Decision Tree for Attrition dataset

* **Code**

import pandas as pd

import numpy as np

from sklearn import tree

from sklearn import preprocessing

dataset=pd.read\_csv("general\_data.csv")

**###selecting columns**

**"""**

'Age', 'DistanceFromHome', 'Gender','JobLevel', 'MaritalStatus', 'MonthlyIncome',

'NumCompaniesWorked', 'Over18', 'PercentSalaryHike', 'StandardHours', 'TotalWorkingYears', 'TrainingTimesLastYear',

'YearsAtCompany', 'YearsSinceLastPromotion', 'YearsWithCurrManager'

**"""**

**#droping Nul rows**

dataset2=dataset.dropna()

**# labelencoder 'Gender','MaritalStatus','Over18',’Attrition’**

#1)

label\_encoder=preprocessing.LabelEncoder()

new\_value = label\_encoder.fit\_transform(dataset2["Gender"])

dataset2["Gender"]=new\_value

#2)

label\_encoder=preprocessing.LabelEncoder()

new\_value= label\_encoder.fit\_transform(dataset2["MaritalStatus"])

dataset2["MaritalStatus"]=new\_value

#3)

label\_encoder=preprocessing.LabelEncoder()

new\_value= label\_encoder.fit\_transform(dataset2["Over18"])

dataset2["Over18"]=new\_value

#4)

label\_encoder=preprocessing.LabelEncoder()

new\_value = label\_encoder.fit\_transform(dataset2["Attrition"])

dataset2["Attrition"]=new\_value

**# Creating model**

tree\_model=tree.DecisionTreeClassifier()

**# Adding all columns for predictions**

predictor=pd.DataFrame([dataset2["Age"], dataset2['DistanceFromHome'], dataset2['Gender'],dataset2['JobLevel'], dataset2['MaritalStatus'],dataset2['MonthlyIncome'],dataset2['NumCompaniesWorked'], dataset2['Over18'], dataset2['PercentSalaryHike'], dataset2['StandardHours'], dataset2['TotalWorkingYears'],dataset2['TrainingTimesLastYear'],dataset2['YearsAtCompany'], dataset2['YearsSinceLastPromotion'], dataset2['YearsWithCurrManager']]).T

tree\_model.fit(X=predictor,y=dataset2["Attrition"])

**# This code was working only on the console**

with open("AttritionDtree.dot","w") as f:

...: f=tree.export\_graphviz(tree\_model,feature\_names=['Age', 'DistanceFromHome', 'Gender','JobLevel', 'MaritalStatus', 'MonthlyIncome',

'NumCompaniesWorked', 'Over18', 'PercentSalaryHike', 'StandardHours', 'TotalWorkingYears', 'TrainingTimesLastYear',

'YearsAtCompany', 'YearsSinceLastPromotion', 'YearsWithCurrManager'],out\_file=f);

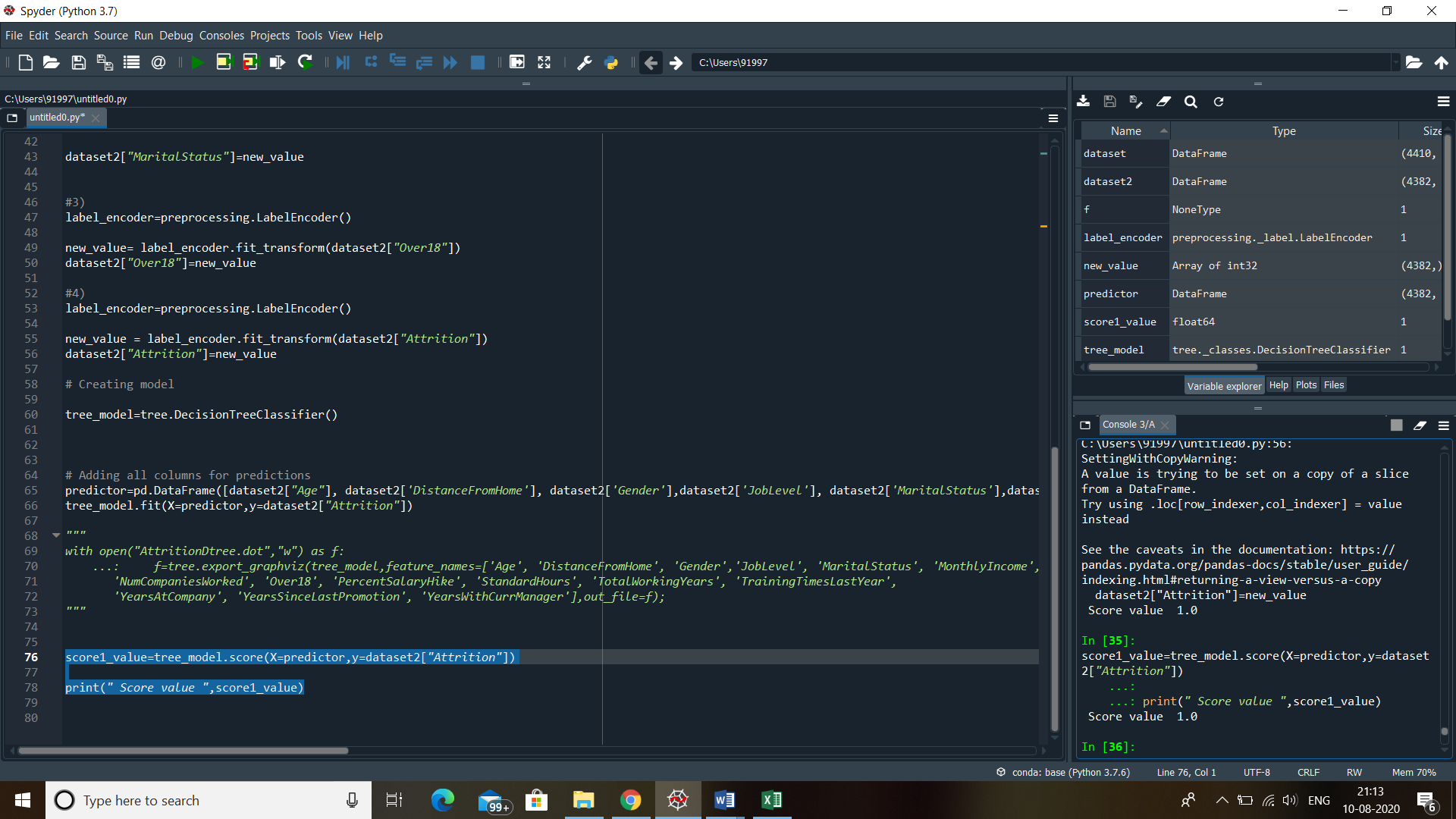
# **Calculate score**

score1\_value=tree\_model.score(X=predictor,y=dataset2["Attrition"])

print(" Score value ",score1\_value)

* **Result**

**Score value we got is = 1**



* **Decision tree info**

**It is given in the file “AttritionDtree”**