

Customer Churn Analysis for Telecommunications Company

Sishir Pandey - Project Manager

Fahim Arman - Data Engineer

Jitesh Akaveeti - Data Analyst (Predictive Modelling)

Chen - Data Analyst (Clustering)

Preeti Khatri - Data Analyst (Predictive Modelling)

Bishesh Aryal - Business Analyst

Final Report

Customer Churn Analysis:

This project aims to predict customer churn and classify customers into meaningful segments to gain deeper insights into churn risk and potential retention strategies. By applying clustering methods and supervised learning to a preprocessed telecom dataset, the study identifies high-risk customer groups, the main factors influencing churn, and actionable, data-driven approaches for improving customer retention.

1. Summary of Key Findings:

From our analysis, we discovered the following main points.

- Customer segmentation using K-Means resulted in three meaningful groups:
 - Loyal and stable customers
 - Value-sensitive customers
 - High-risk new customers
- A supervised machine learning model was trained on the preprocessed dataset to estimate churn probability for individual customers.
- Model evaluation showed strong predictive accuracy, indicating good discriminative ability and suitability for prioritising targeted retention strategies.
- Key factors influencing churn included:
 - Contract type
 - Customer tenure
 - Monthly charges
 - Usage of support services and add-on features
- Findings from the ANN model further confirmed that churn behaviour is shaped by customers' service engagement patterns, billing levels, and contract duration.
- High-risk customer groups, especially those on month-to-month contracts or with higher monthly bills, showed significantly higher churn likelihood.
- Overall, the ANN model effectively captured churn-related patterns, providing valuable insights to guide customer retention and decision-making.

2. Identification of Factors Contributing to Churn and Retention:

Churn Factors:

- Contract Type: Month-to-month customers have higher churn due to flexible exit options and limited commitment.
- Tenure: New customers, especially in their first few months, are more likely to churn as loyalty and trust are not yet established.
- Monthly Charges: High monthly charges relative to perceived value increase the likelihood of leaving.
- Gaps in Support Services: Lack of technical support and online security leads to higher churn and lower satisfaction.
- Limited Bundling: Customers with single services and no bundles are more price-sensitive and easier to lose to competitors.

Retention Factors:

- Long-Term Contracts: One- and two-year contracts reduce churn by encouraging longer-term engagement.
- Bundled Services: Combining multiple services in bundles increases perceived value and reduces cancellations.
- Accessible Support: Reliable and friendly technical/customer support fosters trust and improves retention.
- Customer Engagement: Personalized offers, proactive communication, and loyalty rewards strengthen long-term relationships.

Additional Insights from ANN Model:

- Month-to-month contracts increase churn; long-term contracts promote stability.
- High monthly charges contribute to dissatisfaction and turnover.
- Low engagement (fewer services or lack of add-ons) correlates with higher churn.
- Longer tenure, bundled services, and stable billing patterns enhance retention.
- Service integration and customer loyalty are key factors in maintaining customer relationships.

3. Recommendations for Targeted Retention Strategies:

A. Segment-Specific Strategies:

- Loyal and Stable Customers (Low Risk):
 - Maintain service quality.
 - Offer loyalty bonuses or exclusive upgrades.
 - Request feedback to prevent silent churn.
- Value-Sensitive Customers (Medium Risk):
 - Provide targeted discounts.
 - Optimize plan recommendations.
 - Offer mid-tier bundles balancing cost and value.
- High-Risk New Customers (High Risk):
 - Implement welcome campaigns and early check-ins.
 - Provide short-term incentives to encourage migration to longer-term contracts.

B. Predictive Model-Driven Strategies:

- Early Intervention: Use churn scores to flag high-risk customers for proactive outreach.
- Tailored Offers: Align retention incentives with key churn drivers (e.g., bill credits for high-charge customers, free support upgrades for those lacking technical help).
- CRM Integration: Incorporate churn predictions into the CRM system to prioritize frontline staff outreach.
- Automated Workflows: Trigger email or SMS retention campaigns when churn probability exceeds a set threshold.
- Continuous Monitoring: Regularly review churn metrics and retrain models as new data becomes available.

Additional Recommendations:

- Offer discounts, loyalty benefits, or incentives to high-risk customers for long-term contract upgrades.
- Provide personalized pricing plans or usage-based adjustments for customers with high monthly charges.
- Introduce bundled service packages, promotional upgrades, or personalized service recommendations to boost engagement.
- Implement proactive communication such as early churn-risk notifications, customer support outreach, and satisfaction surveys to address issues before cancellation.

4. Documentation of Limitations and Proposed Solutions:

Limitations:

- Missing Emotional and Qualitative Data: The dataset lacks customer complaints, support interactions, or satisfaction scores, which may influence churn behavior.
- Class Imbalance: Fewer churn cases compared to non-churn cases can bias the model toward predicting the majority class.
- Model Interpretability: High-accuracy models, like tree-based ensembles, can be difficult for non-technical stakeholders to understand.
- Data Quality Issues: Missing values, inconsistent categorical entries, and data transformations can affect performance.
- Overfitting Risk: Complex models may overfit historical patterns without proper regularization and validation.
- Static Snapshot: The dataset represents a single point in time, potentially missing evolving customer behaviors and market trends.
- Hyperparameter Sensitivity: Model performance can vary significantly based on parameter settings.

Proposed Solutions:

- Address Class Imbalance: Use resampling techniques (e.g., SMOTE) or class weighting to handle unequal class distribution.
- Enhance Interpretability: Apply feature importance analysis or SHAP-style explanations for clearer insights.
- Improve Data Quality: Implement automated checks and preprocessing pipelines to handle missing or inconsistent data.
- Prevent Overfitting: Use validation strategies, hyperparameter tuning, and early stopping during training.
- Adapt to Changing Behavior: Conduct time-based validation and regularly retrain the model with new data.
- Optimize Hyperparameters: Use techniques like grid search or cross-validation to stabilize model performance.
- Continuous Monitoring: Track model accuracy over time to ensure reliable long-term predictions.

Conclusion:

The customer churn prediction and segmentation project successfully identified high-risk customers and provided a clear understanding of different customer segments. By combining clustering techniques to group similar customers with supervised learning to predict churn, the company can shift from reactive measures to proactive retention strategies.

As a Business Analyst, the key contribution lies in translating these analytical insights into actionable recommendations. These findings support informed decision-making and help the business improve customer retention, enhance loyalty, and increase long-term customer value.