TELCO CUSTOMER CHURN PREDICTION AND ANALYSIS

TEAM MEMEBERS:

SHAHD MOHAMED MOHI-ELDIEN ELKILANY

ASSIA MOHAMED MOHMAED

HOIDA MOAWAD IBRAHIM

NOSIBA HAMDY MOHAMMED

PROJECT

PROBLEM STATEMENT

Customer churn is a critical challenge for businesses, especially in highly competitive industries such as telecommunications, retail, and subscription-based services. Retaining customers is significantly more cost-effective than acquiring new ones, yet identifying customers at risk of leaving remains a complex task.
The problem lies in predicting churn with high accuracy and understanding the key drivers behind it, enabling businesses to proactively intervene and implement retention strategies. Without an effective solution, businesses face revenue losses, diminished customer loyalty, and reduced market competitiveness.

PROJECT OBJECTIVE

- **Predict Customer Churn:**
  Develop a machine learning model to accurately identify customers likely to leave the business.
- **Understand Churn Drivers:**
  Perform exploratory data analysis to uncover key factors contributing to churn.
- **Enhance Business Decision-Making:**
  Provide actionable insights that enable targeted retention strategies, reducing overall churn rates.
- **Create a Scalable Solution:**
  Deploy the model as an API or a web-based application to ensure usability for stakeholders.
- **Monitor and Optimize:**
  Establish a system for tracking model performance over time and retraining it with new data to maintain accuracy.

DATASET OVERVIEW
TELCO CUSTOMER CHURN

Dataset Overview:

The Telco Customer Churn dataset contains detailed information about customers, including their demographics, service subscriptions, billing details, and churn status. This structured data can be leveraged to identify patterns and relationships that drive churn.

Key Features:

- **Customer Demographics:** Gender, age group, and whether the customer is a senior citizen.
- **Service Details:** Internet service type, additional services (e.g., online security, streaming), contract type.
- **Billing Information:** Monthly charges, total charges, payment method.
- **Target Variable:** Churn (Yes/No).
- **Data Composition:**
  - Total Rows: 7,043
  - Total Columns: 21
- **Churn Distribution:** A moderately imbalanced dataset with ~26% of customers marked as churned.

KEY PATTERNS AND INSIGHTS

TELCO CUSTOMER CHURN

Churn Rates Across Contracts:
Customers on month-to-month contracts churn significantly more than those on longer contracts.

Payment Methods Impact:
Electronic checks correlate with higher churn rates, while automatic payments show lower churn rates.

Monthly Charges:
Higher monthly charges are associated with increased churn, particularly among customers using multiple add-on services.

Tenure:
Longer-tenure customers are less likely to churn, emphasizing loyalty over time.

Service Types:
Fiber optic users exhibit higher churn rates compared to DSL users.

DATA EXPLORATION VISUALIZATIONS SUMMARY

Distributions:

1. **Distribution of Total Charges:**
   - **A histogram shows the spread of Total Charges across customers.**
   - **Most customers have total charges clustered around lower values, with a tail towards higher charges.**
2. **Distribution of Monthly Charges:**
   - **Histogram and box plot reveal a wide range of Monthly Charges, with a median near mid-range values.**
   - **Outliers indicate some customers with significantly higher charges.**
3. **Distribution of Gender:**
   - **Count plots and box plots display a near-even split between genders.**
   - **No major differences in monthly charges by gender.**

DATA EXPLORATION VISUALIZATIONS SUMMARY

Relationships Between Features:

1. **Scatter Plot of Tenure vs. Total Charges:**
   - **Indicates a positive correlation; longer tenure generally leads to higher total charges.**
2. **Scatter Plot of Monthly Charges vs. Total Charges (Colored by Churn Status):**
   - **Shows the relationship between Monthly Charges and Total Charges with churn status differentiation.**
   - **Customers with higher monthly charges tend to churn more frequently.**
3. **Total Charges Over Tenure:**
   - **A line plot or scatter plot highlights how total charges increase over time with tenure.**

DATA EXPLORATION VISUALIZATIONS SUMMARY

Churn-Specific Visualizations:

1. **Box Plot of Monthly Charges by Churn Status:**
   - **Indicates that customers who churn tend to have higher monthly charges on average.**
2. **Violin Plot of Total Charges by Churn Status:**
   - **Shows the distribution and density of total charges for churned and retained customers.**
   - **Churned customers often fall in specific total charge ranges.**
3. **Count Plot of Churn Distribution:**
   - **A bar chart displaying the overall distribution of churned vs. non-churned customers.**

DATA EXPLORATION VISUALIZATIONS SUMMARY

Churn by Categories:

1. **Churn by Contract Type:**
   - **A count plot illustrates higher churn rates among customers with a Month-to-Month Contract compared to One-Year or Two-Year Contracts.**
2. **Churn by Internet Service Type:**
   - **Customers with Fiber Optic Internet Service show higher churn rates compared to DSL or no internet service.**
3. **Churn by Payment Method:**
   - **Customers using Electronic Check as a payment method show significantly higher churn rates compared to other methods like bank transfers or credit cards.**

FEATURE ENGINEERING

Key Features Identified:

- **Customer Tenure:** Duration of the customer's relationship with the company.
- **Monthly Charges:** Recurring charges for services, a key indicator of customer satisfaction.
- **Total Charges:** Cumulative charges, which correlate with tenure and monthly charges.
- **Contract Type:** Indicates customer commitment and correlates with churn likelihood.
- **Payment Method:** Preferred payment method, such as electronic check or credit card, associated with churn trends.
- **Internet Service Type:** Type of internet service (e.g., Fiber Optic, DSL), impacting satisfaction and retention.
- **Additional Services:** Usage of value-added services (e.g., streaming TV, phone).

STUDIO SHODWE

FEATURE ENGINEERING

New Features Created:

- **Average Monthly Spend:** Derived from Total Charges / Tenure, providing insights into spending patterns.
- **Senior Citizen Indicator:** Binary feature for whether the customer is a senior citizen.
- **Churn Score (Categorical):** A binned version of churn likelihood for targeted marketing strategies.
- **Multiple Services Used:** Summing binary flags for additional services to create a service usage score.

FEATURE ENGINEERING

Importance Rankings (Based on Model Results):

1. **OnlineSecurity_Yes**

Indicates whether a customer has online security enabled. A critical factor for churn prediction.

1. **DeviceProtection_No internet service**

Reflects whether customers without internet service opted for device protection.

1. **PaymentMethod_Electronic check**

Highlights the payment method most associated with churn, emphasizing the significance of payment convenience and user experience.

1. **Contract_One year**

Shows how contract duration influences customer loyalty, with one-year contracts being a key variable.

1. **InternetService_Fiber optic**

Indicates how the type of internet service affects churn, with fiber-optic users showing specific patterns.

1. **TotalCharges**

Represents the cumulative charges paid by customers, strongly linked to their likelihood to churn.

1. **Contract_Two year**

Illustrates the impact of longer contracts on churn, often associated with higher retention.

1. **PhoneService**

Indicates whether a customer has phone service, influencing churn differently based on usage and bundling.

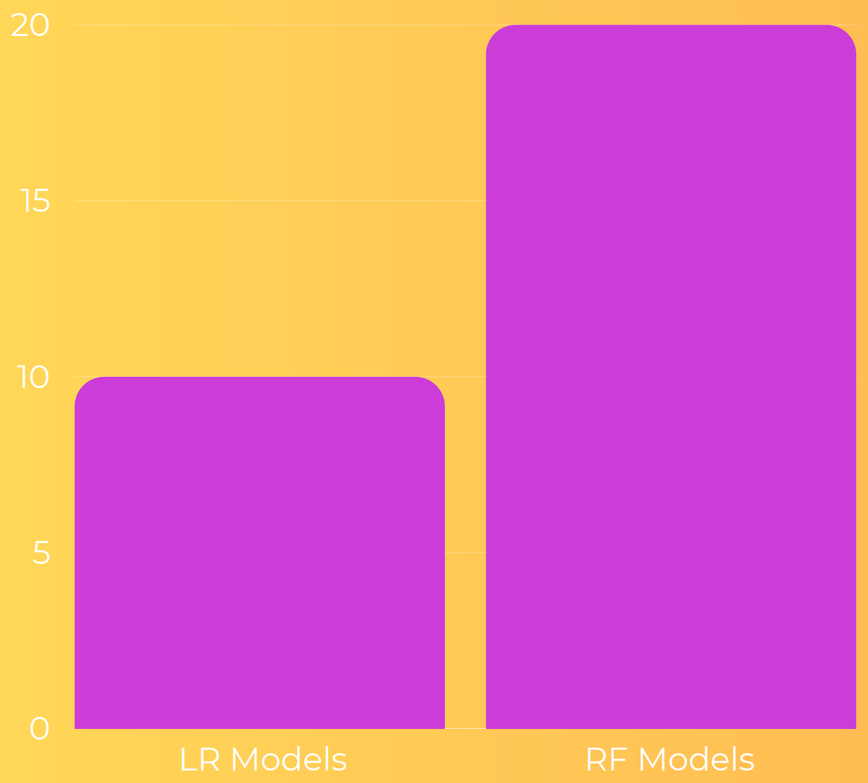
1. **tenure**

Represents the length of time a customer has been with the company, with longer tenures correlating to loyalty.

1. **PaperlessBilling**

Indicates whether customers prefer paperless billing, a modern convenience that may affect retention.

STUDIO SHODWE



MODEL DEVELOPMENT SUMMARY

Best Model Summary

- **Best Model: Random Forest with Min-Max Scaling.**
- **Metrics:**
  - **Accuracy: 0.878**
  - **Precision: 0.836**
  - **Recall: 0.669**
  - **F1 Score: 0.743**
  - **ROC AUC: 0.949**
- **Best Hyperparameters:**
  - **Number of estimators: 200**
  - **Max depth: 30**

STUDIO SHODWE

THANK YOU!

DATA ANALYSIS IS KEY TO BUSINESS GROWTH AND SUCCESS!

WWW.REALLYGREATSITE.COM