Telco Customer Churn Prediction: Final Model Development and Analysis

1. Executive Summary

This report outlines the development and optimization of a machine learning model designed to predict telecom customer churn. The final tuned XGBoost Classifier achieved:

- Recall: 0.80

- Precision: 0.58

- F1-Score: 0.67

- ROC AUC: 0.853

2. Exploratory Data Analysis (EDA): Key Churn Drivers

- 73% Non-Churn, 27% Churn (imbalanced dataset).

- Month-to-month contracts strongly correlate with churn.

- Low tenure customers churn the most.

- Fiber Optic users and customers with high monthly charges show elevated churn rates.

- Senior citizens and customers without partners/dependents churn more often.

3. Model Development and Evaluation Strategy

Boosting models outperformed baseline methods. Evaluation prioritized Recall, Precision, F1-Score, and ROC AUC.

4. Final Model Selection and Optimization

XGBoost hyperparameters:

- learning_rate: 0.1

- max_depth: 5

- n_estimators: 100

Threshold Optimization:

- Best threshold: 0.5503 (Recall: 0.80, Precision: 0.58, F1: 0.67)

5. Conclusion and Recommendations

Deploy the model with threshold = 0.55.

Target month-to-month, low-tenure, high-charge customers for retention.

Monitor false positives against retained revenue.