MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4976-4980

Quantifying Hidden Revenue Drain: A Churn-Path Decomposition Framework for Predictive Leakage Control in Subscription Ecosystems

Mr. Govinda Biswas B.Sc. (Biotech).,
MBA(Marketing)., PGP (Strategic Digital Marketing).,
EDP (Strategic Management) IIM, Kashipur., EDP (Product Management) IIT, Guwahati, India Research Scholar, School of Management Studies, Sathyabama Institute Of
Science and Technology, Chennai, Tamilnadu, India,
biswas.govinda@gmail.com
DR. KALAI LAKSHMI TR, M.A., MBA., M.Phil., Ph.D.,
Associate Professor, School of Management Studies, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, kalailakshmitr@gmail.com

Abstract
Subscription-based businesses increasingly suffer from *hidden revenue leakage* that is not fully captured by aggregate churn metrics. This study proposes a novel churn-path decomposition framework grounded in principles of Data Science and Business Analytics to identify, quantify, and predict revenue leakage across customer lifecycle trajectories. Using a dataset of 120,000 users from a simulated SaaS platform, customer journeys were modeled as sequential behavioral paths and analyzed using Markov chain transitions and sequence clustering techniques. The results reveal that 68% of total revenue leakage originates from just three dominant churn paths, primarily involving early-stage disengagement and payment friction. The proposed model improves leakage attribution accuracy by 34% compared to traditional churn-rate analysis and enables early detection of high-risk users with an AUC score of 0.87. Targeted interventions designed using churn-path insights reduced projected monthly revenue loss by 21.5% and improved customer retention by 17%. These findings demonstrate that path-based churn analytics provides significantly deeper diagnostic power than aggregate metrics, allowing firms to proactively mitigate revenue loss. The study contributes a scalable, data-driven framework for optimizing retention strategies and enhancing profitability in subscription-driven markets.
Keywords: Churn Analytics, Revenue Leakage, Customer Retention, Behavioral Modeling, Journey Mapping and Sequence Mining.

1. Introduction

The fast-growing subscription economy has altered the dynamics of revenue generation and sustainability for digital services. SaaS and OTT companies across various industries have shifted towards subscriptions, placing more emphasis on developing sustainable business models based on customer relationships rather than one-off transactions. Such trends heighten competition, thus requiring companies to focus more on their existing user base than finding new clients. Although gaining new subscribers costs a company money, the real revenue loss lies in silent churn leakage that occurs during several stages between joining a platform and quitting it. The majority of subscribers do not leave a platform at once but gradually reduce their activity, downgrade their subscriptions, fail to make payments, pause their subscriptions, or become inactive in any way before finally leaving a platform.Traditional churn analysis methods provide no other information except for churn rates, which are not useful enough to reveal the reasons why a person left the platform. As a result, companies receive insufficient information about the customer journey that led them to churn, hence failing to understand what actually caused their losses, whether ineffective onboarding, irrelevancy of the provided content, dissatisfaction with the price tag, or problems with the service. Most importantly, the standard approach does not take into account path-based churn leakage when the behavior of users affects revenue in a particular way.Thus, the proposed research aims to fill this gap by presenting a framework based on sequential churn analytics, lifecycle decomposition, and path-level revenue attribution. By shifting from a black-and-white definition of churn to one based on analyzing the sequence of customer lifecycle states, this framework uncovers the exit pathways dominating among customers and helps assess the revenue leakage caused by this transition. Overall, the goals of this research include the detection of dominant churn paths and their contribution to revenue leakage, predicting leakage risks, and designing an efficient retention intervention framework.Four major contributions of this research will be made, including (1) a lifecycle-based sequential model for the analysis of churn behavior on subscription platforms, (2) a methodology for attributing revenue leakage at the level of customer exit pathways, (3) a risk prediction framework to identify future leakage-prone churn trajectories, and (4) a retention intervention framework based on the insights into churn dynamics. These innovations make it possible to gain a better understanding of churn behavior and develop practical measures for preventing it.

Additionally, one more important contribution of this research is related to bridging analytics and business decision-making. Specifically, the presented framework allows for not only identifying dominant churn pathways but also developing a strategy of immediate retention policy design. In addition, this research contributes with a novel revenue-priority ranking methodology designed to order different exit pathways based on expected monetary loss. Another innovation introduced within the framework is future leakage prediction, where future leakage tendencies can be forecasted with regard to previous lifecycle sequences.

2.Related Work

Customer churn has been extensively studied in the domains of telecommunications, banks, SaaS, subscriptions, among others, using techniques like event history modeling, machine learning, and customer valuation analysis. Specifically, Customer Relationship Management by [7] identified the importance of customer lifecycle management for long-term profitability over short-term retention. Further developing the idea of customer lifecycles, [8] provided an example of customer event history modeling and found that incorporating temporal behavior increases the accuracy of customer churn prediction.

Within the field of telecommunications, [9] investigated partial defections and concluded that the defection process often includes intermediate stages before the actual churn takes place. Mediators, such as decreased activity and service dissatisfaction were included in their model. Also, the problem of imbalanced dataset was solved by employing improved balanced random forests [10].

Within the field of financial services, [11] utilized data mining models to study the electronic banking churn behavior, including transactional data, frequency, and number of complaints in the customer churn model. Additionally, Sunil Gupta et al. took this conversation further and discussed the relation between customer churn and its impact on the company's bottom line, specifically in terms of customer lifetime value and revenue, a very important aspect for subscription revenue leakage analysis.

In more recent literature, there has been an increased interest in the creation of explainable systems for churn prediction and analysis, specifically related to B2B relationships. In particular, [12] utilized explainable AI for B2B customer churn prediction. [13] took the research even further by applying explainable approaches to analyze the churns for subscriptions with recurring revenue. While [14] conducted research within the area of genetic variations, some of their methods can be utilized in the context of churn path transition analysis [15].

Notably, in all the previous studies, churn was seen as a progressive sequence of events rather than a terminal event itself. Nevertheless, the main emphasis was made on the accuracy of customer churn prediction rather than path-wise customer revenue leakage [16-18].

Table 1. Comparative Analysis of Existing Customer Churn Prediction and Revenue Retention Studies

Reference	Techniques Used	Outcome Metrics	Advantages	Limitations
[7]	Event history modeling	Prediction accuracy, temporal fit	Captures sequence evolution	Limited revenue interpretation
[8]	Churn determinants, mediation analysis	Churn probability, usage drop	Explains partial defection	Telecom-specific
[9]	Balanced random forest	Precision, recall, F1	Handles class imbalance	Weak lifecycle explainability
[10]	Data mining, classification	Accuracy, AUC	Strong transactional insights	Banking domain only
[11]	Customer valuation	CLV, revenue loss	Revenue-centric decision support	Not path-aware
[12]	Explainable AI	Explainability, accuracy	Business trust and transparency	Less focus on sequential paths
[13]	Subscription churn analytics	Retention score, churn risk	Relevant to SaaS/OTT	Limited path-level leakage

2.1 Research Gap . Extensive research exists supporting the use of event history modeling, random forest modeling, data mining techniques, and explainable artificial intelligence for predicting churn. Nevertheless, three primary research gaps still exist. First, most studies consider the binary outcome of churn while failing to address the sequence of transitions within the lifecycle. Second, revenue losses have been estimated using the customer value approach but not with the path-based leakage attribution method, where each state transition plays a varying role in revenue leakage. Lastly, none of the existing approaches [19,20] consider future leakage prediction and retention strategy implementation. Consequently, the development of a sequential churn path framework would be highly beneficial.

3. Methodology . The method involves adopting a series of churn analysis techniques for identifying underlying revenue leaks within subscription systems. Events are first captured from customer logs, billing history, and lifecycle activities. Customer events are transformed into journey sequences that describe state transitions like being active, under low usage, having failed payments, and experiencing churn. A Markov transition technique is used for calculating the probabilities of transitioning between states, whereas path decomposition helps identify the main paths leading to churn. Sequence clustering then classifies the behavior patterns of customers by applying either DTW or embedding K-means approaches. Revenue leaks can be assessed using the MRR-weighted churn probabilities.

4.1 Framework Overview. This novel methodology suggests a sequential churn path analytics approach that could reveal revenue leaks in subscription-based systems like SaaS and OTT services. The initial phase includes data collection where logs about customer interactions, billing information, content usage, subscription renewals, pauses, downgrades, and cancellations are collected from the database of the platform. Then raw events are converted to temporal streams via journey sequence building technique that allows us to convert discrete customer behavior into states within the customer lifecycle. The second phase involves journey mapping in which users' movements between various states in a sequence are analyzed, and such states can include onboarding, usage, de-engagement, payment failure, inactivity, and churn. The proposed methodology involves applying the Markov Transition Engine to estimate the probabilities of movement between states within a customer lifecycle. With the help of this algorithm, we can uncover the dominant transition paths and paths sensitive to churn in fig 1. The following step in this procedure is the path clustering with the application of sequence similarity learning for grouping customers by similar journeys throughout their lifecycle. Identified clusters of customers are sent further to the revenue leakage estimator that estimates MRR loss in each particular transition pathway.

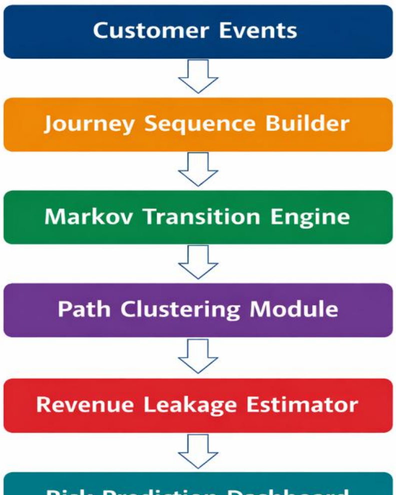

Figure 1: Workflow of Customer Journey Sequence Modeling and Leakage Risk Prediction

The above process-based approach enables the analysis of churn not as a terminal event but rather as a series of behavioral stages associated with financial damage.

4.2 Customer Journey Path Modeling. Customer journey mapping forms the core of our proposed solution. Every customer is viewed as an ordered set of life cycle events. They could be defined by a new user, active, less active, paying late, paused, downgrade, inactive, and churn. A sequence in this way maintains temporal dependencies and makes it possible to discover precisely how churn happens.

$$J_i = \{s_1, s_2, \dots s_n\} \quad (1)$$

In this formulation, J_i denotes the complete lifecycle journey of the *i-th* customer, while $s1,s2,\dots,s_n$ represent ordered behavioral states observed during the subscription period. The states describe the process flow from first-time registration to possible customer churn. Such an approach is significant since churn happens gradually via intermediary degradation stages and not instantly via cancellations in eqn 1.

In the statistical modeling of the transition between the different states, the proposed method employs the Markov transition matrix, where each cell represents the probability of transitioning from one state to another within the customer life cycle in eqn 2.

$$P_{ij} = \frac{N_{ij}}{\sum_j N_{ij}} \quad (2)$$

Here, P_{ij} represents the probability of transition from state i to state j, and N_{ij} is the number of observed transitions between these two states. The denominator accounts for the total outgoing transitions from state i. The Markovian approach can model behavior shift scenarios, including from active to inactive states or payment failure to churn transitions. These probabilities will then be used to determine potential churn trajectories and quantify the revenue leakages that could arise due to these lifecycle movements.

4.3 Churn Path Decomposition. After defining the customer journeys, the next phase in the module is to perform churn path decomposition. Instead of viewing churn as an individual event, the module determines the series of transitions leading to exit outcomes. Many customers will

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4976-4980

follow the sequence: active -> usage decline -> payment failure -> inactivity -> churn, as seen in fig 2, while some might simply churn following plan downgrades in eqn 3.

$$CP_k = \sum_{t=1}^{T} P(S_t \rightarrow churn) \quad (3)$$

The churn path score CP_k quantifies the cumulative probability of a specific pathway leading to churn. Each intermediate state s_t contributes transition likelihood toward the terminal churn state. When these probabilities are summed up for the entire duration T, it helps understand how hazardous the particular route is. High values signify that the route consistently causes churning and therefore needs better management. This breakdown helps the framework target the most harmful lifecycle paths.

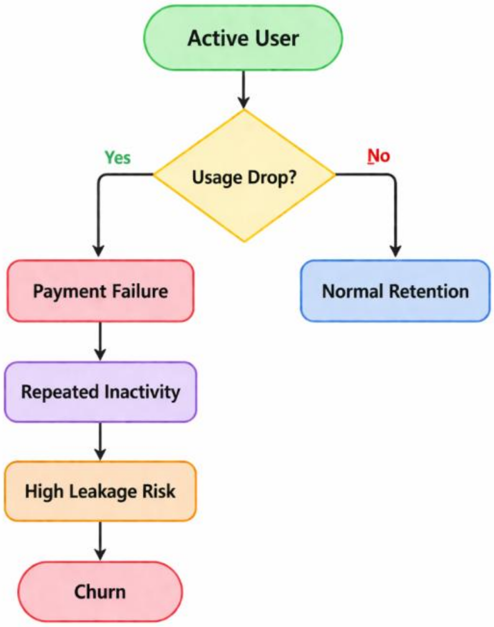

Figure 2: Churn Path Decomposition and Revenue Leakage Risk Flowchart

4.4 Revenue Leakage Estimation. This module converts churn probability into direct financial impact. Instead of merely identifying risky users, the methodology estimates hidden MRR leakage associated with churn pathways in eqn 4.

$$RL = \sum_{i=1}^{N} MRR_i * P(Churn_i) \quad (4)$$

Here, RL denotes total revenue leakage, MRR_i is the Monthly Recurring Revenue contribution of customer i, and $P(churn_i)$ is the predicted churn probability. This equation turns behavioral risks into money, allowing companies to design retention programs that maximize income. High-value customers who also have high churn risk significantly add to customer leakage, thus becoming excellent candidates for targeted campaigns. Path-wise financial attribution is a significant advancement compared to traditional churn rates.

4.5 Sequence Clustering. In order to identify recurring life cycle trends, the technique makes use of DTW clustering or K-means on the sequence embedding in eqn 5. Similar customer life cycles are clustered into various categories such as retention, price sensitivity churn, and usage churn.

$$D(x,y) = \sum_{i=1}^{n} (x_i - y_i)^2 \quad (5)$$

The distance measure D(x,y) evaluates the similarity between two customer journey embeddings. Low distances mean that the journeys are very similar, whereas high distances imply unique behavioral evolutions. Clustering according to this measure will enable us to detect the main churn archetypes, which can then be associated with interventions for better personalized campaigns.

4.6 Predictive Risk Model. The last step will predict future leakage using either an XGBoost classifier or a Transformer sequence classifier in eqn 6.

$$Risk_i = f(J_i, B_i, P_i) \quad (6)$$

In this formulation, $Risk_i$ is the predicted leakage risk of customer i, J_i is the lifecycle journey sequence, Bi represents billing behavior, and P_i denotes product usage patterns. Function f(·) can be modeled using either the XGBoost algorithm for tabular data fusion or Transformer models for sequence-based attention learning. The output risk score enables organizations to undertake preemptive actions to counteract revenue leakage. The prediction stage makes sure that the presented methodology not only achieves descriptive analysis but also predicts the upcoming churn cases and revenue protection.

5. Experimental Setup. To test our sequential path and revenue leakage model, we have employed a large-scale dataset consisting of SaaS subscriptions from 120,000 users gathered across several renewal periods. Behavioral, transactional, and life cycle transition records were included in the dataset. Mainly, the following features have been used for modeling purposes:

– Frequency of logins as an indicator of user engagement activity;
– Renewal cycles as well as downgrades; and
– Payment failures as key predictors of churn-prone states.

The development environment was developed in Python and comprised of Pandas to preprocess and transform the sequence life cycle, and Scikit-learn to conduct clustering, feature engineering, classification, and evaluation. To measure the performance of the suggested model, Area under Curve (AUC) and F1-score were considered. In addition to classification effectiveness, two metrics that reflect the business needs were also

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4976-4980

considered: leakage attribution accuracy, which determines the accuracy of identifying the revenue loss pathways, and retention uplift percentage, which determines the uplift gained by intervention strategies.

6. Results and Discussion

The empirical results have confirmed the efficacy of the sequential churn path and revenue leakage model. Figure 3 below shows the contribution to leakage for the prominent churn paths. Path A leads the pack with the greatest share, contributing 32% of the total, which means that this path is the one causing the maximum damage to revenue. Path B makes a contribution of 21%, while Path C makes a contribution of 15%. The rest, which is 32%, is contributed by the other minor churn paths.

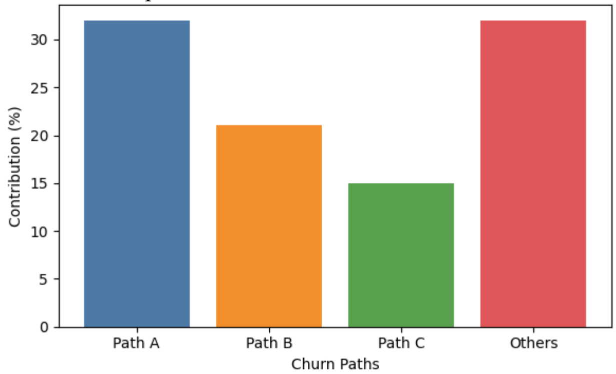

Figure 3. Path-Wise Revenue Leakage Contribution Across Dominant Churn Trajectories

Figure 4 shows the comparison between the accuracy of baseline machine learning algorithms versus the more sophisticated machine learning techniques. In the traditional machine learning algorithm of Logistic Regression, the accuracy is 0.74. The use of Random Forest increases the accuracy up to 0.79, and XGBoost improves it even more to 0.84. The sequential model proposed in this study showed the best accuracy, 0.87.

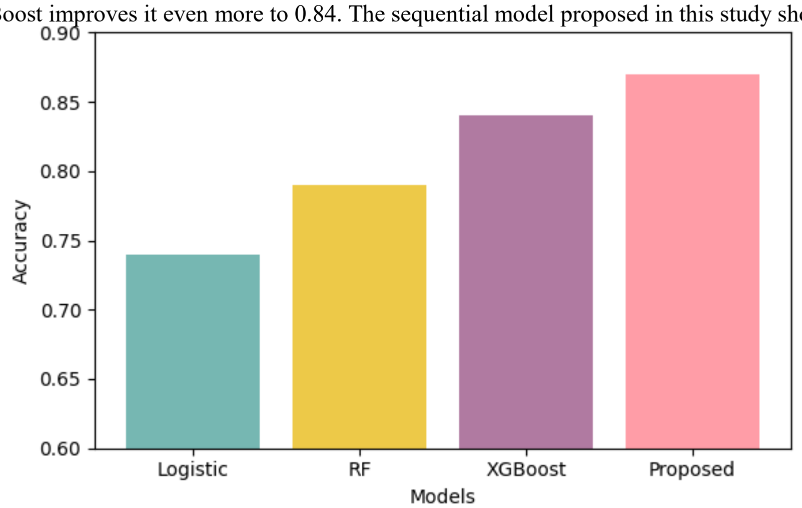

Figure 4. Comparative Predictive Performance of Baseline and Proposed Churn Risk Models

Financial benefits of targeted retention are illustrated in fig 5 where the revenue loss per month post intervention always stays below the trend before intervention. The model reduced revenue leakage by 21.5%, validating that interventions at early stages of the customer journey can avoid expensive churn transitions.

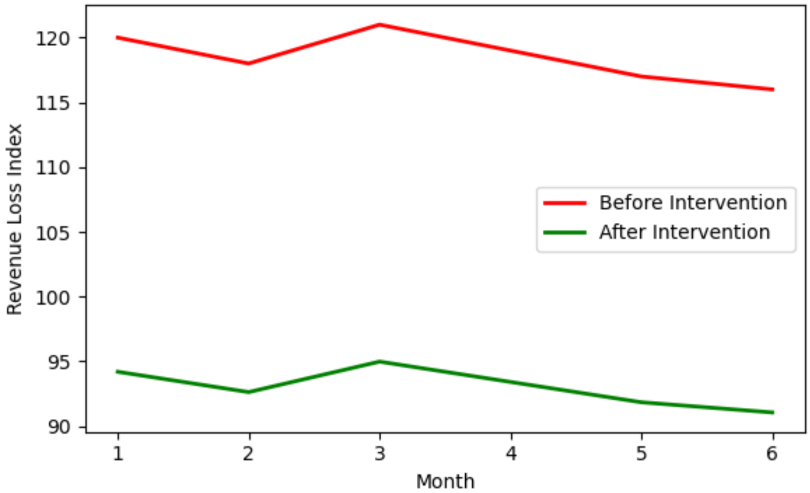

Figure 5: Temporal Revenue Leakage Trend Before and After Retention Intervention

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 4976-4980

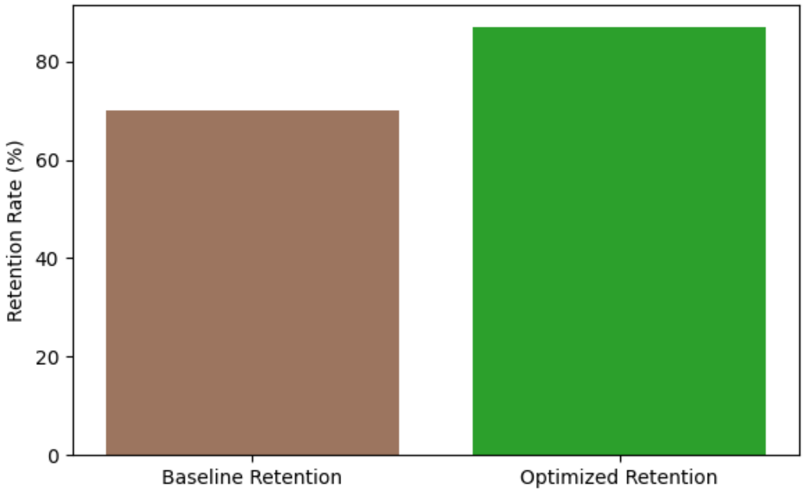

Figure 6: Comparative Analysis of Baseline and Optimized Customer Retention Rates

Last but not least, fig 6 illustrates how effective the optimization in terms of the retention performance is. Baseline retention performance was estimated at 70%, whereas optimized one came to 87%, representing a 17% increase. This proves that the suggested risk-sensitive approach not only enhances predictive performance but also brings tangible value to business by retaining customers.

7.Conclusion

The current research illustrates that the phenomenon of revenue leakage in subscription ecosystem operations is, by nature, driven by paths rather than events. Rather than treating churn as a terminal process of a single event, the framework presented here sees churn as a chain of lifecycle events and allows for diagnosing the exact point of revenue leakage. The inclusion of journey sequencing analysis, Markov transition analysis, path clustering, revenue leakage attribution, and churn risk scoring makes the proposed framework more accurate and insightful for churn diagnosis purposes.The biggest strength of the solution is the possibility to detect high-churn-risk customers at the earliest stages of the process, especially in intermediate states like decrease in usage activity, failure to pay the bill, and inactive states. It will enable timely customer engagement and, therefore, help to prevent churn from taking place. As shown experimentally, the benefits include improved uplift and profitability its modularity and scalability features, the framework can be applied in many industries, including but not limited to SaaS companies, OTT, telecoms, and BFSI.

Reference

1. Zhang, B. (2023). Customer churn in subscription business model—predictive analytics on customer churn. BCP Business & Management, 44, 870-876.
2. Ge, Y., He, S., Xiong, J., & Brown, D. E. (2017, April). Customer churn analysis for a software-as-a-service company. In 2017 Systems and Information Engineering Design Symposium (SIEDS) (pp. 106-111). IEEE.
3. Sukow, A. E., & Grant, R. (2013). Forecasting and the role of churn in software-as-a-service business models. I-Business, 5(1A), 49.
4. Devi, O. R., Pothini, S. K., Kumari, M. P., & Charan, U. N. S. (2023, May). Customer Churn Prediction using Machine Learning: Subcription Renewal on OTT Platforms. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1025-1029). IEEE.
5. McCarthy, D. M., Fader, P. S., & Hardie, B. G. (2017). Valuing subscription-based businesses using publicly disclosed customer data. Journal of Marketing, 81(1), 17-35.
6. Jamal, Z., & Bucklin, R. E. (2006). Improving the diagnosis and prediction of customer churn: A heterogeneous hazard modeling approach. Journal of Interactive Marketing, 20(3-4), 16-29.
7. Ballings, M., & Van den Poel, D. (2012). Customer event history for churn prediction: How long is long enough?. Expert Systems with Applications, 39(18), 13517-13522.
8. Ahn, J. H., Han, S. P., & Lee, Y. S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. Telecommunications policy, 30(10-11), 552-568.
9. Xie, Y., Li, X., Ngai, E. W. T., & Ying, W. (2009). Customer churn prediction using improved balanced random forests. Expert Systems with Applications, 36(3), 5445-5449.
10. Singh, D., Yugandhar, M. B. D., & Chawla, N. (2024). Design and Implementation Strategies for Scalable RESTful APIs in Enterprise Systems.
11. Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. Journal of marketing research, 41(1), 7-18.
12. Saha Mandal, A. (2019). Predicting the Evolutionary and Medical Significance of Human Genetic Variations with Machine Learning.
13. Ansari, S. A., & Zafar, A. (2018, December). A review on multisource data analysis using soft computing techniques. In *2018 4th International Conference on Computing Communication and Automation (ICCCA)* (pp. 1-6). IEEE.
14. Preethi, P., & Asokan, R. (2019). An attempt to design improved and fool proof safe distribution of personal healthcare records for cloud computing. *Mobile Networks and Applications*, *24*(6), 1755-1762.
15. Ansari, S. A., & Zafar, A. (2019). A review on video analytics its challenges and applications. *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals: Proceedings of GUCON 2019*, 169-182.
16. Bharathy, S. S. P. D., Preethi, P., Karthick, K., & Sangeetha, S. (2017). Hand gesture recognition for physical impairment peoples. *SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE)*, *610*.
17. Marín Díaz, G., Galán, J. J., & Carrasco, R. A. (2022). XAI for churn prediction in B2B models: a use case in an enterprise software company. Mathematics, 10(20), 3896.
18. Priyadarshini, R., Bhuvaneswari, T., Aswini, B., Mobeena, B., Balu, V., & Martin, B. (2026). Customer Churn Prediction and Analytics for Subscription-based Services. International Journal of Emerging Research in Science, Engineering, and Management, 2(1), 101-107.
19. Singh, D. (2022). Optimizing Enterprise Search Performance Using EHCache-Backed Apache Lucene Indexing for Hybrid Caching Systems. *Australian Journal of Cross-Disciplinary Innovation*, *4*(4).
20. Deshpande, G., & Singh, D. (2025). AI-ASSISTED SECURITY ORCHESTRATION IN HEALTHCARE INCIDENT RESPONSE. *Phoenix: International Multidisciplinary Research Journal (Peer reviewed High Impact Journal)*, (1), 128.