Project : Telecom Customer Churn

With the rapid development of the telecommunication industry, the service providers are inclined more towards expansion of the subscriber base. To meet the need of surviving in the competitive environment, the retention of existing customers has become a huge challenge. It is stated that the cost of acquiring a new customer is far more than that for retaining the existing one. Therefore, it is imperative for the telecom industries to use advanced analytics to understand consumer behavior and in-turn predict the association of the customers as whether or not they will leave the company

Dataset: Click here for the dataset

Data Description:

Each row represents a customer, each column contains customer's attributes described on the column Metadata.

- **customerID:** A unique identifier for each customer. It's used to track individual customers.
- **gender:** The gender of the customer (e.g., Male or Female).
- SeniorCitizen: Indicates whether the customer is a senior citizen (1 if they are, 0 if they are not).
- **Partner:** Shows if the customer has a partner (Yes or No).
- **Dependents:** Shows if the customer has any dependents (Yes or No).
- **tenure:** The number of months the customer has been with the company.
- **PhoneService:** Indicates whether the customer has phone service (Yes or No).
- **MultipleLines:** Shows if the customer has multiple phone lines (Yes, No, or No phone service).
- **InternetService:** The type of internet service the customer has (e.g., DSL, Fiber optic, or No).
- **OnlineSecurity:** Indicates whether the customer has online security (Yes or No).
- **OnlineBackup:** Indicates whether the customer has online backup (Yes or No).
- **DeviceProtection:** Shows if the customer has device protection (Yes or No).
- **TechSupport:** Indicates whether the customer has tech support (Yes or No).
- **StreamingTV**: Shows if the customer has streaming TV service (Yes or No).
- **StreamingMovies**: Indicates whether the customer has streaming movies service (Yes or No).

- **Contract**: The type of contract the customer has (e.g., Month-to-month, One year, or Two year).
- **PaperlessBilling**: Shows if the customer has opted for paperless billing (Yes or No).
- **PaymentMethod**: The method the customer uses to pay (e.g., Electronic check, Mailed check, Bank transfer, or Credit card).
- **MonthlyCharges**: The amount the customer is billed each month.
- **TotalCharges**: The total amount the customer has been billed since they started the service.
- **Churn**: Indicates whether the customer has left the service (Yes or No).

Sprint 0: Perform EDA With Basic Plots

1. **Understand the Data:** Familiarize yourself with the dataset and its attributes.
2. **Define the Problem Statement:** Describe the primary question(s) this dataset can help you answer.
3. **Data Cleaning and Manipulation:** Handle missing values, correct data types, and resolve other data quality issues.
4. **Identify Data Types:** Classify attributes as discrete, continuous, or ordinal, and identify relationships among them.
5. **Formulate Analysis Questions:** Create specific questions to guide your analysis based on the problem statement.
6. **Visualization:**
    - Explore the distribution of numerical and categorical variables.
    - Identify apparent trends or patterns.
    - Investigate relationships between variables.
    - Identify potential correlations and interactions.
    - Visualize relationships and correlations among multiple variables.

Sprint 1: Preprocessing

1. **Data Preparation:**
    - Split the dataset into training and testing sets.
    - Perform feature engineering.
    - Conduct data preprocessing (e.g., normalization, encoding).

Sprint 2: Model Development

1. **Model Training:** Train machine learning models (e.g., Logistic Regression, Decision Trees, Random Forests).
2. **Model Evaluation:** Evaluate model performance using metrics like accuracy, precision, and recall.

Sprint 3: Model Deployment

- **Deployment:** Deploy the model using your preferred method.

Submission:

- Submit your Python notebook (.ipynb) on the LMS, including a summary of your process and results.