brhama-devuni-manish-assignment-4

September 27, 2023

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```
data = pd.read_csv('WA_Fn-UseC_-HR-Employee-Attrition.csv')
data.head()
```

[2]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	EmployeeNumber	…	RelationshipSatisfaction	StandardHours	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion	YearsWithCurrManager
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences	1	1	…	1	80	0	8	0	1	6			
1	49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences	1	2	…	4	80	1	10	3	3	10			
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other	1	4	…	2	80	0	7	3	3	0			
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences	1	5	…	3	80	0	8	3	3	8			
4	27	No	Travel_Rarely	591	Research & Development	2	1	Medical	1	7	…	4	80	1	6	3	3	2			

```
0                  4                        0                      5
1                  7                        1                      7
2                  0                        0                      0
3                  7                        3                      0
4                  2                        2                      2

[5 rows x 35 columns]
```

[3]:
```
data.tail()
```

[3]:
```
      Age Attrition     BusinessTravel  DailyRate              Department  \
1465   36        No  Travel_Frequently        884  Research & Development
1466   39        No      Travel_Rarely        613  Research & Development
1467   27        No      Travel_Rarely        155  Research & Development
1468   49        No  Travel_Frequently       1023                   Sales
1469   34        No      Travel_Rarely        628  Research & Development

      DistanceFromHome  Education EducationField  EmployeeCount  \
1465                23          2        Medical              1
1466                 6          1        Medical              1
1467                 4          3  Life Sciences              1
1468                 2          3        Medical              1
1469                 8          3        Medical              1

      EmployeeNumber  …  RelationshipSatisfaction StandardHours  \
1465            2061  …                         3            80
1466            2062  …                         1            80
1467            2064  …                         2            80
1468            2065  …                         4            80
1469            2068  …                         1            80

      StockOptionLevel  TotalWorkingYears  TrainingTimesLastYear  \
1465                 1                 17                      3
1466                 1                  9                      5
1467                 1                  6                      0
1468                 0                 17                      3
1469                 0                  6                      3

     WorkLifeBalance  YearsAtCompany YearsInCurrentRole  \
1465               3               5                  2
1466               3               7                  7
1467               3               6                  2
1468               2               9                  6
1469               4               4                  3

      YearsSinceLastPromotion  YearsWithCurrManager
1465                        0                     3
```

```
1466                        1                       7
1467                        0                       3
1468                        0                       8
1469                        1                       2

[5 rows x 35 columns]
```

[4]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1470 entries, 0 to 1469
Data columns (total 35 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Age                       1470 non-null   int64
 1   Attrition                 1470 non-null   object
 2   BusinessTravel            1470 non-null   object
 3   DailyRate                 1470 non-null   int64
 4   Department                1470 non-null   object
 5   DistanceFromHome          1470 non-null   int64
 6   Education                 1470 non-null   int64
 7   EducationField            1470 non-null   object
 8   EmployeeCount             1470 non-null   int64
 9   EmployeeNumber            1470 non-null   int64
 10  EnvironmentSatisfaction   1470 non-null   int64
 11  Gender                    1470 non-null   object
 12  HourlyRate                1470 non-null   int64
 13  JobInvolvement            1470 non-null   int64
 14  JobLevel                  1470 non-null   int64
 15  JobRole                   1470 non-null   object
 16  JobSatisfaction           1470 non-null   int64
 17  MaritalStatus             1470 non-null   object
 18  MonthlyIncome             1470 non-null   int64
 19  MonthlyRate               1470 non-null   int64
 20  NumCompaniesWorked        1470 non-null   int64
 21  Over18                    1470 non-null   object
 22  OverTime                  1470 non-null   object
 23  PercentSalaryHike         1470 non-null   int64
 24  PerformanceRating         1470 non-null   int64
 25  RelationshipSatisfaction  1470 non-null   int64
 26  StandardHours             1470 non-null   int64
 27  StockOptionLevel          1470 non-null   int64
 28  TotalWorkingYears         1470 non-null   int64
 29  TrainingTimesLastYear     1470 non-null   int64
 30  WorkLifeBalance           1470 non-null   int64
 31  YearsAtCompany            1470 non-null   int64
 32  YearsInCurrentRole        1470 non-null   int64
```

```
 33  YearsSinceLastPromotion   1470 non-null   int64
 34  YearsWithCurrManager      1470 non-null   int64
dtypes: int64(26), object(9)
memory usage: 402.1+ KB
```

[5]:
```
data.describe()
```

[5]:

	Age	DailyRate	DistanceFromHome	Education	EmployeeCount	\
count	1470.000000	1470.000000	1470.000000	1470.000000	1470.0	
mean	36.923810	802.485714	9.192517	2.912925	1.0	
std	9.135373	403.509100	8.106864	1.024165	0.0	
min	18.000000	102.000000	1.000000	1.000000	1.0	
25%	30.000000	465.000000	2.000000	2.000000	1.0	
50%	36.000000	802.000000	7.000000	3.000000	1.0	
75%	43.000000	1157.000000	14.000000	4.000000	1.0	
max	60.000000	1499.000000	29.000000	5.000000	1.0	

	EmployeeNumber	EnvironmentSatisfaction	HourlyRate	JobInvolvement	\
count	1470.000000	1470.000000	1470.000000	1470.000000	
mean	1024.865306	2.721769	65.891156	2.729932	
std	602.024335	1.093082	20.329428	0.711561	
min	1.000000	1.000000	30.000000	1.000000	
25%	491.250000	2.000000	48.000000	2.000000	
50%	1020.500000	3.000000	66.000000	3.000000	
75%	1555.750000	4.000000	83.750000	3.000000	
max	2068.000000	4.000000	100.000000	4.000000	

	JobLevel	…	RelationshipSatisfaction	StandardHours	\
count	1470.000000	…	1470.000000	1470.0	
mean	2.063946	…	2.712245	80.0	
std	1.106940	…	1.081209	0.0	
min	1.000000	…	1.000000	80.0	
25%	1.000000	…	2.000000	80.0	
50%	2.000000	…	3.000000	80.0	
75%	3.000000	…	4.000000	80.0	
max	5.000000	…	4.000000	80.0	

	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	\
count	1470.000000	1470.000000	1470.000000	
mean	0.793878	11.279592	2.799320	
std	0.852077	7.780782	1.289271	
min	0.000000	0.000000	0.000000	
25%	0.000000	6.000000	2.000000	
50%	1.000000	10.000000	3.000000	
75%	1.000000	15.000000	3.000000	
max	3.000000	40.000000	6.000000	

```
       WorkLifeBalance  YearsAtCompany  YearsInCurrentRole  \
count      1470.000000     1470.000000         1470.000000
mean          2.761224        7.008163            4.229252
std           0.706476        6.126525            3.623137
min           1.000000        0.000000            0.000000
25%           2.000000        3.000000            2.000000
50%           3.000000        5.000000            3.000000
75%           3.000000        9.000000            7.000000
max           4.000000       40.000000           18.000000

       YearsSinceLastPromotion  YearsWithCurrManager
count              1470.000000           1470.000000
mean                  2.187755              4.123129
std                   3.222430              3.568136
min                   0.000000              0.000000
25%                   0.000000              2.000000
50%                   1.000000              3.000000
75%                   3.000000              7.000000
max                  15.000000             17.000000

[8 rows x 26 columns]
```

[6]:
```
data.isnull().sum()
```

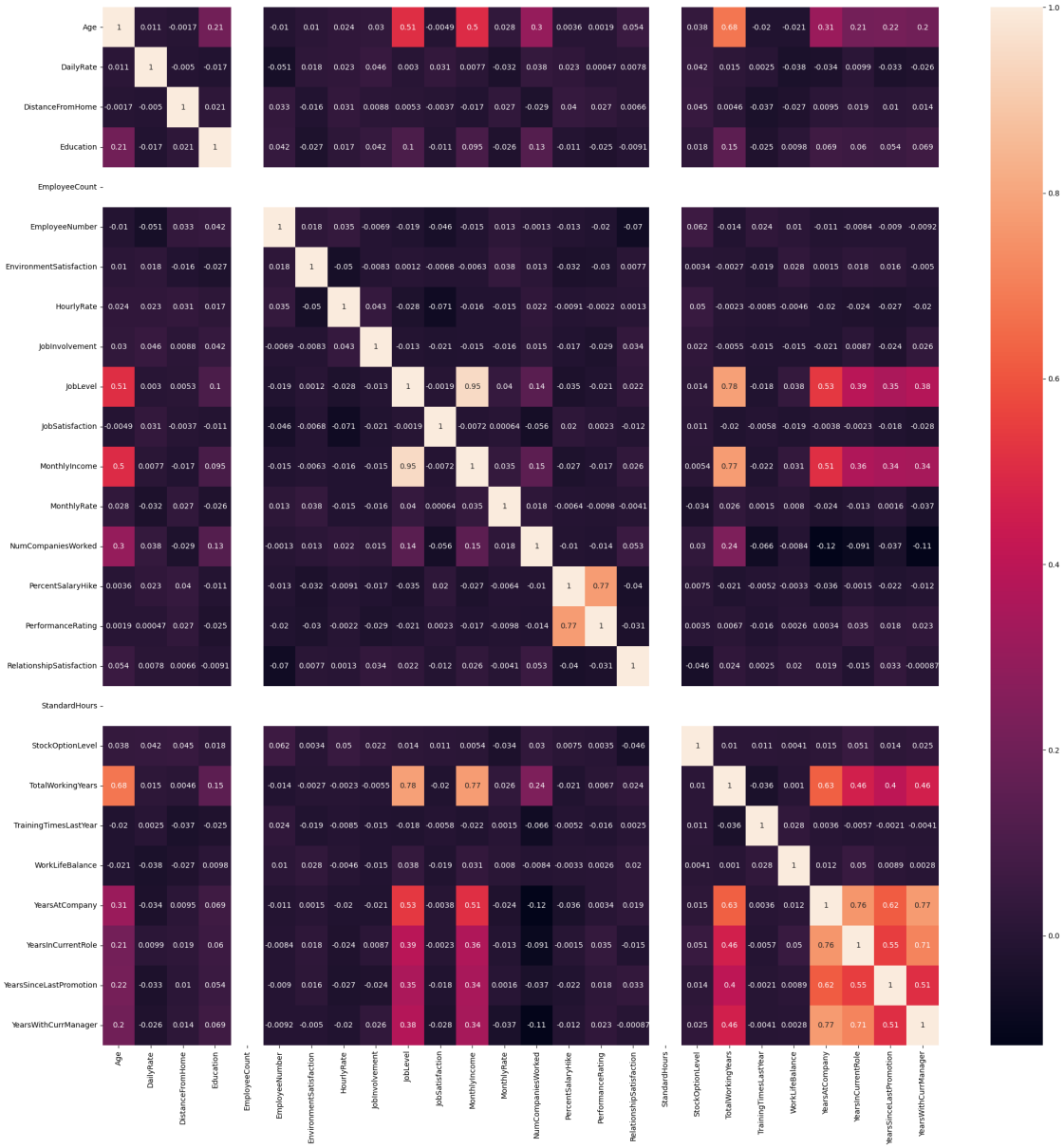
[6]:
```
Age                         0
Attrition                   0
BusinessTravel              0
DailyRate                   0
Department                  0
DistanceFromHome            0
Education                   0
EducationField              0
EmployeeCount               0
EmployeeNumber              0
EnvironmentSatisfaction     0
Gender                      0
HourlyRate                  0
JobInvolvement              0
JobLevel                    0
JobRole                     0
JobSatisfaction             0
MaritalStatus               0
MonthlyIncome               0
MonthlyRate                 0
NumCompaniesWorked          0
Over18                      0
OverTime                    0
```

```
PercentSalaryHike           0
PerformanceRating           0
RelationshipSatisfaction    0
StandardHours               0
StockOptionLevel            0
TotalWorkingYears           0
TrainingTimesLastYear       0
WorkLifeBalance             0
YearsAtCompany              0
YearsInCurrentRole          0
YearsSinceLastPromotion     0
YearsWithCurrManager        0
dtype: int64
```

[7]:
```
cor = data.corr()
```

[8]:
```
fig, ax = plt.subplots(figsize=(25,25))
sns.heatmap(cor, annot=True)
```

[8]:
```
<AxesSubplot:>
```



[ ]:
```
sns.pairplot(data)
```

[ ]: <seaborn.axisgrid.PairGrid at 0x211ab9e6910>

[9]:
```
from sklearn.preprocessing import LabelEncoder
```

[10]:
```
le=LabelEncoder()
```

[11]:
```
data["BusinessTravel"]=le.fit_transform(data["BusinessTravel"])
```

[12]:
```
data["Department"]=le.fit_transform(data["Department"])
```

[13]:
```
data["EducationField"]=le.fit_transform(data["EducationField"])
```

[14]:
```
data["Gender"]=le.fit_transform(data["Gender"])
```

[15]:
```
data["JobRole"]=le.fit_transform(data["JobRole"])
```

[16]:
```
data["MaritalStatus"]=le.fit_transform(data["MaritalStatus"])
```

[17]:
```
data["Over18"]=le.fit_transform(data["Over18"])
```

[18]:
```
data["OverTime"]=le.fit_transform(data["OverTime"])
```

[19]:
```
data.head()
```

[19]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	\
0	41	Yes	2	1102	2	1	
1	49	No	1	279	1	8	
2	37	Yes	2	1373	1	2	
3	33	No	1	1392	1	3	
4	27	No	2	591	1	2	

	Education	EducationField	EmployeeCount	EmployeeNumber	…	\
0	2	1	1	1	…	
1	1	1	1	2	…	
2	2	4	1	4	…	
3	4	1	1	5	…	
4	1	3	1	7	…	

	RelationshipSatisfaction	StandardHours	StockOptionLevel	\
0	1	80	0	
1	4	80	1	
2	2	80	0	
3	3	80	0	
4	4	80	1	

	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	YearsAtCompany	\
0	8	0	1	6	
1	10	3	3	10	
2	7	3	3	0	
3	8	3	3	8	
4	6	3	3	2	

	YearsInCurrentRole	YearsSinceLastPromotion	YearsWithCurrManager
0	4	0	5
1	7	1	7

```
2                    0                          0                          0
3                    7                          3                          0
4                    2                          2                          2

[5 rows x 35 columns]
```

[20]:
```
data.tail()
```

[20]:
```
      Age Attrition  BusinessTravel  DailyRate  Department  DistanceFromHome  \
1465   36        No               1        884           1                23
1466   39        No               2        613           1                 6
1467   27        No               2        155           1                 4
1468   49        No               1       1023           2                 2
1469   34        No               2        628           1                 8

      Education  EducationField  EmployeeCount  EmployeeNumber  …  \
1465          2               3              1            2061  …
1466          1               3              1            2062  …
1467          3               1              1            2064  …
1468          3               3              1            2065  …
1469          3               3              1            2068  …

      RelationshipSatisfaction  StandardHours  StockOptionLevel  \
1465                         3             80                 1
1466                         1             80                 1
1467                         2             80                 1
1468                         4             80                 0
1469                         1             80                 0

      TotalWorkingYears  TrainingTimesLastYear  WorkLifeBalance  \
1465                 17                      3                3
1466                  9                      5                3
1467                  6                      0                3
1468                 17                      3                2
1469                  6                      3                4

      YearsAtCompany  YearsInCurrentRole  YearsSinceLastPromotion  \
1465               5                   2                        0
1466               7                   7                        1
1467               6                   2                        0
1468               9                   6                        0
1469               4                   3                        1

      YearsWithCurrManager
1465                     3
1466                     7
1467                     3
```

```
1468                       8
1469                       2

[5 rows x 35 columns]
```

[21]:
```
X=data.
  ↪drop(columns=["EmployeeNumber","EmployeeCount","StandardHours","Attrition","Over18"],axis=1
```

[22]:
```
y=data["Attrition"]
```

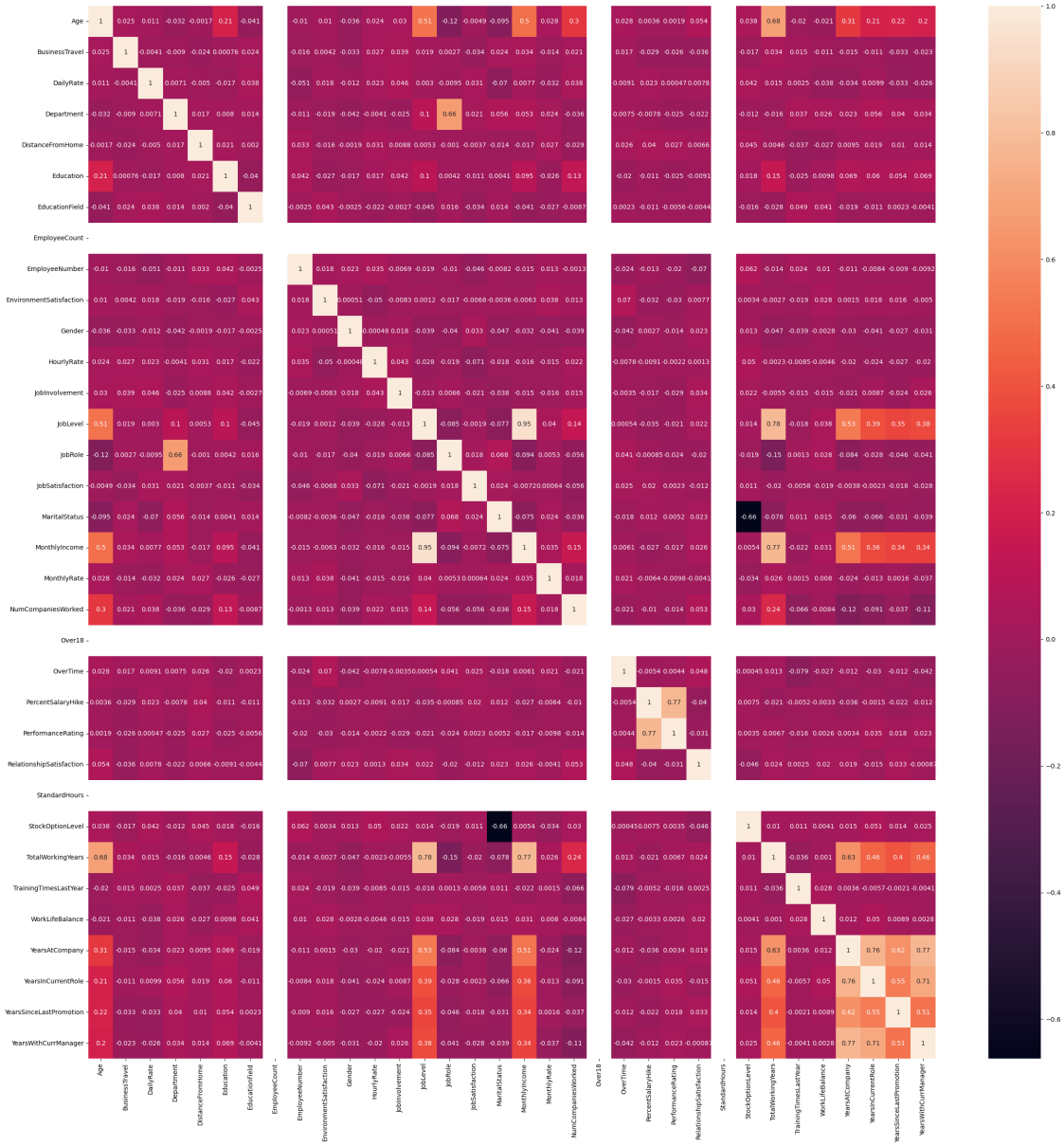
[23]:
```
from sklearn.preprocessing import MinMaxScaler
ms=MinMaxScaler()
```

[24]:
```
X_Scaled=ms.fit_transform(X)
```

[25]:
```
cor=data.corr()
```

[26]:
```
fig, ax = plt.subplots(figsize=(30,30))
sns.heatmap(cor, annot=True)
```

[26]:
```
<AxesSubplot:>
```



[27]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(X_Scaled,y,test_size =0.
↪2,random_state =0)
```

[28]:
```
from sklearn.linear_model import LogisticRegression
classifier = LogisticRegression(random_state=0)
classifier.fit(x_train,y_train)
```

[28]: LogisticRegression(random_state=0)

[29]:
```
from sklearn.metrics import accuracy_score,confusion_matrix
y_pred = classifier.predict(x_test)
cm = confusion_matrix(y_test, y_pred)
print(cm)
accuracy_score(y_test, y_pred)*100
```

```
[[242   3]
 [ 32  17]]
```

[29]: 88.09523809523809

[30]:
```
from sklearn.metrics import␣
↪accuracy_score,confusion_matrix,classification_report,roc_auc_score,roc_curve
```

[31]:
```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

          No       0.88      0.99      0.93       245
         Yes       0.85      0.35      0.49        49

    accuracy                           0.88       294
   macro avg       0.87      0.67      0.71       294
weighted avg       0.88      0.88      0.86       294
```

[32]:
```
from sklearn.tree import DecisionTreeClassifier
dtc=DecisionTreeClassifier()
```

[33]:
```
dtc.fit(x_train,y_train)
```

[33]: DecisionTreeClassifier()

[34]:
```
from sklearn.metrics import accuracy_score,confusion_matrix
y_pred = dtc.predict(x_test)
cm = confusion_matrix(y_test, y_pred)
print(cm)
accuracy_score(y_test, y_pred)*100
```

```
[[206  39]
 [ 35  14]]
```

[34]: 74.82993197278913

[35]:
```
from sklearn import tree
plt.figure(figsize=(25,15))
tree.plot_tree(dtc,filled=True)
```

```
[35]: [Text(0.3274182771381579, 0.9722222222222222, 'X[23] <= 0.038\ngini =
      0.269\nsamples = 1176\nvalue = [988, 188]'),
       Text(0.07894736842105263, 0.9166666666666666, 'X[14] <= 0.75\ngini =
      0.5\nsamples = 78\nvalue = [39, 39]'),
       Text(0.049342105263157895, 0.8611111111111112, 'X[4] <= 0.554\ngini =
      0.426\nsamples = 39\nvalue = [27, 12]'),
       Text(0.03289473684210526, 0.8055555555555556, 'X[13] <= 0.167\ngini =
      0.312\nsamples = 31\nvalue = [25, 6]'),
       Text(0.019736842105263157, 0.75, 'X[14] <= 0.25\ngini = 0.49\nsamples =
      7\nvalue = [3, 4]'),
       Text(0.013157894736842105, 0.6944444444444444, 'X[7] <= 0.333\ngini =
      0.375\nsamples = 4\nvalue = [3, 1]'),
       Text(0.006578947368421052, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue
      = [0, 1]'),
       Text(0.019736842105263157, 0.6388888888888888, 'gini = 0.0\nsamples = 3\nvalue
      = [3, 0]'),
       Text(0.02631578947368421, 0.6944444444444444, 'gini = 0.0\nsamples = 3\nvalue =
      [0, 3]'),
       Text(0.046052631578947366, 0.75, 'X[17] <= 0.056\ngini = 0.153\nsamples =
      24\nvalue = [22, 2]'),
       Text(0.039473684210526314, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue
      = [0, 1]'),
       Text(0.05263157894736842, 0.6944444444444444, 'X[7] <= 0.167\ngini =
      0.083\nsamples = 23\nvalue = [22, 1]'),
       Text(0.046052631578947366, 0.6388888888888888, 'X[22] <= 0.333\ngini =
      0.5\nsamples = 2\nvalue = [1, 1]'),
       Text(0.039473684210526314, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue
      = [1, 0]'),
       Text(0.05263157894736842, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
      [0, 1]'),
       Text(0.05921052631578947, 0.6388888888888888, 'gini = 0.0\nsamples = 21\nvalue
      = [21, 0]'),
       Text(0.06578947368421052, 0.8055555555555556, 'X[19] <= 0.679\ngini =
      0.375\nsamples = 8\nvalue = [2, 6]'),
       Text(0.05921052631578947, 0.75, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
       Text(0.07236842105263158, 0.75, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
       Text(0.10855263157894737, 0.8611111111111112, 'X[9] <= 0.364\ngini =
      0.426\nsamples = 39\nvalue = [12, 27]'),
       Text(0.09210526315789473, 0.8055555555555556, 'X[0] <= 0.369\ngini =
      0.133\nsamples = 14\nvalue = [1, 13]'),
       Text(0.08552631578947369, 0.75, 'gini = 0.0\nsamples = 13\nvalue = [0, 13]'),
       Text(0.09868421052631579, 0.75, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
       Text(0.125, 0.8055555555555556, 'X[18] <= 0.5\ngini = 0.493\nsamples =
      25\nvalue = [11, 14]'),
       Text(0.1118421052631579, 0.75, 'X[6] <= 0.7\ngini = 0.484\nsamples = 17\nvalue
      = [10, 7]'),
       Text(0.10526315789473684, 0.6944444444444444, 'X[2] <= 0.106\ngini =
```

```
0.408\nsamples = 14\nvalue = [10, 4]'),
 Text(0.09868421052631579, 0.6388888888888888, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.1118421052631579, 0.6388888888888888, 'X[25] <= 0.5\ngini =
0.278\nsamples = 12\nvalue = [10, 2]'),
 Text(0.10526315789473684, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11842105263157894, 0.5833333333333334, 'X[2] <= 0.8\ngini =
0.165\nsamples = 11\nvalue = [10, 1]'),
 Text(0.1118421052631579, 0.5277777777777778, 'gini = 0.0\nsamples = 10\nvalue =
[10, 0]'),
 Text(0.125, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.11842105263157894, 0.6944444444444444, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.13815789473684212, 0.75, 'X[29] <= 0.029\ngini = 0.219\nsamples =
8\nvalue = [1, 7]'),
 Text(0.13157894736842105, 0.6944444444444444, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.14473684210526316, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5758891858552632, 0.9166666666666666, 'X[18] <= 0.5\ngini =
0.235\nsamples = 1098\nvalue = [949, 149]'),
 Text(0.3252981085526316, 0.8611111111111112, 'X[25] <= 0.167\ngini =
0.162\nsamples = 798\nvalue = [727, 71]'),
 Text(0.1875, 0.8055555555555556, 'X[2] <= 0.747\ngini = 0.38\nsamples =
47\nvalue = [35, 12]'),
 Text(0.18092105263157895, 0.75, 'X[10] <= 0.5\ngini = 0.463\nsamples =
33\nvalue = [21, 12]'),
 Text(0.15789473684210525, 0.6944444444444444, 'X[4] <= 0.446\ngini =
0.42\nsamples = 10\nvalue = [3, 7]'),
 Text(0.1513157894736842, 0.6388888888888888, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.16447368421052633, 0.6388888888888888, 'X[4] <= 0.714\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.15789473684210525, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.17105263157894737, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.20394736842105263, 0.6944444444444444, 'X[24] <= 0.583\ngini =
0.34\nsamples = 23\nvalue = [18, 5]'),
 Text(0.19078947368421054, 0.6388888888888888, 'X[9] <= 0.107\ngini =
0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.18421052631578946, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.19736842105263158, 0.5833333333333334, 'gini = 0.0\nsamples = 15\nvalue
= [15, 0]'),
 Text(0.21710526315789475, 0.6388888888888888, 'X[26] <= 0.287\ngini =
```

```
0.49\nsamples = 7\nvalue = [3, 4]'),
 Text(0.21052631578947367, 0.5833333333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.2236842105263158, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.19407894736842105, 0.75, 'gini = 0.0\nsamples = 14\nvalue = [14, 0]'),
 Text(0.46309621710526316, 0.8055555555555556, 'X[23] <= 0.975\ngini =
0.145\nsamples = 751\nvalue = [692, 59]'),
 Text(0.4565172697368421, 0.75, 'X[26] <= 0.113\ngini = 0.143\nsamples =
750\nvalue = [692, 58]'),
 Text(0.3223684210526316, 0.6944444444444444, 'X[7] <= 0.167\ngini =
0.218\nsamples = 257\nvalue = [225, 32]'),
 Text(0.26480263157894735, 0.6388888888888888, 'X[29] <= 0.147\ngini =
0.355\nsamples = 65\nvalue = [50, 15]'),
 Text(0.23684210526315788, 0.5833333333333334, 'X[29] <= 0.029\ngini =
0.303\nsamples = 59\nvalue = [48, 11]'),
 Text(0.2138157894736842, 0.5277777777777778, 'X[10] <= 0.5\ngini =
0.463\nsamples = 22\nvalue = [14, 8]'),
 Text(0.20065789473684212, 0.4722222222222222, 'X[9] <= 0.179\ngini =
0.198\nsamples = 9\nvalue = [8, 1]'),
 Text(0.19407894736842105, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.20723684210526316, 0.4166666666666667, 'gini = 0.0\nsamples = 8\nvalue =
[8, 0]'),
 Text(0.22697368421052633, 0.4722222222222222, 'X[9] <= 0.4\ngini =
0.497\nsamples = 13\nvalue = [6, 7]'),
 Text(0.22039473684210525, 0.4166666666666667, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.23355263157894737, 0.4166666666666667, 'X[4] <= 0.286\ngini =
0.346\nsamples = 9\nvalue = [2, 7]'),
 Text(0.22697368421052633, 0.3611111111111111, 'X[2] <= 0.369\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.22039473684210525, 0.3055555555555556, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.23355263157894737, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.24013157894736842, 0.3611111111111111, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.2598684210526316, 0.5277777777777778, 'X[13] <= 0.167\ngini =
0.149\nsamples = 37\nvalue = [34, 3]'),
 Text(0.2532894736842105, 0.4722222222222222, 'X[25] <= 0.5\ngini = 0.5\nsamples
= 6\nvalue = [3, 3]'),
 Text(0.24671052631578946, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.2598684210526316, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.26644736842105265, 0.4722222222222222, 'gini = 0.0\nsamples = 31\nvalue
```

```
= [31, 0]'),
 Text(0.29276315789473684, 0.5833333333333334, 'X[25] <= 0.5\ngini =
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.28618421052631576, 0.5277777777777778, 'X[1] <= 0.75\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.27960526315789475, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.29276315789473684, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.2993421052631579, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.37993421052631576, 0.6388888888888888, 'X[0] <= 0.321\ngini =
0.161\nsamples = 192\nvalue = [175, 17]'),
 Text(0.3223684210526316, 0.5833333333333334, 'X[6] <= 0.1\ngini =
0.294\nsamples = 67\nvalue = [55, 12]'),
 Text(0.3157894736842105, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.32894736842105265, 0.5277777777777778, 'X[25] <= 0.5\ngini =
0.26\nsamples = 65\nvalue = [55, 10]'),
 Text(0.3059210526315789, 0.4722222222222222, 'X[6] <= 0.5\ngini =
0.469\nsamples = 16\nvalue = [10, 6]'),
 Text(0.2993421052631579, 0.4166666666666667, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.3125, 0.4166666666666667, 'X[7] <= 0.833\ngini = 0.444\nsamples =
9\nvalue = [3, 6]'),
 Text(0.3059210526315789, 0.3611111111111111, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.3190789473684211, 0.3611111111111111, 'X[27] <= 0.139\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.3125, 0.3055555555555556, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.3256578947368421, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3519736842105263, 0.4722222222222222, 'X[2] <= 0.037\ngini =
0.15\nsamples = 49\nvalue = [45, 4]'),
 Text(0.34539473684210525, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.35855263157894735, 0.4166666666666667, 'X[2] <= 0.938\ngini =
0.117\nsamples = 48\nvalue = [45, 3]'),
 Text(0.3519736842105263, 0.3611111111111111, 'X[5] <= 0.875\ngini =
0.081\nsamples = 47\nvalue = [45, 2]'),
 Text(0.33881578947368424, 0.3055555555555556, 'X[10] <= 0.167\ngini =
0.043\nsamples = 45\nvalue = [44, 1]'),
 Text(0.33223684210526316, 0.25, 'X[19] <= 0.214\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.3256578947368421, 0.19444444444444445, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.33881578947368424, 0.19444444444444445, 'gini = 0.0\nsamples = 2\nvalue
```

```
= [2, 0]'),
 Text(0.34539473684210525, 0.25, 'gini = 0.0\nsamples = 42\nvalue = [42, 0]'),
 Text(0.3651315789473684, 0.3055555555555556, 'X[13] <= 0.667\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.35855263157894735, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.3717105263157895, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.3651315789473684, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4375, 0.5833333333333334, 'X[6] <= 0.9\ngini = 0.077\nsamples =
125\nvalue = [120, 5]'),
 Text(0.4243421052631579, 0.5277777777777778, 'X[0] <= 0.393\ngini =
0.05\nsamples = 118\nvalue = [115, 3]'),
 Text(0.41776315789473684, 0.4722222222222222, 'X[2] <= 0.956\ngini =
0.185\nsamples = 29\nvalue = [26, 3]'),
 Text(0.41118421052631576, 0.4166666666666667, 'X[29] <= 0.147\ngini =
0.133\nsamples = 28\nvalue = [26, 2]'),
 Text(0.3980263157894737, 0.3611111111111111, 'X[10] <= 0.167\ngini =
0.074\nsamples = 26\nvalue = [25, 1]'),
 Text(0.39144736842105265, 0.3055555555555556, 'X[2] <= 0.216\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.3848684210526316, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.3980263157894737, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.40460526315789475, 0.3055555555555556, 'gini = 0.0\nsamples = 24\nvalue
= [24, 0]'),
 Text(0.4243421052631579, 0.3611111111111111, 'X[14] <= 0.75\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.41776315789473684, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4309210526315789, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4243421052631579, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4309210526315789, 0.4722222222222222, 'gini = 0.0\nsamples = 89\nvalue =
[89, 0]'),
 Text(0.4506578947368421, 0.5277777777777778, 'X[2] <= 0.594\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.4440789473684211, 0.4722222222222222, 'X[13] <= 0.333\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4375, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.4506578947368421, 0.4166666666666667, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4572368421052631, 0.4722222222222222, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.5906661184210527, 0.6944444444444444, 'X[26] <= 0.787\ngini =
0.1\nsamples = 493\nvalue = [467, 26]'),
 Text(0.5563322368421053, 0.6388888888888888, 'X[13] <= 0.5\ngini =
0.094\nsamples = 486\nvalue = [462, 24]'),
```

```
 Text(0.5074013157894737, 0.5833333333333334, 'X[12] <= 0.938\ngini =
0.154\nsamples = 191\nvalue = [175, 16]'),
 Text(0.5008223684210527, 0.5277777777777778, 'X[16] <= 0.481\ngini =
0.145\nsamples = 190\nvalue = [175, 15]'),
 Text(0.4819078947368421, 0.4722222222222222, 'X[29] <= 0.794\ngini =
0.221\nsamples = 95\nvalue = [83, 12]'),
 Text(0.4753289473684211, 0.4166666666666667, 'X[16] <= 0.47\ngini =
0.207\nsamples = 94\nvalue = [83, 11]'),
 Text(0.46875, 0.3611111111111111, 'X[5] <= 0.375\ngini = 0.192\nsamples =
93\nvalue = [83, 10]'),
 Text(0.4440789473684211, 0.3055555555555556, 'X[6] <= 0.9\ngini =
0.363\nsamples = 21\nvalue = [16, 5]'),
 Text(0.4375, 0.25, 'X[15] <= 0.413\ngini = 0.266\nsamples = 19\nvalue = [16,
3]'),
 Text(0.4243421052631579, 0.19444444444444445, 'X[15] <= 0.141\ngini =
0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.41776315789473684, 0.1388888888888889, 'X[1] <= 0.75\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.41118421052631576, 0.08333333333333333, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.4243421052631579, 0.08333333333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4309210526315789, 0.1388888888888889, 'gini = 0.0\nsamples = 14\nvalue =
[14, 0]'),
 Text(0.4506578947368421, 0.19444444444444445, 'X[6] <= 0.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4440789473684211, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.45723684210526316, 0.1388888888888889, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4506578947368421, 0.25, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.4934210526315789, 0.3055555555555556, 'X[27] <= 0.139\ngini =
0.129\nsamples = 72\nvalue = [67, 5]'),
 Text(0.47039473684210525, 0.25, 'X[29] <= 0.206\ngini = 0.444\nsamples =
6\nvalue = [4, 2]'),
 Text(0.46381578947368424, 0.19444444444444445, 'gini = 0.0\nsamples = 3\nvalue
= [3, 0]'),
 Text(0.4769736842105263, 0.19444444444444445, 'X[9] <= 0.686\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.47039473684210525, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.48355263157894735, 0.1388888888888889, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.5164473684210527, 0.25, 'X[9] <= 0.993\ngini = 0.087\nsamples =
66\nvalue = [63, 3]'),
 Text(0.5032894736842105, 0.19444444444444445, 'X[24] <= 0.583\ngini =
0.061\nsamples = 64\nvalue = [62, 2]'),
```

```
 Text(0.4967105263157895, 0.1388888888888889, 'gini = 0.0\nsamples = 51\nvalue =
[51, 0]'),
 Text(0.5098684210526315, 0.1388888888888889, 'X[12] <= 0.812\ngini =
0.26\nsamples = 13\nvalue = [11, 2]'),
 Text(0.5032894736842105, 0.08333333333333333, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.5164473684210527, 0.08333333333333333, 'X[7] <= 0.5\ngini = 0.5\nsamples
= 4\nvalue = [2, 2]'),
 Text(0.5098684210526315, 0.027777777777777776, 'gini = 0.0\nsamples = 2\nvalue
= [0, 2]'),
 Text(0.5230263157894737, 0.027777777777777776, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.5296052631578947, 0.19444444444444445, 'X[11] <= 0.25\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.5230263157894737, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5361842105263158, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4819078947368421, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4884868421052632, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5197368421052632, 0.4722222222222222, 'X[17] <= 0.5\ngini =
0.061\nsamples = 95\nvalue = [92, 3]'),
 Text(0.5131578947368421, 0.4166666666666667, 'gini = 0.0\nsamples = 76\nvalue =
[76, 0]'),
 Text(0.5263157894736842, 0.4166666666666667, 'X[29] <= 0.088\ngini =
0.266\nsamples = 19\nvalue = [16, 3]'),
 Text(0.5131578947368421, 0.3611111111111111, 'X[3] <= 0.75\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.506578947368421, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5197368421052632, 0.3055555555555556, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.5394736842105263, 0.3611111111111111, 'X[15] <= 0.108\ngini =
0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.5328947368421053, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5460526315789473, 0.3055555555555556, 'gini = 0.0\nsamples = 15\nvalue =
[15, 0]'),
 Text(0.5139802631578947, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6052631578947368, 0.5833333333333334, 'X[19] <= 0.036\ngini =
0.053\nsamples = 295\nvalue = [287, 8]'),
 Text(0.5822368421052632, 0.5277777777777778, 'X[28] <= 0.7\ngini =
0.159\nsamples = 46\nvalue = [42, 4]'),
 Text(0.5756578947368421, 0.4722222222222222, 'X[9] <= 0.071\ngini =
```

```
0.124\nsamples = 45\nvalue = [42, 3]'),
 Text(0.5592105263157895, 0.4166666666666667, 'X[17] <= 0.056\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.5526315789473685, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5657894736842105, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5921052631578947, 0.4166666666666667, 'X[23] <= 0.688\ngini =
0.089\nsamples = 43\nvalue = [41, 2]'),
 Text(0.5789473684210527, 0.3611111111111111, 'X[12] <= 0.062\ngini =
0.048\nsamples = 41\nvalue = [40, 1]'),
 Text(0.5723684210526315, 0.3055555555555556, 'X[2] <= 0.487\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.5657894736842105, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.5789473684210527, 0.25, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.5855263157894737, 0.3055555555555556, 'gini = 0.0\nsamples = 37\nvalue =
[37, 0]'),
 Text(0.6052631578947368, 0.3611111111111111, 'X[11] <= 0.625\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.5986842105263158, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6118421052631579, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5888157894736842, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6282894736842105, 0.5277777777777778, 'X[15] <= 0.056\ngini =
0.032\nsamples = 249\nvalue = [245, 4]'),
 Text(0.6118421052631579, 0.4722222222222222, 'X[25] <= 0.5\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.6052631578947368, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.618421052631579, 0.4166666666666667, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.6447368421052632, 0.4722222222222222, 'X[2] <= 0.015\ngini =
0.024\nsamples = 244\nvalue = [241, 3]'),
 Text(0.631578947368421, 0.4166666666666667, 'X[22] <= 0.667\ngini =
0.278\nsamples = 6\nvalue = [5, 1]'),
 Text(0.625, 0.3611111111111111, 'gini = 0.0\nsamples = 5\nvalue = [5, 0]'),
 Text(0.6381578947368421, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6578947368421053, 0.4166666666666667, 'X[21] <= 0.167\ngini =
0.017\nsamples = 238\nvalue = [236, 2]'),
 Text(0.6513157894736842, 0.3611111111111111, 'X[25] <= 0.833\ngini =
0.073\nsamples = 53\nvalue = [51, 2]'),
 Text(0.6381578947368421, 0.3055555555555556, 'X[29] <= 0.088\ngini =
0.041\nsamples = 48\nvalue = [47, 1]'),
 Text(0.631578947368421, 0.25, 'X[0] <= 0.345\ngini = 0.245\nsamples = 7\nvalue
```

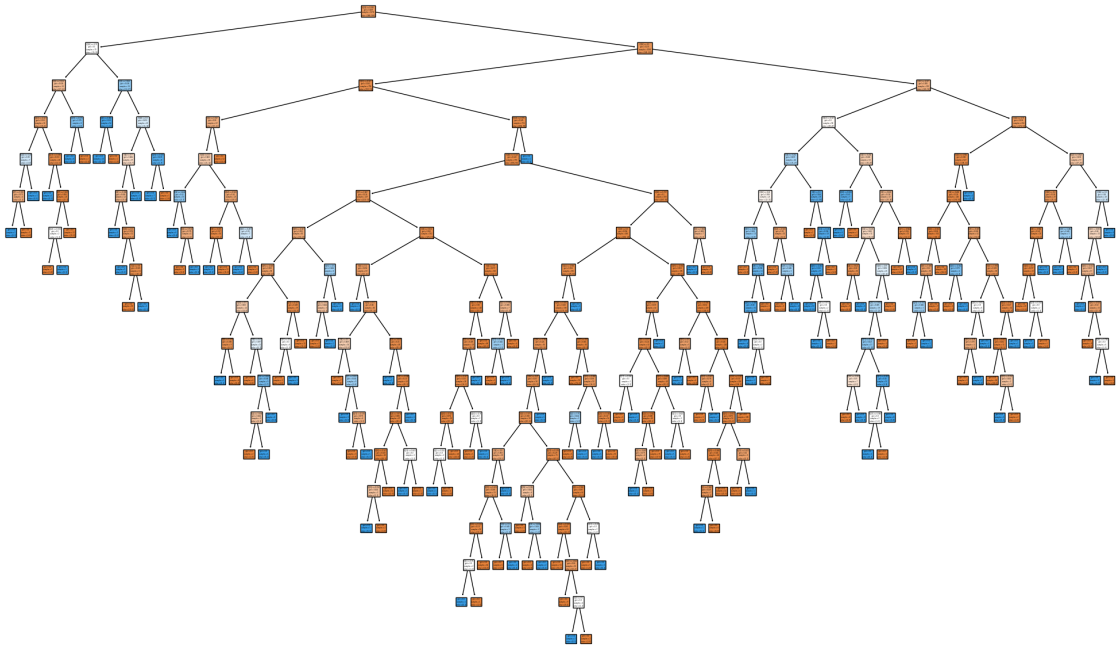
```
= [6, 1]'),
 Text(0.625, 0.19444444444444445, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.6381578947368421, 0.19444444444444445, 'gini = 0.0\nsamples = 6\nvalue =
[6, 0]'),
 Text(0.6447368421052632, 0.25, 'gini = 0.0\nsamples = 41\nvalue = [41, 0]'),
 Text(0.6644736842105263, 0.3055555555555556, 'X[27] <= 0.417\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.6578947368421053, 0.25, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.6710526315789473, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.6644736842105263, 0.3611111111111111, 'gini = 0.0\nsamples = 185\nvalue
= [185, 0]'),
 Text(0.625, 0.6388888888888888, 'X[2] <= 0.366\ngini = 0.408\nsamples =
7\nvalue = [5, 2]'),
 Text(0.618421052631579, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.631578947368421, 0.5833333333333334, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.46967516447368424, 0.75, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.8264802631578947, 0.8611111111111112, 'X[15] <= 0.157\ngini =
0.385\nsamples = 300\nvalue = [222, 78]'),
 Text(0.740953947368421, 0.8055555555555556, 'X[22] <= 0.167\ngini =
0.5\nsamples = 96\nvalue = [49, 47]'),
 Text(0.7072368421052632, 0.75, 'X[4] <= 0.161\ngini = 0.459\nsamples =
42\nvalue = [15, 27]'),
 Text(0.6842105263157895, 0.6944444444444444, 'X[16] <= 0.41\ngini =
0.499\nsamples = 23\nvalue = [12, 11]'),
 Text(0.6710526315789473, 0.6388888888888888, 'X[15] <= 0.061\ngini =
0.426\nsamples = 13\nvalue = [4, 9]'),
 Text(0.6644736842105263, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6776315789473685, 0.5833333333333334, 'X[9] <= 0.993\ngini =
0.298\nsamples = 11\nvalue = [2, 9]'),
 Text(0.6710526315789473, 0.5277777777777778, 'X[16] <= 0.308\ngini =
0.18\nsamples = 10\nvalue = [1, 9]'),
 Text(0.6644736842105263, 0.4722222222222222, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]'),
 Text(0.6776315789473685, 0.4722222222222222, 'X[12] <= 0.5\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.6710526315789473, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6842105263157895, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6842105263157895, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6973684210526315, 0.6388888888888888, 'X[24] <= 0.583\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.6907894736842105, 0.5833333333333334, 'gini = 0.0\nsamples = 7\nvalue =
```

```
[7, 0]'),
 Text(0.7039473684210527, 0.5833333333333334, 'X[19] <= 0.071\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.6973684210526315, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7105263157894737, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.7302631578947368, 0.6944444444444444, 'X[9] <= 0.2\ngini =
0.266\nsamples = 19\nvalue = [3, 16]'),
 Text(0.7236842105263158, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7368421052631579, 0.6388888888888888, 'X[11] <= 0.125\ngini =
0.198\nsamples = 18\nvalue = [2, 16]'),
 Text(0.7302631578947368, 0.5833333333333334, 'X[28] <= 0.433\ngini =
0.111\nsamples = 17\nvalue = [1, 16]'),
 Text(0.7236842105263158, 0.5277777777777778, 'gini = 0.0\nsamples = 15\nvalue =
[0, 15]'),
 Text(0.7368421052631579, 0.5277777777777778, 'X[9] <= 0.371\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.7302631578947368, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.743421052631579, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.743421052631579, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.774671052631579, 0.75, 'X[0] <= 0.202\ngini = 0.466\nsamples = 54\nvalue
= [34, 20]'),
 Text(0.756578947368421, 0.6944444444444444, 'X[10] <= 0.833\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
 Text(0.75, 0.6388888888888888, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.7631578947368421, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7927631578947368, 0.6944444444444444, 'X[2] <= 0.622\ngini =
0.418\nsamples = 47\nvalue = [33, 14]'),
 Text(0.7763157894736842, 0.6388888888888888, 'X[2] <= 0.145\ngini =
0.482\nsamples = 32\nvalue = [19, 13]'),
 Text(0.7631578947368421, 0.5833333333333334, 'X[4] <= 0.821\ngini =
0.18\nsamples = 10\nvalue = [9, 1]'),
 Text(0.756578947368421, 0.5277777777777778, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.7697368421052632, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7894736842105263, 0.5833333333333334, 'X[16] <= 0.87\ngini =
0.496\nsamples = 22\nvalue = [10, 12]'),
 Text(0.7828947368421053, 0.5277777777777778, 'X[25] <= 0.833\ngini =
0.465\nsamples = 19\nvalue = [7, 12]'),
 Text(0.7763157894736842, 0.4722222222222222, 'X[17] <= 0.167\ngini =
```

```
0.415\nsamples = 17\nvalue = [5, 12]'),
 Text(0.7631578947368421, 0.4166666666666667, 'X[19] <= 0.321\ngini =
0.49\nsamples = 7\nvalue = [4, 3]'),
 Text(0.756578947368421, 0.3611111111111111, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.7697368421052632, 0.3611111111111111, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.7894736842105263, 0.4166666666666667, 'X[21] <= 0.333\ngini =
0.18\nsamples = 10\nvalue = [1, 9]'),
 Text(0.7828947368421053, 0.3611111111111111, 'X[17] <= 0.389\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.7763157894736842, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7894736842105263, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7960526315789473, 0.3611111111111111, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]'),
 Text(0.7894736842105263, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.7960526315789473, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.8092105263157895, 0.6388888888888888, 'X[17] <= 0.944\ngini =
0.124\nsamples = 15\nvalue = [14, 1]'),
 Text(0.8026315789473685, 0.5833333333333334, 'gini = 0.0\nsamples = 14\nvalue =
[14, 0]'),
 Text(0.8157894736842105, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9120065789473685, 0.8055555555555556, 'X[14] <= 0.75\ngini =
0.258\nsamples = 204\nvalue = [173, 31]'),
 Text(0.8601973684210527, 0.75, 'X[15] <= 0.992\ngini = 0.138\nsamples =
147\nvalue = [136, 11]'),
 Text(0.8536184210526315, 0.6944444444444444, 'X[4] <= 0.482\ngini =
0.128\nsamples = 146\nvalue = [136, 10]'),
 Text(0.8355263157894737, 0.6388888888888888, 'X[26] <= 0.063\ngini =
0.038\nsamples = 104\nvalue = [102, 2]'),
 Text(0.8289473684210527, 0.5833333333333334, 'X[9] <= 0.193\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.8223684210526315, 0.5277777777777778, 'X[19] <= 0.143\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.8157894736842105, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8289473684210527, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.8355263157894737, 0.5277777777777778, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.8421052631578947, 0.5833333333333334, 'gini = 0.0\nsamples = 94\nvalue =
[94, 0]'),
```

```
 Text(0.8717105263157895, 0.6388888888888888, 'X[7] <= 0.167\ngini =
0.308\nsamples = 42\nvalue = [34, 8]'),
 Text(0.8552631578947368, 0.5833333333333334, 'X[16] <= 0.194\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.8486842105263158, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8618421052631579, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.8881578947368421, 0.5833333333333334, 'X[0] <= 0.393\ngini =
0.229\nsamples = 38\nvalue = [33, 5]'),
 Text(0.875, 0.5277777777777778, 'X[16] <= 0.871\ngini = 0.5\nsamples = 6\nvalue
= [3, 3]'),
 Text(0.868421052631579, 0.4722222222222222, 'X[15] <= 0.35\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.8618421052631579, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.875, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.881578947368421, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.9013157894736842, 0.5277777777777778, 'X[24] <= 0.917\ngini =
0.117\nsamples = 32\nvalue = [30, 2]'),
 Text(0.8947368421052632, 0.4722222222222222, 'X[12] <= 0.812\ngini =
0.062\nsamples = 31\nvalue = [30, 1]'),
 Text(0.8881578947368421, 0.4166666666666667, 'gini = 0.0\nsamples = 28\nvalue =
[28, 0]'),
 Text(0.9013157894736842, 0.4166666666666667, 'X[19] <= 0.214\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.8947368421052632, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9078947368421053, 0.3611111111111111, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.9078947368421053, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8667763157894737, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9638157894736842, 0.75, 'X[12] <= 0.812\ngini = 0.456\nsamples =
57\nvalue = [37, 20]'),
 Text(0.9407894736842105, 0.6944444444444444, 'X[28] <= 0.4\ngini =
0.238\nsamples = 29\nvalue = [25, 4]'),
 Text(0.9276315789473685, 0.6388888888888888, 'X[9] <= 0.964\ngini =
0.142\nsamples = 26\nvalue = [24, 2]'),
 Text(0.9210526315789473, 0.5833333333333334, 'X[20] <= 0.5\ngini =
0.077\nsamples = 25\nvalue = [24, 1]'),
 Text(0.9144736842105263, 0.5277777777777778, 'gini = 0.0\nsamples = 23\nvalue =
[23, 0]'),
 Text(0.9276315789473685, 0.5277777777777778, 'X[23] <= 0.263\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
```

```
 Text(0.9210526315789473, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9342105263157895, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9342105263157895, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9539473684210527, 0.6388888888888888, 'X[19] <= 0.286\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.9473684210526315, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.9605263157894737, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9868421052631579, 0.6944444444444444, 'X[28] <= 0.1\ngini =
0.49\nsamples = 28\nvalue = [12, 16]'),
 Text(0.9802631578947368, 0.6388888888888888, 'X[10] <= 0.833\ngini =
0.48\nsamples = 20\nvalue = [12, 8]'),
 Text(0.9736842105263158, 0.5833333333333334, 'X[26] <= 0.013\ngini =
0.415\nsamples = 17\nvalue = [12, 5]'),
 Text(0.9671052631578947, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.9802631578947368, 0.5277777777777778, 'X[16] <= 0.505\ngini =
0.32\nsamples = 15\nvalue = [12, 3]'),
 Text(0.9736842105263158, 0.4722222222222222, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.9868421052631579, 0.4722222222222222, 'X[16] <= 0.706\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.9802631578947368, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.993421052631579, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.9868421052631579, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.993421052631579, 0.6388888888888888, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]')]
```



[36]:
```
from sklearn.model_selection import GridSearchCV
parameter={
 'criterion':['gini','entropy'],
  'splitter':['best','random'],
  'max_depth':[1,2,3,4,5,6,7,8,9,10],
  'max_features':['auto', 'sqrt', 'log2']

}
```

[37]:
```
grid_search=GridSearchCV(estimator=dtc,param_grid=parameter,cv=5,scoring="accuracy")
```

[38]:
```
grid_search.fit(x_train,y_train)
```

[38]:
```
GridSearchCV(cv=5, estimator=DecisionTreeClassifier(),
             param_grid={'criterion': ['gini', 'entropy'],
                         'max_depth': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
                         'max_features': ['auto', 'sqrt', 'log2'],
                         'splitter': ['best', 'random']},
             scoring='accuracy')
```

[39]:
```
grid_search.best_params_
```

[39]:
```
{'criterion': 'gini',
 'max_depth': 4,
 'max_features': 'auto',
```

```
 'splitter': 'best'}
```

[40]:
```
dtc_cv=DecisionTreeClassifier(criterion= 'entropy',
 max_depth= 4,
 max_features= 'sqrt',
 splitter= 'best')
dtc_cv.fit(x_train,y_train)
```

[40]:
```
DecisionTreeClassifier(criterion='entropy', max_depth=4, max_features='sqrt')
```

[41]:
```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

          No       0.85      0.84      0.85       245
         Yes       0.26      0.29      0.27        49

    accuracy                           0.75       294
   macro avg       0.56      0.56      0.56       294
weighted avg       0.76      0.75      0.75       294
```

[42]:
```
from sklearn.ensemble import RandomForestClassifier
classifier = RandomForestClassifier(n_estimators = 1000, criterion = 'entropy',␣
 ↪random_state = 0)
classifier.fit(x_train, y_train)
```

[42]:
```
RandomForestClassifier(criterion='entropy', n_estimators=1000, random_state=0)
```

[43]:
```
from sklearn.metrics import confusion_matrix, accuracy_score
y_pred = classifier.predict(x_test)
cm = confusion_matrix(y_test, y_pred)
print(cm)
accuracy_score(y_test, y_pred)
```

```
[[243   2]
 [ 41   8]]
```

[43]:
```
0.8537414965986394
```

[44]:
```
from sklearn.ensemble import RandomForestClassifier
```

[45]:
```
rfc=RandomForestClassifier()
```

[46]:
```
forest_params = [{'max_depth': list(range(10, 15)), 'max_features':␣
 ↪list(range(0,14))}]
```

[47]:
```
rfc_cv=GridSearchCV(rfc,param_grid=forest_params,cv=10,scoring="accuracy")
```

[ ]:
```
rfc_cv.fit(x_train,y_train)
```

[ ]:
```
print(classification_report(y_test,y_pred))
```