Telecommunication Case Study

Business Understanding

The telecommunications sector has become one of the main industries in developed countries. The technical progress and the increasing number of operators raised the level of competition. Companies are working hard to survive in this competitive market depending on multiple strategies. Three main strategies have been proposed to generate more revenues:
1. Acquire new customers
2. Upsell the existing customers
3. Increase the retention period of customers
However, comparing these strategies taking the value of return on investment (RoI) of each into account has shown that the third strategy is the most profitable strategy, proving that retaining an existing customer costs much lower than acquiring a new one, in addition to being considered much easier than the upselling strategy. To apply the third strategy, companies have to decrease the potential of customer's churn, known as "the customer movement from one provider to another".

Customer churn is a considerable concern in service sectors with high competitive services. On the other hand, predicting the customers who are likely to leave the company will represent a potentially large additional revenue source if it is done in the early phase.

Many research confirmed that machine learning technology is highly efficient to predict this situation. This technique is applied through learning from previous data.

Customer Churn is one of the most important and challenging problems for businesses such as Credit Card companies, cable service providers, SASS and telecommunication companies worldwide. Even though it is not the most fun to look at, customer churn metrics can help businesses improve customer retention.

Customer Churn Dataset

Domain - Telecommunication
Dataset - Click here to download the dataset: **churn_dataset.csv**

The data set includes information about:

1. Customers who left within the last month – the column is called Churn
2. Services that each customer has signed up for – phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
3. Customer account information – how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges
4. Demographic info about customers – gender, age range, and if they have partners and dependents

SPRINT 1 - Exploratory Data Analysis

Task - This is an open ended question. Kindly apply all your knowledge to perform an exploratory data analysis on the given dataset. It is known that the target variable is **Churn**.

Write proper conclusions and provide recommendations to the telecom company based on the insights.

References -
https://www.kaggle.com/code/prathameshgadekar/world-population-eda-with-world-map-visualization

SPRINT 2 - Data Preparation and Model Building

Problem Statement - Given various features about a customer like <u>Gender, SeniorCitizen, Partner, Dependents etc..</u> , predict if the customer will <u>churn or not</u>.

Task - Prepare the Data and build a model to predict the churn of a customer.

Step - 1: Load the data
Step - 2: Document the below mentioned points properly:
    - Identify the input and output/target variables.
    - Identify the type of ML Task.
    - Identify the Evaluation Metric.
        - For regression task - Mean Absolute Error
        - For classification task - Accuracy
Step - 3: Split the dataset into Training and Testing (recommended 75:25 split).
Step - 4: Data preparation on train data:
    - For Numerical Variables - Standardization or Normalization (Fit and Transform)
    - For Categorical - LabelEncoding or OneHotEncoding (Choose wisely)
Step - 5: Data preparation on test data:
    - For Numerical Variables - Standardization (Transform)
    - For Categorical - LabelEncoding or OneHotEncoding (Choose wisely)
Step - 6: Model Training Phase - Use all the algorithms mentioned below to train separate models:
    - KNN
    - Logistic Regression
    - Support Vector Machines
    - Decision Trees
    - Random Forest
Step - 7: Predict and evaluate each model separately using the correct evaluation metric. Use metrics.accuracy_score(actual, predict).
Step - 8: Display a plot which shows all the algorithms applied along with the accuracies achieved. **Write your conclusion on the best algorithm for Churn Prediction**.