KINGS ENGINEERING COLLEGE
Chennai-Bangalore Highway, Irungattukottai, Sriperumbudur, Chennai - 602 117.
Ph: 044 - 71224401 – 08. Fax: 044 - 71224410
ACCREDITED WITH NAAC, APPROVED BY AICTE & AFFILIATED TO ANNA UNIVERSITY

Department of Artificial Intelligence & Data Science

College Code: 2108

Project Domain: Data Analysis

Project Title : Customer Churn Prediction

Project Mentor: Mr. Sundarajan

Team Members:

1) **Jeredson Daniel Raj S (210821243023)**
2) **Kumaran M (210821243032)**
3) **Anbumani (210821243005)**
4) **Venkat Prasad (210821243061)**

Abstract:

"Customer churn prediction is a data-driven approach that leverages machine learning to forecast when customers are likely to discontinue their engagement with a business. By analyzing historical customer data, such as transaction history, demographics, and behavior patterns, predictive models can identify potential churners. This empowers businesses to take proactive measures to retain valuable customers, enhance customer satisfaction, and optimize marketing strategies, ultimately reducing revenue loss and maintaining a loyal customer base.

Detailed Steps

1) Download the dataset from kaggle
   https://www.kaggle.com/datasets/blastchar/telco-customer-churn

2) Go through the dataset and understand the features and their trends.

3) Perform feature extraction - This approach is used to extract the features that are relevant to the Customer Churn Prediction. **Sci-kit Learn** library in python will be helpful in this process.

4) Choose the Deep Learning model that will help us to predict the Customer Churn using the extracted features. **Ski Kit Learn** library has prebuilt Deep learning models that can be simply fit into our data and get the model trained

5) Split the data into training and test data so that the model uses certain data for training purposes and after training the model can be evaluated using the test data. This can be done through **the train_test_split()** function from sklearn library

6) After the model is trained, find its accuracy using functions like mean_squared_error, accuracy or any other similar application.

Procedure Involved

DATA COLLECTION:

The first process in the project is to collect dataset about the Customer Churn and the data that shows about the Customer Churns in an organization. The main things that we need to be careful of while collecting is

- Data should be prevalent
- Data should be complete
- Data should be in convert form to perform analyze.

DATA CLEANING:

Once we collected data, the next process is to clean the data. We have process the data may contain some unwanted data or irrelevant data in it .so, we need to clean all these data to get the better analyze. Some of the important steps in data cleaning

- Looking for any missing data
- Looking for duplicate data
- Looking for irrelevant data

DATA PREPROCESSING:

Before we start to analyze the data, we need to preprocess the data. It is the process of making the data or transforming the data into the form that we need for analysis. It includes changing the dimensions of data.

DATA ANALYSIS:

Once the data is cleaned and preprocessed, we can now proceed with analysis of the data. In analyzing the data, we perform some of the statistical procedures to get to know about the Customer Churn and also the clear perspective of customer’s detail we need.

DATA VISUALIZATION:

It is the process where we can able to communicate with data. Visualization provides us the analysis in a way we can interpret at make decision. Some of the best ways to visualize data is by plotting it in graphs, charts etc..