Customer Churn Analyzer

A Data-Driven Approach to Predicting and
Reducing Customer Churn

Agenda

- 1. Business Problem & Objective
- 2. Data Overview & Preparation
- 3. Key Insights from Data Analysis
- 4. Model Development & Performance
- 5. Key Churn Drivers
- 6. Recommendations & Next Steps

Business Problem & Project Objective

- **Business Problem:**
  - High customer churn directly impacts revenue, as acquiring new customers is far more expensive than retaining existing ones.

- **Project Objective:**
  - Build a predictive model to identify customers at high risk of churning and understand the key drivers to inform proactive retention strategies.

Data Overview & Preparation

- **Dataset:**
  - Customer Churn.csv (3150 rows, 14 columns)

-

  Data Cleaning Steps:
  - Standardized column names for consistency.
  - Removed 300 duplicate rows.
  - Confirmed no missing values were present.

Key Insights from Data Analysis

- **Churn Distribution:**
  - Data is imbalanced (15.6% churn rate). This was handled using `class_weight='balanced'` during modeling to prevent bias.

- **Key Correlations:**
  - 'Complains' and 'Subscription_Length' show the strongest correlation with Churn.

- **Usage Patterns:**
  - Churners exhibit significantly lower call duration ('Seconds_of_Use') and SMS frequency.

Model Development & Selection

- **Models Tested:**
  - Logistic Regression, Decision Tree, and Random Forest.

- **Optimization:**
  - GridSearchCV was used to tune the Random Forest model's hyperparameters, optimizing for F1-score to handle the class imbalance.

- **Final Model Selection:**
  - The Optimized Random Forest was selected due to its superior performance on the F1-score and overall accuracy.

Final Model Performance

- **Optimized Random Forest Metrics:**
  - Accuracy: 96.1%
  - **F1-Score (for Churn class): 0.870**
  - AUC (Area Under ROC Curve): 0.983

Key Drivers of Customer Churn

- The most influential factors predicting churn were:
  - **1. Subscription Length (Higher impact)**
  - **2. Complains**
  - 3. Seconds of Use
  - 4. Frequency of SMS
  - 5. Distinct Called Numbers

Actionable Business Recommendations

- **1. Onboarding & Loyalty:**
  - Target new customers (low subscription length) with enhanced onboarding and early loyalty rewards.

- **2. Proactive Complaint Management:**
  - Implement a high-priority alert system for any customer complaint, triggering a retention specialist to engage.

- **3. Usage-Based Re-engagement:**
  - Automate outreach campaigns for customers whose call or SMS usage drops below a certain threshold.

- **4. Tariff Plan Optimization:**
  - Review and A/B test Tariff Plan 2 to improve its value proposition and reduce its associated higher churn rate.

Conclusion & Next Steps

- **Conclusion:**
  - The Optimized Random Forest model provides a reliable tool for churn prediction. The key churn drivers are clear indicators of customer disengagement and dissatisfaction.

- **Next Steps:**
  - Deploy the model via the Streamlit app for business user access.
  - Integrate the model with CRM systems for real-time scoring and automated retention alerts.
  - Continuously monitor and retrain the model with new data to maintain its accuracy.

THANK YOU