**Titanic for Decisiontree**

**Import packages:**

import pandas as pd

import numpy as np

from sklearn import tree

from sklearn import preprocessing

**Import data:**

data=pd.read\_csv("train.csv")

**Describe:**

data.describe

data.dtypes

#convert catagorical value as numarical

label\_encoder=preprocessing.LabelEncoder()

data["Sex"]=label\_encoder.fit\_transform(data["Sex"])

data["Fare"]=np.round(data["Fare"])

data["Age"]=np.round(data["Age"])

**Fit model:**

fea=["Age","Fare","Sex"]

finmodel=tree.DecisionTreeClassifier(max\_depth=8,max\_leaf\_nodes=10)

finmodel.fit(X=data[fea],y=data["Survived"])

finmodel.score(X=data[fea],y=data["Survived"])

Out[1]: 0.8008998875140607

**Predict:**

testdata=pd.read\_csv("test.csv")

testdata["Sex"]=label\_encoder.fit\_transform(testdata["Sex"])

testdata["Fare"]=np.round(testdata["Fare"])

testdata["Age"]=np.round(testdata["Age"])

prefe=["Age","Fare","Sex"]

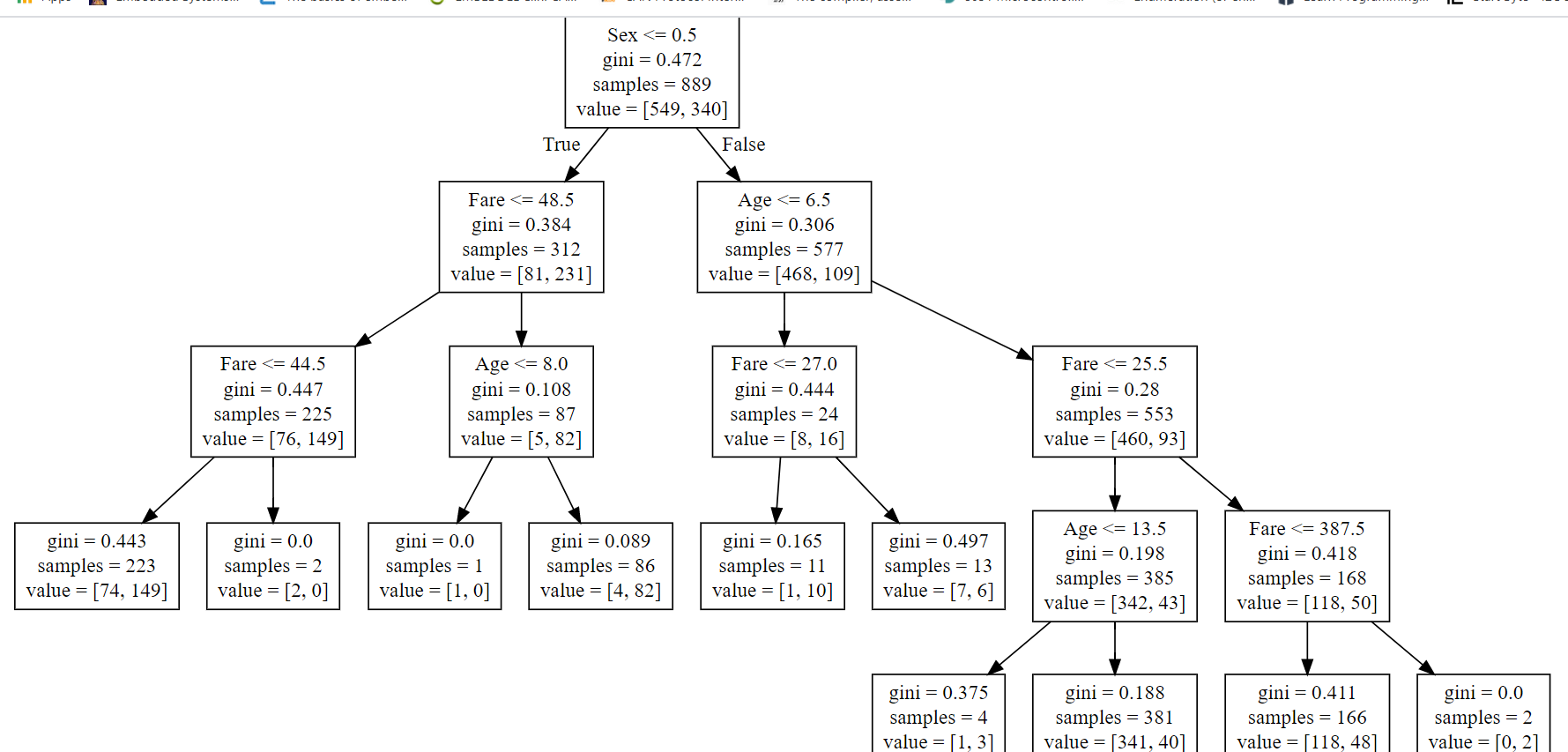
testpre=finmodel.predict(X=testdata[prefe])

fdataf=pd.DataFrame({"Age":testdata["Age"],"Survived":testpre})

**Pictorial format:**

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

f=tree.export\_graphviz(finmodel,feature\_names=prefe,out\_file=f);



**Rules:**

Here if peoples whon have sex<=0.5=False&age<=6.5=False&Fare<=25.5=False&Fare<=387.5.they would get survived 1.

Here if peoples whon have sex<=0.5=True&Fare<=48.5=False&Age<=8.0=True.they would get survived 0.