International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Churn Prediction for OTT/Tele-Communication

Prof Tushar.P. Sharma[1], Bhagyashri Nandu Jadhav[2]

Assistant Professor[1], Student[2]

[1,2]Department of Computer Science K.B.H.S.S. Trusts Indira College Malegaon, Maharashtra

---

Abstract - Operating a subscription-based business entails addressing diverse customer preferences, often leading to subscription cancellations. Thus, accurate churn forecasting becomes imperative, as even minor changes can have significant repercussions. Without prior notification of impending cancellations, sellers are unable to take preventive measures to retain customers. Consequently, this study endeavors to devise a churn prediction model, anticipating subscription cancellations and offering incentives to retain subscribers. This not only yields substantial cost savings but also generates additional revenue streams for online businesses. The primary objective is to assess various models for accurately predicting active churners. Unlike existing systems that reactively address cancellations, this study seeks to proactively anticipate and mitigate churn. By comparing established machine learning techniques, the research aims to enhance prediction accuracy, thereby providing a more effective solution to this prevalent challenge in subscription-based enterprises.

Keywords: *Customer Relationship Management, Customer Retention, Machine Learning, Churn Analysis*

1. INTRODUCTION

In today's dynamic landscape of Over-The-Top (OTT) and telecommunication services, understanding customer behavior is paramount for sustained success. Predicting and addressing churn, where customers terminate their subscriptions or cease service usage, is crucial for maintaining business sustainability. This project aims to tackle this challenge by crafting a churn prediction model tailored to the OTT and telecommunication sector.

Churn prediction is essential for proactive customer management, allowing businesses to anticipate and intervene before customers churn. By harnessing advanced analytics and machine learning techniques, this project seeks to equip OTT and telecommunication firms with insights to identify at-risk customers, implement targeted retention strategies, and foster long-term customer loyalty.

Through this project documentation, we will delve into the methodology, data exploration, model development, evaluation metrics, and practical implications of the churn prediction system. By sharing our approach and findings, we aspire to enrich the collective knowledge base of customer relationship management in the OTT and telecommunication realm, assisting companies in enhancing customer satisfaction and maximizing business outcomes.

2. LITERATURE REVIEW

M.A.H. Farquad proposed a hybrid approach to address the limitations of the conventional SVM model, which often generates a black box model, lacking human-understandable insights gained during training. The hybrid approach consists of three phases: firstly, SVM-RFE (SVM-recursive feature elimination) is utilized to reduce the feature set. Next, the dataset with reduced features is used to construct an SVM model, and support vectors are extracted. Finally, rules are generated using Naive Bayes Tree (NBTree), a combination of Decision tree with a naive Bayesian Classifier. The study employs a bank credit card customer dataset from the Business Intelligence Cup 2004, characterized by high imbalance.

Wouter Verbeke introduced the application of Ant-Miner+ and ALBA algorithms on a publicly available churn prediction dataset to develop accurate and interpretable classification rule-sets for churn prediction models. Ant-Miner+ is a data mining method based on Ant Colony Optimization principles, offering high accuracy and comprehensibility, while ALBA combines the predictive accuracy of a non-linear support vector machine model with the comprehensibility of rule sets.

Ning Lu proposed the utilization of boosting algorithms to enhance a customer churn prediction model, separating customers into two clusters based on the weight assigned by the boosting algorithm. Logistic regression serves as the basis learner, with a churn prediction model constructed for each cluster. The experimental findings highlight the effectiveness of boosting algorithms in separating churn data compared to a single logistic regression model.

Ssu-Han Chen employed a novel mechanism based on the gamma Cumulative (CUSUM) chart, monitoring individual customers' Inter Arrival Time (IAT) using a finite mixture model to design the reference value and decision interval. Recency, a complementary time interval variable, is integrated into the model to track recent login behavior. The results demonstrate the superiority of the gamma CUSUM chart in accuracy rate and Average Time to Signal (ATS) compared to exponential CUSUM.

3. PROBLEM STATEMENT:

Customer retention stands as a pivotal growth pillar for products adopting a subscription-based business model. Within the highly competitive Software as a Service (SaaS) industry, where clients have abundant provider options, even within a single product category, a single negative experience can prompt a consumer to discontinue their subscription. The departure of dissatisfied customers, especially in large numbers and at a swift pace, can result in substantial financial losses and reputational damage. To tackle this issue, we will explore methods and recommended procedures for gathering information on client-brand relationships, identifying customer behavior traits associated with churn, and selecting top-performing machine learning models. Current systems often lack the granularity and predictive capacity required to pinpoint individual customers at risk of churning, relying instead on broad segmentation and generic customer management approaches.

By addressing these challenges, the proposed project aims to equip OTT and telecommunication companies with the tools and insights necessary to elevate customer satisfaction, optimize revenue streams, and maintain a competitive advantage in the ever-evolving digital services landscape.

4. AIM & OBJECTIVES

1. Our project will predict Churn rate.
2. The company will learn about its weaknesses.
3. Company will be able to improve the customer satisfaction rate.
4. By identifying potential churners, the company can proactively address issues, enhancing customer loyalty and reducing revenue loss.
5. Prediction of churn facilitates the development of personalized retention strategies, fostering stronger relationships with customers.

5. PROPOSED SYSTEM

Churn Prediction for OTT/Tele-Communication Systems introduction:

The proposed system aims to transform churn prediction methods within the OTT (Over-The-Top) and telecommunication industry by leveraging advanced analytics and machine learning techniques. At its core, it addresses the critical challenge of customer churn, where subscribers terminate their service or subscription, leading to revenue loss and operational disruptions. By accurately forecasting churn and implementing targeted retention strategies, the system empowers OTT and telecommunication companies to proactively manage customer relationships, boost satisfaction levels, and optimize business outcomes. Central to the system is the use of comprehensive datasets covering customer interactions, subscription details, usage patterns, and churn history. Through meticulous data preprocessing and feature engineering, actionable insights are extracted, enabling the development of robust churn prediction models. These models, driven by machine learning algorithms like logistic regression, decision trees, random forests, or gradient boosting, identify individual customers at risk of churn with high accuracy.

Additionally, the system emphasizes seamless integration of churn prediction capabilities into existing IT infrastructure. This integration enables real-time prediction and automation of retention strategies, such as personalized offers and targeted marketing campaigns. By leveraging data-driven decision-making, companies can anticipate customer behavior, mitigate churn risk, and foster long-term loyalty. Ultimately, the proposed system represents a paradigm shift in customer relationship management for OTT and telecommunication companies. By merging cutting-edge technology with domain-specific expertise, it seeks to redefine industry standards, drive innovation, and position companies for sustained growth and competitiveness in an evolving marketplace.

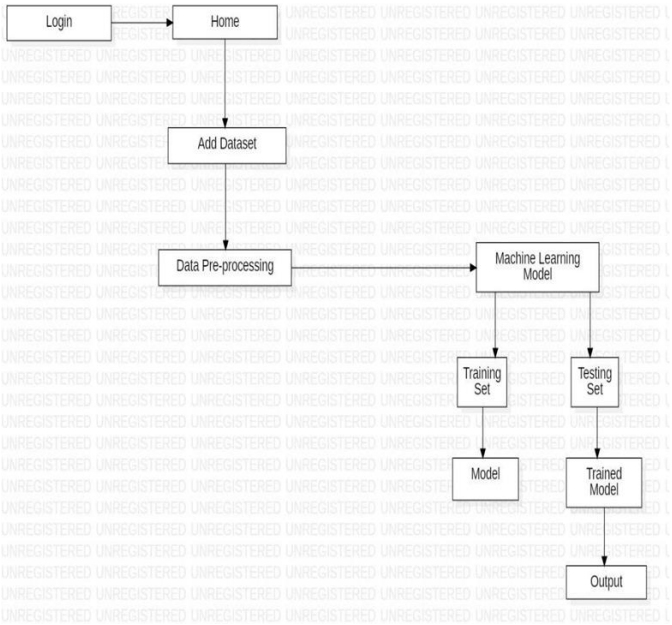

Fig. Churn Prediction for OTT/Tele-Communication System

The major components of the Churn Prediction for OTT/Tele-Communication systems are:

1) Gather relevant data from OTT and telecommunication providers, including customer demographics, subscription details, usage patterns, and churn history.
2) Clean and preprocess the data to handle missing values, outliers, and inconsistencies, ensuring data quality and integrity.

3) Extract meaningful features from the data, such as customer tenure, subscription plan attributes, usage frequency, and engagement metrics.
4) Build and train machine learning models using techniques like logistic regression, decision trees, random forests, or gradient boosting to predict churn likelihood.
5) Evaluate model performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to determine effectiveness in identifying churners.
6) Integrate the churn prediction model into existing IT infrastructure, ensuring seamless deployment and real-time prediction capabilities.
7) Design targeted retention strategies based on churn predictions, such as personalized offers, loyalty programs, and proactive customer outreach.
8) Continuously monitor model performance in production, gather feedback, and iterate on the model to improve accuracy and reliability over time.
9) Document the entire project including data sources, methodologies, model development process, evaluation results, and implementation details.
10) Assess the impact of the churn prediction system on key business metrics such as customer retention rates, revenue growth, and customer satisfaction scores.

6.. WORKING

Churn prediction for OTT (Over-The-Top) services and telecom companies involves using data analytics and machine learning techniques to identify customers who are likely to stop using the service. Here is a detailed explanation of the process:

1. **Data Collection**
   a) Customer Demographics: Age, gender, location, etc.
   b) Usage Patterns: Frequency and duration of service usage, types of content consumed (for OTT), call and data usage (for telecom).
   c) Interaction Data: Customer service interactions, complaint logs.
   d) Subscription Details: Plan type, renewal dates, payment history.
   e) Behavioral Data: Login times, device types, content preferences.
   f) Churn History: Data on previous churners, reasons for churn if available.
2. **Data Preprocessing**
   a) Data Cleaning: Handling missing values, outlier detection, and noise reduction.
   b) Feature Engineering: Creating new features from raw data, such as average usage per month, number of logins per week, etc.
   c) Normalization/Scaling: Ensuring numerical features are on the same scale.
   d) Encoding Categorical Variables: Converting categorical data into numerical format using techniques like one-hot encoding.
3. **Exploratory Data Analysis (EDA)**
   a) Statistical Analysis: Understanding distributions, mean, median, etc.
   b) Visualizations: Histograms, box plots, scatter plots to identify trends and correlations.
4. **Model Building**
   a) Choosing Algorithms: Common algorithms include Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, and Neural Networks.
   b) Training the Model: Using historical data to train the model on identifying patterns that lead to churn.
   c) Feature Selection: Identifying the most significant features that impact churn prediction.
5. **Model Evaluation**
   a) Validation Techniques: Cross-validation, splitting data into training and testing sets.
   b) Performance Metrics: Accuracy, Precision, Recall, F1-Score, ROC-AUC curve to evaluate the model's effectiveness.
6. **Deployment**
   a) Real-Time Prediction: Implementing the model to predict churn on new data as it comes in.
   b) Integration: Incorporating the prediction system into the existing CRM or analytics platforms.
7. **Actionable Insights**
   a) Customer Retention Strategies: Developing targeted campaigns, personalized offers, and proactive customer support to retain high-risk customers.
   b) Continuous Monitoring: Regularly updating the model with new data and retraining to maintain accuracy.
   c) Example Workflow
   d) Data Collection: Collect data from various sources such as CRM systems, user activity logs, customer support records, etc.
   e) Data Preprocessing: Clean and preprocess data to handle missing values, normalize features, and encode categorical variables.
   f) EDA: Perform exploratory data analysis to understand the data distribution and correlations.
   g) Feature Engineering: Create new features that might help in better predicting churn.
   h) Model Training: Use machine learning algorithms to train the model on historical data.
   i) Model Evaluation: Evaluate the model using metrics like accuracy, precision, recall, and the ROC-AUC curve.

j) Deployment: Deploy the model to a production environment to start predicting churn.
k) Insights and Actions: Use the predictions to drive customer retention strategies and continuously monitor the model performance.
l) Conclusion

Churn prediction helps OTT and telecom companies to proactively manage customer relationships by identifying those at risk of leaving and implementing targeted retention strategies. This process leverages data analytics and machine learning to turn raw data into actionable insights, ultimately improving customer retention and business outcomes.

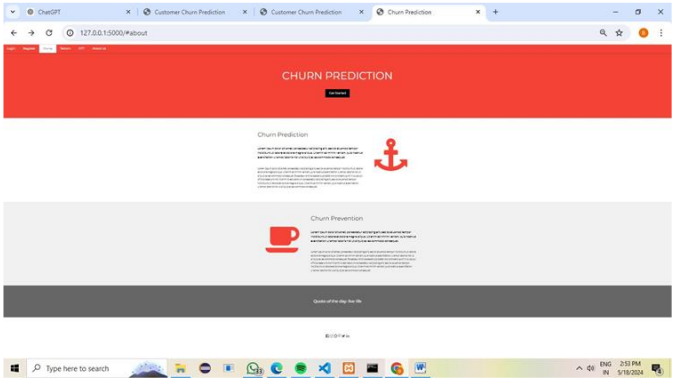

Fig1. Home Page

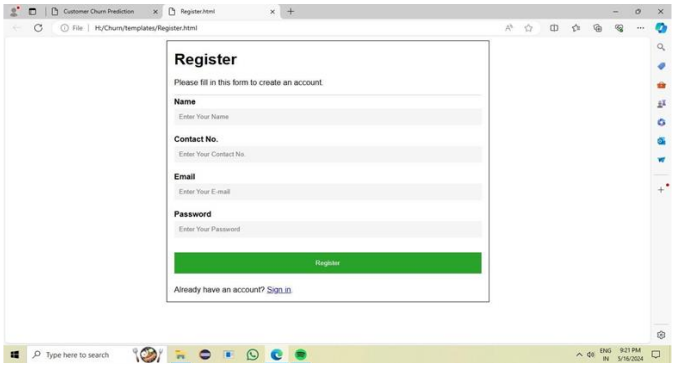

Fig2. Sign Up Page

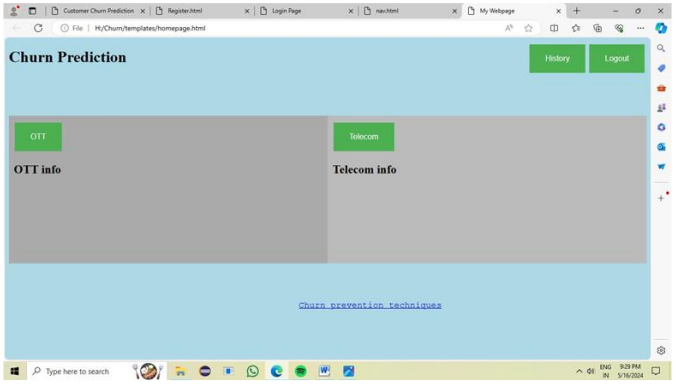

Fig3. Churn Prediction

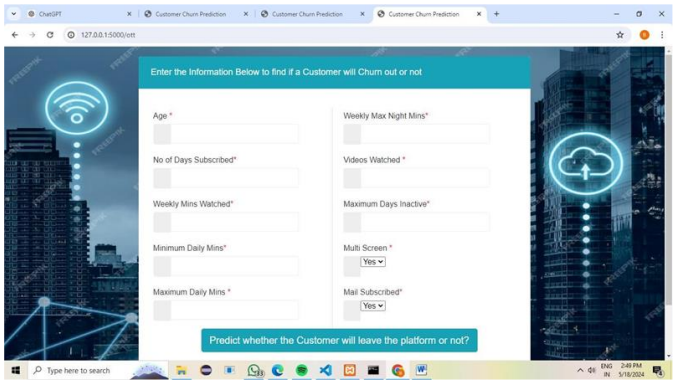

Fig4. Result Page

7. ADVANTAGES & APPLICATIONS

7.1. Advantages: -

1. Enhanced Customer Retention: By accurately predicting which customers are likely to churn, service providers can implement proactive measures to retain them. This may include targeted offers, personalized recommendations, or improved customer service, ultimately leading to higher retention rates and increased customer lifetime value.
2. Cost Savings: Acquiring new customers can be significantly more expensive than retaining existing ones. By reducing churn through targeted retention strategies, service providers can save on customer acquisition costs and improve overall profitability.
3. Improved Customer Satisfaction: By addressing the needs and preferences of at-risk
4. customers before they churn, service providers can enhance overall customer satisfaction. Personalized offers and proactive communication demonstrate a commitment to customer care, fostering loyalty and positive brand perception.
5. Optimized Resource Allocation: Churn prediction allows service providers to allocate resources more efficiently by focusing retention efforts on customers with the highest churn probability. This ensures that resources such as marketing budgets, customer support staff, and retention incentives are deployed where they can have the greatest impact.
6. Data-Driven Decision Making: Churn prediction relies on advanced analytics and machine learning techniques to analyze large volumes of customer data. By leveraging data driven insights, service providers can make informed decisions about retention strategies, product offerings, and customer engagement initiatives, leading to more effective business outcomes.

7. Competitive Advantage: In today's competitive OTT/Telecommunication market, the ability to accurately predict and prevent churn can be a significant differentiator. Service providers that proactively address churn and prioritize customer retention are better positioned to succeed in a crowded marketplace, gaining a competitive edge over rivals.
8. Long-Term Revenue Growth: By reducing churn and increasing customer retention, service providers can foster long-term revenue growth. Loyal customers are more likely to purchase additional products or services, upgrade their subscriptions, and recommend the service to others, driving sustainable revenue streams over time.

7.2. Applications: -

Gather data on customer demographics, subscription details, usage patterns, and interaction history.

8. CONCLUSION

As a result of the discussion above, it can be said that every organization, regard- less of its type, needs to be concerned about customer churn. Customer retention is the process of preserving a customer's loyalty by comprehending their needs and meeting them appropriately. The organizational management will be helped by a strong churn prediction model to anticipate customer churn. Support vector machines can be useful for predicting the churn rate, depending on the complex data of the industry. The concept of customer retention as well as the churn prediction were the main topics of the aforementioned report. Along with the algorithm, the use of support vector machines to improve the churn prediction process has also been covered here.

REFERENCES

[1]Saran Kumar A,Chandrakala D "A Survey on Customer Churn Prediction using Machine Learning Techniques" November 2016,International Journal of ComputerApplications,154(10):1316,DOI:10.5120/ijca2016912237 volume Article number: 28 (2019).

[2]Abdelrahim Kasem Ahmad, Assef Jafar & Kadan Aljoumaa "Customer churn prediction in telecom using machine learning in big data platform" Journal of Big Data, Article number: 28 (2019).

[3]Damandeep Singh , Vansh ,Dr. M. Kanchana, Associate Professor, "Survey Paper on Churn Prediction on Telecom", SRM, India .

[4] Nasebah Almufadi , Ali Mustafa Qamar, Rehan Ullah Khan, Mohamed Tahar Ben Othman, "Deep Learning-based Churn Prediction of Telecom Subscribers", International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 12 (2019),pp. 2743-2748.

[5]N Lakshmi Kalyani and Kolla Bhanu Prakash, "Soil Color as a Measurement for Estimation of Fertility using Deep Learning Techniques" International Journal of Advanced Computer Science and Applications(IJACSA), 13(5), 2022.

[6]Kriti, "Customer Churn: A Study of Factors Affecting Customer Churn using Machine Learning", Iowa State University Ames, Iowa 2019 .

[7]Adnan Amin , Feras Al-Obeidat , Babar Awais Adnan , Jonathan Loo , Sajid Anwar, "Customer churn prediction in telecommunication industry under uncertain situation", Center for Excellence in Information Technology, Institute of Management Sciences, Peshawar, 25000 Pakistan. College of Technological Innovation, Zayed University,144534 Abu Dhabi, United Arab Emirates. Computing and Communication Engineering, University West London.

[8]Pronay Ghosh, "Project report on customer churn prediction using supervised machine learning" .

[9]Peddarapu Rama Krishna, Pothuraju Rajarajeswari, "Early Detection Of Melanoma Skin Cancer Using Efficient Netb6", International Conference on Advanced Computing and Communication Systems (ICACCS), Vol 1 & pg.01-05,07-Jun-2022.

[10]Nurulhuda Mustafa, Siti Fatimah Abdul Razak, "Customer churn prediction for telecommunication industry": A Malaysian Case Study.

[11]K.Sandhya Rani,Shaik Thaslima ,N.G.L Prasanna R.Vindhya ,P. Srilakshmi ."Analysis of Customer Churn Prediction in Telecom Industry Using Logistic Regression International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ,Volume-9,July 2021, https://doi.org/10.21276/ijircst.2021.9.4.6.

[12]Srinivas Kolli, Peddarapu Rama Krishna and Parvathala Balakesava Reddy, A Novel NLP And Machine Learning Based Text Extraction
Approach From Online News Feed, ARPN Journal of Engineering and Applied Sciences, Vol. 16, No. 6, pg no's:679-685,Mar-2021.

[13]Ming Zhao ,Qingjun Zeng , Ming Chang , Qian Tong , and Jiafu Su "A Prediction Model of Customer Churn considering Customer Value: An Empirical Research of Telecom Industry in China", Research Center for Economy of Upper Reaches of the Yangtze
River, Chongqing Technology and Business University, Chongqing 400067, China.

[14]Benjamin Ghaffari & Yasin Osman, "Customer churn prediction using machine learning Benjamin Ghaffari & Yasin Osman, A study in the B2B subscription based service context" .

[15]Praveen Lalwani, Manas Kumar Mishra, Jasroop Singh Chadha, Pratyush Sethi, "Customer churn prediction system: a machine learning approach".