

Customer Churn Analysis and Prediction – Documentation

Objective:

The objective of this project is to analyse customer churn in a telecommunications company and develop predictive models to understand the factors influencing customer churn. The goal is to provide insights that can help the company reduce churn rates and improve customer retention.

Design Thinking Process:

1. **Define:** Understand the problem of customer churn in the telecommunications industry and define the objectives of the analysis, including identifying key factors affecting churn and building predictive models.
2. **Research:** Collect and prepare the dataset containing customer information, services, and churn status. Research the data for any issues like missing values, data types, and outliers.
3. **Ideate:** Develop a plan for exploratory data analysis (EDA) to understand the data and identify insights. Consider the design of Python code for analysis and visualization.
4. **Prototype:** Build Python code for EDA, data visualization, and predictive modelling. Integrate IBM Cognos for data visualization.
5. **Test:** Evaluate the effectiveness of the Python code and IBM Cognos visualizations in providing valuable insights.
6. **Implement:** Deploy the code and visualization tools for ongoing analysis.

Development Phases:

1. **Data Collection:**
 - Collect customer data, including services, contract details, payment methods, demographic information, and churn status.
 - The data used was collected from the online sources.
 - Perform data cleaning, addressing missing values and data type issues.
2. **Data Visualization:**
 - Use IBM Cognos for initial data visualization, providing a high-level overview of key metrics.
 - Create visualizations to understand customer demographics, contract details, and payment methods.
3. **Exploratory Data Analysis (EDA) - Python:**
 - Perform detailed EDA using Python to gain deeper insights into the data.
 - Analyse data statistics, data types, and the impact of customer services on churn.
4. **Feature Engineering:**
 - Encode categorical features and apply feature scaling as necessary.
 - Create one-hot encoding for categorical features with multiple categories.

5. Model Selection and Evaluation - Python:

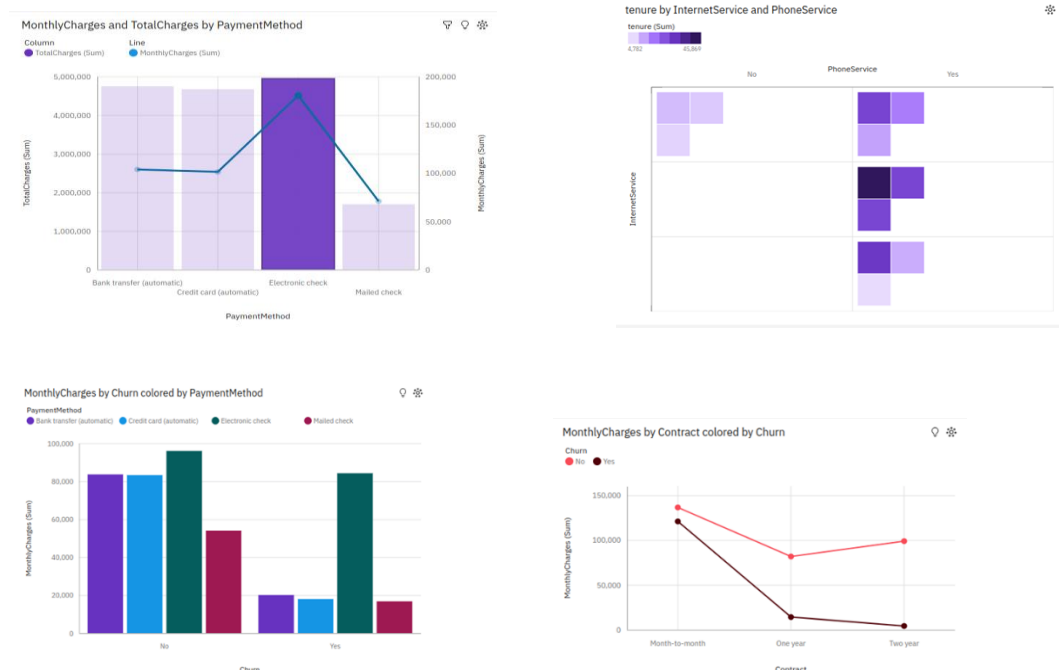
- Develop Python code to split the data into training and testing sets.
- Implement and evaluate different models, including K-Nearest Neighbours, Logistic Regression, and a Neural Network.
- Assess model performance using metrics like accuracy, ROC curves, AUC scores, precision-recall curves, and F1 scores.

6. Insights and Recommendations:

- Present the insights from the analysis, including factors affecting churn, feature importance, and model performance.
- Provide recommendations for the telecommunications company to improve customer retention.

Data Visualization Using IBM Cognos:

IBM Cognos is utilized to provide an initial overview of key metrics, customer demographics, contract details, and payment methods. It helps in presenting high-level visualizations and summaries of the data. Some of the visualizations done by this tool are:



Python Code Integration:

- Python is used for in-depth analysis, feature engineering, and model development.
- Python code is integrated into the project to perform exploratory data analysis and build predictive models.

How Insights Improve User Experience

1. The insights gained from the analysis can help the telecommunications company understand the reasons behind customer churn.
2. By identifying specific customer demographics, services, and contract details that are associated with higher churn rates, the company can tailor its marketing and customer retention strategies.
3. Understanding the impact of payment methods on churn can lead to improvements in payment processing, which can enhance the customer experience.
4. The predictive models can be used to proactively identify customers at risk of churn and implement targeted retention efforts, ultimately leading to a better user experience for retained customers.