Logistic regression

import pandas as pd

data=pd.read\_csv(r"C:\Users\ADMIN\Downloads\telecom\_customer\_churn (1).csv")

data.head()

data=data.drop(["Churn Category","Churn Reason"],axis=1)

numerics = ['int16', 'int32', 'int64', 'float16', 'float32', 'float64','uint8']

odf = data.select\_dtypes(exclude=numerics)

for i in odf:

print(i," ",data[i].unique())

data = data.drop(columns=["Customer ID","City","Zip Code","Latitude","Longitude"],axis=1)

col=["Avg Monthly Long Distance Charges","Avg Monthly GB Download"]

data[col]=data[col].fillna(0)

data=data.dropna()

data.info()

for i in data.select\_dtypes(include="object"):

if(len(data[i].unique())<3):

state = pd.get\_dummies(data[i],drop\_first=True,prefix=i)

data = pd.concat([data,state],axis=1)

data = data.drop([i],axis=1)

else:

state = pd.get\_dummies(data[i],prefix=i)

data = pd.concat([data,state],axis=1)

data = data.drop([i],axis=1)

data.info()

x=data.drop("Customer Status", axis =1).values

y=data.loc[:,"Customer Status"].values

print(x[0])

print(y)

from sklearn.model\_selection import train\_test\_split

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y, test\_size=0.30, random\_state=5)

from sklearn.preprocessing import StandardScaler

ss=StandardScaler()

x\_train=ss.fit\_transform(x\_train)

x\_test=ss.fit\_transform(x\_test)

from sklearn.linear\_model import LogisticRegression

reg=LogisticRegression()

reg.fit(x\_train,y\_train)

y\_pred=reg.predict(x\_test)

from sklearn.metrics import confusion\_matrix,accuracy\_score,classification\_report

confusion\_matrix(y\_test, y\_pred)

accuracy\_score(y\_test, y\_pred)

print(classification\_report(y\_test, y\_pred))