Project Title-Customer Churn Prediction

# Introduction:

Customer churn prediction is a vital component for customer relationship management and business analysis. It involves the usage of strategies like “Data Analysis” and “Machine Learning” to forecast when and why the customers may stop using a product or service.

The primary goal of customer churn prediction is to identify at risk customers early on and take proactive measures to retain them.By analyzing historical data of the customers and various business factors that may influence businesses can develop many predictive strategies which may help them in maintaining their customers and will probably help in increasing their customer services.

# About Phase4:

The phase 4 is all about data preprocessing of the given csv file for the purpose of performing various operations such as analysis, exploratory data analysis and visualizing of the dataset.

# Phase 4 of Customer churn prediction:

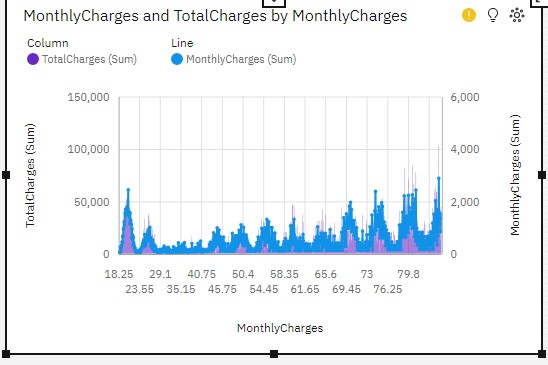
The phase 4 of the project customer churn prediction refers to visualizing of the data using the “IBM Cognos Tool”. The various charts displayed in this document are Bar chart,Pie chart,Line chart and scatter plot.

# About IBM Cognos:

The IBM Cognos tool is used for analyzing the files such as csv files and other files to visualize data from them.

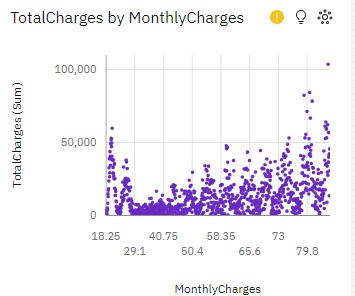
# 1)Line and column chart:

The line and column chart is the combined version of column chart and the line chart.In the below chart it depicts the relation of Total charges with monthly charges.



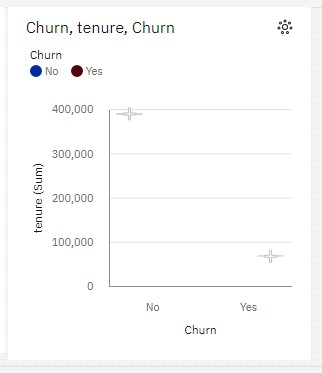
# 2)Packed bubble chart:

The packed bubble chart is a modified version of the scatterplot and point chart.



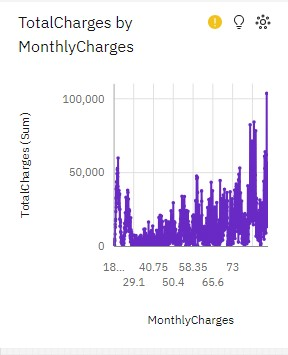
# 3)Boxplot:

A box plot, also known as a whisker plot, is a graphical representation of the distribution of a dataset.



# 4)Bubble chart:

The bubble chart is a modified version of scatterplot but only uses scatter function.



# Performance Metrics:

Accuracy is a common metric used to evaluate the performance of machine learning models, especially in classification tasks. It measures the proportion of correctly classified instances out of the total instances in a dataset.

The test accuracy is found out to be,Test Accuracy=0.89