purushothamreddy-assignment-4

September 30, 2023

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

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

```
[16]:    Age Attrition     BusinessTravel  DailyRate              Department  \
      0   41       Yes      Travel_Rarely       1102                   Sales
      1   49        No  Travel_Frequently        279  Research & Development
      2   37       Yes      Travel_Rarely       1373  Research & Development
      3   33        No  Travel_Frequently       1392  Research & Development
      4   27        No      Travel_Rarely        591  Research & Development

         DistanceFromHome  Education EducationField  EmployeeCount  EmployeeNumber  \
      0                 1          2  Life Sciences              1               1
      1                 8          1  Life Sciences              1               2
      2                 2          2          Other              1               4
      3                 3          4  Life Sciences              1               5
      4                 2          1        Medical              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]
```

[17]: `data.head()`

```
[17]:    Age Attrition      BusinessTravel  DailyRate                Department  \
      0   41       Yes       Travel_Rarely       1102                     Sales
      1   49        No  Travel_Frequently        279  Research & Development
      2   37       Yes       Travel_Rarely       1373  Research & Development
      3   33        No  Travel_Frequently       1392  Research & Development
      4   27        No       Travel_Rarely        591  Research & Development

         DistanceFromHome  Education EducationField  EmployeeCount  EmployeeNumber  \
      0                 1          2  Life Sciences              1               1
      1                 8          1  Life Sciences              1               2
      2                 2          2          Other              1               4
      3                 3          4  Life Sciences              1               5
      4                 2          1        Medical              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]
```

[18]: `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
```

[19]: `data.describe()`

```
[19]:                 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]
```

[20]: `data.isnull().any()`

```
[20]: Age                        False
      Attrition                  False
      BusinessTravel             False
      DailyRate                  False
      Department                 False
      DistanceFromHome           False
      Education                  False
      EducationField             False
      EmployeeCount              False
      EmployeeNumber             False
      EnvironmentSatisfaction    False
      Gender                     False
      HourlyRate                 False
      JobInvolvement             False
      JobLevel                   False
      JobRole                    False
      JobSatisfaction            False
      MaritalStatus              False
      MonthlyIncome              False
      MonthlyRate                False
      NumCompaniesWorked         False
      Over18                     False
      OverTime                   False
      PercentSalaryHike          False
      PerformanceRating          False
      RelationshipSatisfaction   False
      StandardHours              False
      StockOptionLevel           False
      TotalWorkingYears          False
      TrainingTimesLastYear      False
```

```
WorkLifeBalance                False
YearsAtCompany                 False
YearsInCurrentRole             False
YearsSinceLastPromotion        False
YearsWithCurrManager           False
dtype: bool
```

[21]:
```python
cor = data.corr()
```

```
<ipython-input-21-06847dd9a2e1>:1: FutureWarning: The default value of
numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only
to silence this warning.
  cor = data.corr()
```

[22]:
```python
ax = plt.subplots(figsize=(20,20))
sns.heatmap(cor, annot=True)
```

[22]: <Axes: >
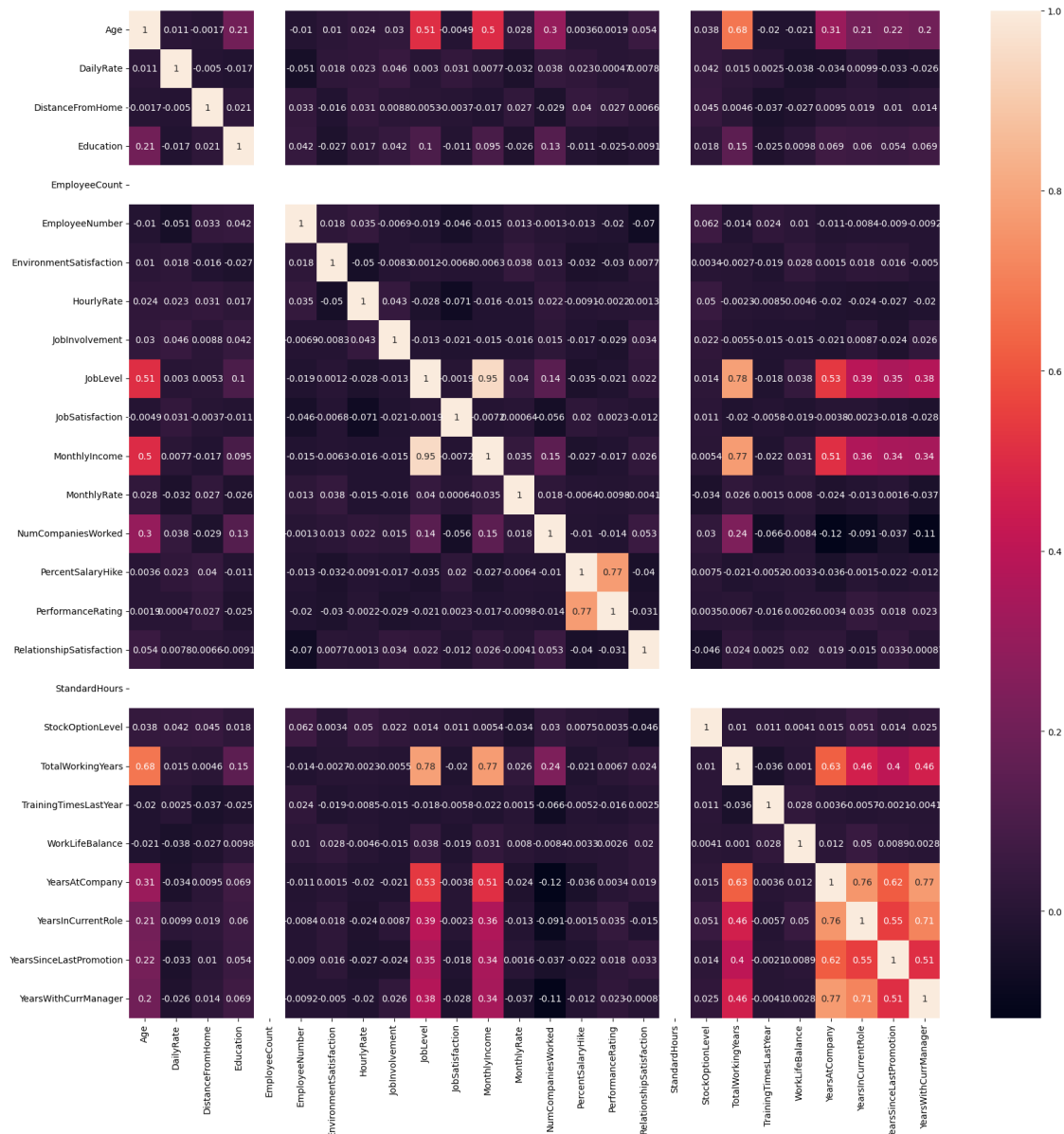
```python
[23]: from sklearn.preprocessing import LabelEncoder
```

```python
[24]: le=LabelEncoder()
```

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

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

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

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

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

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

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

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

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

[33]:

	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]
```

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

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

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

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

```
[38]: 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)
```

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

```
[39]: LogisticRegression(random_state=0)
```

```
[40]: 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]]
```

```
[40]: 88.09523809523809
```

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

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

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

```
[43]: DecisionTreeClassifier()
```

```
[44]: 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]
 [ 31  18]]
```

[44]: 76.19047619047619

[45]:
```python
from sklearn import tree
plt.figure(figsize=(25,15))
tree.plot_tree(dtc,filled=True)
```

[45]: [Text(0.32511340725806454, 0.9722222222222222, 'x[23] <= 0.038\ngini =
0.269\nsamples = 1176\nvalue = [988, 188]'),
 Text(0.07741935483870968, 0.9166666666666666, 'x[14] <= 0.75\ngini =
0.5\nsamples = 78\nvalue = [39, 39]'),
 Text(0.04838709677419355, 0.8611111111111112, 'x[4] <= 0.554\ngini =
0.426\nsamples = 39\nvalue = [27, 12]'),
 Text(0.03225806451612903, 0.8055555555555556, 'x[13] <= 0.167\ngini =
0.312\nsamples = 31\nvalue = [25, 6]'),
 Text(0.01935483870967742, 0.75, 'x[14] <= 0.25\ngini = 0.49\nsamples = 7\nvalue
= [3, 4]'),
 Text(0.012903225806451613, 0.6944444444444444, 'x[2] <= 0.175\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.0064516129032258064, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.01935483870967742, 0.6388888888888888, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.025806451612903226, 0.6944444444444444, 'gini = 0.0\nsamples = 3\nvalue
= [0, 3]'),
 Text(0.04516129032258064, 0.75, 'x[17] <= 0.056\ngini = 0.153\nsamples =
24\nvalue = [22, 2]'),
 Text(0.03870967741935484, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.05161290322580645, 0.6944444444444444, 'x[7] <= 0.167\ngini =
0.083\nsamples = 23\nvalue = [22, 1]'),
 Text(0.04516129032258064, 0.6388888888888888, 'x[12] <= 0.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.03870967741935484, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.05161290322580645, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.05806451612903226, 0.6388888888888888, 'gini = 0.0\nsamples = 21\nvalue
= [21, 0]'),
 Text(0.06451612903225806, 0.8055555555555556, 'x[19] <= 0.679\ngini =
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.05806451612903226, 0.75, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.07096774193548387, 0.75, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
```

Text(0.1064516129032258, 0.8611111111111112, 'x[9] <= 0.364\ngini =
0.426\nsamples = 39\nvalue = [12, 27]'),
 Text(0.09032258064516129, 0.8055555555555556, 'x[0] <= 0.369\ngini =
0.133\nsamples = 14\nvalue = [1, 13]'),
 Text(0.08387096774193549, 0.75, 'gini = 0.0\nsamples = 13\nvalue = [0, 13]'),
 Text(0.0967741935483871, 0.75, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.12258064516129032, 0.8055555555555556, 'x[18] <= 0.5\ngini =
0.493\nsamples = 25\nvalue = [11, 14]'),
 Text(0.10967741935483871, 0.75, 'x[6] <= 0.7\ngini = 0.484\nsamples = 17\nvalue
= [10, 7]'),
 Text(0.1032258064516129, 0.6944444444444444, 'x[2] <= 0.106\ngini =
0.408\nsamples = 14\nvalue = [10, 4]'),
 Text(0.0967741935483871, 0.6388888888888888, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.10967741935483871, 0.6388888888888888, 'x[25] <= 0.5\ngini =
0.278\nsamples = 12\nvalue = [10, 2]'),
 Text(0.1032258064516129, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11612903225806452, 0.5833333333333334, 'x[16] <= 0.892\ngini =
0.165\nsamples = 11\nvalue = [10, 1]'),
 Text(0.10967741935483871, 0.5277777777777778, 'gini = 0.0\nsamples = 10\nvalue
= [10, 0]'),
 Text(0.12258064516129032, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11612903225806452, 0.6944444444444444, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.13548387096774195, 0.75, 'x[4] <= 0.839\ngini = 0.219\nsamples =
8\nvalue = [1, 7]'),
 Text(0.12903225806451613, 0.6944444444444444, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.14193548387096774, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5728074596774193, 0.9166666666666666, 'x[18] <= 0.5\ngini =
0.235\nsamples = 1098\nvalue = [949, 149]'),
 Text(0.3165826612903226, 0.8611111111111112, 'x[25] <= 0.167\ngini =
0.162\nsamples = 798\nvalue = [727, 71]'),
 Text(0.18387096774193548, 0.8055555555555556, 'x[2] <= 0.747\ngini =
0.38\nsamples = 47\nvalue = [35, 12]'),
 Text(0.1774193548387097, 0.75, 'x[10] <= 0.5\ngini = 0.463\nsamples = 33\nvalue
= [21, 12]'),
 Text(0.15483870967741936, 0.6944444444444444, 'x[4] <= 0.446\ngini =
0.42\nsamples = 10\nvalue = [3, 7]'),
 Text(0.14838709677419354, 0.6388888888888888, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.16129032258064516, 0.6388888888888888, 'x[21] <= 0.333\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.15483870967741936, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =

[0, 1]'),
 Text(0.16774193548387098, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.2, 0.6944444444444444, 'x[24] <= 0.583\ngini = 0.34\nsamples = 23\nvalue
= [18, 5]'),
 Text(0.1870967741935484, 0.6388888888888888, 'x[28] <= 0.333\ngini =
0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.18064516129032257, 0.5833333333333334, 'gini = 0.0\nsamples = 15\nvalue
= [15, 0]'),
 Text(0.1935483870967742, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.2129032258064516, 0.6388888888888888, 'x[26] <= 0.287\ngini =
0.49\nsamples = 7\nvalue = [3, 4]'),
 Text(0.2064516129032258, 0.5833333333333334, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.21935483870967742, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.19032258064516128, 0.75, 'gini = 0.0\nsamples = 14\nvalue = [14, 0]'),
 Text(0.44929435483870966, 0.8055555555555556, 'x[23] <= 0.975\ngini =
0.145\nsamples = 751\nvalue = [692, 59]'),
 Text(0.4428427419354839, 0.75, 'x[26] <= 0.113\ngini = 0.143\nsamples =
750\nvalue = [692, 58]'),
 Text(0.30483870967741933, 0.6944444444444444, 'x[7] <= 0.167\ngini =
0.218\nsamples = 257\nvalue = [225, 32]'),
 Text(0.2532258064516129, 0.6388888888888888, 'x[29] <= 0.147\ngini =
0.355\nsamples = 65\nvalue = [50, 15]'),
 Text(0.23225806451612904, 0.5833333333333334, 'x[29] <= 0.029\ngini =
0.303\nsamples = 59\nvalue = [48, 11]'),
 Text(0.20967741935483872, 0.5277777777777778, 'x[10] <= 0.5\ngini =
0.463\nsamples = 22\nvalue = [14, 8]'),
 Text(0.1967741935483871, 0.4722222222222222, 'x[9] <= 0.179\ngini =
0.198\nsamples = 9\nvalue = [8, 1]'),
 Text(0.19032258064516128, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.2032258064516129, 0.4166666666666667, 'gini = 0.0\nsamples = 8\nvalue =
[8, 0]'),
 Text(0.22258064516129034, 0.4722222222222222, 'x[9] <= 0.4\ngini =
0.497\nsamples = 13\nvalue = [6, 7]'),
 Text(0.2161290322580645, 0.4166666666666667, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.22903225806451613, 0.4166666666666667, 'x[4] <= 0.286\ngini =
0.346\nsamples = 9\nvalue = [2, 7]'),
 Text(0.22258064516129034, 0.3611111111111111, 'x[16] <= 0.591\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.2161290322580645, 0.3055555555555556, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.22903225806451613, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =

[0, 1]'),
 Text(0.23548387096774193, 0.3611111111111111, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.25483870967741934, 0.5277777777777778, 'x[13] <= 0.167\ngini =
0.149\nsamples = 37\nvalue = [34, 3]'),
 Text(0.24838709677419354, 0.4722222222222222, 'x[25] <= 0.5\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.24193548387096775, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.25483870967741934, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.26129032258064516, 0.4722222222222222, 'gini = 0.0\nsamples = 31\nvalue
= [31, 0]'),
 Text(0.27419354838709675, 0.5833333333333334, 'x[21] <= 0.667\ngini =
0.444\nsamples = 6\nvalue = [2, 4]'),
 Text(0.267741935483871, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.2806451612903226, 0.5277777777777778, 'x[5] <= 0.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.27419354838709675, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.2870967741935484, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3564516129032258, 0.6388888888888888, 'x[0] <= 0.321\ngini =
0.161\nsamples = 192\nvalue = [175, 17]'),
 Text(0.3096774193548387, 0.5833333333333334, 'x[6] <= 0.1\ngini =
0.294\nsamples = 67\nvalue = [55, 12]'),
 Text(0.3032258064516129, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.3161290322580645, 0.5277777777777778, 'x[25] <= 0.5\ngini =
0.26\nsamples = 65\nvalue = [55, 10]'),
 Text(0.3, 0.4722222222222222, 'x[9] <= 0.679\ngini = 0.469\nsamples = 16\nvalue
= [10, 6]'),
 Text(0.29354838709677417, 0.4166666666666667, 'x[6] <= 0.4\ngini =
0.444\nsamples = 9\nvalue = [3, 6]'),
 Text(0.2870967741935484, 0.3611111111111111, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.3, 0.3611111111111111, 'x[2] <= 0.126\ngini = 0.245\nsamples = 7\nvalue
= [1, 6]'),
 Text(0.29354838709677417, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.3064516129032258, 0.3055555555555556, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.3064516129032258, 0.4166666666666667, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.33225806451612905, 0.4722222222222222, 'x[2] <= 0.037\ngini =
0.15\nsamples = 49\nvalue = [45, 4]'),

Text(0.3258064516129032, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3387096774193548, 0.4166666666666667, 'x[2] <= 0.938\ngini =
0.117\nsamples = 48\nvalue = [45, 3]'),
 Text(0.33225806451612905, 0.3611111111111111, 'x[5] <= 0.875\ngini =
0.081\nsamples = 47\nvalue = [45, 2]'),
 Text(0.3193548387096774, 0.3055555555555556, 'x[10] <= 0.167\ngini =
0.043\nsamples = 45\nvalue = [44, 1]'),
 Text(0.31290322580645163, 0.25, 'x[14] <= 0.75\ngini = 0.444\nsamples =
3\nvalue = [2, 1]'),
 Text(0.3064516129032258, 0.19444444444444445, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.3193548387096774, 0.19444444444444445, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3258064516129032, 0.25, 'gini = 0.0\nsamples = 42\nvalue = [42, 0]'),
 Text(0.34516129032258064, 0.3055555555555556, 'x[17] <= 0.278\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.3387096774193548, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.35161290322580646, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.34516129032258064, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4032258064516129, 0.5833333333333334, 'x[6] <= 0.9\ngini =
0.077\nsamples = 125\nvalue = [120, 5]'),
 Text(0.3903225806451613, 0.5277777777777778, 'x[0] <= 0.393\ngini =
0.05\nsamples = 118\nvalue = [115, 3]'),
 Text(0.38387096774193546, 0.4722222222222222, 'x[2] <= 0.956\ngini =
0.185\nsamples = 29\nvalue = [26, 3]'),
 Text(0.3774193548387097, 0.4166666666666667, 'x[10] <= 0.167\ngini =
0.133\nsamples = 28\nvalue = [26, 2]'),
 Text(0.36451612903225805, 0.3611111111111111, 'x[19] <= 0.214\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.3580645161290323, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.3709677419354839, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3903225806451613, 0.3611111111111111, 'x[29] <= 0.147\ngini =
0.074\nsamples = 26\nvalue = [25, 1]'),
 Text(0.38387096774193546, 0.3055555555555556, 'gini = 0.0\nsamples = 24\nvalue
= [24, 0]'),
 Text(0.3967741935483871, 0.3055555555555556, 'x[5] <= 0.25\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.3903225806451613, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.4032258064516129, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.3903225806451613, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3967741935483871, 0.4722222222222222, 'gini = 0.0\nsamples = 89\nvalue =
[89, 0]'),

```
Text(0.4161290322580645, 0.5277777777777778, 'x[2] <= 0.594\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.4096774193548387, 0.4722222222222222, 'x[23] <= 0.137\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4032258064516129, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4161290322580645, 0.4166666666666667, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.42258064516129035, 0.4722222222222222, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.5808467741935484, 0.6944444444444444, 'x[26] <= 0.787\ngini =
0.1\nsamples = 493\nvalue = [467, 26]'),
 Text(0.5423387096774194, 0.6388888888888888, 'x[13] <= 0.5\ngini =
0.094\nsamples = 486\nvalue = [462, 24]'),
 Text(0.4846774193548387, 0.5833333333333334, 'x[12] <= 0.938\ngini =
0.154\nsamples = 191\nvalue = [175, 16]'),
 Text(0.4782258064516129, 0.5277777777777778, 'x[16] <= 0.481\ngini =
0.145\nsamples = 190\nvalue = [175, 15]'),
 Text(0.4596774193548387, 0.4722222222222222, 'x[16] <= 0.47\ngini =
0.221\nsamples = 95\nvalue = [83, 12]'),
 Text(0.4532258064516129, 0.4166666666666667, 'x[29] <= 0.794\ngini =
0.207\nsamples = 94\nvalue = [83, 11]'),
 Text(0.4467741935483871, 0.3611111111111111, 'x[5] <= 0.375\ngini =
0.192\nsamples = 93\nvalue = [83, 10]'),
 Text(0.42258064516129035, 0.3055555555555556, 'x[6] <= 0.9\ngini =
0.363\nsamples = 21\nvalue = [16, 5]'),
 Text(0.4161290322580645, 0.25, 'x[15] <= 0.413\ngini = 0.266\nsamples =
19\nvalue = [16, 3]'),
 Text(0.4032258064516129, 0.19444444444444445, 'x[17] <= 0.056\ngini =
0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.3967741935483871, 0.1388888888888889, 'x[19] <= 0.107\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.3903225806451613, 0.08333333333333333, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4032258064516129, 0.08333333333333333, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4096774193548387, 0.1388888888888889, 'gini = 0.0\nsamples = 14\nvalue =
[14, 0]'),
 Text(0.4290322580645161, 0.19444444444444445, 'x[14] <= 0.25\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.42258064516129035, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.43548387096774194, 0.1388888888888889, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4290322580645161, 0.25, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.47096774193548385, 0.3055555555555556, 'x[27] <= 0.139\ngini =
0.129\nsamples = 72\nvalue = [67, 5]'),
```

```
 Text(0.4483870967741935, 0.25, 'x[29] <= 0.206\ngini = 0.444\nsamples =
6\nvalue = [4, 2]'),
 Text(0.44193548387096776, 0.19444444444444445, 'gini = 0.0\nsamples = 3\nvalue
= [3, 0]'),
 Text(0.45483870967741935, 0.19444444444444445, 'x[9] <= 0.686\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4483870967741935, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4612903225806452, 0.1388888888888889, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4935483870967742, 0.25, 'x[2] <= 0.958\ngini = 0.087\nsamples =
66\nvalue = [63, 3]'),
 Text(0.4806451612903226, 0.19444444444444445, 'x[24] <= 0.583\ngini =
0.061\nsamples = 64\nvalue = [62, 2]'),
 Text(0.47419354838709676, 0.1388888888888889, 'gini = 0.0\nsamples = 52\nvalue
= [52, 0]'),
 Text(0.4870967741935484, 0.1388888888888889, 'x[12] <= 0.812\ngini =
0.278\nsamples = 12\nvalue = [10, 2]'),
 Text(0.4806451612903226, 0.08333333333333333, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.4935483870967742, 0.08333333333333333, 'x[26] <= 0.163\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4870967741935484, 0.027777777777777776, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.5, 0.027777777777777776, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.5064516129032258, 0.19444444444444445, 'x[0] <= 0.655\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.5, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.5129032258064516, 0.1388888888888889, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4596774193548387, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4661290322580645, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4967741935483871, 0.4722222222222222, 'x[17] <= 0.5\ngini =
0.061\nsamples = 95\nvalue = [92, 3]'),
 Text(0.49032258064516127, 0.4166666666666667, 'gini = 0.0\nsamples = 76\nvalue
= [76, 0]'),
 Text(0.5032258064516129, 0.4166666666666667, 'x[29] <= 0.088\ngini =
0.266\nsamples = 19\nvalue = [16, 3]'),
 Text(0.49032258064516127, 0.3611111111111111, 'x[10] <= 0.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.4838709677419355, 0.3055555555555556, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4967741935483871, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5161290322580645, 0.3611111111111111, 'x[15] <= 0.108\ngini =
```

0.117\nsamples = 16\nvalue = [15, 1]'),
 Text(0.50967741935548387, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5225806451612903, 0.3055555555555556, 'gini = 0.0\nsamples = 15\nvalue =
[15, 0]'),
 Text(0.49112903225806454, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6, 0.5833333333333334, 'x[19] <= 0.036\ngini = 0.053\nsamples =
295\nvalue = [287, 8]'),
 Text(0.5774193548387097, 0.5277777777777778, 'x[28] <= 0.7\ngini =
0.159\nsamples = 46\nvalue = [42, 4]'),
 Text(0.5709677419354838, 0.4722222222222222, 'x[23] <= 0.688\ngini =
0.124\nsamples = 45\nvalue = [42, 3]'),
 Text(0.5548387096774193, 0.4166666666666667, 'x[9] <= 0.071\ngini =
0.089\nsamples = 43\nvalue = [41, 2]'),
 Text(0.5419354838709678, 0.3611111111111111, 'x[12] <= 0.688\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.535483870967742, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5483870967741935, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.567741935483871, 0.3611111111111111, 'x[12] <= 0.062\ngini =
0.048\nsamples = 41\nvalue = [40, 1]'),
 Text(0.5612903225806452, 0.3055555555555556, 'x[16] <= 0.346\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.5548387096774193, 0.25, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.567741935483871, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.5741935483870968, 0.3055555555555556, 'gini = 0.0\nsamples = 37\nvalue =
[37, 0]'),
 Text(0.5870967741935483, 0.4166666666666667, 'x[7] <= 0.333\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.5806451612903226, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5935483870967742, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5838709677419355, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6225806451612903, 0.5277777777777778, 'x[15] <= 0.056\ngini =
0.032\nsamples = 249\nvalue = [245, 4]'),
 Text(0.6064516129032258, 0.4722222222222222, 'x[15] <= 0.054\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.6, 0.4166666666666667, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.6129032258064516, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6387096774193548, 0.4722222222222222, 'x[2] <= 0.015\ngini =
0.024\nsamples = 244\nvalue = [241, 3]'),
 Text(0.6258064516129033, 0.4166666666666667, 'x[22] <= 0.667\ngini =

17

0.278\nsamples = 6\nvalue = [5, 1]'),
 Text(0.6193548387096774, 0.3611111111111111, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.632258064516129, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6516129032258065, 0.4166666666666667, 'x[21] <= 0.167\ngini =
0.017\nsamples = 238\nvalue = [236, 2]'),
 Text(0.6451612903225806, 0.3611111111111111, 'x[25] <= 0.833\ngini =
0.073\nsamples = 53\nvalue = [51, 2]'),
 Text(0.632258064516129, 0.3055555555555556, 'x[29] <= 0.088\ngini =
0.041\nsamples = 48\nvalue = [47, 1]'),
 Text(0.6258064516129033, 0.25, 'x[16] <= 0.824\ngini = 0.245\nsamples =
7\nvalue = [6, 1]'),
 Text(0.6193548387096774, 0.19444444444444445, 'gini = 0.0\nsamples = 6\nvalue =
[6, 0]'),
 Text(0.632258064516129, 0.19444444444444445, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6387096774193548, 0.25, 'gini = 0.0\nsamples = 41\nvalue = [41, 0]'),
 Text(0.6580645161290323, 0.3055555555555556, 'x[28] <= 0.367\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.6516129032258065, 0.25, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.6645161290322581, 0.25, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.6580645161290323, 0.3611111111111111, 'gini = 0.0\nsamples = 185\nvalue
= [185, 0]'),
 Text(0.6193548387096774, 0.6388888888888888, 'x[2] <= 0.366\ngini =
0.408\nsamples = 7\nvalue = [5, 2]'),
 Text(0.6129032258064516, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.6258064516129033, 0.5833333333333334, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.4557459677419355, 0.75, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.8290322580645161, 0.8611111111111112, 'x[15] <= 0.157\ngini =
0.385\nsamples = 300\nvalue = [222, 78]'),
 Text(0.7443548387096774, 0.8055555555555556, 'x[22] <= 0.167\ngini =
0.5\nsamples = 96\nvalue = [49, 47]'),
 Text(0.7096774193548387, 0.75, 'x[4] <= 0.161\ngini = 0.459\nsamples =
42\nvalue = [15, 27]'),
 Text(0.6838709677419355, 0.6944444444444444, 'x[16] <= 0.41\ngini =
0.499\nsamples = 23\nvalue = [12, 11]'),
 Text(0.6645161290322581, 0.6388888888888888, 'x[15] <= 0.061\ngini =
0.426\nsamples = 13\nvalue = [4, 9]'),
 Text(0.6580645161290323, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6709677419354839, 0.5833333333333334, 'x[24] <= 0.25\ngini =
0.298\nsamples = 11\nvalue = [2, 9]'),
 Text(0.6645161290322581, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),

Text(0.6774193548387096, 0.5277777777777778, 'x[0] <= 0.214\ngini =
0.18\nsamples = 10\nvalue = [1, 9]'),
 Text(0.6709677419354839, 0.4722222222222222, 'x[28] <= 0.1\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.6645161290322581, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6774193548387096, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6838709677419355, 0.4722222222222222, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]'),
 Text(0.7032258064516129, 0.6388888888888888, 'x[7] <= 0.167\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.6967741935483871, 0.5833333333333334, 'x[26] <= 0.138\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.6903225806451613, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.7032258064516129, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7096774193548387, 0.5833333333333334, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.7354838709677419, 0.6944444444444444, 'x[23] <= 0.35\ngini =
0.266\nsamples = 19\nvalue = [3, 16]'),
 Text(0.7290322580645161, 0.6388888888888888, 'x[9] <= 0.2\ngini =
0.198\nsamples = 18\nvalue = [2, 16]'),
 Text(0.7225806451612903, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7354838709677419, 0.5833333333333334, 'x[28] <= 0.433\ngini =
0.111\nsamples = 17\nvalue = [1, 16]'),
 Text(0.7290322580645161, 0.5277777777777778, 'gini = 0.0\nsamples = 15\nvalue =
[0, 15]'),
 Text(0.7419354838709677, 0.5277777777777778, 'x[14] <= 0.75\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.7354838709677419, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7483870967741936, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7419354838709677, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7790322580645161, 0.75, 'x[0] <= 0.202\ngini = 0.466\nsamples =
54\nvalue = [34, 20]'),
 Text(0.7612903225806451, 0.6944444444444444, 'x[10] <= 0.833\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
 Text(0.7548387096774194, 0.6388888888888888, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.7677419354838709, 0.6388888888888888, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.7967741935483871, 0.6944444444444444, 'x[2] <= 0.622\ngini =

0.418\nsamples = 47\nvalue = [33, 14]'),
 Text(0.7806451612903226, 0.6388888888888888, 'x[2] <= 0.145\ngini =
0.482\nsamples = 32\nvalue = [19, 13]'),
 Text(0.7677419354838709, 0.5833333333333334, 'x[15] <= 0.068\ngini =
0.18\nsamples = 10\nvalue = [9, 1]'),
 Text(0.7612903225806451, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7741935483870968, 0.5277777777777778, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.7935483870967742, 0.5833333333333334, 'x[16] <= 0.87\ngini =
0.496\nsamples = 22\nvalue = [10, 12]'),
 Text(0.7870967741935484, 0.5277777777777778, 'x[25] <= 0.833\ngini =
0.465\nsamples = 19\nvalue = [7, 12]'),
 Text(0.7806451612903226, 0.4722222222222222, 'x[17] <= 0.167\ngini =
0.415\nsamples = 17\nvalue = [5, 12]'),
 Text(0.7677419354838709, 0.4166666666666667, 'x[19] <= 0.321\ngini =
0.49\nsamples = 7\nvalue = [4, 3]'),
 Text(0.7612903225806451, 0.3611111111111111, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.7741935483870968, 0.3611111111111111, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.7935483870967742, 0.4166666666666667, 'x[3] <= 0.25\ngini =
0.18\nsamples = 10\nvalue = [1, 9]'),
 Text(0.7870967741935484, 0.3611111111111111, 'x[2] <= 0.241\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.7806451612903226, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7935483870967742, 0.3055555555555556, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8, 0.3611111111111111, 'gini = 0.0\nsamples = 8\nvalue = [0, 8]'),
 Text(0.7935483870967742, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.8, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.8129032258064516, 0.6388888888888888, 'x[9] <= 0.064\ngini =
0.124\nsamples = 15\nvalue = [14, 1]'),
 Text(0.8064516129032258, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8193548387096774, 0.5833333333333334, 'gini = 0.0\nsamples = 14\nvalue =
[14, 0]'),
 Text(0.9137096774193548, 0.8055555555555556, 'x[14] <= 0.75\ngini =
0.258\nsamples = 204\nvalue = [173, 31]'),
 Text(0.8629032258064516, 0.75, 'x[15] <= 0.992\ngini = 0.138\nsamples =
147\nvalue = [136, 11]'),
 Text(0.8564516129032258, 0.6944444444444444, 'x[4] <= 0.482\ngini =
0.128\nsamples = 146\nvalue = [136, 10]'),
 Text(0.8387096774193549, 0.6388888888888888, 'x[26] <= 0.063\ngini =
0.038\nsamples = 104\nvalue = [102, 2]'),

```
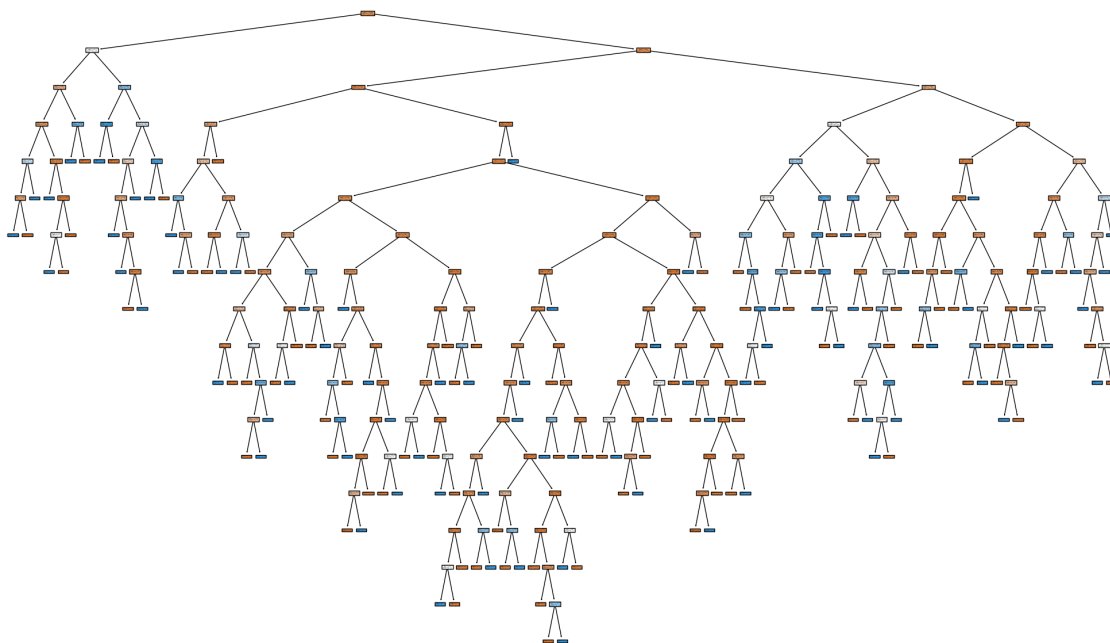 Text(0.832258064516129, 0.5833333333333334, 'x[9] <= 0.193\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.8258064516129032, 0.5277777777777778, 'x[24] <= 0.417\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.8193548387096774, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.832258064516129, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.8387096774193549, 0.5277777777777778, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.8451612903225807, 0.5833333333333334, 'gini = 0.0\nsamples = 94\nvalue =
[94, 0]'),
 Text(0.8741935483870967, 0.6388888888888888, 'x[7] <= 0.167\ngini =
0.308\nsamples = 42\nvalue = [34, 8]'),
 Text(0.8580645161290322, 0.5833333333333334, 'x[16] <= 0.194\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.8516129032258064, 0.5277777777777778, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.864516129032258, 0.5277777777777778, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.8903225806451613, 0.5833333333333334, 'x[0] <= 0.393\ngini =
0.229\nsamples = 38\nvalue = [33, 5]'),
 Text(0.8774193548387097, 0.5277777777777778, 'x[4] <= 0.821\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.8709677419354839, 0.4722222222222222, 'x[2] <= 0.311\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.864516129032258, 0.4166666666666667, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8774193548387097, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.8838709677419355, 0.4722222222222222, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.9032258064516129, 0.5277777777777778, 'x[24] <= 0.917\ngini =
0.117\nsamples = 32\nvalue = [30, 2]'),
 Text(0.896774193548387, 0.4722222222222222, 'x[12] <= 0.812\ngini =
0.062\nsamples = 31\nvalue = [30, 1]'),
 Text(0.8903225806451613, 0.4166666666666667, 'gini = 0.0\nsamples = 28\nvalue =
[28, 0]'),
 Text(0.9032258064516129, 0.4166666666666667, 'x[26] <= 0.125\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.896774193548387, 0.3611111111111111, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9096774193548387, 0.3611111111111111, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.9096774193548387, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8693548387096774, 0.6944444444444444, 'gini = 0.0\nsamples = 1\nvalue =
```

[0, 1]'),
 Text(0.964516129032258, 0.75, 'x[12] <= 0.812\ngini = 0.456\nsamples =
57\nvalue = [37, 20]'),
 Text(0.9419354838709677, 0.6944444444444444, 'x[28] <= 0.4\ngini =
0.238\nsamples = 29\nvalue = [25, 4]'),
 Text(0.9290322580645162, 0.6388888888888888, 'x[9] <= 0.964\ngini =
0.142\nsamples = 26\nvalue = [24, 2]'),
 Text(0.9225806451612903, 0.5833333333333334, 'x[20] <= 0.5\ngini =
0.077\nsamples = 25\nvalue = [24, 1]'),
 Text(0.9161290322580645, 0.5277777777777778, 'gini = 0.0\nsamples = 23\nvalue =
[23, 0]'),
 Text(0.9290322580645162, 0.5277777777777778, 'x[0] <= 0.333\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.9225806451612903, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9354838709677419, 0.4722222222222222, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9354838709677419, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9548387096774194, 0.6388888888888888, 'x[1] <= 0.25\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.9483870967741935, 0.5833333333333334, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9612903225806452, 0.5833333333333334, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.9870967741935484, 0.6944444444444444, 'x[28] <= 0.1\ngini =
0.49\nsamples = 28\nvalue = [12, 16]'),
 Text(0.9806451612903225, 0.6388888888888888, 'x[10] <= 0.833\ngini =
0.48\nsamples = 20\nvalue = [12, 8]'),
 Text(0.9741935483870968, 0.5833333333333334, 'x[26] <= 0.013\ngini =
0.415\nsamples = 17\nvalue = [12, 5]'),
 Text(0.967741935483871, 0.5277777777777778, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.9806451612903225, 0.5277777777777778, 'x[16] <= 0.505\ngini =
0.32\nsamples = 15\nvalue = [12, 3]'),
 Text(0.9741935483870968, 0.4722222222222222, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.9870967741935484, 0.4722222222222222, 'x[16] <= 0.706\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.9806451612903225, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.9935483870967742, 0.4166666666666667, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.9870967741935484, 0.5833333333333334, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.9935483870967742, 0.6388888888888888, 'gini = 0.0\nsamples = 8\nvalue =
[0, 8]')]

```
[46]: 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']


      }
```

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

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

```
[49]: grid_search.best_params_
```

```
[49]: {'criterion': 'gini',
       'max_depth': 4,
       'max_features': 'auto',
       'splitter': 'random'}
```

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

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

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

```
              precision    recall  f1-score   support

          No       0.87      0.84      0.85       245
         Yes       0.32      0.37      0.34        49

    accuracy                           0.76       294
   macro avg       0.59      0.60      0.60       294
weighted avg       0.78      0.76      0.77       294
```

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

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

```
[53]: 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]]
```

```
[53]: 0.8537414965986394
```

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

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

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

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

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

/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py:378: FitFailedWarning:
50 fits failed out of a total of 700.
The score on these train-test partitions for these parameters will be set to
nan.

If these failures are not expected, you can try to debug them by setting
error_score='raise'.

Below are more details about the failures:
----
50 fits failed with the following error:
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-
packages/sklearn/model_selection/_validation.py", line 686, in _fit_and_score
    estimator.fit(X_train, y_train, **fit_params)
  File "/usr/local/lib/python3.10/dist-packages/sklearn/ensemble/_forest.py",
line 340, in fit
    self._validate_params()
  File "/usr/local/lib/python3.10/dist-packages/sklearn/base.py", line 600, in
_validate_params
    validate_parameter_constraints(
  File "/usr/local/lib/python3.10/dist-
packages/sklearn/utils/_param_validation.py", line 97, in
validate_parameter_constraints
    raise InvalidParameterError(
sklearn.utils._param_validation.InvalidParameterError: The 'max_features'
parameter of RandomForestClassifier must be an int in the range [1, inf), a
float in the range (0.0, 1.0], a str among {'log2', 'sqrt', 'auto' (deprecated)}
or None. Got 0 instead.

  warnings.warn(some_fits_failed_message, FitFailedWarning)
/usr/local/lib/python3.10/dist-packages/sklearn/model_selection/_search.py:952:
UserWarning: One or more of the test scores are non-finite: [nan
0.85034043 0.85375924 0.85798928 0.86224105 0.85882949
 0.85882949 0.86137911 0.86477618 0.85714182 0.86307403 0.85796755
 0.85542518 0.85543966   nan 0.84949297 0.85715631 0.8596842
 0.85885122 0.85541793 0.85542518 0.85712009 0.85967695 0.8622483
 0.85371578 0.86476894 0.85880776 0.85798928   nan 0.84949297
 0.85885122 0.85885122 0.85713458 0.85969144 0.86137187 0.86055338
 0.85969144 0.8605389 0.86477618 0.86136462 0.85884398 0.8613936
        nan 0.85120962 0.85884398 0.85545415 0.85712734 0.86054614
 0.86223381 0.86563813 0.85541069 0.85713458 0.85882949 0.85541069
 0.8553962 0.86138635   nan 0.85203535 0.85800377 0.85884398
 0.8605389 0.86138635 0.86138635 0.8596842 0.86307403 0.85969868
 0.86394321 0.86221932 0.86137911 0.8596842]
  warnings.warn(

[60]: GridSearchCV(cv=10, estimator=RandomForestClassifier(),
             param_grid=[{'max_depth': [10, 11, 12, 13, 14],
                          'max_features': [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
                                           12, 13]}],
             scoring='accuracy')

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

	precision	recall	f1-score	support
No	0.86	0.99	0.92	245
Yes	0.80	0.16	0.27	49
accuracy			0.85	294
macro avg	0.83	0.58	0.59	294
weighted avg	0.85	0.85	0.81	294