IBM- NAAN MUDHALVAN

DAC_Phase 5

Name: AnbuMani K
Register Number: 210821243032
College: Kings Engineering College
Domain: Data Analytics with cognos
Project: Customer Churn Prediction

Program:

```
import pandas as pd
from matplotlib import pyplot as plt
import numpy as np
df = pd . read_csv(r"Customer Churn.csv")
df.sample(5)
df.drop('customerID' , axis='columns' , inplace = True)
df.dtypes
df.TotalCharges.values
df.MonthlyCharges.values
pd.to_numeric(df.TotalCharges, errors='coerce').isnull()
df1 = df[df.TotalCharges!= ' ']
df1

df1.shape
df1.dtypes
```

```
df1.TotalCharges = pd.to_numeric(df1.TotalCharges)
df1.TotalCharges.dtypes
tenure_churn_no = df1[df1.Churn=="No"].tenure
tenure_churn_yes = df1[df1.Churn=="Yes"].tenure

plt.hist([tenure_churn_yes,tenure_churn_no],
label=["Churn=Yes","Churn=No"])
plt.legend()

def unique_col_values(df):
    for col in df:
        if df[col].dtypes == 'object':
            print(f'{col}:{df[col].unique()}')

unique_col_values(df1)

df1.replace("No internet service", "No" , inplace =True)
df1.replace("No phone service", "No" , inplace =True)

unique_col_values(df1)

yes_no_columns = ['Partner','Dependents',
'PhoneService','MultipleLines','OnlineSecurity',
'OnlineBackup',
'DeviceProtection',
'TechSupport',
'StreamingTV',
'StreamingMovies',
'Contract',
'PaperlessBilling','Churn']

for col in yes_no_columns:
```

```
    df1[col] . replace({'Yes':1 , 'No':0},inplace=True)

for col in df1:
    print(f'{col}:{df1[col].unique()}')

df1['gender'].replace({'Female': 1 ,
'Male':0},inplace=True)
df1.gender.unique()

df2=pd.get_dummies(data=df1,columns=["InternetService","C
ontract","PaymentMethod"])
df2.columns

df2.sample(4)

df2.dtypes

col_to_scale = ['tenure', 'MonthlyCharges' ,
"TotalCharges"]

from sklearn . preprocessing import MinMaxScaler

scaler = MinMaxScaler()

df2[col_to_scale]  =
scaler.fit_transform(df2[col_to_scale])

df2.sample (4)

for col in df2:
    print(f'{col}:{df2[col].unique()}')
```

```
X= df2.drop('Churn',axis='columns')
Y=df2['Churn']

from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test =
train_test_split(X,Y,test_size=0.2,random_state=5)

X_train.shape
X_test.shape

X_train[:10]
len(X_train.columns)

import tensorflow as tf
from tensorflow import keras

model = keras.Sequential([
    keras.layers.Dense(20,input_shape=(26,),
activation='relu'),
    keras.layers.Dense(1, activation='sigmoid'),
])
model.compile(optimizer='adam',
              loss='binary_crossentropy',
              metrics=['accuracy'])
model.fit(X_train, Y_train, epochs=100)
```

1. Project Objective, Design Thinking Process, and Development Phases:

<u>Objective:</u> The main goal of our project is to predict when customers are likely to leave our business and, more importantly, to find ways to prevent that from happening.

<u>Design Thinking Process:</u>

a. Empathize: To kick things off, we put ourselves in our customers' shoes, trying to understand what drives them to leave. We collected data, like customer feedback and their history with us.

b. Define: With this understanding, we defined our problem clearly: customer churn. We set specific goals and metrics to measure our success.

c. Ideate: We brainstormed various techniques and data sources that might help us predict churn effectively.

d. Prototype: We created a predictive model using the chosen techniques.

e. Test: We assessed how well our model performs using past data, so we can make improvements.

f. Implement: Once our model is solid, we integrate it into our business processes and gather feedback for further refinement.

Development Phases:

> **Data Collection and Preprocessing:** We gathered historical data on our customers, including things like their demographics, purchase history, how they've interacted with us, and whether they've churned. We cleaned and organize this data.
>
> **Feature Engineering:** We pinpointed the most important features that influence churn, such as how long a customer has been with us, how often they interact with us, and their satisfaction scores.
>
> Model Selection: Based on our needs, we'll choose a suitable technique like logistic regression, decision trees, or something more advanced like neural networks.
>
> **Training and Validation:** We'll teach our model on part of the data and evaluate its performance using data we've kept aside. We'll fine-tune it if necessary.

Data Visualization: We'll use tools like IBM Cognos to create visualizations, making it easier to grasp data patterns and model results.
Deployment: Our predictive model will be put to work, helping us predict which customers are likely to churn in real time.
Monitoring and Continuous Improvement: We won't stop here. We'll constantly monitor how our model performs and make adjustments as customer behavior changes.

2. Analysis Objectives, Data Collection, Data Visualization, and Predictive Modeling:
Analysis Objectives:
a. Identify Churn Factors: We aim to find out what's causing customers to leave. Are there specific patterns or trends?
b. Predict Churn: We want to build a model that can forecast which customers are likely to churn in the future.
Data Collection Process:

- We'll collect data from a variety of sources, including our customer relationship management systems, sales records, customer support interactions, and feedback surveys.

Data Visualization using IBM Cognos:

- We'll utilize IBM Cognos to generate interactive dashboards and reports. These visuals will help us understand customer behavior, spot trends, and highlight crucial metrics.

Predictive Modeling:

- We'll choose the right algorithm (like logistic regression, random forests) and feed it historical customer data to train it.
- The data will be split into training and testing sets to assess how well our model works.
- We'll measure its performance using metrics such as accuracy, precision, recall, and F1-score.
- We'll fine-tune our model and tweak its settings as needed.

3. How Insights and Prediction Models Help Reduce Customer Churn:
Insights and prediction models are incredibly useful for businesses looking to cut down on customer churn:

a. Early Warning System: Our model will act as an early warning system, flagging customers at risk of leaving so we can take action to retain them.
b. Personalized Retention Strategies: The insights gained will help us understand why customers leave. We can then craft personalized strategies for each customer, making them more likely to stay.
c. Resource Allocation: Instead of using the same approach for everyone, we can allocate resources more efficiently, focusing on those at higher risk of churning.
d. Product and Service Improvement: Insights can reveal where we're falling short, allowing us to improve our offerings.
e. Customer Satisfaction: By addressing customer concerns and pain points, we can increase overall satisfaction, making it less likely that they'll leave.
f. Revenue Growth: Retaining more customers means increased revenue and better profitability for our business.