Telco Customer Churn: Executive Summary and Strategic Recommendations

Date: August 20, 2025
Prepared by: Tharusha Rasath Hemachandra

1. Executive Summary

Customer churn represents a critical threat to sustained revenue and market position in the highly competitive telecommunications industry. This report provides a comprehensive analysis of customer churn behavior, leverages a predictive machine learning model to identify high-risk customers, and outlines a data-driven strategy to mitigate churn and enhance customer lifetime value.

The analysis reveals that churn is not a random event but is heavily influenced by specific customer behaviors and service attributes. The key drivers are predominantly related to customer tenure, contract type, and service utilization, with a significant portion of churn occurring early in the customer lifecycle. My predictive model successfully identifies customers at risk of churn with high accuracy, enabling proactive and targeted retention efforts.

I recommend a three-pronged strategic approach focused on:

1. **Targeted Retention Campaigns:** Implement distinct, personalized campaigns for different customer risk segments to maximize the return on investment (ROI).
2. **Product and Service Re-evaluation:** Address the core pain points identified in the analysis, such as the friction associated with short-term contracts and high monthly charges.
3. **Operationalizing the Data Pipeline:** Integrate the predictive system into daily business operations to ensure a continuous, data-driven approach to customer retention.

By shifting from a reactive to a proactive retention strategy, the company can save a significant portion of its at-risk annual revenue and foster a more loyal, profitable customer base.

2. Project Overview & Methodology

The objective of this project was to move beyond a simple analysis of historical data and build a scalable, end-to-end solution for customer churn management. My methodology was twofold:

- **Exploratory Data Analysis (EDA) and Business Insights:** I first performed an in-depth analysis of a dataset containing information on customer demographics, account details, and service usage. This phase, documented in a series of Jupyter notebooks, identified key behavioral and product-related factors that drive customer churn.
- **Predictive Modeling and Pipeline Development:** I then developed a machine learning model to predict which customers are most likely to churn. Crucially, this effort was built on a modular, test-driven Python data pipeline. This architecture ensures that the entire process—from data ingestion and cleaning to feature engineering and model scoring—is automated, reproducible, and ready for deployment in a production environment. This structured approach provides a robust foundation for a long-term business solution.

3. Key Findings & Analysis

My analysis uncovered several critical insights that should guide strategic decision-making.

3.1 Churn is a Significant and Ongoing Problem

The overall churn rate in my customer base stands at approximately **26.5%**, representing a substantial portion of our revenue at risk each year. This is in line with or slightly higher than industry benchmarks.

3.2 Key Drivers of Churn

I identified the following as the most significant predictors of customer churn:

- **Contract Type:** Customers on month-to-month contracts have an extremely high churn rate. Their lack of commitment makes them highly susceptible to competitor offers and service issues. Conversely, customers with one- and two-year contracts show significantly higher loyalty.
- **Tenure:** The probability of churn is highest in the first six months of a customer's tenure. This suggests that the onboarding experience, initial service quality, and new-customer engagement are critical moments for retention.
- **Monthly Charges:** Higher monthly charges are correlated with an increased likelihood of churn, particularly when combined with a short tenure. Customers with high bills early on are likely to be seeking better value elsewhere.
- **Fiber Optic Internet Service:** While a premium offering, customers with fiber optic service have a higher churn rate than those with DSL. This may be due to unmet performance expectations or a more competitive market for high-speed services.
- **Paperless Billing:** Customers who opt for paperless billing are more likely to churn. This may not be a causal factor but could serve as a behavioral signal that these customers are more digitally savvy and likely to manage their accounts online and compare services with ease.

4. Model Performance and Business Impact Analysis

A key component of this project was rigorously evaluating multiple machine learning models to find the one that delivers the most business value. I recognized that a simple accuracy score can be misleading for this type of problem due to the imbalance between churned and non-churned customers. Therefore, I focused on a comprehensive set of metrics and a business-specific cost-benefit analysis.

4.1 Model Performance Comparison

I compared several models, including Logistic Regression, Decision Tree, Random Forest, XGBoost, and CatBoost, in both basic and hyper-parameter-tuned versions. The following table summarizes their performance across key evaluation metrics.

Model	Accuracy	Precision	Recall	F1	ROC-AUC	CV-F1
Logistic Regression	0.7395	0.5060	**0.7861**	0.6157	0.8407	0.7822
Decision Tree	0.7353	0.5012	0.5588	0.5284	0.6787	0.7865

Random Forest Basic	0.7672	0.5646	0.5374	0.5507	0.8184	0.8465
XGBoost Basic	0.7686	0.5635	0.5695	0.5665	0.8164	0.8232
CatBoost Basic	**0.7871**	**0.6022**	0.5829	0.5924	0.8333	0.8237
Random Forest Tuned	0.7658	0.5604	0.5455	0.5528	0.8201	**0.8487**
XGBoost Tuned	0.7764	0.5598	0.7380	**0.6367**	**0.8439**	0.8337
CatBoost Tuned	0.7764	0.5751	0.6043	0.5893	0.8399	0.8296

4.2 The Importance of Threshold Optimization

The most crucial step was optimizing the classification threshold. A standard threshold of 0.5 would classify a customer as at-risk if the model's prediction was 50% or higher. However, given the business objective of catching as many churners as possible (high **Recall**) without wasting too many resources on false alarms (good **Precision**), I fine-tuned this threshold.

This process led to a significant improvement in business value. For example, by adjusting the threshold for the **CatBoost Tuned** model to **0.100**, its performance metrics shifted to:

- **Optimal F1-Score:** 0.554
- **Optimal Precision:** 0.390
- **Optimal Recall:** 0.952

This adjustment allowed the model to correctly identify over **95%** of the customers who actually churned, which is a massive win for proactive retention. The analysis also showed a **Net Benefit of $269,200** by implementing a strategy based on this model, demonstrating a positive ROI.

4.3 Final Best Model Selection

Based on this in-depth analysis and threshold optimization, the **CatBoost Tuned** model was selected as the best overall choice. Its ability to achieve a superior balance between high recall (capturing most churners) and a strong F1-score makes it the most effective tool for this business problem.

The key insights for selecting this model were:

- The significant class imbalance in the data (2.8:1 ratio) makes standard accuracy a poor metric.
- **Recall** is the most critical metric because a lost customer is costlier than a wasted retention effort.
- **Threshold optimization** is essential to maximize business value and effectively address class imbalance.
- The model delivers a positive net benefit, making its implementation a financially sound decision.

5. Strategic Recommendations

Based on my findings, I propose the following strategic recommendations to effectively combat customer churn.

5.1 Implement a Targeted Retention Program

A one-size-fits-all approach is inefficient and costly. I recommend tailoring retention efforts to each risk segment.

- **For High-Risk Customers:** Initiate proactive, personalized outreach within the first 60-90 days of their service. Offer incentives to upgrade to a 1-year contract, such as a one-time bill credit, a discount on monthly charges, or a free hardware upgrade.
- **For Medium-Risk Customers:** Engage these customers with targeted value-based offers. This could include a limited-time discount on a service they don't currently have (e.g., streaming TV) or a loyalty reward for their tenure.
- **For Low-Risk Customers:** Focus on a low-cost, high-value loyalty program. Provide occasional, non-monetary rewards such as early access to new features, exclusive content, or "thank you" messages to reinforce their loyalty and maintain a positive relationship.

5.2 Improve the Onboarding and Early-Tenure Experience

Given that churn is highest among new customers, it is crucial to make the initial experience as seamless and valuable as possible.

- **Onboarding:** Develop a structured onboarding process that includes a welcome call, a guided tour of their account and services, and a follow-up check-in to address any initial issues.
- **Contract and Pricing Transparency:** Clearly communicate the benefits of longer-term contracts and incentivize them with competitive pricing. The data suggests that making the leap from a month-to-month plan to a long-term contract is a key milestone for retention.

5.3 Operationalize the Predictive Pipeline

The modular Python data pipeline I built is a significant business asset. I recommend operationalizing it to enable continuous churn management.

- **Daily Predictions:** Integrate the pipeline to run automatically on a daily or weekly basis, scoring the entire customer base to identify new at-risk customers in real time.
- **Automated Triggers:** Connect the model's output to your marketing and customer relationship management (CRM) systems to trigger automated, personalized communications or alerts for customer service agents.

6. Conclusion

By leveraging this data-driven approach, I can move beyond generalized assumptions and make precise, impactful decisions to reduce customer churn. The strategic recommendations outlined in this report are designed to be both effective and financially sound, transforming the churn problem from a cost center into a strategic advantage. This project provides not only a powerful predictive tool but a foundational data system that can be expanded to optimize other critical business functions, such as upselling, cross-selling, and customer service.