```python
import pandas as pd

df = pd.read_csv("https://raw.githubusercontent.com/AmenaNajeeb/Data/master/data.csv")

df.head()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embark
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	

```python
df = df.drop(["PassengerId","Name","Ticket","Cabin"],axis=1)

df.isnull().sum()
```

```
Survived      0
Pclass        0
Sex           0
Age         177
SibSp         0
Parch         0
Fare          0
Embarked      2
dtype: int64
```

```python
df = df.dropna(how="any")

df.isnull().sum()
```

```
Survived    0
Pclass      0
Sex         0
Age         0
SibSp       0
Parch       0
Fare        0
Embarked    0
dtype: int64
```

```python
df.dtypes
```

```
Survived      int64
Pclass        int64
Sex          object
Age         float64
SibSp         int64
Parch         int64
Fare        float64
Embarked     object
dtype: object
```

```python
from sklearn import preprocessing

le = preprocessing.LabelEncoder()

df["Sex"]=le.fit_transform(df["Sex"])

df["Embarked"]=le.fit_transform(df["Embarked"])

df.dtypes
```

```
Survived      int64
Pclass        int64
Sex           int64
Age         float64
SibSp         int64
Parch         int64
Fare        float64
Embarked      int64
dtype: object
```

```python
y=df["Survived"]
x=df.drop(["Survived"],axis=1)
```

```python
from sklearn.model_selection import train_test_split
```

```python
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.15,random_state=5)
```

```python
from sklearn.tree import DecisionTreeClassifier
```

```python
model = DecisionTreeClassifier(max_leaf_nodes=10) #max_leaf_nodes = 10
```

```python
model.fit(x_train,y_train) #trained
```

```
▼          DecisionTreeClassifier
DecisionTreeClassifier(max_leaf_nodes=10)
```

```python
y_pred=model.predict(x_test)
```

```python
model.score(x_train,y_train)
```

```
0.8330578512396695
```

```python
model.score(x_test,y_test)
```

```
0.8598130841121495
```

```python
from sklearn.metrics import confusion_matrix
```

```python
confusion_matrix(y_test,y_pred)
```

```
array([[64,  4],
       [11, 28]])
```

```python
from sklearn.metrics import classification_report
```

```python
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

           0       0.85      0.94      0.90        68
           1       0.88      0.72      0.79        39

    accuracy                           0.86       107
   macro avg       0.86      0.83      0.84       107
weighted avg       0.86      0.86      0.86       107
```

```python
from sklearn import tree
```
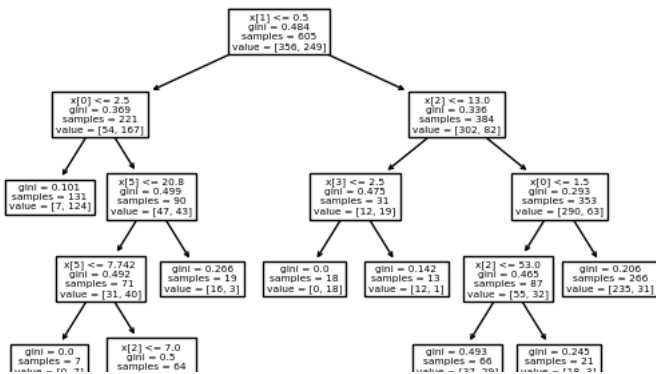
```python
tree.plot_tree(model)
```

```
[Text(0.4230769230769231, 0.9166666666666666, 'x[1] <= 0.5\ngini = 0.484\nsamples = 605\nvalue = [356,
249]'),
 Text(0.15384615384615385, 0.75, 'x[0] <= 2.5\ngini = 0.369\nsamples = 221\nvalue = [54, 167]'),
 Text(0.07692307692307693, 0.5833333333333334, 'gini = 0.101\nsamples = 131\nvalue = [7, 124]'),
 Text(0.23076923076923078, 0.5833333333333334, 'x[5] <= 20.8\ngini = 0.499\nsamples = 90\nvalue = [47,
43]'),
 Text(0.15384615384615385, 0.4166666666666667, 'x[5] <= 7.742\ngini = 0.492\nsamples = 71\nvalue =
[31, 40]'),
 Text(0.07692307692307693, 0.25, 'gini = 0.0\nsamples = 7\nvalue = [0, 7]'),
 Text(0.23076923076923078, 0.25, 'x[2] <= 7.0\ngini = 0.5\nsamples = 64\nvalue = [31, 33]'),
 Text(0.15384615384615385, 0.08333333333333333, 'gini = 0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.3076923076923077, 0.08333333333333333, 'gini = 0.497\nsamples = 56\nvalue = [30, 26]'),
 Text(0.3076923076923077, 0.4166666666666667, 'gini = 0.266\nsamples = 19\nvalue = [16, 3]'),
 Text(0.6923076923076923, 0.75, 'x[2] <= 13.0\ngini = 0.336\nsamples = 384\nvalue = [302, 82]'),
 Text(0.5384615384615384, 0.5833333333333334, 'x[3] <= 2.5\ngini = 0.475\nsamples = 31\nvalue = [12,
19]'),
 Text(0.46153846153846156, 0.4166666666666667, 'gini = 0.0\nsamples = 18\nvalue = [0, 18]'),
 Text(0.6153846153846154, 0.4166666666666667, 'gini = 0.142\nsamples = 13\nvalue = [12, 1]'),
 Text(0.8461538461538461, 0.5833333333333334, 'x[0] <= 1.5\ngini = 0.293\nsamples = 353\nvalue = [290,
32]'),
 Text(0.6923076923076923, 0.25, 'gini = 0.493\nsamples = 66\nvalue = [37, 29]'),
 Text(0.8461538461538461, 0.25, 'gini = 0.245\nsamples = 21\nvalue = [18, 3]'),
 Text(0.9230769230769231, 0.4166666666666667, 'gini = 0.206\nsamples = 266\nvalue = [235, 31]')]
```