



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 9.935**

**Volume 15, Issue 5, May 2026**

# Machine Learning Based Customer Churn Prediction: A Review

Avadhute Vaishnavi Rameshrao, Pankaj J. Kokane

Department of Computer Engineering, M.S. Bidve Engineering College, Latur, Maharashtra, India

**ABSTRACT:** Research regarding the interpretability and feature importance of customer churn prediction is rare compared to study on its performance. Nevertheless, knowing what type of customer information are the most significant for predicting churn is critical for firms to layout cost-efficient investment strategies. In this work, feature importance for churn prediction in telecom industry is investigated using a dataset with nearly 40 customer features, and 3 interpretable, tree-based machine learning models. In particular, the data is trained in random forest, decision tree and extra tree classifier algorithms, and uses their built-in feature importance functions to interpret the significance of each feature. The feature importance scores from all three models are compared to analyze which features are the most significant overall. As the result, all models yield optimal prediction accuracy, with decision tree model's slightly lower than the others. The resulting top features from all three models have large overlaps, indicating all models are suitable for the purpose of this study. Random forest and extra tree classifier, as the more complicated models, produce the most satisfactory results, which demonstrate that features related to customer's tenure, consumption in the firm and referral amount are the most important to make accurate churn prediction. This analysis result is reasonable since all these features imply customers' personal preference towards the company.

**KEYWORDS:** Machine Learning, Feature Extraction, Image Global Features, Classification.

## I. INTRODUCTION

Customer churn, the phenomenon where customers discontinue using a company's product or service, poses a critical challenge for businesses across industries. In an increasingly competitive market, retaining existing customers is not only more cost-effective than acquiring new ones but also essential for sustaining long-term growth and profitability. Understanding and predicting customer churn allows companies to implement targeted strategies to enhance customer satisfaction, loyalty, and ultimately, retention.

The churn rate refers to the rate at which subscribers or customers stop transacting with your business. Simply put, they are subscribers who cancelled their subscriptions or customers who did not return to your store.

A higher churn rate means more customers are leaving your business. In contrast, a lower churn rate means retaining more customers than you already have.

## II. LITERATURE SURVEY

ArneshTelukdarie, MangalisoMaduna, [1] adopted methodology involves four different phases, namely, data preparation and processing, feature extraction, dataset and testing, performance evaluation.

B. Prabadevi, R. Shalini [2] authors' described that data included in this investigation includes all customer data going back about nine months before the churn. The accuracy of the aforementioned algorithms are 83.9%, 82.6%, 82.9% and 78.1% respectively.

SuhelMalik, SiddhartRunwalet al [3], This paper provides an overview of customer churn prediction with a focus on machine learning algorithms such as logistic regression, decision trees, random forests, and artificial neural networks.

R. Srinivasan, D. Rajeswarie al [4], This paper discusses the various ML algorithms used to construct the churn model that helps telecom operators to predict customers who are likely to churn. The experimental results are compared to predict the best model among various techniques.

Pulin Yanget al [5], this paper, firstly, the data set is imported, and then the data visualization analysis is carried out. Then, the random forest model, SVM model and GBDT model are introduced for comparison. Experiments show that random forest has better classification performance than other methods.

Ke Peng, Yen Peng, Wengaug Liet al [6], Using Shapley values, we explain how each feature affects the model results, and analyze the features that have a high impact on the model prediction, such as the total number of transactions in the past year, the amount of transactions in the past year.

Peddarapu Rama Krishnaet al [7], The primary goal of this research work is to analyze different models for predicting the active churners with high accuracy. In existing systems, the service providers track down the clients before they leave in order to solve this problem. This study has compared the well-known machine learning techniques to solve the problem and also predict the results in a more accurate way.

Yana Fareniuk et al [8], The key factors influencing the customer churn are identified and form a basis for future prediction of customer outflow and optimization of company’s services. Implementation of customer churn prediction models will help to maintain customer loyalty, reduce customer outflow and increase business results.

OleksandrDluhopolskyet al [9], The research goals are to compare different approaches and methods for customer churn prediction and construct different Data Science models to classify customers according to the probability of their churn from the company’s client base and predict potential customers who could stop to use the company’s services.

Jainam Det al [10], The main intention of this paper is to specify the process of designing the churn prediction model, its application and causes, challenges and problems for designing the model and subsequently, the ways through which the churn rate can be ameliorated.

Paper title	Publication Year	Advantages	Limitation
Predictive Analytics	-	Basic concepts of prediction models	Not telecom specific
Customer Attrition	-	Explains churn concept clearly	No technical models
Supervised Learning	-	Strong ML basics for prediction	General theory only
Churn Analysis of Telecom Company	2022	Practical telecom case study	Limited technical depth
Customer Churn Prediction using Interpretable ML — LiwenOu	2023	Transparent and explainable models	Slightly lower accuracy
Customer Churn Prediction by Hybrid	2009	Better prediction using hybrid NN	High computation, old method

Neural Networks — Chih-Fong & Lu			
Clustering Analysis for Silent Telecom Customers — Qiu et al.	2020	Good customer segmentation	Not full churn prediction
Comprehensive Churn Models using Rule Induction — Verbeke et al.	2011	Easy to understand rules	May reduce accuracy
Survey on Customer Churn Prediction — Umayaparvathi & Iyakutti	2016	Good literature review	No new model
Customer Churn Prediction in Telecommunications — Bingquan et al.	2012	Strong ML comparison	No deep learning methods

**LIMITATION**

- ⌊ Focus mainly on accuracy, less on interpretability.
- ⌊ Limited exploration of feature importance.
- ⌊ Results specific to telecom industry, not easily generalizable.
- ⌊ Short historical datasets restrict long-term insights.
- ⌊ Balancing prediction accuracy with cost efficiency remains challenging.

**III. RESEARCH CHALLENGE**

- **Cost Efficiency:** Retaining existing customers is significantly more cost-effective than acquiring new ones. Predicting churn allows businesses to allocate resources more efficiently towards customer retention strategies.
- **Revenue Preservation:** Customer churn directly impacts a company’s revenue. By predicting churn, businesses can implement timely interventions to prevent revenue loss and sustain financial growth.
- **Customer Lifetime Value (CLV):** Enhancing customer retention directly contributes to increasing the lifetime value of each customer. Predicting churn helps in focusing efforts on customers with high CLV, ensuring long-term profitability.

## IV. CONCLUSION

As we continue to develop our Customer Churn Prediction model, we are focused on refining our approach to ensure it meets the needs of our intended users. Currently, we are analysing various algorithms required for predicting valuable insights into customer behaviors, which can be crucial for retention strategies. Once completed, we anticipate that our model will be a powerful tool for organizations to proactively manage customer relationships and reduce churn outcomes.

## REFERENCES

1. Predictive analytics, Retrieved from [http://en.wikipedia.org/wiki/Predictive\\_analytics](http://en.wikipedia.org/wiki/Predictive_analytics)
2. Customer attrition, Retrieved from [http://en.wikipedia.org/wiki/Customer\\_attrition](http://en.wikipedia.org/wiki/Customer_attrition)
3. Supervised learning Retrieved from [http://en.wikipedia.org/wiki/Supervised\\_learning](http://en.wikipedia.org/wiki/Supervised_learning)
4. <https://www.analyticsvidhya.com/blog/2022/01/churn-analysis-of-a-telecom-company/>  
Customer Churn Prediction based on Interpretable ML algorithms. LiwenOu | Year: 2023 | Conference Paper | Publisher: IEEE
5. <https://ieeexplore.ieee.org/document/10549047/>
6. T. Chih-Fong and Y. H. Lu, "Customer churn prediction by hybrid neural networks", Expert Systems with Applications, vol. 36, no. 10, pp. 12547-12553, 2009. CrossRef Google Scholar
7. Y. Qiu et al., "Clustering Analysis for Silent Telecom Customers Based on K-means++", 2020 IEEE 4th Information Technology Networking Electronic and Automation Control Conference (ITNEC), vol. 1, 2020. View Article
8. W. Verbeke et al., "Building comprehensible customer churn prediction models with advanced rule induction techniques", Expert systems with applications, vol. 38, no. 3, pp. 2354-2364, 2011. CrossRef Google Scholar
9. V. Umayaparvathi and K. Iyakutti, "A survey on customer churn prediction in telecom industry: Datasets methods and metrics", International Research Journal of Engineering and Technology (IRJET) 3.04, 2016. Google Scholar
10. H. Bingquan, M. T. Kechadi and B. Buckley, "Customer churn prediction in telecommunications", Expert Systems with Applications, vol. 39, no. 1, pp. 1414-1425, 2012.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details