DECODING CHURN

A Data-Driven Strategy for Customer Retention

Prepared by : Winnie Njoroge

BUSINESS UNDERSTANDING: THE PROBLEM

⚠ REVENUE LOSS

SyriaTel is losing revenue due to customer churn. This is a significant challenge because **acquiring new customers is more expensive than retaining existing ones**.

Current retention strategies are reactive, often addressing the issue only after the customer has already decided to leave.

◎ THE NEED

A data-driven approach is needed to identify customers at **high risk of churning** before they leave.

This will allow the marketing and retention teams to apply targeted strategies more effectively, optimizing the budget and saving revenue.

PROJECT OBJECTIVES

PRIMARY GOAL

Maximize Recall.

Our priority is to catch as many potential churners as possible. We prioritize minimizing "False Negatives" (missed churners) over precision, as losing a customer is more costly than a retention offer.

SECONDARY GOALS

Identify Key Features Uncover drivers like pricing and service quality to inform retention strategy.

Compare Models Evaluate algorithms to balance the Precision-Recall tradeoff effectively.

Assess Performance Assess model performance using business-relevant evaluation metrics.

DATA OVERVIEW: PROFILE & BALANCE

3,333

TOTAL CUSTOMERS

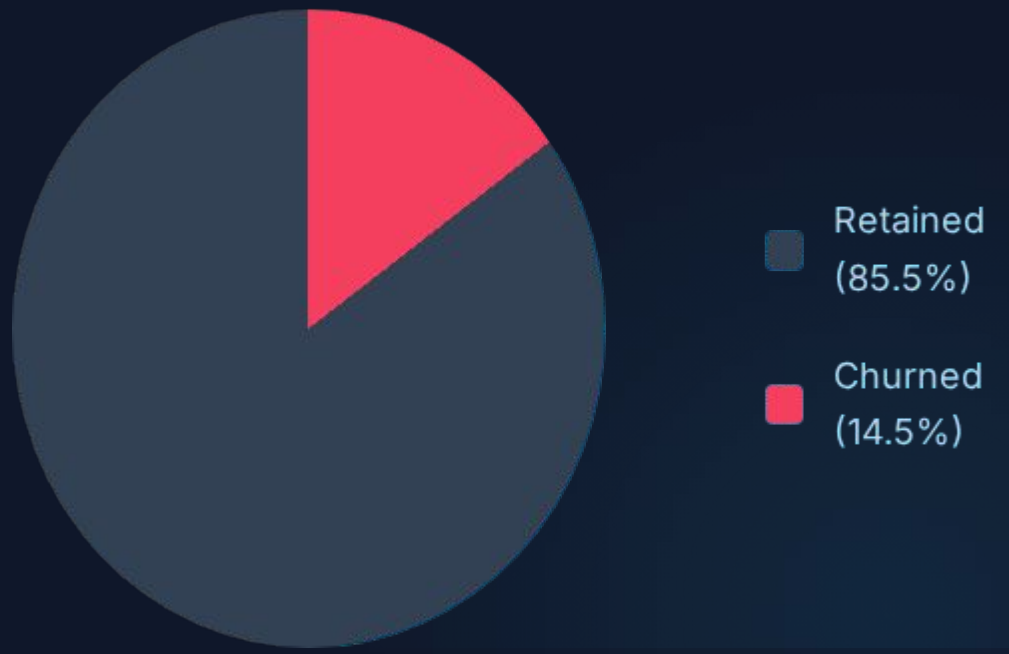
Complete historical usage records

21

FEATURES RECORDED

Including call minutes, plans, & service calls

> **The Churn Reality:** Only 14.5% of customers churned. This "Class Imbalance" makes accuracy a misleading metric for success.



BEHAVIORAL FEATURE ENGINEERING

To improve predictive power, we moved beyond raw data to create advanced behavioral features:

Total Minutes This represents the customer's overall reliance on the network regardless of time.

Total Charges This serves as a proxy for price sensitivity—higher spenders are often the first to churn if value drops.

Peak Period Identified the specific time of day (Day, Evening, Night, or Intl) where the customer has the highest usage intensity.

Day-Night Ratio Captured the balance between peak and off-peak usage to detect unusual lifestyle patterns or outliers.

Intl Usage Flag A binary indicator for high-intensity international users, helping segment customers with specific global connectivity needs.

METHODOLOGY: THE "ANTI-LEAKAGE" STRATEGY

PREPROCESSING

Applied **StandardScaler** and **OneHotEncoder** to ensure the model treats all numerical magnitudes and categories fairly without bias.

SMOTE STRATEGY

Used **Synthetic Minority Over-sampling Technique** to "teach" the model what a churner looks like by creating synthetic data points.

INTEGRITY CHECK

Crucially, SMOTE was applied **only to the training data**. The test set remains pure to ensure our performance results are realistic.

MODEL COMPARISON RESULTS

LOGISTIC REGRESSION

The Baseline. Provided a good starting point but struggled to capture complex, non-linear customer behaviors.

DECISION TREE

The Learner. Better at capturing patterns but highly prone to "overfitting"—memorizing the data rather than learning from it.

RANDOM FOREST

The Champion. An Ensemble Model that combines multiple perspectives to provide the most stable, reliable, and high-recall predictions.

PERFORMANCE RESULTS

THE OUTCOME

Model: Final Hypertuned Random Forest (Threshold: 0.65)

Our optimized model successfully identifies a significant majority of customers who are about to leave.

- **Recall Score: ~66%**
  (Captures approx. 3 out of 5 churners)

- **F1 Score: ~74%**
  (Balanced precision and recall)

RECALL MAXIMIZATION

By prioritizing Recall, we accept some "False Alarms" to ensure we don't miss the expensive loss of a customer.

TOP DRIVERS OF CHURN (FEATURE IMPORTANCE)

Insight: Customers with high bills who have called support 3+ times are in the "Danger Zone".

RECOMMENDATIONS FOR RETENTION

High-Usage Discounts Proactively offer loyalty discounts to the "High Charge" segment identified by the model. Don't wait for them to complain.

The "3-Call" Rule Flag any customer who makes their 3rd service call for immediate follow-up by a senior retention specialist.

International Plan Audit Review the value proposition of the international plan. It is currently a "churn magnet," suggesting pricing or connection quality issues.

CONCLUSION & NEXT STEPS

ADDRESS KEY DRIVERS OF CHURN

A/B test retention offers (e.g., 10% discount vs. free data) to minimize churn.

ACTION PLAN

Deploy the model to flag customers with >3 service calls and offer proactive "Usage Discounts" to high spenders.

"Retention starts with Detection"

THANK YOU

A detailed analysis can be found on my

github:https://github.com/NjorogeWinnie/Telecom-churn-analysis

---

Winnie Njoroge

Data Scientist

njorogewinnie573@gmail.com