

# Customer Churn Prediction Using Machine Learning Algorithms

Shantanu Sonune<sup>1</sup>, Abhijeet Mohite<sup>1</sup>, Anjali Borhude<sup>1</sup>, Yash Patil<sup>1</sup>, Prof. Anuradha Thorat<sup>2</sup>

<sup>1</sup>B.E(Information Technology), Savitribai Phule Pune University, Zeal College of Engineering and Research, Pune, Maharashtra, India

<sup>2</sup>Asst.Professor. Department of Information Technology, Zeal College of Engineering and Research, Pune, Maharashtra, India

## ARTICLE INFO

### Article History:

Accepted: 20 Feb 2024

Published: 03 March 2024

### Publication Issue :

Volume 11, Issue 2

March-April-2024

### Page Number :

28-31

## ABSTRACT

In today's highly competitive industries, retaining customers is vital for sustaining business growth and profitability. Customer churn, the phenomenon where customers switch from one service provider to another, poses a significant challenge for different companies. Predicting churn can help these companies take proactive measures to retain valuable customers. This study explores the application of machine learning algorithms for predicting customer churn in different industries like Telecom Industry, IT and Banking Sector. Additionally, the research contributes to the existing body of knowledge in the field of customer churn prediction, showcasing the potential of machine learning algorithms in addressing complex business challenges.

The project begins with data collection and preprocessing, involving the extraction and cleaning of relevant features from diverse sources such as customer interactions, transaction history, and demographic information. The predictive modeling phase employs state-of-the-art machine learning algorithms, including but not limited to logistic regression, decision trees and random forest, feature engineering is employed to enhance the model's ability to capture nuanced customer behavior. The dataset is split into training and testing sets to evaluate model performance accurately.

The outcomes of this project have practical implications for businesses aiming to proactively manage customer retention strategies. By identifying potential churners early, companies can implement targeted interventions, personalized marketing strategies, and loyalty programs to mitigate the risk of customer loss and foster long-term relationships.

**Keywords :** Customer Churn Prediction, Industry, Machine Learning Algorithms, Predictive Analytics

## I. INTRODUCTION

In today's dynamic and fiercely competitive industries, retaining customers is paramount for sustaining business growth and ensuring profitability. With a plethora of options available to consumers, telecom companies face the significant challenge of minimizing customer churn, the phenomenon where existing customers switch to competitors' services. Customer churn not only leads to revenue loss but also incurs substantial acquisition costs as companies must invest in attracting new customers to offset the losses. To address this challenge, predictive analytics and machine learning algorithms have emerged as powerful tools that enable telecom companies to foresee customer churn and take proactive measures to retain valuable subscribers. By leveraging historical customer data and employing advanced analytical techniques, telecom companies can gain valuable insights into customer behaviour, preferences, and patterns, allowing them to predict churn and implement targeted retention strategies. The study also investigates the crucial factors contributing to customer churn, ranging from customer demographics to usage patterns, customer service interactions, and pricing plans. By discerning the key drivers of churn, telecom companies can tailor their retention efforts to address these specific issues. Moreover, the interpretability of machine learning models will be explored to provide actionable insights into why customers churn, empowering telecom companies to make informed business decisions.

### A. Motivation

Customer churn is a critical challenge for businesses across industries, leading to revenue loss and reduced customer lifetime value. Developing a Customer Churn Prediction model using Machine Learning can help companies identify at-risk customers early, allowing them to implement targeted retention

strategies, improve customer satisfaction, and enhance their bottom line.

### B. Objective

The objectives of the research is to suggest a system that predicts customer churning, Cost Reduction: Acquiring new customers is typically more expensive than retaining existing ones. By identifying customers at risk of churning, businesses can take proactive measures to retain them, saving costs associated with customer acquisition.

Revenue Retention: Reducing churn means retaining a significant portion of your revenue. Predictive models help companies take actions to prevent customer defection and maintain a steady stream of income.

Improved Customer Satisfaction: Understanding the factors that lead to churn can help businesses improve their services and customer experience. By addressing the root causes, they can enhance customer satisfaction and loyalty.

Marketing and Communication Optimization: Churn prediction allows businesses to tailor marketing efforts and communication to individual customer needs. This personalization can increase customer engagement and reduce the likelihood of churn.

Product Development: Churn analysis can provide insights into customer preferences and pain points, guiding product development and enhancement.

## II. METHODS AND MATERIAL

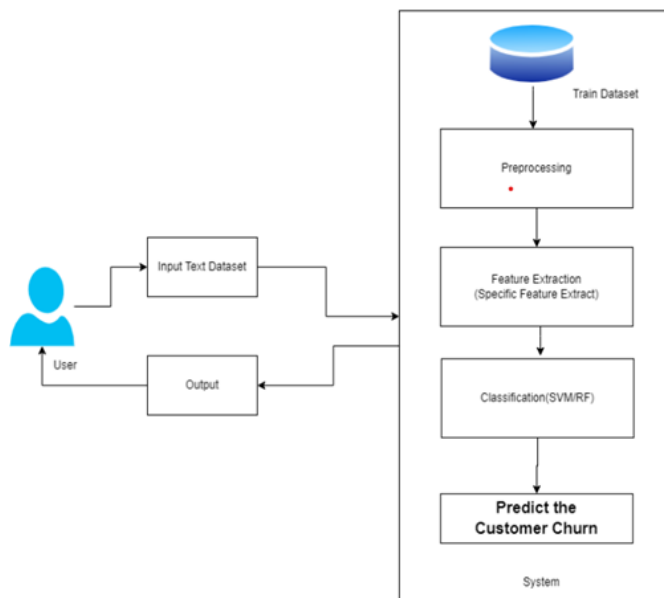
### A. System Architecture

To perform customer churn prediction using ML algorithms, businesses typically follow these steps:

- 1) Data Collection: Gather relevant customer data, including demographics, usage patterns,

transaction history, customer interactions, and customer feed

- 2) Data Preprocessing: Clean and preprocess the data to handle missing values, outliers, and ensure it's suitable for machine learning.
- 3) Feature Engineering: Create meaningful features that capture customer behavior and characteristics, such as customer lifetime value, customer satisfaction scores, and more.
- 4) Model Selection: Choose appropriate ML algorithms, such as logistic regression, decision trees, random forests, support vector machines, or neural networks, based on the data and business requirements.



- 5) Training and Validation: Split the data into training and validation sets to train the model and assess its performance. Techniques like cross-validation can help evaluate different models.
- 6) Prediction: Deploy the trained model to make churn predictions for current customers.
- 7) Action Planning: Based on model predictions, take proactive measures to retain customers at risk of churning. This might include personalized offers, loyalty programs, or improved customer support.
- 8) Monitoring: Continuously monitor the model's performance and update it as necessary to adapt to changing customer behaviours.

## B. Module

**Admin:** In this module, the Admin has to log in by using valid user name and password. After login successful he can do some operations such as View All Users and Authorize.

**View and Authorize Users:** In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

**View Charts results:** View All Products Search Ratio, View All Keyword Search Results, View All Product Review Rank Results.

**End User:** In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will best or to the database. After registration successful. He has to login by using authorized user name and password. Once Login is successful user will do some operations like Manage Account.

## C. Data flow Diagram

In Data flow diagram we show that flow of data in our system in DFD0 we show that base DFD in which rectangle present input as well as output and circle show our system, In DFD1 we show actual input and actual output of system input of our system is text or image and output is rumour detected likewise in DFD 2 we present operation of user as well as admin.

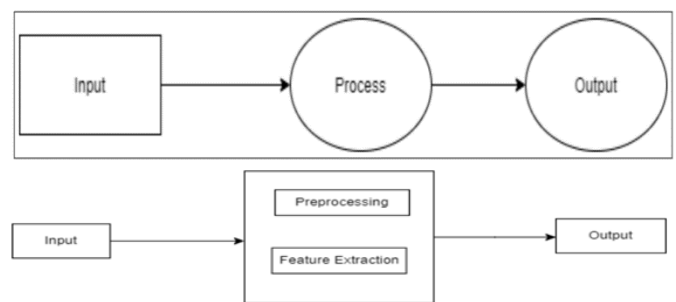


Figure 4.3: Data Flow(1) diagram

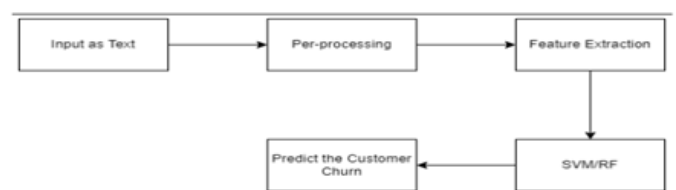
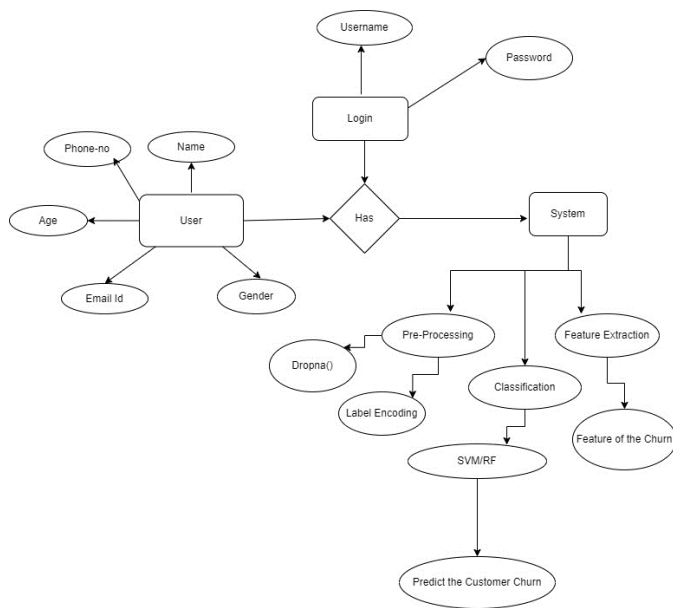


Figure 4.4: Data Flow(2) diagram.

### D. ER Digram



## III. Algorithms

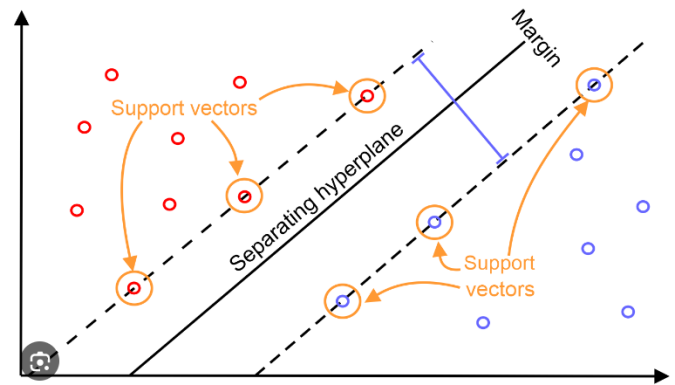
Machine learning algorithms are a specific set of algorithms that allow computers to learn patterns and make predictions or decisions based on data.

### A. SVM

SVMs can model non-linear decision boundaries effectively using techniques like the kernel trick, which transforms data into a higher-dimensional space where non-linear relationships become linear. SVMs are less prone to overfitting, especially when compared to some other machine learning algorithms.

Steps:

- Step 1: Load the important libraries.
- Step 2: Import dataset and extract the X variables and Y separately.
- Step 3: Divide the dataset into train and test.
- Step 4: Initializing the SVM classifier model
- Step 5: Fitting the SVM classifier model.
- Step 6: Coming up with predictions.
- Step 7: Evaluating model's performance

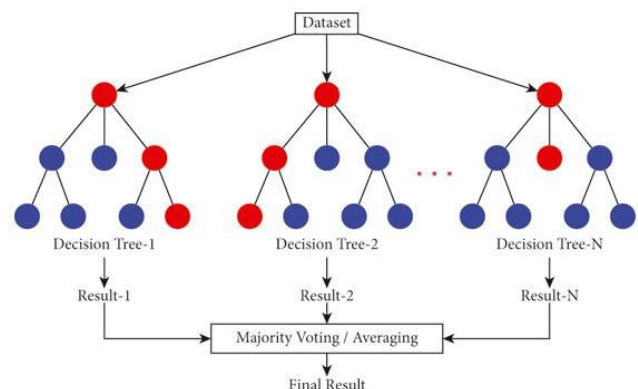


### B. Random Forest

RF is an ensemble learning method that combines multiple decision trees to make predictions. RF can capture these non-linear relationships effectively. RF tends to produce stable and consistent results across different runs and datasets, which is important in healthcare where reproducibility is critical

Steps:

- Step 1: Importing and processing the data.
- Step 2: Training the random forest classifier.
- Step 3: Testing the prediction accuracy.
- Step 4: Visualizing the results of the classifier.



## IV. CONCLUSION

Customer churn prediction in various industries using machine learning algorithms is a vital and feasible initiative that offers significant benefits to telecom companies. By leveraging advanced analytical techniques, different service providers can proactively identify customers at risk of churning, allowing for targeted retention strategies and improved customer satisfaction. By investing in this technology, service providers can not only reduce churn rates but also

foster stronger customer relationships, ensuring their long-term success in a highly competitive industry.

## V. REFERENCES

- [1] W. M. Van der Aalst, "Process modeling and analysis," in *Process Mining: Data Science in Action*, 2nd ed. Berlin, Germany: Springer, 2016, ch. 3, pp. 55–88.
- [2] S. Sakr, Z. Maamar, A. Awad, B. Benatallah, and W. M. P. van der Aalst, "Business process analytics and big data systems: A roadmap to bridge the gap," *IEEE Access*, vol. 6, pp. 77308–77320, 2018
- [3] Z. Tangyuan and S. Moro, "Research trends in customer churn prediction: A data mining approach," in *Proc. World Conf. Inf. Syst. Technol.*, 2021, pp. 227–237.
- [4] A. D. Caigny, K. Coussement, and K. W. D. Bock, "A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees," *Eur. J. Oper. Res.*, vol. 269, pp. 760–772, Sep. 2018.
- [5] N. Alboukaey, A. Joukhadar, and N. Ghneim, "Dynamic behavior based churn prediction in mobile telecom," *Expert Syst. Appl.*, vol. 162, Dec. 2020, Art. no. 113779
- [6] K. Viol, H. Schöller, A. Kaiser, C. Fartacek, W. Aichhorn, and G. Schiepek, "Detecting pattern transitions in psychological time series—A validation study on the pattern transition detection algorithm (PTDA)," *PLoS ONE*, vol. 17, no. 3, Mar. 2022, Art. no. e0265335.
- [7] K. Viol, H. Schöller, A. Kaiser, C. Fartacek, W. Aichhorn, and G. Schiepek, "Detecting pattern transitions in psychological time series—A validation study on the pattern transition detection algorithm (PTDA)," *PLoS ONE*, vol. 17, no. 3, Mar. 2022, Art. no. e0265335.
- [8] P. Lalwani, M. K. Mishra, J. S. Chadha, and P. Sethi, "Customer churn prediction system: A machine learning approach," *Computing*, vol. 104, no. 2, pp. 271–294, Feb. 2022.
- [9] G. Mohammadi, R. Tavakkoli-Moghaddam, and M. Mohammadi, "Hierarchical neural regression models for customer churn prediction," *J. Eng.*, vol. 2013, pp. 1–9, Feb. 2013.
- [10] Z. Chen, S. Zhang, S. McClean, B. Allan, and I. Kegel, "Sequence mining TV viewing data using embedded Markov modelling," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI)*, Oct. 2021, pp. 665–670

### Cite this article as :

Shantanu Sonune, Abhijeet Mohite, Anjali Borhude, Yash Patil, Prof. Anuradha Thorat, "Customer Churn Prediction Using Machine Learning Algorithms", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 11 Issue 2, pp. 28-31, March-April 2024. Available at doi : <https://doi.org/10.32628/IJSRST52411140>  
Journal URL : <https://ijsrst.com/IJSRST52411140>