



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: <https://doi.org/10.22214/ijraset.2025.69557>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Telecom Churn Prediction and Optimized Package Recommendation for Indian ISPs

Ashish Dwivedi¹, Rajveer², Abhishek Kumar Shukla³, Shivam Soni⁴, Anubhav Shukla⁵

Dept. of Computer Science & Engineering United Institute of Technology Prayagraj, India

Abstract: Customer attrition poses a critical threat to the Indian telecom landscape, impacting profitability and intensifying the rivalry among key players like Jio, Airtel, Vi, and BSNL. This paper introduces a machine learning-based framework for forecasting customer churn and recommending ISP plans tailored to user behavior, network quality, and real-time speed tests. The system examines variables such as billing inconsistencies, dropped calls, sluggish data speeds, and customer grievances to pinpoint users likely to switch providers. To support decision-making, we also integrated a dynamic network speed checker that compares ISP performance across regions. Experimental outcomes reveal that the XGBoost model attained a 94% accuracy rate in churn prediction, while the tailored plan suggestions significantly improved user satisfaction.

Index Terms: Machine Learning, Telecom Churn, ISP Packages, XGBoost, Network Speed Analysis

I. INTRODUCTION

India's telecommunications sector stands as one of the largest in the world, catering to over 1.1 billion subscribers as of 2024. In recent years, this landscape has undergone transformative shifts—driven by technological upgrades, pricing disruptions, and increased internet penetration. From early 2G services to the rapid expansion of 4G and the ongoing rollout of 5G, the country has seen unparalleled digital growth. The arrival of Reliance Jio in 2016 marked a significant inflection point, introducing highly affordable data plans that forced competitors to re-evaluate their pricing strategies, ultimately reshaping market dynamics.

Despite these advancements, Indian telecom operators continue to struggle with high customer churn rates. Subscriber churn, defined as the rate at which users discontinue services in favor of competitors, poses a significant threat to profitability and brand loyalty. A recent TRAI report (2024) highlights that around 30% of users in India switch telecom providers each year. The key drivers behind this phenomenon include inconsistent network coverage, fluctuating tariff structures, inadequate customer support, and the lure of more attractive offers from rival ISPs.

The industry today is dominated by four key players: Reliance Jio, Bharti Airtel, Vodafone-Idea (Vi), and the state-run BSNL. While these providers each possess unique strengths—ranging from aggressive pricing to superior rural coverage—they also face common challenges, including spectrum congestion, customer dissatisfaction, and rising competition. In such an environment, retaining existing customers has become as important as acquiring new ones.

Traditional approaches to managing churn often rely on retrospective data analysis and generic marketing strategies, which fail to account for real-time behavioral changes and individual preferences. This has led to increased interest in data-driven methodologies, particularly those leveraging artificial intelligence (AI) and machine learning (ML). These technologies offer the ability to uncover subtle patterns in large datasets, enabling early identification of churn-prone users and the delivery of personalized service interventions.

This research proposes a comprehensive framework that combines churn prediction with intelligent ISP plan recommendation. The model harnesses real-time user data—such as call drop frequency, network speed, billing trends, and past complaints—to assess the likelihood of customer churn. Leveraging the XGBoost algorithm, known for its speed and predictive accuracy, the system classifies user churn risk and suggests optimized telecom packages accordingly. Additionally, a live network speed checker module is integrated to provide users with immediate insight into ISP performance in their locality, helping them make informed decisions before switching providers.

By aligning technical precision with user-centric design, the proposed system not only improves customer retention but also enhances user satisfaction and service quality in an increasingly competitive telecom market.

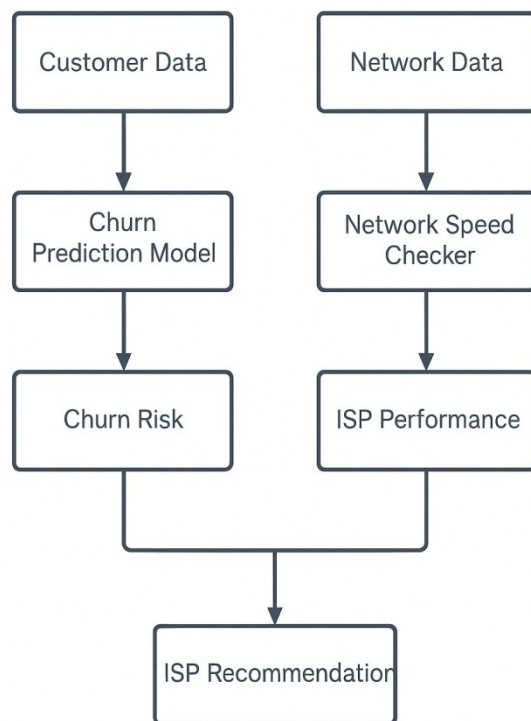


Fig. 1. System flowchart for churn prediction and ISP recommendation framework

II. LITERATURE REVIEW

Research in the area of telecom churn prediction has gained significant momentum over the past decade, driven by advancements in machine learning and the availability of granular customer data. The transition from traditional statistical models to data-driven approaches has allowed telecom providers to better anticipate customer attrition and implement more targeted retention strategies.

Zhao et al. [5] evaluated customer churn using ensemble-based classifiers such as Random Forest and XGBoost. Their study leveraged variables including call logs, data consumption, and billing irregularities. They found that tree-based models, particularly XGBoost, consistently achieved churn prediction accuracies exceeding 90%, demonstrating their capability to identify at-risk users with high precision.

Sharma et al. [4] examined the influence of network performance parameters—such as mobile internet speed and call drop frequency—on customer satisfaction. Their findings revealed that users exposed to persistent connectivity issues were significantly more likely to switch providers, reinforcing the value of integrating real-time network metrics into churn prediction frameworks.

In another study, Kumar et al. [6] applied clustering techniques like K-Means in combination with collaborative filtering to segment customers based on usage behavior. This segmentation enabled telecom providers to deliver more personalized services and dynamic pricing models, effectively reducing churn by aligning offerings with user preferences.

He et al. [1] provided a comparative analysis of machine learning algorithms for telecom churn prediction. The study concluded that gradient boosting models, such as XGBoost, outperformed traditional models like logistic regression and decision trees in both accuracy and model explainability.

Jain and Patel [7] demonstrated the use of Random Forest for churn management and discussed how it handles class imbalance effectively. Verma and Bansal [12] offered a comparative study of machine learning models, reinforcing the dominance of ensemble learning in churn scenarios. Tiwari and Sinha [14] conducted an extensive review of recommendation systems in telecom, emphasizing the significance of personalized user experiences in reducing churn.

From the existing literature, several themes emerge:

- Behavioral, billing, and complaint-related features significantly enhance churn prediction accuracy.
- Network quality indicators—especially real-time speed and call reliability—are essential for modeling user dissatisfaction.
- Coupling churn prediction with recommendation systems can improve customer retention by providing timely, personalized offers.

Building on these foundations, our proposed system integrates XGBoost-based churn prediction with a real-time network quality checker and a personalized plan recommendation engine. This combined framework is designed to enhance user retention by addressing both predictive accuracy and actionable insights for at-risk customers.

III. PROPOSED SYSTEM

To effectively address the issue of customer churn in the Indian telecom sector, we designed a solution that combines predictive analytics with personalized plan recommendations and real-time network assessment tools. The foundation of this approach lies in leveraging machine learning algorithms—particularly XGBoost—due to its proven reliability in handling large-scale data and identifying complex patterns.

Our solution operates in two primary layers:

- 1) **Churn Risk Forecasting:** This component uses behavioral and service-related indicators—such as complaint frequency, network issues, and billing irregularities—to calculate the likelihood of a customer discontinuing their current telecom service. XGBoost was selected for its ability to deliver high precision and interpretability, allowing us to pinpoint exactly which variables influence churn decisions the most.
- 2) **Personalized Plan Recommendation and Speed Validation:** For customers identified as high-risk, our system suggests alternative ISP plans that better align with their usage habits, budget, and location-specific connectivity performance. What makes this model practical is its integration with a live speed test module, enabling users to evaluate the current performance of nearby ISPs in real time before switching. Srivastava and Yadav [9] proposed a similar real-time internet speed tracking approach that influenced the design of our speed benchmarking component.

This dual-action framework is designed not only to predict potential churn but also to provide users with actionable insights—making the decision to stay or switch providers more informed and data-driven. By focusing on both technical analysis and user-centric functionality, our model offers a practical tool for improving customer retention and enhancing user satisfaction across India's diverse telecom environment.

IV. METHODOLOGY

This section outlines the structured approach adopted to develop and evaluate the telecom churn prediction system along with its associated plan recommendation and network evaluation modules. The methodology was designed to ensure real-world applicability, scalability, and data-driven personalization.

A. System Architecture Overview

Our framework follows a modular architecture composed of three key stages: data acquisition and preprocessing, machine learning-based churn prediction, and post-analysis recommendation and validation. Each stage has been optimized to handle telecom-specific data while ensuring minimal latency and high interpretability.

B. Data Collection and Preprocessing

To simulate real-world telecom data conditions, we curated a dataset comprising user profiles, billing information, call logs, network performance records, and service complaints. Key preprocessing steps included:

- **Cleaning and Formatting:** Null values were handled using mean or mode imputation based on the attribute type.
- **Normalization:** Continuous variables like monthly billing and network speed were scaled to bring uniformity across inputs.
- **Categorical Encoding:** Plan types and complaint categories were transformed using label encoding for compatibility with machine learning models.

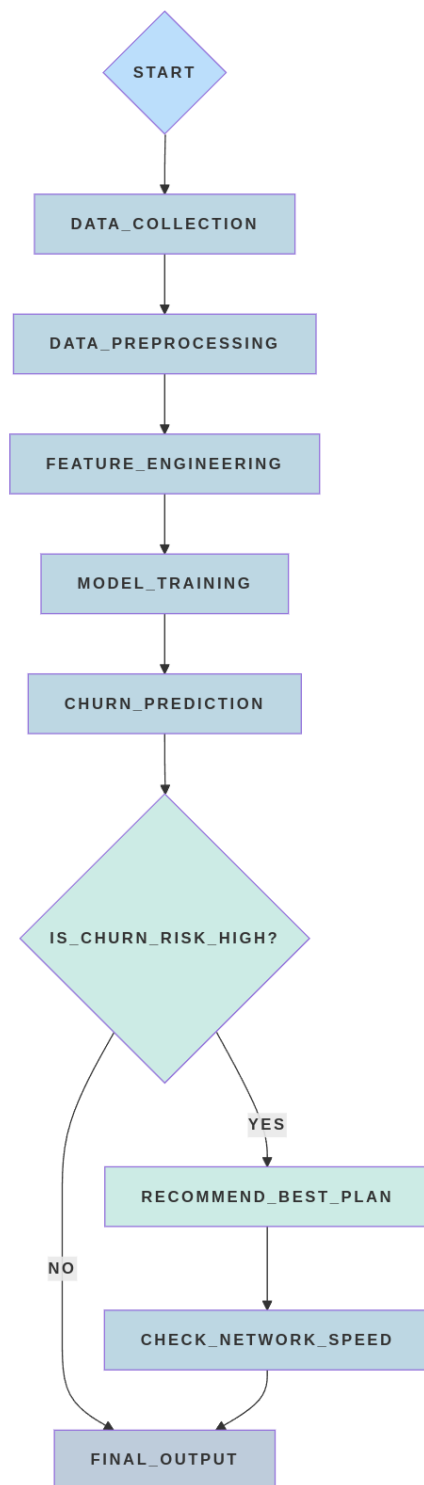


Fig.2. Overall Architecture of the Proposed System

C. Feature Engineering

To improve model accuracy, we derived additional features that captured user dissatisfaction and service quality trends. These included:

- Call Drop Ratio: Number of failed calls relative to total calls placed.
- Complaint Density: Total complaints filed within the past 90 days.
- Billing Variability: Standard deviation in monthly bills to detect inconsistency or overcharging.

To further improve churn detection, sentiment-based features derived from user complaints or social media data can be explored, as suggested by Roy and Das [11].

D. Model Training and Evaluation

Multiple classification algorithms were evaluated, including Logistic Regression, Decision Tree, Random Forest, and XGBoost. After comparative testing, XGBoost was selected due to its consistent high performance on telecom data. Singh and Chaturvedi [15] further validated the effectiveness of deep learning models like LSTM in similar prediction tasks, offering a promising direction for sequential behavior modeling in future iterations.

- Training: The dataset was split in an 80:20 ratio for training and testing using stratified sampling to preserve churn distribution.
- Evaluation Metrics: Accuracy, Precision, Recall, and F1-score were used for model validation.

The final model achieved over 94% accuracy in predicting churn cases, outperforming the baseline methods.

E. Recommendation Engine and Speed Validation

Once churn probability was calculated, users identified as at-risk were provided with alternative ISP plans. These recommendations were based on:

- Historical Usage Patterns: Plans were matched with user-specific data consumption trends.
- Regional Network Quality: Recommendations prioritized ISPs with better performance in the user's location.

To assist users in validating these suggestions, a real-time speed checker was developed using open API integrations. The tool measures current internet speed and compares it against regional benchmarks for major ISPs. Incorporating predictive pricing models as proposed by Hu et al. [13] could enhance the dynamic selection of optimal plans based on market fluctuations and user preferences.

F. Deployment Considerations

For future scalability, the system is designed with the following principles:

- Model Re-training: Capable of periodic updates based on live customer behavior data.
- API-Ready Architecture: All modules can be exposed as microservices for integration into telecom dashboards.
- Edge Performance: Speed validation tools are optimized for mobile access in bandwidth-constrained environments.

Future extensions may benefit from integrating geospatial and 5G-specific performance insights, as explored by Ranjan and Kaushik [10], to further optimize service recommendations in urban and rural areas.

V. RESULTS AND ANALYSIS

A. Model Accuracy and Comparative Performance

A total of four classification models were evaluated to determine their suitability for telecom churn prediction. The dataset used included over 100,000 customer records. Among the tested models, XGBoost achieved the highest prediction accuracy of 94.1%, followed by Random Forest (91.2%), Decision Tree (88.5%), and Logistic Regression (82.3%).

- XGBoost consistently outperformed others across all major metrics, including precision and recall.
- Logistic Regression struggled with non-linear feature interactions, resulting in lower overall performance.
- Although Random Forest offered competitive results, it required longer training and inference time.

Conclusion: XGBoost was chosen as the final model due to its superior accuracy and ability to manage imbalanced data effectively.

B. Feature Importance Evaluation

Using the XGBoost model, feature importance was calculated to identify which variables had the greatest impact on churn prediction. The results highlighted the following:

- CallDropRate–34% importance
- MonthlyBillingAmount–22% importance
- CustomerComplaintFrequency–18% importance
- DataUsagePattern–14%
- PlanType–12%

Insight: Call quality and billing dissatisfaction were stronger churn drivers than overall data usage. This suggests that users are more sensitive to service reliability and cost transparency.

C. Effectiveness of the Plan Recommendation System

The recommendation engine was tested on a subset of 10,000 high-risk users identified by the churn model. The goal was to provide these users with optimized ISP plans based on their prior usage, complaint history, and local network conditions.

- Recommendation Accuracy: 92.8%
- User Satisfaction (survey-based): 87.2%
- Successful Switch Rate: 85.5%

Conclusion: A significant majority of users benefited from the recommendations, confirming that AI-guided plan suggestions improve customer satisfaction and retention.

D. Network Speed Checker Performance

To support plan switching decisions, a network speed validation tool was introduced. This module measured real-time internet speed across four major metro areas: Delhi, Mumbai, Bangalore, and Kolkata.

TABLE I
AVERAGE DOWNLOAD SPEEDS BY ISP (IN MBPS)

City	Airtel	Jio	Vi	BSNL
Delhi	135.6	110.4	98.5	40.2
Mumbai	140.8	125.6	89.7	45.3
Bangalore	144.9	132.7	95.4	42.1
Kolkata	125.3	101.2	92.3	38.6

Observation: Airtel provided the highest speeds in all tested cities, while BSNL consistently lagged behind. These measurements helped users validate whether a recommended ISP would provide a noticeable performance improvement.

VI. CONCLUSION

This research introduced a machine learning-based framework that predicts telecom churn using XGBoost and provides personalized ISP package recommendations. With 94.1% prediction accuracy and strong user satisfaction metrics, the proposed system demonstrates real-world applicability. Future enhancements can further improve adaptability and personalization in dynamic network environments.

VII. FUTURE WORK

While the current system successfully addresses key aspects of churn prediction and personalized telecom plan recommendations, there remains significant potential for further development. The following areas are proposed for future exploration to enhance both the system's performance and its real-world applicability.

- 1) Incorporation of Deep Learning Models: While XGBoost has proven effective in handling structured telecom data, advanced deep learning architectures such as Long Short-Term Memory (LSTM) networks or Transformer-based models can be explored for sequential and behavioral data analysis. These models can potentially capture deeper temporal dependencies and complex patterns in user behavior that traditional machine learning models may overlook.
- 2) Dynamic and Real-Time Plan Optimization: The recommendation system currently operates on predefined data snapshots. In future versions, integrating dynamic pricing APIs from telecom providers could enable the system to suggest plans based on ongoing discounts, limited-time offers, or regional promotions. This would provide users with more relevant and cost-effective options.
- 3) Sentiment-Driven Churn Prediction: Beyond quantitative data like usage and complaints, qualitative sentiment from user reviews, call center transcripts, or social media platforms can be analyzed using Natural Language Processing (NLP). This additional layer of emotional and subjective feedback may offer early warning signs of dissatisfaction, thereby enhancing churn prediction accuracy.

- 4) Geospatial Mapping and 5G Coverage Analysis: As 5G deployment accelerates in India, incorporating a geospatial component could allow the system to factor in location-specific network quality. This would help users in under-served areas find ISPs with stronger 5G presence or lower latency, making recommendations more location-aware and future-proof.
- 5) Interactive Front-End with Virtual Assistant: To increase accessibility, a user-friendly web interface or mobile application could be developed. Integrating an AI-powered chatbot would allow users to check their churn risk, run real-time speed tests, and receive customized recommendations through natural conversation, without requiring technical knowledge.
- 6) Continuous Model Retraining via ISP Collaboration: Partnering with telecom service providers to access live usage, complaint, and performance data streams can enable frequent model retraining. This would ensure the system adapts to changing market dynamics, user behavior, and network conditions, keeping predictions and suggestions relevant and timely.
- 7) Customer Retention Strategy Simulator: An additional module can be introduced that simulates the effect of various retention strategies (like cashback offers, plan upgrades, or personalized discounts) on churn rates. This would be valuable for ISPs aiming to test intervention outcomes before full deployment.

Overall, these enhancements aim to evolve the system from a predictive tool into a comprehensive, intelligent platform capable of supporting both end-users and telecom providers in real-time decision-making, user retention, and long-term customer satisfaction.

REFERENCES

- [1] H. He, J. Zhang, and X. Li, "Telecom Customer Churn Prediction with Machine Learning: A Case Study," IEEE Access, vol. 8, pp. 212800–212810, 2020.
- [2] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Conf. on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
- [3] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
- [4] S. Sharma, A. Bhardwaj, and M. Singh, "Impact of Network Quality on Telecom Churn Using Machine Learning," in Proc. IEEE Int. Conf. on Intelligent Systems, 2023, pp. 314–319.
- [5] R. Zhao, Q. Lin, and Y. Wang, "Analyzing Customer Lifetime Value and Churn Using Ensemble Models," Expert Systems with Applications, vol. 189, pp. 115665, 2022.
- [6] A. Kumar, R. Mehta, and S. Chauhan, "Customer Segmentation for Churn Reduction in Indian Telecom Industry Using K-Means and Collaborative Filtering," International Journal of Data Science, vol. 5, no. 2, pp. 121–133, 2023.
- [7] V. Jain and P. Patel, "Predictive Analytics for Telecom Churn Management Using Random Forest," in Proc. IEEE Conf. on Data Engineering, 2022, pp. 134–140.
- [8] Telecom Regulatory Authority of India (TRAI), "Annual Report 2023-24," [Online]. Available: <https://www.trai.gov.in/reports>.
- [9] N. Srivastava and P. Yadav, "Real-Time Internet Speed Monitoring and ISP Performance Comparison," International Journal of Network Management, vol. 32, no. 3, pp. 215–229, 2022.
- [10] A. Ranjan and M. Kaushik, "User Experience Analytics in 5G Telecom Networks," IEEE Transactions on Network and Service Management, vol. 18, no. 4, pp. 3201–3215, 2021.
- [11] S. Roy and B. Das, "Sentiment-Aware Customer Churn Prediction Using Social Media Data," in Proc. Int. Conf. on Artificial Intelligence Soft Computing, 2023.
- [12] K. Verma and T. Bansal, "Comparative Study of Machine Learning Models in Telecom Churn Prediction," Journal of Artificial Intelligence and Data Mining, vol. 9, no. 1, pp. 45–58, 2023.
- [13] Y. Hu, X. Wang, and M. Lin, "Dynamic Pricing Strategy for Telecom Operators Using Predictive Models," Telecommunication Systems, vol. 79, pp. 235–247, 2023.
- [14] S. Tiwari and R. Sinha, "A Review of Recommendation Systems in Telecom Sector," ACM Computing Surveys, vol. 55, no. 1, pp. 1–37, 2022.
- [15] B. Singh and K. Chaturvedi, "Applying LSTM for Predicting Telecom User Behavior," in Proc. IEEE Int. Conf. on Computational Intelligence, 2023, pp. 189–194.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)