OPEN ACCESS

Volume: 5

Issue: 1

Month: January

Year: 2026

ISSN: 2583-7117

Published: 23.01.2026

Citation:

Md Shahib Khan, Sandeep Kumar Sahu "Analyzing Customer Churn in Telecommunications: Insights from Data Patterns and Trends" International Journal of Innovations in Science Engineering and Management, vol. 5, no. 1, 2026, pp. 73-80.

DOI:

10.69968/ijisem.2026v5i173-80

Analyzing Customer Churn in Telecommunications: Insights from Data Patterns and Trends

Md Shahib Khan[1], Sandeep Kumar Sahu[2]

[1]*Research Scholar, SAM global university*

[2]*Assistant Professor, SAM global university*

Abstract

Customer churn has become a major issue for the telecom industry as the cost of getting new customers is much higher than that of keeping the existing ones. The patterns of churn are not straightforward and are affected by demographic, behavioral, and regional factors, which in turn interact in subtle ways. A thorough and descriptive analysis is carried out in this research using a real-world telecom dataset to identify the impact that a person's age, gender, state or region, tenure, calls made, SMS usage and data consumption have on a person's decision to churn. The dataset goes through cleaning, processing and enrichment with engineered features in the form of tenure and age groups which allows for a deeper analytical insight. This study that focuses on three core research questions uncovers the following very clear churn patterns: young and mid-age groups have a higher propensity to churn, low-usage customers quit faster, and some states have consistently high churn rates due to regional competition. The results emphasize the significance of having the basic descriptive analytics before moving on to predictive modeling and also point out the demographic, behavioral, geographic dimensions which are vital for telecom retention strategies.

Keywords; Customer churn, telecom analytics, descriptive analysis, usage patterns, tenure behavior, customer retention.

INTRODUCTION

Telecom companies are grappling with a constantly changing customer base primarily due to competition, poor customer service, price changes, and the ease of mobile number portability. All these factors combined, the industry turns out to be very fluid where customers can easily switch their providers. Therefore, churn analysis has to be done accurately and continuously to keep the company profits long and healthy. For every subscriber retained, income lost from the newly acquired ones can be considered; hence the process of customer loss and early sign of churn detection is very important [1]. The traditional churn detection systems which depended largely on fixed rules and human evaluations have lost their value in the present digital ecosystem. The modern customers are changing their usage habits, demands, and service interactions very quickly, thus requiring data-driven techniques that will be able to gain deeper insights into customer behavior [2], [3].

Numerous studies reveal how demographic factors play a critical role in determining churn behavior. The variable of age has been reputed as the most potent factor in this case, with the youth being the most vulnerable to the changing of the provider due to their experimenting nature, their being technologically advanced and promotional offers having more impact on them. The youth are more likely to compare services actively and have less patience with service-related problems. On the other hand, the elderly customers are generally more loyal unless the provider has very severe service issues or they are very dissatisfied [4]. Churn trends based on gender are less uniform but still appear in the literature, as it is believed that preferences for communication, service expectations, and usage habits might be slightly different for men and women in certain geographical and socioeconomic areas [5]. By having these demographic clarifications, the telecommunication giants can pinpoint the exact customer segments that need the most attention in their retention strategies.

Besides demographic characteristics, behavior concerning usage has been continually one of the essential factors in churn evaluation process. A lot of research supports the idea that among all the changes in usage activity, especially reductions in call frequency, SMS volume, and data consumption, the lower usage activity is the strongest and most immediate sign of customer disengagement [6]. Customers who show such behavioral patterns are often the ones who would soon be classified as churned, thus, monitoring such metrics is vital for the proactive detection of churn cases. On the other hand, it is also possible for high data or call usage customers to leave the company due to network issues or high rates which implies that both low and high usage patterns need to be interpreted considering the context. The continuous backing of usage indicators' predictive power in different studies emphasizes their role not only in churn analysis but also in churn understanding [7].

Churn behavior within telecoms companies is further complicated by the regional differences. A number of factors, such as the availability of the network, different pricing strategies in regions, the population density, competition locally, and the socio-economic conditions, are responsible for the huge differences in churn rates even within one state or city. For instance, in some areas the churn rates are higher because of the aggressive competitor expansion or network problems while others enjoy more stable customer bases where brand loyalty is determined by the presence of the brand, and the reliability of the service [8]. Such geographical disparities call for localized churn analysis and not for the assumption of a uniform pattern across the entire customer base.

Factors mentioned above prompt the current research to apply a structured descriptive analytic approach in order to tackle customer churn through three main research questions. First of all, demographic and socioeconomic characteristics are looked at and their effect on likelihood of churn is measured. The second one looks at usage behaviors, such as calls made, SMS sent, and data consumption to see if there is a difference between churned and non-churned customers. The third one looks at regional and temporal churn trends and assesses how churn varies with states and telecom partners over time. The study by answering these questions provides complete understanding of churn behavior and also lays down the knowledge base which will be explored for predictive modeling and targeted retaining customers strategies in the future.

Related Work

Customer churn is a subject that has been investigated for ages in the telecommunications industry, as the direct impact of customer churn on the operator's revenues, customer lifetime value, and market position is very significant. The literature has been, moreover, quite vocal in stating that churn dynamics analysis is the main and the only way for a telecom operator to improve retaining and servicing quality. Initially, the churn problem was stated to be complicated with different multi-dimensional factors such as demographic, behavioral, and even regional market aspects, thus, the whole matter was heavily relying on descriptive analytics, which was to be considered as the first step in the churn research pipeline providing the interpretive layer that was necessary before deploying advanced machine learning models [9].

The demographic factors have always been the subjects of most of the research done on customer churn. Age, for example, is usually found to be the main underlying reason for differences in customer behavior with the result that the younger generation is always associated with high churn rates as they are always after the best deal and also very open to changing their service provider thus making their customer loyalty very short. These findings have been further confirmed by the reviews of other studies which recognize that younger demographics are the most responsive to promotional stimuli and therefore the most likely to switch to the providers who are offering the best value for money [10]. On the other hand, gender has been another factor but the literature is inconclusive about its role, mainly because of the influence of outside factors like cultural norms, service availability, and regional economic conditions. In some cases, studies suggest that women might have slightly different patterns of communication and, hence, implying different churn drivers, while others claim that when considering usage behavior the gender difference in churn is very small [11].

One of the major factors that have been fundamental in the analysis of telecom churn is the behavioral usage patterns. It is a fact that research has proved several times that the decrease in call activity, SMS use, and data consumption are the signs of disconnection and the collapse of service dependency. Customers who have shown a decrease in usage are more on the verge of leaving the company in the following billing cycles. This makes these behaviors very important for churn forecasting models. Data mining and machine learning studies have additionally demonstrated that behavioral characteristics are often more predictive than just demographic factors [12]. Cross-operator comparisons among telecom providers strengthen the claim that usage-based indicators such as data consumption patterns and call duration metrics are still the main sources in understanding churn behavior across different customer segments [13]. These customer behavioral patterns also indicate the changing trend of communication, especially with the advent of smartphones and digital services that keep reshaping customer expectations.

Churn analysis is made even more complicated by the inclusion of regional and socioeconomic factors. One of the main reasons for the differences in churn rates over different geographical areas is the different levels of network coverage, market competition and the economic conditions of the states. A number of studies that have prevalence on regional churn indicate that the telecommunication market is not one overall unit and factors like the rural–urban divide, the density of the network and the regional discounts on services play a great role in determining customer loyalty [14]. In addition, some research indicates that certain states or provinces have a constant high churn rate due to the lack of good network infrastructure or too much market competition, while economically stable or urbanized areas might have good service continuity and thus low churn rates [15]. The differences in the churn rates call for the use of state-level or region-specific churn models, as opposed to the application of the nationwide generalized predictions.

Nowadays, machine learning has emerged as the primary technology in churn analysis and research, providing better precision and non-linear relationship modeling. Various methods like logistic regression, decision trees, random forests, and even deep neural networks are used in the analysis of telecom churn datasets. A number of researches claim that the use of ensemble learning approaches leads to better performance as these methods are able to combine and fine-tune multiple weak learners [16]. Integration of clustering techniques with supervised learning has been one of the most successful hybrid models in partitioning the customer base and detecting the at-risk groups based on the same behavioral or demographic traits. Data mining–inspired strategies also offer controlled ways of recognizing customer state changes and showing how the service migration decision is influenced by factors [17].

In the past few years, the use of advanced analytical techniques like survival analysis has become one of the methods to represent the time factor in churn. Survival-based models not only determine whether a customer is going to churn or not but also give a probable date for the churn event, hence, providing the telecom operators with useful information about the timing for intervention [18]. Time-to-event analyses are mainly employed for classifying customers according to their duration of stay and spotting the points of change in customer satisfaction. On the other hand, real-time churn scoring models that can help telecom operators instantly identify the churn risk have been created by using live data from customer interactions, network incidents, and service usage logs [19]. The mentioned developments are all part of the proactive retention strategies which involve giving timely incentives, marketing targeted, and, customer engagement personalized—thus, making it all easier for the customers and businesses.

To sum up, the current literature shows that customer turnover is a complex phenomenon and it is affected by various factors such as changes in demographics, customer behavior, location and time of the year. The combination of classic descriptive analytics and state-of-the-art machine learning techniques has pushed the boundaries of the discipline considerably, but basic exploratory analysis is still a must. This study goes along with these directions of research by looking at demographic trends, differences in usage behavior and regional churn patterns and then moving on to predictive modeling applications.

METHODOLGY

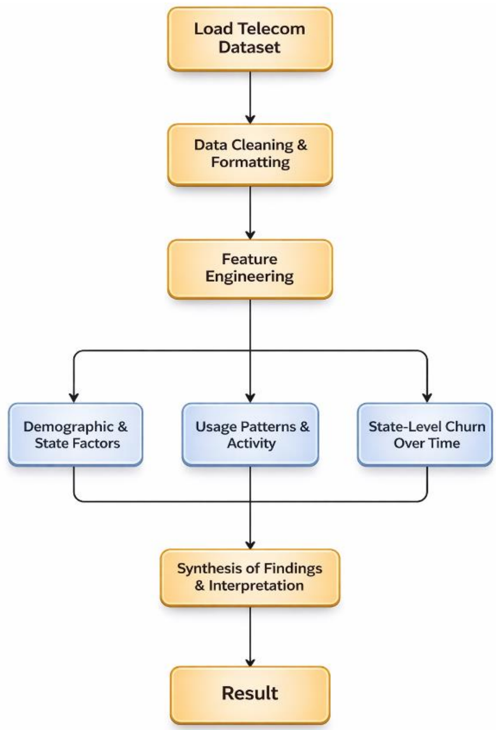

Figure 1 Methodology Flow Chart

The research methodology is built to address three main research questions and each question is answered through a structured set of analytical procedures. The procedures commence with the preparation of the dataset, continue with feature engineering and exploratory analysis that are in line with demographic, behavioral and regional aspects of telecom churn

Data Loading and Initial Inspection

Initially, the dataset is populated in the analytical setting, and its organization, types of columns, and summary statistics are scrutinized. The step consists of looking at the data's shape, checking the sample rows, viewing the distributions of numbers, and going through the details of the categorical and numerical features. Missing data is detected, counted, and ranked according to the extent of cleaning required.

Table 1 Dataset

Cust-Id	T_Partner	Gen	Age	State	Date_Of_Registration	Num_Dependents	Estimated_Salary	Calls_Made	Sms Sent	Data Used	Churn
1	Reliance Jio	F	25	Karnataka	01-01-2020	4	124962	44	45	-361	0
2	Reliance Jio	F	55	Mizoram	01-01-2020	2	130556	62	39	5973	0
3	Vodafone	F	57	Arunachal Pradesh	01-01-2020	0	148828	49	24	193	1
4	BSNL	M	46	Tamil Nadu	01-01-2020	1	38722	80	25	9377	1
5	BSNL	F	26	Tripura	01-01-2020	2	55098	78	15	1393	0

Data Cleaning and Formatting

Data cleaning processes are implemented after the initial inspection; the dataset then being prepared for analysis. Missing values are considered and solved if necessary. The churn variable undergoes standardization and is binary coded, thus implying the same representation to the following grouping and analysis. The date_of_registration field undergoes transformation into a datetime format for calculating tenure. The columns get standardization in terms of formatting to avoid variations in the dataset.

Feature Engineering

New variables are created to increase the analytical depth. Age groups are formed demographically by segmenting age values into meaningful categories. Customer tenure in months is calculated based on the registration date. The created variables allow the continuous data to be transformed into interpretable categories that facilitate demographic and behavioral analysis. The categorical variables such as gender, state and telecom partner are prepared for the grouping operations as well.

Research Question 1: Demographic and Socioeconomic Factors

In this study, to explore the impact of demographic and socioeconomic characteristics on churn, the dataset was first divided according to age group, gender and state. Subsequently, for every segment, the total customers, churned customers and the churn rate were determined. The outcomes were presented through heatmaps, bar charts and state-wise segment distributions. This method has been able to reveal areas of high-churn by age groups, differences in churn by gender and states with high churn, thus aiding the detection of demographic and socio-economic drivers.

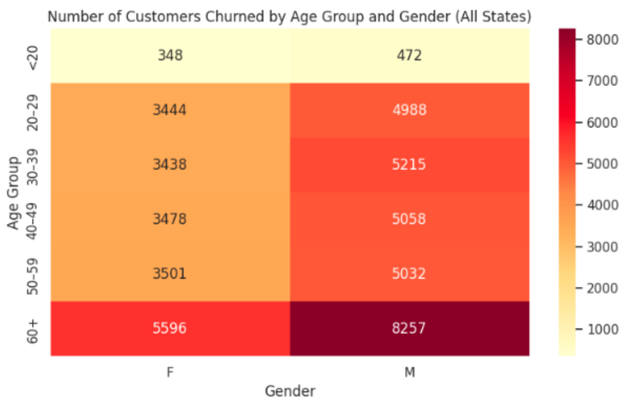

Figure 2 Number of Customers Churned by Age Group and Gender (All States)

E. Research Question 2: Usage Patterns and Activity Levels

The second research question investigates the link between customer usage behavior and churn. In this regard, only the dataset that comprises churned customers is considered for analysis. Segmentation of the customers based on their tenure at the company is done, and for each segment, average calls, SMS and data usage are calculated, thereby showing the changing engagement over time. The usage patterns are represented graphically through line plots.

Customers showing high risk of churn are singled out by selecting those whose usage metrics are in the bottom quartile, thus giving a glimpse into behavioral disengagement before the actual churn.

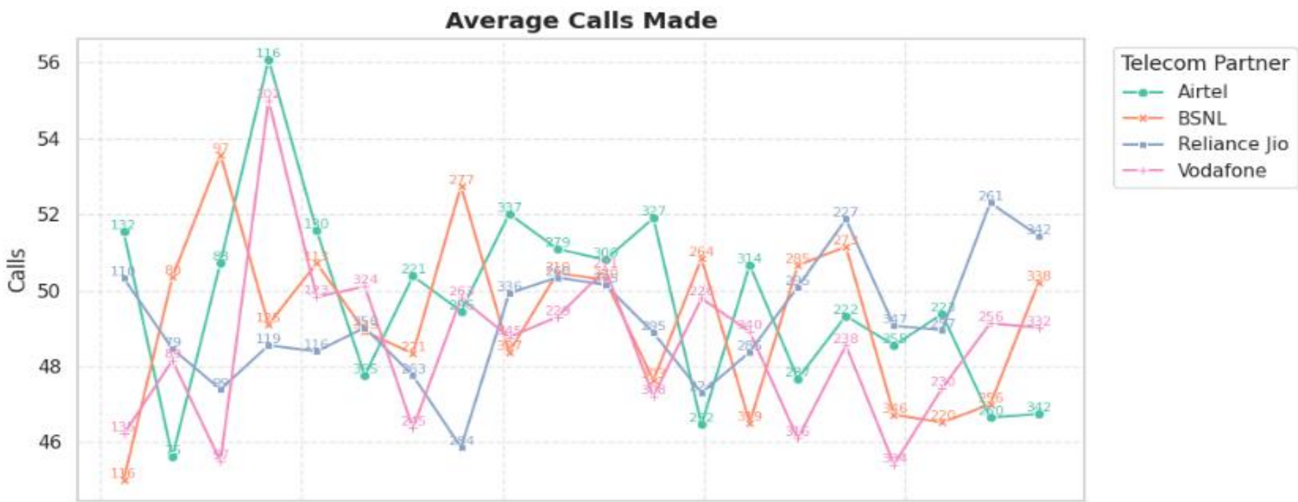

Figure 3 Average Calls Made

Research Question 3: State-Level Churn Over Time

The third research question is about to discuss variances of churn in different states and with different telecom partners over time. The registration year is taken from the dataset, which allows grouping of data temporally. Customers who have churned are categorized by year, state and partner, this creates time-series views of churn trends. Visualizations show the partner-specific churn patterns in different states from one year to another. On the other hand, the trend lines with annotations show the states with increasing or decreasing churn which indicates the changes in the regional market. This step-by-step approach brings together demographic segmentation, behavioral trend analysis, and temporal-regional mapping for a comprehensive churn investigation framework. Thus, it enables thorough answering of each research question and insights that are directly generated from the dataset and analytical methods employed.

RESULTS AND DISCUSSION

The analysis results show numerous significant trends that clarify customer churn behaviors in the telecom industry. The discovery of each heading is presented in paragraph format below as you requested.

A. Demographic Factors Influencing Churn

The demographic analysis shows that the separation of customers into different age groups yields a different picture of churn rates, with the highest churn rate being at the young adult and early middle-age customer categories. Customers aged 20 to 29 years old are likely to exhibit higher churn rates and the reason could be that they are more sensitive to price or it could be that attractive promotional offers from competing providers are available which they want to take advantage of. The 30-39 age group also switches over considerably, which means that working professionals may change their service providers when they face poor network performance or find more attractive data plans in other places. The differences based on gender are small, but in some states, there is a little more churn among women customers which might be the result of local competition or factors like availability of the service etc. The demographic distribution by state indicates that there are certain areas where the churn rate is consistently high, especially in places where there is less network and more competition between telecom companies.

Usage Behavior Analysis

One of the most powerful evidence-based indicators for customer churn is usage behavior. The churning customers are usually very less active in using the telecom services compared to the others who are still with the provider. The frequency of their calls is less, the number of sent messages is negligible, and their data usage is only a fraction of that of the loyal customers. This implies that disengaged customers would probably leave as they have fewer ties with the telecom service. The usage patterns based on customer tenure also exhibit strong connections with churn, as customers with short tenure usually demonstrate limited-service usage. The inverse relationship of tenure and usage suggests that new customers, who are still judging the quality of the service, may change their service provider instantly if they do not find the service to their satisfaction.

Customers with longer tenure indicate very high usage which is a sign of one being familiar with the service hence loyalty. The smooth and steady upward slope of usage with the increase of tenure is an evidence for the belief that long-term customers are the major contributors to the stability of the telecom industry.

High-Risk Churn Segment Identification

High-risk segments were identified and among them, it was revealed that churn was mainly among the customers who belonged to the lowest quartile service usage. The customers who seldom call, send SMS or consume data are the main part of the churn among the customers. Their low engagement signifies that they are not very dependent on the service, thus making them easily switch to other providers based on the price or promotional offers. These indicators coupled with short tenure turn out to be very predictive of the customer being at risk of churn. Customers with both low usage and short tenure are considered the most at-risk and need special attention through targeted retention strategies if telecom operators are to manage to reduce overall churn. Moreover, the existence of such a distinct cluster emphasizes the importance of descriptive analytics in the process of early-stage churn detection.

Table 2 High-Risk Churn Customer Summary

Tenure Bin	Avg Calls	Avg SMS	Avg Data	Customer Count
(-0.0478, 4.78]	37.72	19.79	4206.32	468
(4.78, 9.56]	42.67	18.51	3938.18	491
(9.56, 14.34]	40.04	19.46	4136.15	1115
(14.34, 19.12]	39.43	19.48	4192.56	1205
(19.12, 23.9]	40.52	18.86	4071.07	1362
(23.9, 28.68]	41.20	18.90	4099.11	1404
(28.68, 33.46]	39.69	19.42	4036.25	1309
(33.46, 38.24]	40.75	19.16	4089.46	1276
(38.24, 43.02]	38.28	19.33	4266.80	1354
(43.02, 47.8]	39.91	19.30	4043.16	1376

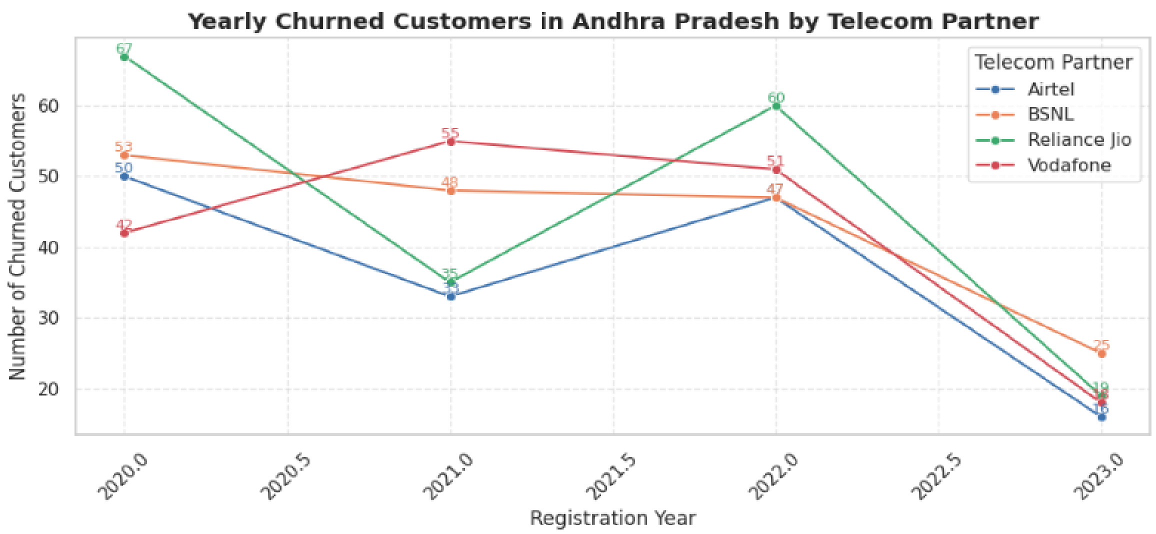

Figure 4 Yearly Churned Customers

Figure .5 Yearly Churned Customers in Andhra Pradesh by Telecom Partner Regional and Partner-Based Churn Trends

The analysis of churn at a regional level carried out throughout the states shows that different geographical areas have quite distinct characteristics. Geographic regions characterized by the fierce competition among the telecom partners usually reveal higher churn rates, conversely, the less competitive and more stable areas have lower churn rates. The year of registration being the basis of the temporal analysis of churn, it is evident that the states from which the customers migrate have had their trends declining because of such factors as efficient network infrastructure and improved customer service. On the other hand, the other parts of the world still have changeable churn levels that are usually the result of aggressive promotional activities, or, simply put, customers' lack of service reliability. The companies competing in the telecom sector serve as a good indicator of the fact that no single operator has the majority of the market in all regions. An operator can be very successful in big cities but on the other hand very unsuccessful in rural states, while another one can have similar churn patterns throughout several years. These differences between regions do not only highlight the need for localized business strategies but also the necessity of state-specific retention campaigns.

CONCLUSION

The researchers present a detailed descriptive analysis of the telecom customer churn by looking into the demographic, behavioral and regional aspects. It has been shown that churn is not a random occurrence, but rather a consequence of the factors that can be identified. The younger customers and the ones in mid-career positions are the main groups that are subject to the highest churn rates, while low usage is an indicator of early detachment from the service. The telecom industry in the states marked by a high level of competition or poor network quality is suffering from the highest customer churn, while the partners of telecom companies show a different performance in each region. The insights obtained through this research will serve as a strong basis for future efforts in predictive modeling. Operators in the telecom industry will be able to implement more targeted retention strategies by knowing the demographic, behavioral, and regional aspects. The future work could consist of combining machine learning models, creating churn prediction pipelines, and carrying out survival analysis to determine when the churn will occur. This descriptive research brings attention to the need of carrying out feature-rich exploratory data analysis in the telecom churn research and also affirms its necessity in the data-driven customer retention framework development process.

REFERANCE

[1] M. Bagri, et al., "Churn analysis in telecommunication industry," 2018 International Conference on Automation and Computational Engineering (ICACE), IEEE, 2018.

[2] M. Al-Mashraie, S. H. Chung, and H. W. Jeon, "Customer switching behavior analysis in the telecommunication industry via push-pull-mooring framework: A machine learning approach," Computers & Industrial Engineering, vol. 144, 2020, p. 106476.

[3] Dahiya, Kiran, and Surbhi Bhatia. "Customer churn analysis in telecom industry." 2015 4th International conference on reliability, infocom technologies and optimization (ICRITO)(trends and future directions). IEEE, 2015.

[4] Kayaalp, Fatih. "Review of Customer Churn Analysis Studies in Telecommunications Industry." Karaelmas Science & Engineering Journal/Karaelmas Fen ve Mühendislik Dergisi 7.2 (2017).

[5] Ahn, Jae-Hyeon, Sang-Pil Han, and Yung-Seop Lee. "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry." Telecommunications policy 30.10-11 (2006): 552-568.

[6] Gürsoy, Umman Şimşek. "Customer churn analysis in telecommunication sector." İstanbul Üniversitesi İşletme Fakültesi Dergisi 39.1 (2010): 35-49.

[7] Mbarek, Rahmoun, and Ya Baeshen. "Telecommunications customer churn and loyalty intention." (2019).

[8] Pandey, Yash, Roshan Jha, and U. Umamaheswari. "Customer churn analysis in telecom organization." Journal of Positive School Psychology (2022): 5475-5488.

[9] Ribeiro, Hugo, et al. "Determinants of churn in telecommunication services: a systematic literature review." Management Review Quarterly 74.3 (2024): 1327-1364.

[10] Dash, Manoj Kumar, and Susmita Dash. "A systematic review of literatures on customer Churn Analysis in the Telecommunication Industry." Orissa Commerce Association (2019): 24.

[11] Zdziebko, Tomasz, et al. "Optimizing customer retention in the telecom industry: A fuzzy-based churn modeling with usage data." Electronics 13.3 (2024): 469.

[12] Masarifoglu, Melik, and Ali Hakan Buyuklu. "Applying survival analysis to telecom churn data." American Journal of Theoretical and Applied Statistics 8.6 (2019): 261-275.

[13] Edwine, Nabahirwa, et al. "Detecting the risk of customer churn in telecom sector: a comparative study." Mathematical Problems in Engineering 2022.1 (2022): 8534739.

[14] Keramati, Abbas, and Seyed MS Ardabili. "Churn analysis for an Iranian mobile operator." Telecommunications Policy 35.4 (2011): 344-356.

[15] Bilişik, Özge Nalan, and Damla Tuğba Sarp. "Analysis of customer churn in telecommunication industry with machine learning methods." Düzce Üniversitesi Bilim ve Teknoloji Dergisi 11.4 (2023): 2185-2208.

[16] Pejić Bach, Mirjana, Jasmina Pivar, and Božidar Jaković. "Churn management in telecommunications: Hybrid approach using cluster analysis and decision trees." Journal of Risk and Financial Management 14.11 (2021): 544.

[17] binti Oseman, Khalida, Norazrina Abu Haris, and Faizin bin Abu Bakar. "Data mining in churn analysis model for telecommunication industry." Journal of Statistical Modeling and Analytics Vol 1.19-27 (2010).

[18] Rahman, Hafiz, et al. "IBM Telco Customer Churn Prediction with Survival Analysis." Proceedings of the International Conference on Advanced Technology and Multidiscipline (ICATAM 2024). Vol. 245. Springer Nature, 2024.

[19] Tamuka, Nyashadzashe, and Khulumani Sibanda. "Real time customer churn scoring model for the telecommunications industry." 2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC). IEEE, 2020.

.