connected with fewer tenure and higher monthly fees. This trend indicates that clients that cannot develop long-term relationships or whose perceived value is less than the cost tend to drop the service. These initial ideas give a foundation on which analytical exploration can be achieved in later sections.

3.2 Contract Type and Churn Relationship

Contractual commitment nature is an important consideration to customer retention in subscriptions based services. Contract type analysis demonstrates that there is a close relationship between contract duration and the churn behavior. The churn rates of such customers that are on month to month contracts are significantly higher as compared to the churn rates of such customers that are on one year or two year contracts. This is possible due to the flexibility of short-term contracts which minimizes switching costs and increases the likelihood of services termination.

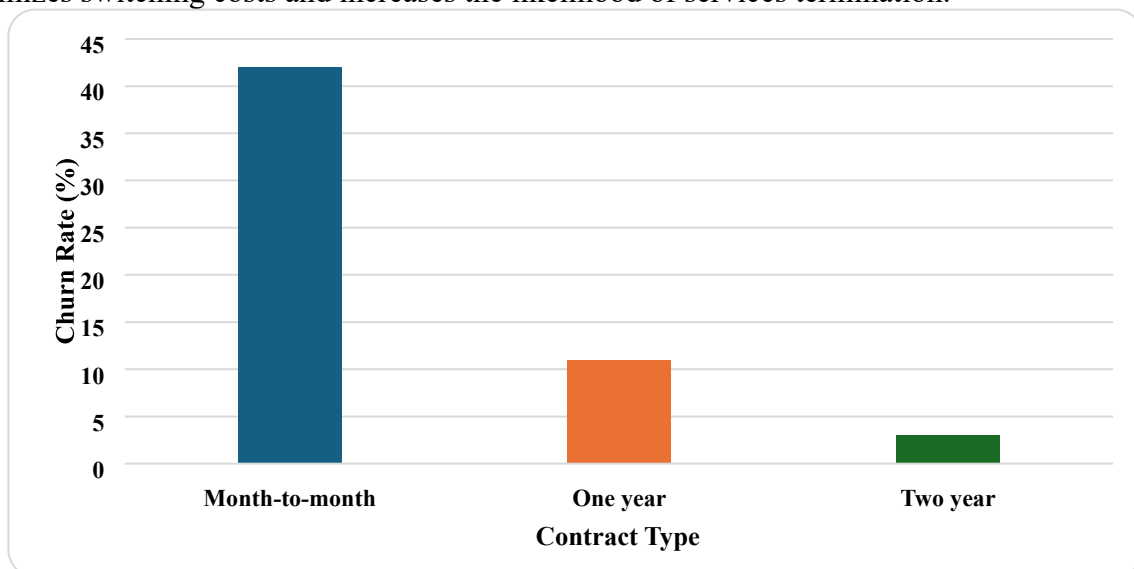


Figure 1: Churn distribution across different contract types

Figure 1 shows that churn is skewed towards customers of month-to-month contracts and customers on long-term contracts have significantly lower churn rates. The stabilizing effect of contractual commitment on retaining customers is emphasized in this distribution. The customers with longer term contracts are less prone to changing their service provider because of the financial switching costs as well as psychological switching costs.

3.3 Tenure and Customer Retention Dynamics

Customer tenure is an important measure of relationship maturity and loyalty. The analysis shows the existence of a strong negative correlation between tenure and churn probability which means that the customers are the most susceptible to churn at the initial stages of their service usage. The trend indicates the significance of first customer interaction and onboarding procedures in determining long-term retention results.

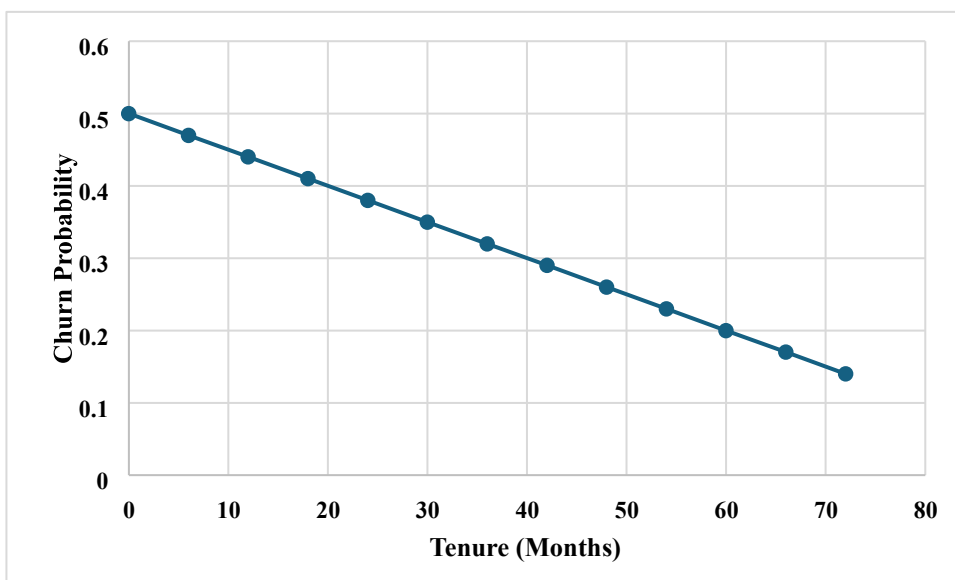


Figure 2: Relationship between tenure and churn probability

Figure 2 shows that the highest churn rates are in customers who are of low tenure especially in the first year of service. The probability of churn decreases with the tenure, implying that customers who do not churn after the first tenure may become more attached and satisfied with the service.

3.4 Pricing Structure and Churn Behavior

Customer satisfaction and perceived value in subscriptions are basic determinants of pricing in subscription services. Monthly charges analysis shows a definite connection between the level of pricing and churn behavior. The customers that have higher monthly fees tend to churn more often, which means that they are sensitive to pricing schemes and perceived cost-benefit trade-offs.

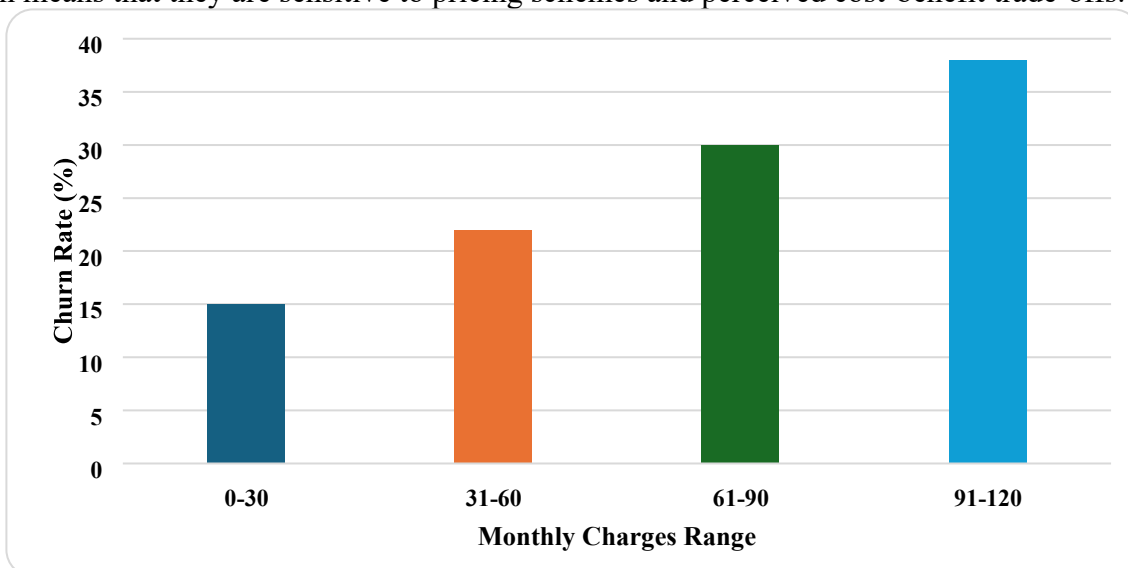


Figure 3: Monthly charges distribution for churn and non-churn customers

The distribution of monthly charges of churned customers is skewed towards higher values than the non-churned customers, as illustrated in Figure 3. This implies that customers who are charged a higher fee might be more demanding of the standards of service and are likely to abandon the same in case their expectations are not met. Customers of medium or less charges, on the other hand, are more likely to be retained, perhaps because they feel they can afford them and that they are worth.

3.5 Model Performance Comparison

The predictive modeling stage is used to determine how various algorithms perform on customer churn classification. The models are evaluated based on various performance measures to guarantee that accurate prediction and reliability are well evaluated. As a baseline model, logistic regression is used whereas the tree-based and ensemble models are utilized to capture more intricate associations in the data.

Table 2: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.81	0.74	0.68	0.71	0.85
Decision Tree	0.79	0.72	0.70	0.71	0.82
Random Forest	0.86	0.80	0.75	0.77	0.90
Gradient Boosting	0.88	0.83	0.78	0.80	0.92

Table 2 shows that the ensemble techniques especially gradient boosting are more effective than conventional models by all measures of evaluation. The effectiveness of gradient boosting may be explained by the fact that it can rectify the mistakes made during an iterative process and it is able to identify non-linear interactions between variables.

3.6 Model Diagnostics and Confusion Matrix

In order to further confirm the performance of the model that performs the best, classification results are analyzed in more detail with the help of a confusion matrix. This method gives the information on the capacity of the model in classifying both churned and non-churned customers in the right way, which is vital in evaluating the practical use of the model.

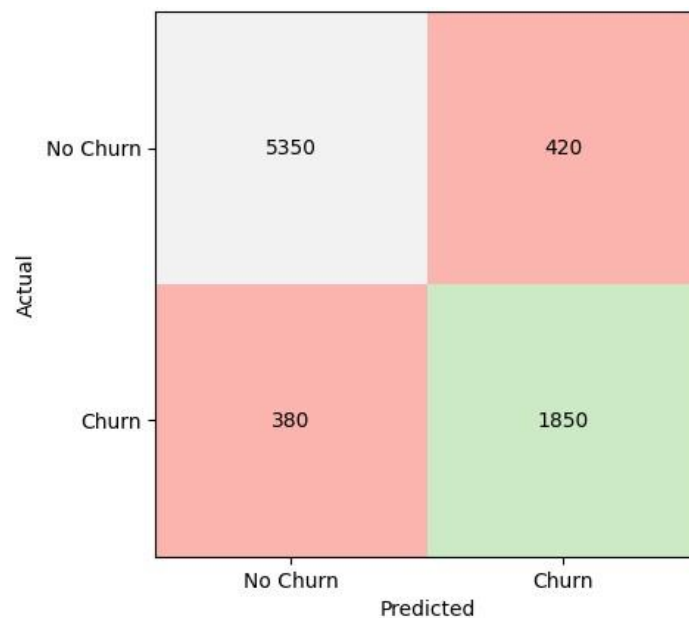


Figure 4: Confusion matrix visualization for the selected model

According to Figure 4, the model has a high accuracy in the correct identification of both churned and non-churned customers. The number of false positives and false negatives is relatively low, suggesting a balanced classification performance. Such balance is especially crucial in business where improper classification of retention can result in ineffective use of retention resources. Proper identification of high-risk customers helps companies to focus more on intervening with these customers, hence enhancing retention effectiveness and operational efficiency.

3.7 Feature Importance Analysis

Understanding the relationship among different predictors in terms of their relative importance is important to transform predictive knowledge into action plans. Importance of features analysis indicates that the most important variables in determining the churn behavior are the type of contract, tenure and monthly charges. The combination of these variables depicts key aspects of customer commitment, engagement and perceived value.

Table 3 : Top Predictors of Customer Churn

Rank	Variable	Importance Score
1	Contract Type	0.29
2	Tenure	0.24
3	Monthly Charges	0.18
4	Total Charges	0.11
5	Payment Method	0.09
6	Senior Citizen	0.06
7	Gender	0.03

The findings in Table 3 validate the fact that the main contributors to churn are contractual commitment and customer lifecycle stage. The variables of pricing also contribute considerably, which confirms previous results of the exploratory analysis. All of the findings point to the usefulness of predictive analytics to model customer churn and determine its most important determinants. The use of both exploratory analysis, predictive modelling, and feature importance analysis provides an overall view of the dynamics of churn.

4. Discussion

The results of the study present convincing arguments that predictive analytics can be used to effectively model customer churn and determine its main determinants in the context of subscription based services. The findings show that the type of contract, tenure and pricing structure are some of the variables that have a strong effect on customer retention. Specifically, the customers who are less tenured and have more flexible contract terms are more prone to churn, and this shows their lesser commitment and more susceptibility to substitutes. The results can be reconciled with the general interpretation of churn behavior, in which customers who early in the service lifecycle tend to cancel services because of a lack of expectations or inadequate interest. The predictive models are useful in the determination of high-risk customer segments because of their capability to demonstrate such patterns. Moreover, the findings support the need to balance predictive accuracy with interpretability since decision-makers need to have clear understanding of the churn drivers in order to develop cost-effective retention strategies (Coussement & Benoit, 2021).

This study found similar results to other studies that highlight the usefulness of machine learning methods in predicting churn. This high performance of ensemble models was seen in this study and is indicative of the trend towards the increasing adoption of more sophisticated means of analysis to enhance predictive performance. Similar results have been found in previous studies indicating that hybrid and ensemble models are more effective than traditional models because they are able to generate complex, non-linear customer data relationships (De Caigny et al., 2018). Moreover, the significance of data preparation and feature engineering identified in this study is consistent with previous studies that have shown that preprocessing is a crucial factor in enhancing model performance. High-quality data transformation and features selection allow models to be more representative of the relevant patterns, thus improving predictive accuracy (Coussement et al., 2017). The research also corroborates the findings of systematic reviews indicating that customer churn

prediction is a fast-paced area with a trend towards an ever-growing methodological complexity and a stronger focus on machine learning methods (De & Prabu, 2022). These similarities affirm the fact that the current results have a strong foundation in the current literature. The applied meaning of the results is especially important to the organizations that work in the fields of services based on subscriptions. By determining the key churn factors, companies can come up with specific retention measures to help minimize the loss of customers. As an example, the high impact of the type of contract implies that contract retention can be promoted by motivating the customers to sign more prolonged contracts, as this will raise switching costs and commitment.

Likewise, the correlation between tenure and churn suggests the need of customer engagement at an early stage. Organizations need to work on streamlining the onboarding and providing value in the early stages of the customer lifecycle. It has been found that the combination of predictive analytics and customer segmentation and recommendation systems can greatly improve the retention performance through personalized intervention (Jahan and Sanam, 2024). In addition, pricing should also be appropriately tuned to ensure profitability and customer satisfaction. Customers who have higher monthly rates might need extra value propositions or incentives to be maintained in their loyalty. The firms can distribute their resources more effectively and proactively retaining the customers by predicting such customers with the help of predictive models. The findings also point to the growing importance of the state-of-the-art machine learning techniques in churn prediction. The excellent results of ensemble and boosting models demonstrate that they are able to handle the complex data sets and can demonstrate the complex relationship between variables. It is consistent with the recent studies, which reveal that deep learning and composite models can be used to enhance predictive accuracy, and high-dimensional patterns of data (Khattak et al., 2023). Meanwhile, the growing complexity of these models brings up some issues with interpretability and practical applicability. Although highly accurate models are preferable, they must be practical in the real world by providing practical information. Thus, predictive performance and model transparency have to be balanced with the transparency of such systems so that decision-makers could interpret and rely on the results produced by them (Lalwani et al., 2022).

The results of this paper are particularly relevant to the sphere of subscription-based services where the customer behaviour is based on the constant usage and regular payment habits. The results connected to the flexibility of the contract and sensitivity to prices are indicative of the general trends in the telecommunications industry, streaming services, and online shops. These markets are competitive as well, which implies that customers that have a chance to utilize different service providers are likely to turn to the alternatives based on the perceived value (Mohan and Jadhav, 2022). Predictive analytics is an important instrument in these settings to comprehend customer preferences and predict churn behavior. With the help of data-driven insights, organizations will be in a position to develop more responsive and adaptive business strategies that will respond to the changing customer expectations.

In terms of management, the research presents good advice on how to incorporate predictive analytics in the strategy decision-making process. High-risk customer identification also implies that companies can implement certain interventions that will allow them to retain more customers and distribute resources more effectively. Such strategy can be especially applied in competitive markets, where retention gains can be significantly increased even by minor improvement in gains (Gordini and Veglio, 2017). Moreover, the results emphasize the need to match predictive models and marketing and customer relationship management strategies. The studies have shown that predictive analytics utilized along with marketing programs can be used to make retention efforts more effective by allowing to be more precise in targeting and also communicate more one-on-one.

5. Conclusion

The usefulness of predictive analytics to analyze customer churn and determine important factors that impact customer retention within subscription-based services. The analysis using structured customer data identifies factors like the type of contract, tenure, and pricing framework to be critical influencers of churn behavior. The use of machine learning models also proves that high-level methods of

analysis can be used to greatly improve the predictive power and get a more comprehensive insight into the decision-making behavior of the customers. The results highlight the need to incorporate the use of data-driven methods in business strategies to actively manage customer relationships. The predictive models can assist the organization to identify the high-risk customers and implement certain retention measures thereby increasing customer satisfaction and reducing attrition. Specifically, concentrating on the customer interaction at the initial stage and maximizing the price and agreement provisions could cause significant retention results. On balance, the research makes a contribution to both academic and practical fields as it fills the gap between prediction modeling and strategic decisionmaking. It highlights the increasing applicability of analytics to business performance improvement and assists the creation of sustainable customer retention systems in an ever-competitive and data-driven market environment.

References

1. Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, 6(1), 1-24.
2. Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, 94, 290-301.
3. Backiel, A., Baesens, B., & Claeskens, G. (2016). Predicting time-to-churn of prepaid mobile telephone customers using social network analysis. *Journal of the Operational Research Society*, 67(9), 1135-1145.
4. Chang, V., Hall, K., Xu, Q. A., Amao, F. O., Ganatra, M. A., & Benson, V. (2024). Prediction of customer churn behavior in the telecommunication industry using machine learning models. *Algorithms*, 17(6), 231.
5. Coussement, K., & Benoit, D. F. (2021). Interpretable data science for decision making. *Decision Support Systems*, 150, 113664.
6. Coussement, K., Lessmann, S., & Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, 95, 27-36.
7. De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European journal of operational research*, 269(2), 760772.
8. De Caigny, A., Coussement, K., De Bock, K. W., & Lessmann, S. (2020). Incorporating textual information in customer churn prediction models based on a convolutional neural network. *International Journal of Forecasting*, 36(4), 1563-1578.
9. De, S., & Prabu, P. (2022). Predicting customer churn: A systematic literature review. *Journal of Discrete Mathematical Sciences and Cryptography*, 25(7), 1965-1985.
10. Faritha Banu, J., Neelakandan, S., Geetha, B. T., Selvalakshmi, V., Umadevi, A., & Martinson, E. O. (2022). Artificial intelligence based customer churn prediction model for business markets. *Computational Intelligence and Neuroscience*, 2022(1), 1703696.
11. Gordini, N., & Veglio, V. (2017). Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. *Industrial Marketing Management*, 62, 100-107.
12. Jahan, I., & Sanam, T. F. (2024). A comprehensive framework for customer retention in E-commerce using machine learning based on churn prediction, customer segmentation, and recommendation: I. Jahan, TF Sanam. *Electronic Commerce Research*, 1-44.
13. Khattak, A., Mehak, Z., Ahmad, H., Asghar, M. U., Asghar, M. Z., & Khan, A. (2023). Customer churn prediction using composite deep learning technique. *Scientific Reports*, 13(1), 17294.
14. Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. *Computing*, 104(2), 271-294.

15. Mohan, M., & Jadhav, A. (2022). Predicting customer churn on OTT platforms: customers with subscription of multiple service providers. *Journal of Information and Organizational Sciences*, 46(2), 433-451.
16. Poudel, S. S., Pokharel, S., & Timilsina, M. (2024). Explaining customer churn prediction in telecom industry using tabular machine learning models. *Machine learning with applications*, 17, 100567.
17. Rajamohamed, R., & Manokaran, J. (2018). Improved credit card churn prediction based on rough clustering and supervised learning techniques. *Cluster Computing*, 21(1), 65-77.
18. Shahabikargar, M., Beheshti, A., Zhang, X., Foo, J., & Jolfaei, A. (2026). A comprehensive survey on customer churn analysis studies. *Journal of Information and Telecommunication*, 10(1), 24-70.
19. Sikri, A., Jameel, R., Idrees, S. M., & Kaur, H. (2024). Enhancing customer retention in telecom industry with machine learning driven churn prediction. *Scientific Reports*, 14(1), 13097.
20. Vijaya, J., & Sivasankar, E. (2018). Computing efficient features using rough set theory combined with ensemble classification techniques to improve the customer churn prediction in telecommunication sector. *Computing*, 100(8), 839-860.
21. Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, J. R., Ambadekar, S. P., & Gawande, S. H. (2024). Customer churn prediction in telecom sector using machine learning techniques. *Results in Control and Optimization*, 14, 100342.
22. Ziya. (2025). *E-commerce user behavior and transaction dataset* [Data set]. Kaggle.
<https://www.kaggle.com/datasets/ziya07/e-commerce-user-behavior-and-transaction-dataset>