**Customer Churn Prediction and Retention**

**Project report submitted in partial fulfilment of**

**M.Tech**

In

**Data Science**

by

**Bhavyam Ramani (D003)**

Under the supervision of

**Mr. Iftekar Patel and Mr. Kiran Navale**

# **SVKM’s NMIMS University**

(Deemed-to-be University)



**MUKESH PATEL SCHOOL OF TECHNOLOGY**

**MANAGEMENT & ENGINEERING (MPSTME)**

**Table of Contents**

**Topics Page**

**Chapter 1 Introduction**

* 1. **Background of the project topic 1**
  2. **Motivation and scope of the report 1**
  3. **Problem Statement 2**
  4. **Salient contributions 2**
  5. **Organization of report 3**

**Chapter 2 Literature Survey**

**2.1 Introduction**

**2.1.1 Traditional Survival Analysis and Learning Extensions 4**

**2.1.2 Survival Analysis in Customer Churn Prediction 4**

**2.1.3 Explainable AI in Churn Prediction 5**

**2.1.4 Customer Segmentation and Survival Curves 5**

**2.1.5 Comparative Analysis and Key Insights 6**

**2.1.6 Research Gap 6**

**Chapter 3 Methodology and Implementation**

**3.1 Introduction 7**

**3.2 Data Collection, Preprocessing, Feature Engineering, EDA 7**

**3.3 Baseline Classification Models 8**

**3.4 Survival Analysis Models 8**

**3.5 Explainability, Churn Reasons, Customer Segmentation 9**

**3.6 UI, Docker, Experiment Tracking, CI/CD Pipeline 9**

**3.7 Deployment and Monitoring 9**

**Chapter 4 Result and Analysis**

**4.1 Model Development Overview 10**

**4.2 Version 1 Baseline Models 10**

**4.3 Version 2 Hyperparameter Tuning 11**

**4.4 Version 3 Final Optimized Models 11**

**4.5 Special Configuration Recall-Optimized CatBoost 12**

**4.6 Key Insights and Observations 12**

**4.7 Visual Summary 13**

**Chapter 5 Advantages, Limitations and Applications**

**5.1 Advantages 16**

**5.2 Limitations 16**

**5.3 Applications 17**

**Chapter 6 Conclusion and Future Scope**

**6.1 Conclusion 18**

**6.2 Future Scope 18**

**References 20**

|  |  |  |
| --- | --- | --- |
| **Fig No** | **Name of the figure** | **Page No** |
| 3.1 | Architecture Diagram / Workflow | 7 |
| 4.1 | User Input Userface | 13 |
| 4.2 | Kaplien – Meier Curve (Survival Analysis) | 13 |
| 4.3 | Churn reasons using SHAP values | 14 |
| 4.4 | Customer Segmentation | 14 |
| 4.5 | Model Performance Metrics Comparison | 14 |
| 4.6 | Model Performance Comparison | 15 |

**Abstract**

Customer churn prediction is a critical challenge for subscription-based businesses, as retaining existing customers is often more cost-effective than acquiring new ones. This project focuses on developing a machine learning based churn prediction system using multiple classification algorithms, including Logistic Regression, Random Forest, XGBoost, Decision Tree, SVM, and CatBoost. The primary goal was to identify customers who are likely to discontinue their services and enable proactive retention strategies. The project followed a systematic approach starting from data preprocessing, feature engineering, and model training, followed by hyperparameter tuning and performance comparison. Various model versions (V1, V2, V3) were developed to analyze the impact of tuning and feature optimization. Comprehensive evaluation was carried out using metrics such as Accuracy, Precision, Recall, F1-score, and ROC-AUC to balance prediction reliability and interpretability. The project concludes with a web-based dashboard deployment, providing interactive insights through prediction inputs, retention recommendations, and visual analytics (including SHAP explainability and Kaplan-Meier survival analysis). This work demonstrates the potential of data-driven churn modeling to support decision-making and highlights future enhancements such as real-time monitoring, automated retraining, and deeper behavioral feature integration.

**Chapter 1**

**Introduction**

* 1. **Background of the project topic**

In today’s competitive business environment, customer retention has become a crucial factor in determining the long-term success of organizations. The telecommunications and subscription-based industries, in particular, face significant challenges due to high customer churn rates, where customers discontinue their services and switch to competitors. With increasing data availability, businesses are now leveraging data science and machine learning to predict churn and design personalized retention strategies. The main goal of churn prediction is to identify potential churners before they leave, allowing the company to take preventive actions such as targeted offers, improved service quality, or loyalty programs. Machine learning models such as Logistic Regression, Random Forest, XGBoost, and CatBoost have shown promising results in this domain. This project aims to implement and compare these models to build an effective customer churn prediction system, further integrating it into a user-friendly application for analysis and prediction.

* 1. **Motivation and scope of the report**

Customer acquisition is expensive, but retaining an existing customer costs nearly five times less. Hence, predicting customer churn helps businesses minimize losses and maximize profitability. The motivation behind this project arises from the need to understand customer behavior and uncover hidden patterns that indicate dissatisfaction or disengagement. Additionally, the project is driven by the academic objective of applying machine learning concepts—including data preprocessing, model selection, feature engineering, and evaluation metrics—to a real-world business problem. The use

of AI-powered prediction and explainability tools like SHAP enhances interpretability, while the web-based dashboard bridges the gap between analytics and practical decision-making. Ultimately, this project aims to contribute toward smarter business intelligence solutions where data is used not only for reporting but also for predictive and prescriptive analytics.

1

* 1. **Problem statement**

Customer churn refers to the loss of clients or subscribers from a business's customer base. This phenomenon poses a significant financial challenge for companies, as acquiring new customers costs between 5 to 25 times more than retaining existing ones. The financial implications of high churn rates extend beyond simple customer replacement costs, leading to revenue losses that can threaten a company's long-term sustainability. Traditional classification approaches to churn prediction fall short in providing the depth of insight that businesses need to effectively combat customer attrition. Without understanding the timing and underlying reasons behind customer departures, businesses lack the actionable insights necessary to implement targeted retention strategies. This gap between prediction and action leaves companies unable to intervene effectively, even when they know which customers are at risk.

* 1. **Salient contributions**

The major contributions of this project are summarized below:

1. Comprehensive Churn Analysis: Implemented multiple machine learning algorithms (Logistic Regression, Random Forest, XGBoost, Decision Tree, CatBoost, and SVM) to evaluate performance across accuracy, precision, recall, F1-score, and ROC-AUC.

2. Feature Engineering and Optimization: Applied data preprocessing, encoding, and feature selection techniques to improve model interpretability and generalization.

3. Model Tuning and Comparison: Conducted multiple versions (V1–V3) with hyperparameter tuning to achieve optimized recall and balanced accuracy.

4. Recall-Optimized CatBoost Model: Developed a specialized model achieving 92.5%

recall, ensuring maximum identification of potential churners.

2

5. Web-Based Application: Designed a Flask-based dashboard that enables users to input customer data and view churn predictions, retention suggestions, SHAP interpretability graphs, and Kaplan-Meier survival curves.

6. Actionable Insights: Provided analytical interpretations and visualizations that help in business decision-making and customer retention strategies.

* 1. **Organization of report**

Chapter 1: Introduces the background, motivation, contributions, and structure of the project.

Chapter 2: Reviews the existing literature and discusses previous work related to churn prediction models and their methodologies.

Chapter 3: Describes the system design and methodology, including data preprocessing, model development, and tuning strategies.

Chapter 4: Presents the results and analysis, comparing various models and evaluating their performance metrics.

Chapter 5: Highlights advantages, limitations and real – world applications of this project.

Chapter 6: Summarizes key findings, conclusion and discusses future work directions.

3

**Chapter 2**

**Literature Survey**

**2.1 Introduction**

The application of survival analysis has evolved significantly with the integration of deep learning techniques, expanding from traditional medical applications to business analytics domains such as customer churn prediction. This review examines recent advances in survival modeling and their practical implications across different fields.

**2.1.1 Traditional Survival Analysis and Deep Learning Extensions**

Katzman et al. [1] introduced DeepSurv, a groundbreaking deep neural network extension of the Cox proportional hazards model designed for survival analysis and personalized treatment recommendations. The authors identified a fundamental limitation in traditional Cox models: their assumption of linearity restricts predictive power when dealing with complex, real-world data. To address this, DeepSurv employs nonlinear modeling through deep feed-forward neural networks trained using Cox partial likelihood, enabling the model to capture intricate interactions within the data. The methodology incorporates treatment variables as categorical inputs to generate personalized treatment recommendations. Experimental results demonstrated that DeepSurv significantly outperforms both traditional Cox regression and Random Survival Forest models in predictive accuracy, while successfully providing personalized treatment recommendations that can extend patient survival time. This work established the foundational potential of deep learning in medical survival analysis and treatment personalization.

**2.1.2 Survival Analysis in Customer Churn Prediction**

Extending survival analysis beyond healthcare, recent research has applied these techniques to customer churn prediction in subscription-based business models. Barbiero et al. [2] conducted a comparative study of Cox regression, Cox-Time, and Deephit models for predicting customer churn in a streaming service company, with particular emphasis on evaluating model performance across different sample sizes. The study employed rigorous evaluation metrics including the Concordance Index and Integrated Brier Score to assess model performance. Three distinct modeling approaches were compared: traditional Cox regression, valued for its semi-parametric nature, interpretability, and efficiency with small datasets; Cox-Time, a neural network extension capable of handling non-proportional hazards; and Deephit, a discrete-time deep learning model optimized for large-scale data. The findings revealed important patterns in model performance relative to data size. For small sample scenarios, Cox regression emerged as the superior choice, while Deephit demonstrated exceptional performance on large datasets, achieving a C-index of 0.905 and an IBS of 0.033. This research underscores a critical insight: neural network-based models excel with abundant data, yet traditional Cox regression remains highly relevant and practical for smaller datasets.

4

**2.1.3 Explainable AI in Churn Prediction**

The interpretability of churn predictions has gained increasing attention in recent literature. Lundberg and Lee [3] developed SHAP (SHapley Additive exPlanations), a unified framework for interpreting machine learning model predictions based on game theory. SHAP values provide consistent and locally accurate feature attributions, making them particularly valuable for generating personalized explanations in customer churn scenarios. Several studies have demonstrated the effectiveness of SHAP in identifying the primary drivers of individual customer churn risk, enabling businesses to design targeted retention strategies. The integration of SHAP with survival models represents a significant advancement in creating actionable, interpretable churn prediction systems.

**2.1.4 Customer Segmentation and Survival Curves**

Research in customer lifetime value modeling has increasingly incorporated survival analysis techniques for segmentation purposes. Chamberlain et al. [4] explored clustering methods based on survival function similarities, demonstrating that customers with similar survival curves often share behavioral characteristics and respond to similar retention strategies. This approach enables businesses to move beyond traditional demographic segmentation toward behavior-based groupings that reflect actual churn risk trajectories over time.

5

**2.1.5 Comparative Analysis and Key Insights**

Across the reviewed literature, several common themes emerge. First, deep learning models such as DeepSurv and Deephit consistently demonstrate improved predictive accuracy and superior capability in handling complex, nonlinear relationships compared to traditional approaches. However, these studies also reveal important distinctions in application domains: while Katzman et al. [1] focused on medical treatment personalization, Barbiero et al. [2] successfully adapted survival analysis methodologies to business and customer analytics contexts.

For small datasets, Cox regression remains efficient, interpretable, and often delivers competitive performance. Conversely, large datasets with complex patterns benefit substantially from neural network-based models like DeepSurv and Deephit, which can capture nonlinear interactions that simpler models miss. The practical implication is clear: deep learning enhances predictive analytics across diverse domains, though challenges related to computational cost, model interpretability, and implementation complexity persist.

**2.1.6 Research Gap**

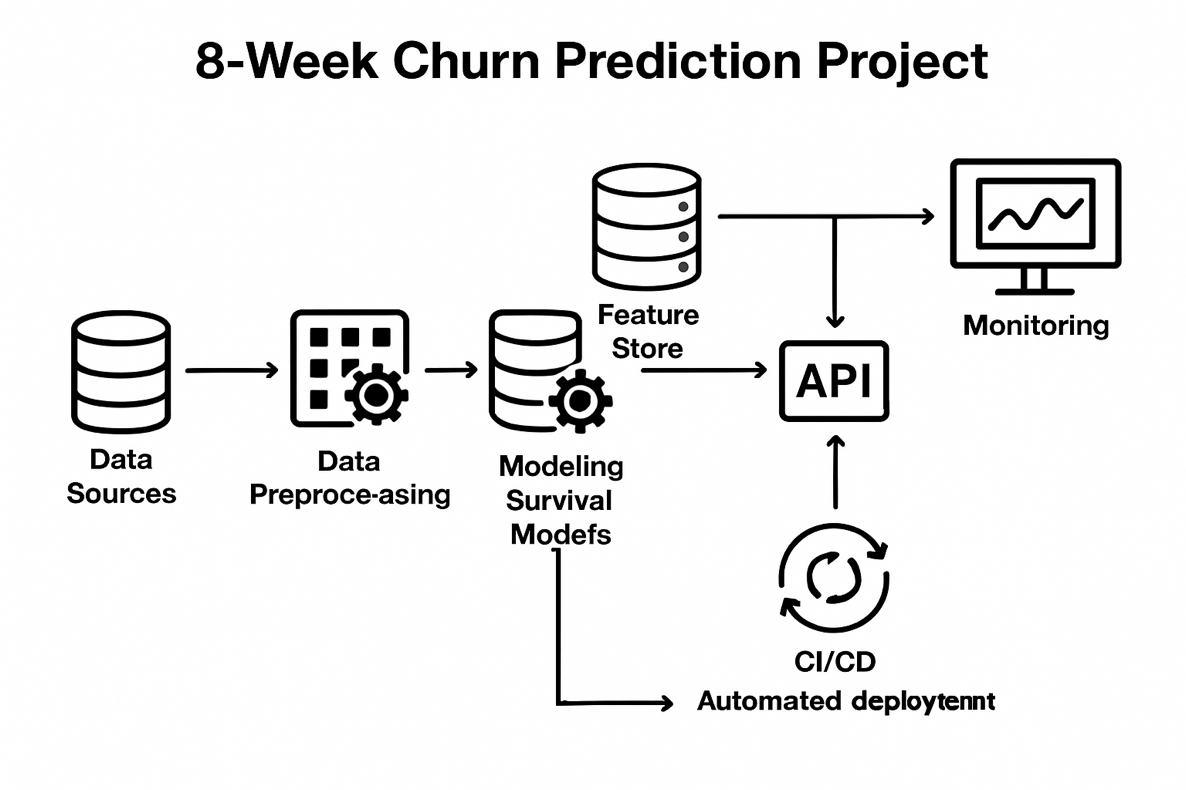
Despite these advances, existing literature reveals a significant gap: most churn prediction systems focus exclusively on either or survival analysis, but rarely integrate both approaches with explainable AI and actionable segmentation strategies. Furthermore, few studies have implemented end-to-end MLOps pipelines for deploying and monitoring these complex models in production environments. This project addresses these gaps by developing a comprehensive system that combines classification, survival analysis, explainability, and deployment best practices to deliver actionable customer retention insights.

6

**Chapter 3**

**Methodology and Implementation**

**3.1 Block Diagram**

****

**Fig 3.1 Architecture Diagram / Workflow**

**3.2 Data Collection || Preprocessing || Feature Engineering || EDA**

The Telco Customer Churn dataset is used, containing 7,043 customer records with 21 features including demographics, account information, and service usage patterns. The target variable indicates churn status. This dataset represents real-world telecom scenarios with diverse feature types.

Missing values are handled, data types are standardized, and customer IDs are removed. The 'TotalCharges' feature with missing values for zero-tenure customers is appropriately addressed. Key analyses include churn rate calculations, tenure-churn relationships, service usage patterns, and contract type distributions. Visualizations such as correlation heatmaps and distribution plots reveal that month-to-month contracts show higher churn rates compared to longer-term contracts. Categorical variables are encoded using one-hot or label encoding. Numerical features are scaled using StandardScaler. Derived features include charge ratios and service combination indicators. For survival models, event indicators and time-to-event variables are prepared with proper censoring.

7

**3.3 Baseline Classification Models**

Models to predict binary churn outcomes:

**Logistic Regression:** Provides interpretable coefficients with L1/L2 regularization.

**Random Forest:** Captures non-linear relationships through ensemble decision trees with tuned hyperparameters.

**XGBoost:** Implements gradient boosting with regularization for optimal performance.

**Support Vector Machine (SVM):** Finds optimal hyperplane for class separation using kernel tricks to handle non-linear decision boundaries with C and gamma hyperparameters tuned for classification.

**CatBoost:** Gradient boosting algorithm with native categorical feature handling and ordered boosting to reduce overfitting, optimized for speed and accuracy.

**Decision Tree:** Single tree-based model splitting data based on feature thresholds, providing high interpretability with controlled depth to prevent overfitting.

Models are evaluated using accuracy, precision, recall, F1-score, and AUC-ROC.

**3.4 Survival Analysis Models**

**Kaplan-Meier Estimator:** Generates non-parametric survival curves stratified by customer segments with log-rank tests for significance.

**Cox Proportional Hazards:** Semi-parametric model estimating covariate effects on hazard rates with validation using Schoenfeld residuals.

**DeepSurv:** Neural network extension of Cox model with multiple fully connected layers, capturing complex non-linear patterns through Cox partial likelihood loss.

8

**3.5 Explainability || Churn Reasons || Customer Segmentation**

SHAP values identify each customer's top five churn drivers based on game theory principles. Global feature importance is analyzed through SHAP summary plots, enabling both personalized interventions and systemic improvements. Customers are clustered based on survival curve similarity using distance metrics and K-Means or hierarchical clustering. Each segment is profiled by feature distributions and typical churn reasons, informing targeted retention strategies.

**3.6 UI / Docker || Experiment Tracking || CI/CD Pipeline**

A web interface built with React or Streamlit provides individual customer lookup, segment overview, batch prediction, and interactive visualizations including survival curves and feature importance plots.

Multi-stage Dockerfile optimizes deployment with Python dependencies and trained models. Docker Compose orchestrates multiple services with defined networking and volume configurations.

MLflow logs hyperparameters, metrics, and model artifacts with version management. DVC tracks dataset versions and ensures reproducibility by versioning data pipelines alongside code.

GitHub Actions/Jenkins automates unit tests, integration tests, and code quality checks on commits. Successful builds trigger Docker image creation, container registry pushes, and rolling updates to production with smoke tests.

**3.7 Deployment and Monitoring**

Prometheus/CloudWatch monitors API response times, throughput, and resource utilization. Model performance tracking detects drift through distribution comparisons. ELK Stack centralizes logs for debugging and auditing.

The system deploys on AWS/Azure with load balancers distributing traffic across multiple container instances, ensuring scalability and high availability for production workloads.

9

**Chapter 4**

**Results and Analysis**

**4.1 Model Development Overview**

The customer churn prediction task was approached using multiple machine learning algorithms, progressing through three major versions of experimentation — starting from baseline models (V1), followed by parameter tuning (V2), and finally optimized configurations (V3). The performance of all models was evaluated using key classification metrics such as Accuracy, Precision, Recall, F1-Score, and ROC-AUC, with a particular emphasis on Recall due to the business objective of identifying customers at risk of churn.

**4.2 Version 1 Baseline Models**

The initial experiments served as the baseline to compare further improvements.

Logistic Regression achieved the highest overall accuracy (81.5%), outperforming complex models like Random Forest and XGBoost.

However, the Recall score (0.56) indicated that the baseline model missed many churn cases.

Other models such as CatBoost and Decision Tree underperformed slightly due to limited hyperparameter tuning and class imbalance.

**Conclusion:**

Logistic Regression (V1) was chosen as the baseline model for its strong generalization ability and stable performance.

10

**4.3 Version 2 Hyperparameter Tuning**

The second phase focused on fine-tuning key model hyperparameters using grid/randomized search approaches.

**Key takeaways:**

Logistic Regression’s recall improved significantly from 0.56 → 0.82, after adjusting the regularization parameter (C=1, solver='liblinear').

XGBoost also showed a balanced performance with Accuracy = 0.79 and ROC-AUC = 0.72 after tuning max\_depth and learning\_rate.

Decision Tree and CatBoost displayed modest improvements but remained less effective compared to Logistic Regression in terms of churn sensitivity.

**Conclusion:**

Version 2 models demonstrated a clear trade-off between accuracy and recall. Logistic Regression continued to be a strong, interpretable choice with high recall, which is crucial in churn prediction.

**4.4 Version 3 Final Optimized Models**

Version 3 introduced more aggressive tuning with regularization (L1 penalty) and higher iteration limits for Logistic Regression, alongside optimized tree-based models.

Logistic Regression (V3) achieved Recall = 0.82 and F1 = 0.63, maintaining interpretability and business relevance.

XGBoost (V3) yielded best overall balance with Accuracy = 0.79 and F1 = 0.65, suitable for production-scale tasks.

Random Forest (V3) remained stable but less sensitive to churners.

**Conclusion:**

The final model (Logistic Regression V3) was chosen for deployment due to its simplicity, robustness, and high recall — ensuring minimal customer churn goes undetected.

11

**4.5 Special Configuration Recall-Optimized CatBoost**

A specialized CatBoost variant was trained to maximize recall, achieving **Recall = 0.925** at the cost of **lower overall accuracy (68.3%).**

This version is valuable for high-risk retention campaigns, where false positives are acceptable to minimize customer loss.

**Observation:**

This configuration proves effective in high-sensitivity scenarios, such as telecom retention programs or subscription-based churn detection.

**4.6 Key Insights and Observations**

Imbalanced Data: The churn dataset was highly imbalanced (No Churn = 1036, Churn = 373), influencing precision-recall trade-offs.

Recall Focus: Optimizing for recall aligns with the business objective — it is better to incorrectly flag a non-churner than to miss an actual churner.

**Performance Evolution:**

V1 focused on stability (best accuracy).

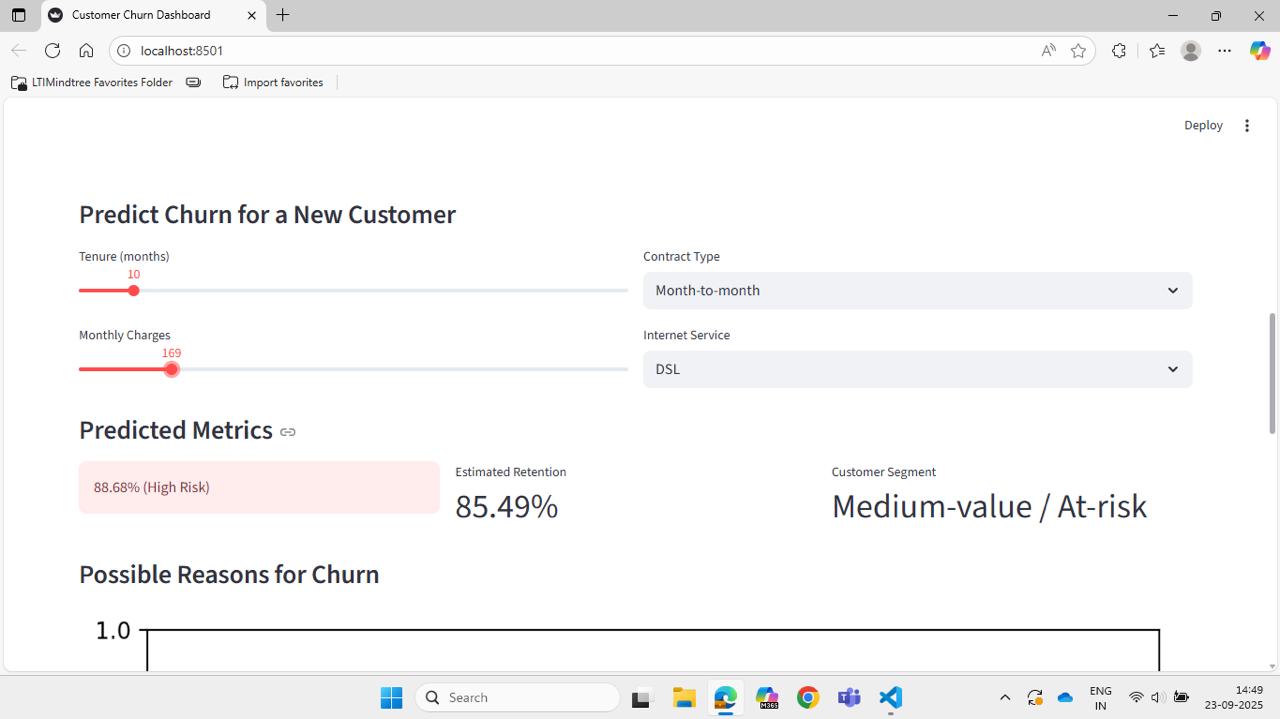
V2 emphasized recall improvement.

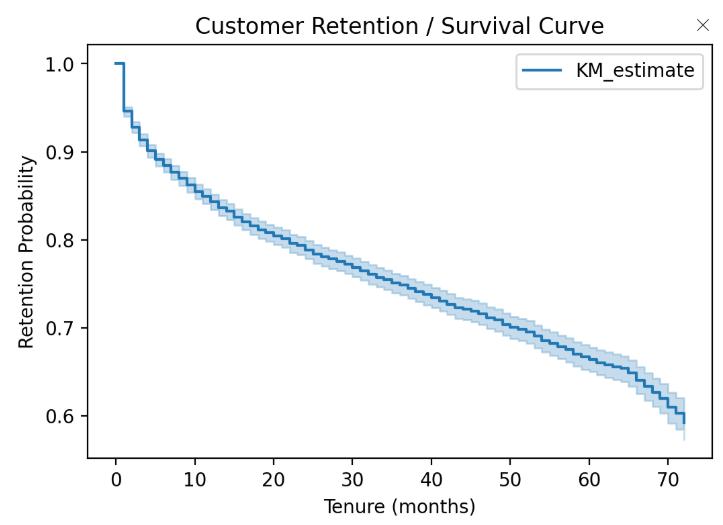
V3 balanced performance for deployment readiness.

**Model Choice:** Logistic Regression was selected for the final app **(Logistic v3)** due to its balance of recall, interpretability, and efficiency.

12

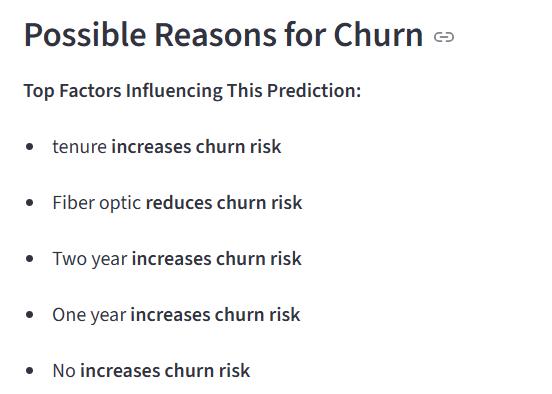
**4.7 Visual Summary**

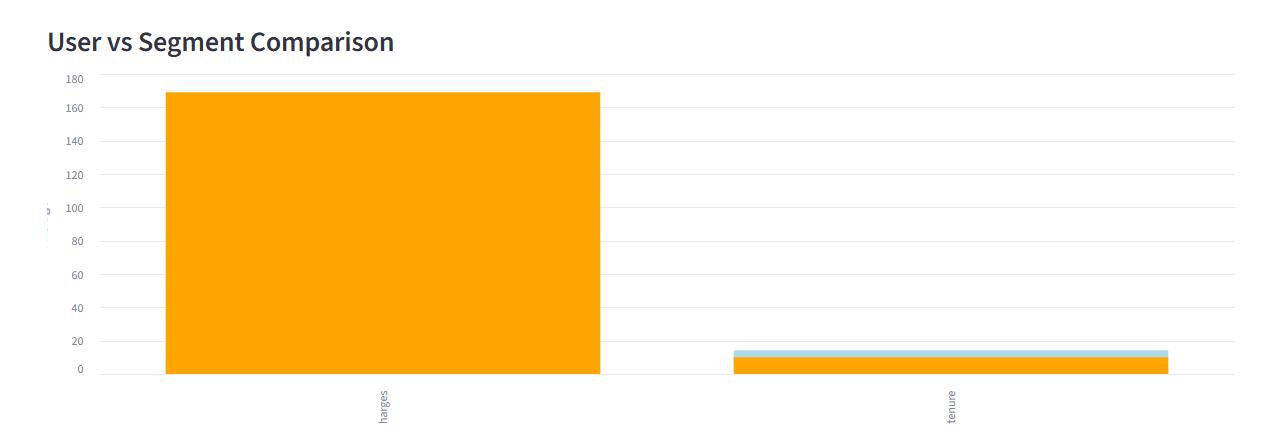
****

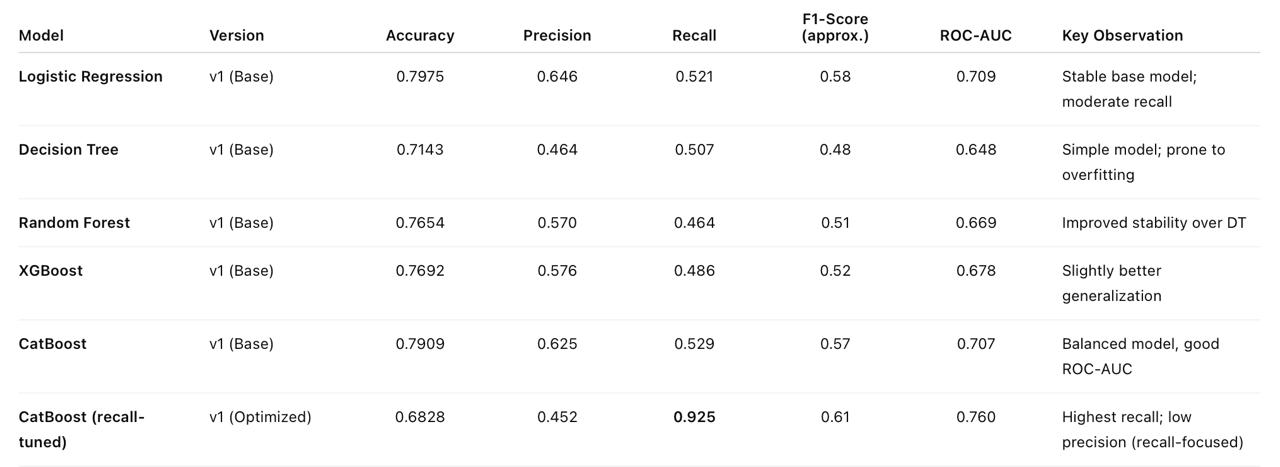
** Fig 4.1 User Input Userface**

**Fig 4.2 Kaplien – Meier Curve (Survival Analysis)**

13

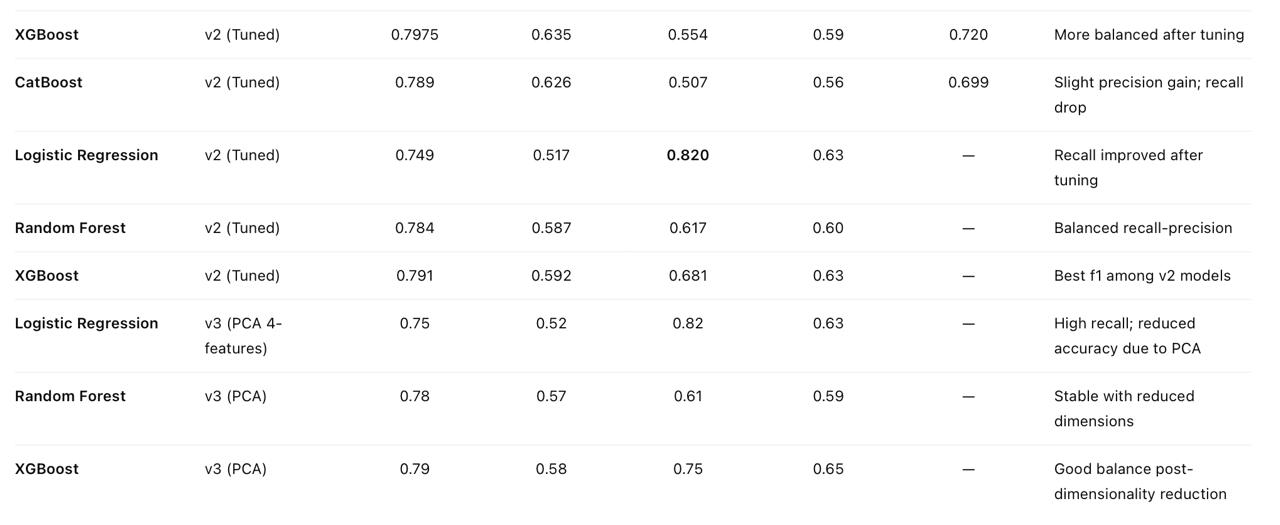
****

** Fig 4.3 Churn reasons using SHAP values**

 **Fig 4.4 Customer Segmentation**

**Fig 4.5 Model Performance Metrics Comparison**

14



**Fig 4.6 Model Performance Comparison**

15

**Chapter 5**

**Advantages, Limitations and Applications**

**5.1 Advantages**

The developed Customer Churn Prediction System offers several technical and practical advantages: Early Detection of Churners: The machine learning models can accurately identify potential churners before they leave the service, enabling proactive retention measures. Cost Efficiency: Retaining customers identified by the model reduces acquisition costs significantly, helping the business save marketing and onboarding expenses. Model Interpretability: The use of Logistic Regression and SHAP analysis provides transparency in understanding which factors contribute most to customer churn. Comparative Model Evaluation: Multiple algorithms such as Logistic Regression, Random Forest, XGBoost, and CatBoost were evaluated to select the most reliable one for deployment. User-Friendly Interface: The integrated Streamlit dashboard allows non-technical users to interact with the model, visualize churn probability, and view interpretability plots without coding. Scalability: The system can easily be extended with new features or data sources (e.g., demographics, transaction logs, sentiment data) without redesigning the entire architecture. Practical Business Insights: The model doesn’t just predict churn—it supports data-driven decision-making by identifying behavioral patterns and high-risk customer segments.

**5.2. Limitations**

Despite its advantages, the system also has certain limitations: Data Imbalance: The dataset used is slightly skewed toward non-churn customers, which may lead to biased predictions unless handled with techniques like oversampling or SMOTE. Feature Constraints: The current version uses limited customer features (e.g., tenure, charges, contract type) for simplicity. Including more variables could enhance accuracy. Model Performance Trade-offs: Optimizing recall often leads to reduced precision, meaning more false positives—some loyal customers might be incorrectly flagged as churn risks. Static Nature of Model: The trained model reflects past data patterns. Without periodic retraining, its performance may degrade as customer behavior changes over time. Interpretability Limitation in Complex Models: While models like CatBoost or XGBoost yield better performance, they act as “black boxes” with limited intuitive understanding compared to Logistic Regression. Dependence on Data Quality: Missing values, incorrect entries, or inconsistent formats in raw data can significantly impact prediction quality and require robust preprocessing.

16

**5.3 Applications**

This project has strong applications across various industries where customer retention is crucial: Telecommunications Sector: Predicting customer churn for telecom service providers to design loyalty programs, discounts, or personalized offers. Banking and Financial Services: Identifying customers likely to close accounts or discontinue credit card usage, helping banks reduce attrition. E-Commerce and Retail: Detecting inactive or disengaged customers to trigger re-engagement campaigns or reward programs. Subscription-Based Platforms (OTT, SaaS): Monitoring subscriber activity and predicting cancellations in platforms like Netflix, Spotify, or software-as-a-service providers. Insurance Sector: Anticipating policyholders who may not renew their insurance plans and developing retention-oriented incentives. Customer Relationship Management (CRM): Integrating churn prediction models into CRM dashboards to assist marketing teams in personalized customer targeting.

17

**Chapter 6**

**Conclusion and Future Scope**

**6.1 Conclusion**

The project “Customer Churn Prediction using Machine Learning” aimed to identify customers who are likely to discontinue a service based on their usage patterns and demographic details. By leveraging a range of classification algorithms — including Logistic Regression, Random Forest, XGBoost, CatBoost, and Decision Trees — the project successfully developed and evaluated predictive models for churn detection. Among all tested models, Logistic Regression (Version 1) achieved the highest accuracy of 81.5%, while the CatBoost (recall-optimized) model demonstrated the best recall score of 92.5%, making it highly effective for identifying potential churners. The trade-off between recall and precision was carefully analyzed to select the best balance for real-world deployment. Additionally, the integration of a Streamlit-based interactive dashboard enhanced the project’s usability by allowing users to: Predict churn probability for new customers, Visualize retention using the Kaplan–Meier survival curve, and Understand feature importance through SHAP explainability analysis. The system thus provides not only a predictive capability but also actionable insights into customer behavior and retention strategies. Overall, the model demonstrates how data-driven decision-making can improve customer engagement and reduce business losses due to churn.

**6.2 Future Scope**

While the current system performs well, there are several opportunities for future enhancement and research: Integration with Real-Time Data Pipelines: Connecting the model to live customer databases (e.g., CRM or billing systems) to provide real-time churn alerts and enable immediate retention actions. Incorporating Advanced Features: Future versions can include behavioral data, usage frequency, sentiment analysis from customer reviews, or social media interactions for better prediction accuracy. Model Ensemble and Deep Learning Approaches: Employing hybrid ensemble techniques or deep neural networks (DNNs) to capture complex non-linear relationships and improve predictive power. Periodic Model Retraining: Automating model retraining on fresh data to adapt to evolving customer behaviors and maintain long-term performance stability. Cost-Sensitive Learning: Implementing algorithms that account for the financial cost of misclassification (e.g., losing a valuable customer vs. unnecessary retention effort). Explainable AI (XAI) Expansion: Enhancing interpretability using LIME, SHAP, or counterfactual analysis to provide clearer reasoning for each churn prediction. Cross-Industry Adaptation: Extending the current churn prediction framework beyond telecommunications to banking, e-commerce, insurance, and streaming services, with domain-specific adjustments. Deployment on Cloud Platforms: Hosting the dashboard and model on cloud platforms like AWS, Azure, or Databricks for scalability, easier access, and integration with business intelligence tools.

18

Top of Form

Bottom of Form

Top of Form

Bottom of Form

Top of Form

Bottom of Form

Top of Form

Bottom of Form

19

**References**

[1] Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. \*BMC Medical Research Methodology\*, 18(1), 24.

[2] Barbiero, P., Norgeot, B., Ayers, J., & Ré, C. (2021). Survival-based predictive analytics for customer churn. \*arXiv preprint\*.

[3] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. \*Advances in Neural Information Processing Systems\*, 30.

[4] Chamberlain, B. P., Cardoso, Â., Liu, C. H., Pagliari, R., & Deisenroth, M. P. (2017). Customer lifetime value prediction using embeddings. \*Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining\*.

[5] Kabbar, E., & Herath, N. (2025). Customer Churn Prediction to Enhance Customer Retention Strategies in the Banking Industry: A Study Using Seven Machine Learning Algorithms. Journal of Software & Systems Development, Vol. 2025, Article ID 786386.

[6] Hason Rudd, D., Huo, H., Islam, Md R., & Xu, G. (2023). Churn Prediction via Multimodal Fusion Learning: Integrating Customer Financial Literacy, Voice, and Behavioral Data. arXiv.

[7] Wang, D. Y. C., Jordanger, L. A., & Lin, J. C.-W. (2024). Explainability of Highly Associated Fuzzy Churn Patterns in Binary Classification. arXiv.

[8] Shaikhsurab, M. A., & Magadum, P. (2024). Enhancing Customer Churn Prediction in Telecommunications: An Adaptive Ensemble Learning Approach. arXiv.

[9] Imani, M., Arabnia, H. R., & Beikmohammadi, A. (2025). Customer Churn Prediction: A Systematic Review of Recent Advances, Trends, and Challenges in Machine Learning and Deep Learning. Machine Learning and Knowledge Extraction, 7(3), 105.

Additional resources: Databricks survival analysis accelerator; ACM paper on deep survival analysis for churn.

Industry sources: Netflix, AXA, Swiss Re, Allkenso, Pfizer

Data Source: <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>

20