SYRIATEL CUSTOMER CHURN PREDICTION PREDICTIVE ANALYTICS FOR CUSTOMER RETENTION

FAITH MUMBI GITHAIGA

JUNE 2025

OVERVIEW

- **Objective:** Predict whether a SyriaTel customer is likely to churn using historical data.
- **Problem Type:** Binary classification
- **Target Audience:** Business stakeholders aiming to reduce customer churn and improve retention.

BUSINESS UNDERSTANDING

- **Challenge:** High customer churn impacts revenue and increases acquisition costs.
- **Goal:** Use predictive analytics to identify patterns leading to churn.
- **Value:** Enables proactive retention strategies, increasing customer lifetime value.

DATA UNDERSTANDING

- **Dataset:** 3333 customer records and 21 features
- **Features:** Demographics, usage behavior, service plans, customer service calls
- **Target Variable:** churn (0 = No, 1 = Yes)

DATA EXPLORATION

- **Univariate Analysis:** Distribution of churn, service usage.
- **Bivariate Analysis:** Relationship between churn and international/voice mail plans.
- **Insights:** Churned customers had higher day-time charges and more customer service calls.

DATA PREPROCESSING

- Handled missing values and data consistency
- Encoded categorical features using Label Encoding
- Normalized numerical variables with StandardScaler
- Train-test split (80-20 ratio)

MODELING

- Models used:
  - Logistic Regression
  - Decision Tree

- Decision Tree gave the best performance in terms of balance and interpretability

EVALUATION METRICS

Metric	Logistic Regression	Decision Tree
Accuracy	86.5%	90.9%
Precision	57.8%	69.0%
Recall	25.5%	67.6%
F1 Score	35.4%	68.3%
ROC AUC Score	81.4%	81.2%

- **Decision Tree outperforms Logistic Regression** in Precision, Recall, F1 Score, and Accuracy.

- **High Recall (67.6%)** means it detects most customers likely to churn.

- Logistic Regression performs worse in identifying churners (low Recall: 25.5%).

FINDINGS

•Machine learning models can effectively predict which customers are likely to churn based on service usage and account features.

• Achieved **90.9% accuracy** and **68.3% F1 Score**, indicating a good balance between precision and recall.

•Despite a high ROC AUC (81.4%), it had a **low recall (25.5%)**, meaning it missed many actual churn cases.
•Factors like **Monthly Charges**, **Contract Type**, **Tenure**, and **Customer Service Calls** contributed heavily to churn prediction.

•The Decision Tree's ability to capture **67.6% of churners** enables proactive customer retention strategies.

RECOMMENDATIONS

•Use the model to identify customers with a high likelihood of churning and intervene with personalized retention offers (discounts, loyalty rewards, or improved service).

•Since service-related features influence churn, invest in improving the customer experience, especially around billing and technical support.

•Focus on factors like **monthly charges**, **contract type**, and **tenure** to monitor and act on early signs of dissatisfaction.

•Retrain the model regularly with new data to adapt to changing customer behavior and improve prediction accuracy.

•Experiment with retention campaigns using model predictions and track which interventions lead to reduced churn.

NEXTSTEPS

- Set up automated churn alerts
- Continuously retrain model with fresh data
- Monitor performance and feedback for improvement

THANK YOU
QUESTIONS?
CONTACT: GFAITHMUMBI@GMAIL.COM