assignment-4

September 28, 2023

Importing libraries

[103]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Importing Dataset

[104]:
```
df = pd.read_csv("/content/WA_Fn-UseC_-HR-Employee-Attrition.csv")
```

[105]:
```
df.head()
```

[105]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	EmployeeNumber	…	RelationshipSatisfaction	StandardHours	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	YearsAtCompany
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                   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]
```

[106]:
```
df.shape
```

[106]:
```
(1470, 35)
```

[107]:
```
df.Attrition.value_counts()
```

[107]:
```
No     1233
Yes     237
Name: Attrition, dtype: int64
```

[108]:
```
df.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
```

[109]:
```
df.describe()
```

[109]:

	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]
```

Checking for Null values

[110]:
```
df.isnull().any()
```

[110]:
```
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
```

[111]:
```
df.isnull().sum()
```

[111]:
```
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
```

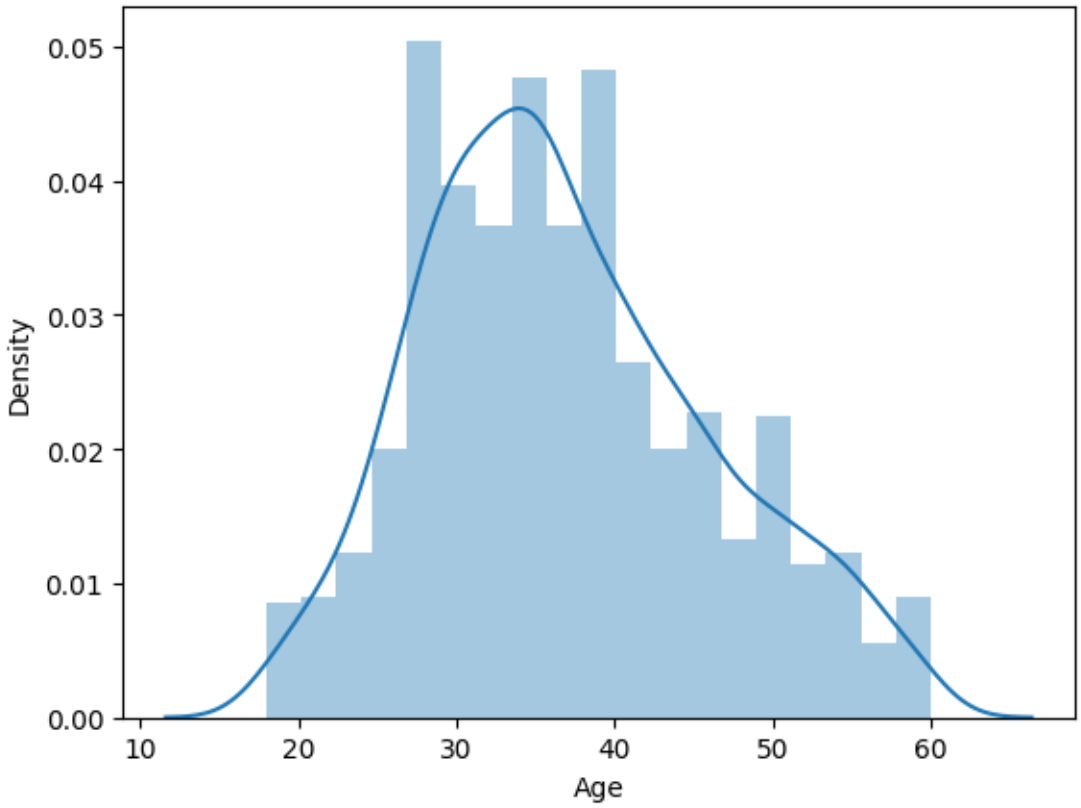
```
JobLevel                    0
JobRole                     0
JobSatisfaction             0
MaritalStatus               0
MonthlyIncome               0
MonthlyRate                 0
NumCompaniesWorked          0
Over18                      0
OverTime                    0
PercentSalaryHike           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
```

Data Visualization

[112]:
```
sns.distplot(df['Age'])
```

```
<ipython-input-112-0fafe04ea3f6>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['Age'])
```

[112]:
```
<Axes: xlabel='Age', ylabel='Density'>
```



[113]:
```
df.corr()
```

```
<ipython-input-113-2f6f6606aa2c>: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.
  df.corr()
```

[113]:

	Age	DailyRate	DistanceFromHome	Education	\
Age	1.000000	0.010661	-0.001686	0.208034	
DailyRate	0.010661	1.000000	-0.004985	-0.016806	
DistanceFromHome	-0.001686	-0.004985	1.000000	0.021042	
Education	0.208034	-0.016806	0.021042	1.000000	
EmployeeCount	NaN	NaN	NaN	NaN	
EmployeeNumber	-0.010145	-0.050990	0.032916	0.042070	
EnvironmentSatisfaction	0.010146	0.018355	-0.016075	-0.027128	
HourlyRate	0.024287	0.023381	0.031131	0.016775	
JobInvolvement	0.029820	0.046135	0.008783	0.042438	
JobLevel	0.509604	0.002966	0.005303	0.101589	
JobSatisfaction	-0.004892	0.030571	-0.003669	-0.011296	
MonthlyIncome	0.497855	0.007707	-0.017014	0.094961	

MonthlyRate	0.028051	-0.032182	0.027473	-0.026084
NumCompaniesWorked	0.299635	0.038153	-0.029251	0.126317
PercentSalaryHike	0.003634	0.022704	0.040235	-0.011111
PerformanceRating	0.001904	0.000473	0.027110	-0.024539
RelationshipSatisfaction	0.053535	0.007846	0.006557	-0.009118
StandardHours	NaN	NaN	NaN	NaN
StockOptionLevel	0.037510	0.042143	0.044872	0.018422
TotalWorkingYears	0.680381	0.014515	0.004628	0.148280
TrainingTimesLastYear	-0.019621	0.002453	-0.036942	-0.025100
WorkLifeBalance	-0.021490	-0.037848	-0.026556	0.009819
YearsAtCompany	0.311309	-0.034055	0.009508	0.069114
YearsInCurrentRole	0.212901	0.009932	0.018845	0.060236
YearsSinceLastPromotion	0.216513	-0.033229	0.010029	0.054254
YearsWithCurrManager	0.202089	-0.026363	0.014406	0.069065

	EmployeeCount	EmployeeNumber
Age	NaN	-0.010145
DailyRate	NaN	-0.050990
DistanceFromHome	NaN	0.032916
Education	NaN	0.042070
EmployeeCount	NaN	NaN
EmployeeNumber	NaN	1.000000
EnvironmentSatisfaction	NaN	0.017621
HourlyRate	NaN	0.035179
JobInvolvement	NaN	-0.006888
JobLevel	NaN	-0.018519
JobSatisfaction	NaN	-0.046247
MonthlyIncome	NaN	-0.014829
MonthlyRate	NaN	0.012648
NumCompaniesWorked	NaN	-0.001251
PercentSalaryHike	NaN	-0.012944
PerformanceRating	NaN	-0.020359
RelationshipSatisfaction	NaN	-0.069861
StandardHours	NaN	NaN
StockOptionLevel	NaN	0.062227
TotalWorkingYears	NaN	-0.014365
TrainingTimesLastYear	NaN	0.023603
WorkLifeBalance	NaN	0.010309
YearsAtCompany	NaN	-0.011240
YearsInCurrentRole	NaN	-0.008416
YearsSinceLastPromotion	NaN	-0.009019
YearsWithCurrManager	NaN	-0.009197

	EnvironmentSatisfaction	HourlyRate	JobInvolvement
Age	0.010146	0.024287	0.029820
DailyRate	0.018355	0.023381	0.046135
DistanceFromHome	-0.016075	0.031131	0.008783

Education	-0.027128	0.016775	0.042438
EmployeeCount	NaN	NaN	NaN
EmployeeNumber	0.017621	0.035179	-0.006888
EnvironmentSatisfaction	1.000000	-0.049857	-0.008278
HourlyRate	-0.049857	1.000000	0.042861
JobInvolvement	-0.008278	0.042861	1.000000
JobLevel	0.001212	-0.027853	-0.012630
JobSatisfaction	-0.006784	-0.071335	-0.021476
MonthlyIncome	-0.006259	-0.015794	-0.015271
MonthlyRate	0.037600	-0.015297	-0.016322
NumCompaniesWorked	0.012594	0.022157	0.015012
PercentSalaryHike	-0.031701	-0.009062	-0.017205
PerformanceRating	-0.029548	-0.002172	-0.029071
RelationshipSatisfaction	0.007665	0.001330	0.034297
StandardHours	NaN	NaN	NaN
StockOptionLevel	0.003432	0.050263	0.021523
TotalWorkingYears	-0.002693	-0.002334	-0.005533
TrainingTimesLastYear	-0.019359	-0.008548	-0.015338
WorkLifeBalance	0.027627	-0.004607	-0.014617
YearsAtCompany	0.001458	-0.019582	-0.021355
YearsInCurrentRole	0.018007	-0.024106	0.008717
YearsSinceLastPromotion	0.016194	-0.026716	-0.024184
YearsWithCurrManager	-0.004999	-0.020123	0.025976

	JobLevel	…	RelationshipSatisfaction	\
Age	0.509604	…	0.053535	
DailyRate	0.002966	…	0.007846	
DistanceFromHome	0.005303	…	0.006557	
Education	0.101589	…	-0.009118	
EmployeeCount	NaN	…	NaN	
EmployeeNumber	-0.018519	…	-0.069861	
EnvironmentSatisfaction	0.001212	…	0.007665	
HourlyRate	-0.027853	…	0.001330	
JobInvolvement	-0.012630	…	0.034297	
JobLevel	1.000000	…	0.021642	
JobSatisfaction	-0.001944	…	-0.012454	
MonthlyIncome	0.950300	…	0.025873	
MonthlyRate	0.039563	…	-0.004085	
NumCompaniesWorked	0.142501	…	0.052733	
PercentSalaryHike	-0.034730	…	-0.040490	
PerformanceRating	-0.021222	…	-0.031351	
RelationshipSatisfaction	0.021642	…	1.000000	
StandardHours	NaN	…	NaN	
StockOptionLevel	0.013984	…	-0.045952	
TotalWorkingYears	0.782208	…	0.024054	
TrainingTimesLastYear	-0.018191	…	0.002497	
WorkLifeBalance	0.037818	…	0.019604	

YearsAtCompany	0.534739	…	0.019367
YearsInCurrentRole	0.389447	…	-0.015123
YearsSinceLastPromotion	0.353885	…	0.033493
YearsWithCurrManager	0.375281	…	-0.000867

	StandardHours	StockOptionLevel	TotalWorkingYears	\
Age	NaN	0.037510	0.680381	
DailyRate	NaN	0.042143	0.014515	
DistanceFromHome	NaN	0.044872	0.004628	
Education	NaN	0.018422	0.148280	
EmployeeCount	NaN	NaN	NaN	
EmployeeNumber	NaN	0.062227	-0.014365	
EnvironmentSatisfaction	NaN	0.003432	-0.002693	
HourlyRate	NaN	0.050263	-0.002334	
JobInvolvement	NaN	0.021523	-0.005533	
JobLevel	NaN	0.013984	0.782208	
JobSatisfaction	NaN	0.010690	-0.020185	
MonthlyIncome	NaN	0.005408	0.772893	
MonthlyRate	NaN	-0.034323	0.026442	
NumCompaniesWorked	NaN	0.030075	0.237639	
PercentSalaryHike	NaN	0.007528	-0.020608	
PerformanceRating	NaN	0.003506	0.006744	
RelationshipSatisfaction	NaN	-0.045952	0.024054	
StandardHours	NaN	NaN	NaN	
StockOptionLevel	NaN	1.000000	0.010136	
TotalWorkingYears	NaN	0.010136	1.000000	
TrainingTimesLastYear	NaN	0.011274	-0.035662	
WorkLifeBalance	NaN	0.004129	0.001008	
YearsAtCompany	NaN	0.015058	0.628133	
YearsInCurrentRole	NaN	0.050818	0.460365	
YearsSinceLastPromotion	NaN	0.014352	0.404858	
YearsWithCurrManager	NaN	0.024698	0.459188	

	TrainingTimesLastYear	WorkLifeBalance	\
Age	-0.019621	-0.021490	
DailyRate	0.002453	-0.037848	
DistanceFromHome	-0.036942	-0.026556	
Education	-0.025100	0.009819	
EmployeeCount	NaN	NaN	
EmployeeNumber	0.023603	0.010309	
EnvironmentSatisfaction	-0.019359	0.027627	
HourlyRate	-0.008548	-0.004607	
JobInvolvement	-0.015338	-0.014617	
JobLevel	-0.018191	0.037818	
JobSatisfaction	-0.005779	-0.019459	
MonthlyIncome	-0.021736	0.030683	
MonthlyRate	0.001467	0.007963	

NumCompaniesWorked	-0.066054	-0.008366
PercentSalaryHike	-0.005221	-0.003280
PerformanceRating	-0.015579	0.002572
RelationshipSatisfaction	0.002497	0.019604
StandardHours	NaN	NaN
StockOptionLevel	0.011274	0.004129
TotalWorkingYears	-0.035662	0.001008
TrainingTimesLastYear	1.000000	0.028072
WorkLifeBalance	0.028072	1.000000
YearsAtCompany	0.003569	0.012089
YearsInCurrentRole	-0.005738	0.049856
YearsSinceLastPromotion	-0.002067	0.008941
YearsWithCurrManager	-0.004096	0.002759

	YearsAtCompany	YearsInCurrentRole	\
Age	0.311309	0.212901	
DailyRate	-0.034055	0.009932	
DistanceFromHome	0.009508	0.018845	
Education	0.069114	0.060236	
EmployeeCount	NaN	NaN	
EmployeeNumber	-0.011240	-0.008416	
EnvironmentSatisfaction	0.001458	0.018007	
HourlyRate	-0.019582	-0.024106	
JobInvolvement	-0.021355	0.008717	
JobLevel	0.534739	0.389447	
JobSatisfaction	-0.003803	-0.002305	
MonthlyIncome	0.514285	0.363818	
MonthlyRate	-0.023655	-0.012815	
NumCompaniesWorked	-0.118421	-0.090754	
PercentSalaryHike	-0.035991	-0.001520	
PerformanceRating	0.003435	0.034986	
RelationshipSatisfaction	0.019367	-0.015123	
StandardHours	NaN	NaN	
StockOptionLevel	0.015058	0.050818	
TotalWorkingYears	0.628133	0.460365	
TrainingTimesLastYear	0.003569	-0.005738	
WorkLifeBalance	0.012089	0.049856	
YearsAtCompany	1.000000	0.758754	
YearsInCurrentRole	0.758754	1.000000	
YearsSinceLastPromotion	0.618409	0.548056	
YearsWithCurrManager	0.769212	0.714365	

	YearsSinceLastPromotion	YearsWithCurrManager
Age	0.216513	0.202089
DailyRate	-0.033229	-0.026363
DistanceFromHome	0.010029	0.014406
Education	0.054254	0.069065

```
EmployeeCount                      NaN                NaN
EmployeeNumber               -0.009019          -0.009197
EnvironmentSatisfaction       0.016194          -0.004999
HourlyRate                   -0.026716          -0.020123
JobInvolvement               -0.024184           0.025976
JobLevel                      0.353885           0.375281
JobSatisfaction              -0.018214          -0.027656
MonthlyIncome                 0.344978           0.344079
MonthlyRate                   0.001567          -0.036746
NumCompaniesWorked           -0.036814          -0.110319
PercentSalaryHike            -0.022154          -0.011985
PerformanceRating             0.017896           0.022827
RelationshipSatisfaction      0.033493          -0.000867
StandardHours                      NaN                NaN
StockOptionLevel              0.014352           0.024698
TotalWorkingYears             0.404858           0.459188
TrainingTimesLastYear        -0.002067          -0.004096
WorkLifeBalance               0.008941           0.002759
YearsAtCompany                0.618409           0.769212
YearsInCurrentRole            0.548056           0.714365
YearsSinceLastPromotion       1.000000           0.510224
YearsWithCurrManager          0.510224           1.000000

[26 rows x 26 columns]
```

df.drop([“Age”,“DailyRate”,“DistanceFromHome”,“Education”,“EmployeeCount”,“EmployeeNumber”,“Environm

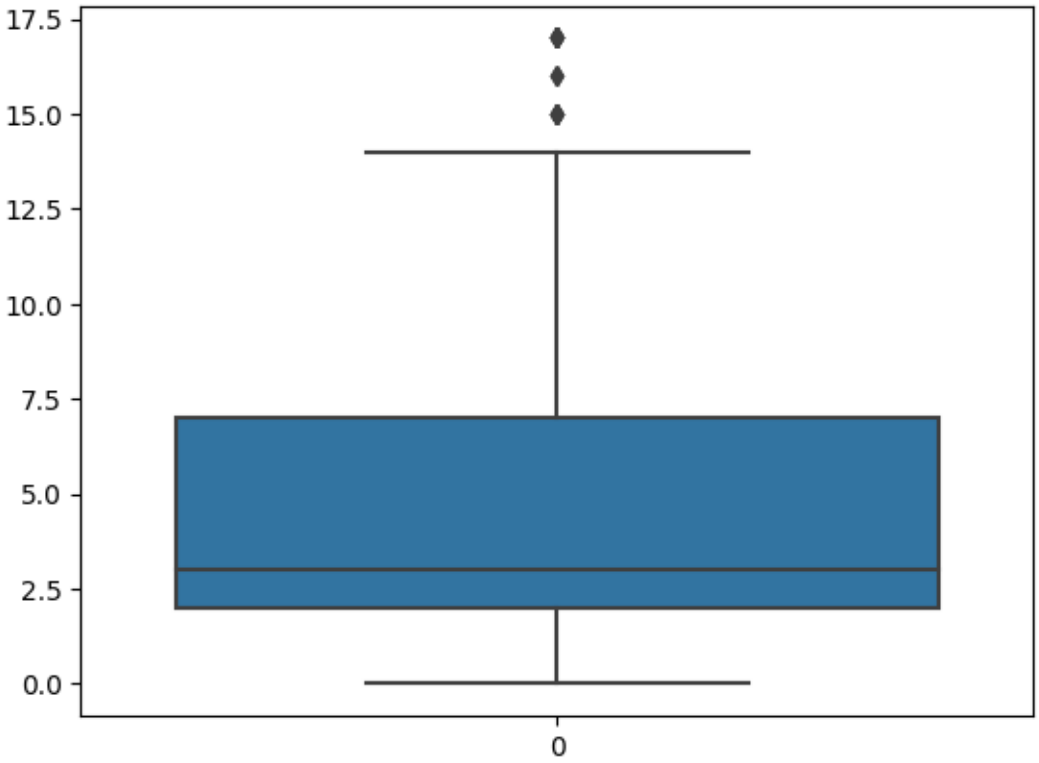
[114]:
```
df.corr()
plt.subplots(figsize=(20,15))
sns.heatmap(df.corr(),annot=True)
```

```
<ipython-input-114-0df31041a0eb>: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.
  df.corr()
<ipython-input-114-0df31041a0eb>:3: 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.
  sns.heatmap(df.corr(),annot=True)
```

[114]: <Axes: >

[115]:
```
sns.boxplot(df.YearsWithCurrManager)
```

[115]: <Axes: >



[116]:
```
q1 = df.YearsWithCurrManager.quantile(0.25)
q3 = df.YearsWithCurrManager.quantile(0.75)
```

[117]:
```
IQR =q3-q1
```

[118]:
```
upper_limit =q3+1.5*IQR
```

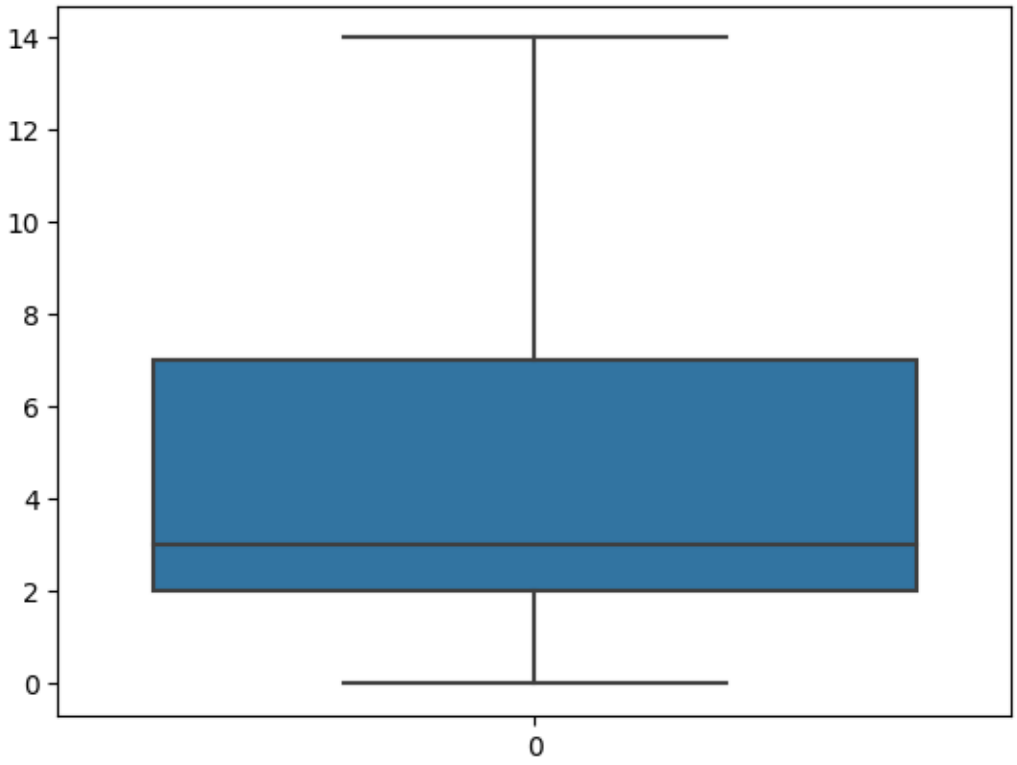
[119]:
```
upper_limit
```

[119]: 14.5

[120]:
```
df = df[df.YearsWithCurrManager<upper_limit]
```

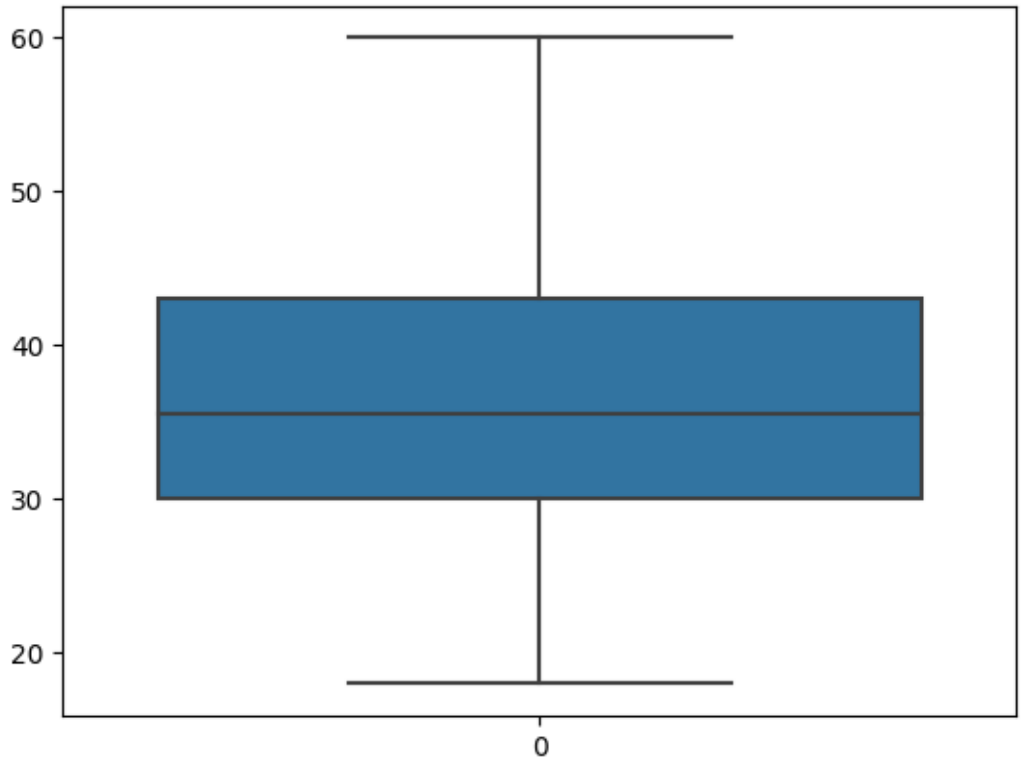
[121]:
```
sns.boxplot(df.YearsWithCurrManager)
```

[121]: <Axes: >



[122]:
```
sns.boxplot(df.Age)
```

[122]: <Axes: >



[123]:
```
df.head()
```

[123]:

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

Label Encoding

[124]:
```
print(df.dtypes)
```

```
Age                        int64
Attrition                 object
BusinessTravel            object
DailyRate                  int64
Department                object
DistanceFromHome           int64
Education                  int64
EducationField            object
EmployeeCount              int64
EmployeeNumber             int64
EnvironmentSatisfaction    int64
Gender                    object
HourlyRate                 int64
JobInvolvement             int64
JobLevel                   int64
JobRole                   object
JobSatisfaction            int64
MaritalStatus             object
MonthlyIncome              int64
MonthlyRate                int64
NumCompaniesWorked         int64
Over18                    object
OverTime                  object
PercentSalaryHike          int64
PerformanceRating          int64
RelationshipSatisfaction   int64
StandardHours              int64
```

```
StockOptionLevel             int64
TotalWorkingYears            int64
TrainingTimesLastYear        int64
WorkLifeBalance              int64
YearsAtCompany               int64
YearsInCurrentRole           int64
YearsSinceLastPromotion      int64
YearsWithCurrManager         int64
dtype: object
```

[125]:
```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
df.Attrition=le.fit_transform(df.Attrition)
df.BusinessTravel=le.fit_transform(df.BusinessTravel)
df.Department=le.fit_transform(df.Department)
df.EducationField=le.fit_transform(df.EducationField)
df.Gender=le.fit_transform(df.Gender)
df.JobRole=le.fit_transform(df.JobRole)
df.MaritalStatus=le.fit_transform(df.MaritalStatus)
df.Over18=le.fit_transform(df.Over18)
df.OverTime=le.fit_transform(df.OverTime)


df.head()
```

[125]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	\
0	41	1	2	1102	2	1	
1	49	0	1	279	1	8	
2	37	1	2	1373	1	2	
3	33	0	1	1392	1	3	
4	27	0	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]
```

[126]:
```python
print(df.dtypes)
```

```
Age                          int64
Attrition                    int64
BusinessTravel               int64
DailyRate                    int64
Department                   int64
DistanceFromHome             int64
Education                    int64
EducationField               int64
EmployeeCount                int64
EmployeeNumber               int64
EnvironmentSatisfaction      int64
Gender                       int64
HourlyRate                   int64
JobInvolvement               int64
JobLevel                     int64
JobRole                      int64
JobSatisfaction              int64
MaritalStatus                int64
MonthlyIncome                int64
MonthlyRate                  int64
NumCompaniesWorked           int64
Over18                       int64
OverTime                     int64
PercentSalaryHike            int64
PerformanceRating            int64
RelationshipSatisfaction     int64
StandardHours                int64
StockOptionLevel             int64
TotalWorkingYears            int64
TrainingTimesLastYear        int64
```

```
WorkLifeBalance             int64
YearsAtCompany              int64
YearsInCurrentRole          int64
YearsSinceLastPromotion     int64
YearsWithCurrManager        int64
dtype: object
```

Separating Dependent and Independent Variables

[127]:
```
df.
↪drop(["Age","DailyRate","DistanceFromHome","Education","EmployeeCount","EmployeeNumber","En
```

[128]:
```
df
```

[128]:

	Attrition	BusinessTravel	Department	EducationField	Gender	JobRole	\
0	1	2	2	1	0	7	
1	0	1	1	1	1	6	
2	1	2	1	4	1	2	
3	0	1	1	1	0	6	
4	0	2	1	3	1	2	
…	…	…	…	…	…	…	
1465	0	1	1	3	1	2	
1466	0	2	1	3	1	0	
1467	0	2	1	1	1	4	
1468	0	1	2	3	1	7	
1469	0	2	1	3	1	2	

	MaritalStatus	MonthlyIncome	Over18	OverTime	TotalWorkingYears	\
0	2	5993	0	1	8	
1	1	5130	0	0	10	
2	2	2090	0	1	7	
3	1	2909	0	1	8	
4	1	3468	0	0	6	
…	…	…	…	…	…	
1465	1	2571	0	0	17	
1466	1	9991	0	0	9	
1467	1	6142	0	1	6	
1468	1	5390	0	0	17	
1469	1	4404	0	0	6	

	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion	\
0	6	4	0	
1	10	7	1	
2	0	0	0	
3	8	7	3	
4	2	2	2	
…	…	…	…	
1465	5	2	0	

```
1466                7                    7                          1
1467                6                    2                          0
1468                9                    6                          0
1469                4                    3                          1

      YearsWithCurrManager
0                        5
1                        7
2                        0
3                        0
4                        2
…                      …
1465                     3
1466                     7
1467                     3
1468                     8
1469                     2

[1456 rows x 15 columns]
```

[129]:
```
df.head()
```

[129]:
```
   Attrition  BusinessTravel  Department  EducationField  Gender  JobRole  \
0          1               2           2               1       0        7
1          0               1           1               1       1        6
2          1               2           1               4       1        2
3          0               1           1               1       0        6
4          0               2           1               3       1        2

   MaritalStatus  MonthlyIncome  Over18  OverTime  TotalWorkingYears  \
0              2           5993       0         1                  8
1              1           5130       0         0                 10
2              2           2090       0         1                  7
3              1           2909       0         1                  8
4              1           3468       0         0                  6

   YearsAtCompany  YearsInCurrentRole  YearsSinceLastPromotion  \
0               6                   4                        0
1              10                   7                        1
2               0                   0                        0
3               8                   7                        3
4               2                   2                        2

   YearsWithCurrManager
0                     5
1                     7
2                     0
```

```
3                     0
4                     2
```

[130]:
```
x=df.iloc[:,1:15]
x.head()
```

[130]:
```
   BusinessTravel  Department  EducationField  Gender  JobRole  MaritalStatus  \
0               2           2               1       0        7              2
1               1           1               1       1        6              1
2               2           1               4       1        2              2
3               1           1               1       0        6              1
4               2           1               3       1        2              1

   MonthlyIncome  Over18  OverTime  TotalWorkingYears  YearsAtCompany  \
0           5993       0         1                  8               6
1           5130       0         0                 10              10
2           2090       0         1                  7               0
3           2909       0         1                  8               8
4           3468       0         0                  6               2

   YearsInCurrentRole  YearsSinceLastPromotion  YearsWithCurrManager
0                   4                        0                     5
1                   7                        1                     7
2                   0                        0                     0
3                   7                        3                     0
4                   2                        2                     2
```

[131]:
```
y = df["Attrition"]
y.head()
```

[131]:
```
0    1
1    0
2    1
3    0
4    0
Name: Attrition, dtype: int64
```

Feature Scaling

[132]:
```
from sklearn.preprocessing import MinMaxScaler
ms=MinMaxScaler()
x_scaled=pd.DataFrame(ms.fit_transform(x),columns=x.columns)
```

[133]:
```
x_scaled
```

[133]:
```
      BusinessTravel  Department  EducationField  Gender  JobRole  \
0                1.0         1.0             0.2     0.0    0.875
```

```
1                0.5         0.5             0.2     1.0    0.750
2                1.0         0.5             0.8     1.0    0.250
3                0.5         0.5             0.2     0.0    0.750
4                1.0         0.5             0.6     1.0    0.250
…              …           …               …       …      …
1451             0.5         0.5             0.6     1.0    0.250
1452             1.0         0.5             0.6     1.0    0.000
1453             1.0         0.5             0.2     1.0    0.500
1454             0.5         1.0             0.6     1.0    0.875
1455             1.0         0.5             0.6     1.0    0.250

      MaritalStatus  MonthlyIncome  Over18  OverTime  TotalWorkingYears  \
0               1.0       0.262454     0.0       1.0              0.200
1               0.5       0.217009     0.0       0.0              0.250
2               1.0       0.056925     0.0       1.0              0.175
3               0.5       0.100053     0.0       1.0              0.200
4               0.5       0.129489     0.0       0.0              0.150
…             …             …      …       …                …
1451            0.5       0.082254     0.0       0.0              0.425
1452            0.5       0.472986     0.0       0.0              0.225
1453            0.5       0.270300     0.0       1.0              0.150
1454            0.5       0.230700     0.0       0.0              0.425
1455            0.5       0.178778     0.0       0.0              0.150

      YearsAtCompany  YearsInCurrentRole  YearsSinceLastPromotion  \
0              0.150            0.222222                 0.000000
1              0.250            0.388889                 0.066667
2              0.000            0.000000                 0.000000
3              0.200            0.388889                 0.200000
4              0.050            0.111111                 0.133333
…              …                  …                    …
1451           0.125            0.111111                 0.000000
1452           0.175            0.388889                 0.066667
1453           0.150            0.111111                 0.000000
1454           0.225            0.333333                 0.000000
1455           0.100            0.166667                 0.066667

      YearsWithCurrManager
0                 0.357143
1                 0.500000
2                 0.000000
3                 0.000000
4                 0.142857
…                    …
1451              0.214286
1452              0.500000
1453              0.214286
```

```
1454              0.571429
1455              0.142857

[1456 rows x 14 columns]
```

Splitting Data into Train and Test

[134]:
```
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)
```

[135]:
```
x_train.shape,x_test.shape,y_train.shape,y_test.shape
```

[135]:
```
((1164, 14), (292, 14), (1164,), (292,))
```

[136]:
```
x_train.head()
```

[136]:

	BusinessTravel	Department	EducationField	Gender	JobRole	\
1399	1.0	0.5	0.8	1.0	0.25	
1308	1.0	0.5	0.2	0.0	0.25	
679	1.0	0.5	1.0	1.0	0.25	
638	1.0	0.5	1.0	1.0	0.50	
247	1.0	0.5	0.2	1.0	0.75	

	MaritalStatus	MonthlyIncome	Over18	OverTime	TotalWorkingYears	\
1399	0.0	0.156293	0.0	1.0	0.175	
1308	1.0	0.057715	0.0	0.0	0.175	
679	1.0	0.103423	0.0	0.0	0.025	
638	0.5	0.521011	0.0	0.0	0.400	
247	1.0	0.070090	0.0	0.0	0.150	

	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion	\
1399	0.050	0.111111	0.000000	
1308	0.050	0.111111	0.133333	
679	0.025	0.000000	0.000000	
638	0.325	0.555556	0.266667	
247	0.150	0.277778	0.066667	

	YearsWithCurrManager
1399	0.142857
1308	0.000000
679	0.000000
638	0.571429
247	0.357143

Model Building (Logistic Regression)

[137]:
```
from sklearn.linear_model import LogisticRegression
model=LogisticRegression()
```

[138]:
```
model.fit(x_train,y_train)
```

[138]:
```
LogisticRegression()
```

[139]:
```
pred=model.predict(x_test)
```

[140]:
```
pred
```

[140]:
```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0])
```

[141]:
```
y_test
```

[141]:
```
517     0
973     0
236     1
697     0
35      0
       ..
315     0
1440    0
41      0
1253    0
76      0
Name: Attrition, Length: 292, dtype: int64
```

[142]:
```
df
```

[142]:

	Attrition	BusinessTravel	Department	EducationField	Gender	JobRole	\
0	1	2	2	1	0	7	
1	0	1	1	1	1	6	
2	1	2	1	4	1	2	

```
3              0                1            1                1        0         6
4              0                2            1                3        1         2
…            …                …          …                …      …      …
1465           0                1            1                3        1         2
1466           0                2            1                3        1         0
1467           0                2            1                1        1         4
1468           0                1            2                3        1         7
1469           0                2            1                3        1         2

      MaritalStatus  MonthlyIncome  Over18  OverTime  TotalWorkingYears  \
0                 2           5993       0         1                  8
1                 1           5130       0         0                 10
2                 2           2090       0         1                  7
3                 1           2909       0         1                  8
4                 1           3468       0         0                  6
…               …             …      …       …                …
1465              1           2571       0         0                 17
1466              1           9991       0         0                  9
1467              1           6142       0         1                  6
1468              1           5390       0         0                 17
1469              1           4404       0         0                  6

      YearsAtCompany  YearsInCurrentRole  YearsSinceLastPromotion  \
0                  6                   4                        0
1                 10                   7                        1
2                  0                   0                        0
3                  8                   7                        3
4                  2                   2                        2
…                …                   …                      …
1465               5                   2                        0
1466               7                   7                        1
1467               6                   2                        0
1468               9                   6                        0
1469               4                   3                        1

      YearsWithCurrManager
0                        5
1                        7
2                        0
3                        0
4                        2
…                      …
1465                     3
1466                     7
1467                     3
1468                     8
1469                     2
```

```
[1456 rows x 15 columns]
```

[143]:
```
model.predict(ms.transform([[2,2,1,0,7,2,5993,0,1,8,6,4,0,5]]))
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but MinMaxScaler was fitted with feature names
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but LogisticRegression was fitted with feature
names
  warnings.warn(
```

[143]:
```
array([0])
```

Evaluation of Classification Model

[144]:
```
from sklearn.metrics import␣
↪accuracy_score,confusion_matrix,classification_report,roc_auc_score,roc_curve
```

[145]:
```
accuracy_score(y_test,pred)
```

[145]:
```
0.8698630136986302
```

[146]:
```
confusion_matrix(y_test,pred)
```

[146]:
```
array([[245,   4],
       [ 34,   9]])
```

[147]:
```
#predicted no     predicted yes

#Actual No      245=TN                  4=FP
#Actual yes     34=FN                   9=TP
```

[148]:
```
(245+9)/292 #accuracy
```

[148]:
```
0.8698630136986302
```

[149]:
```
print(classification_report(y_test,pred))
```

```
              precision    recall  f1-score   support

           0       0.88      0.98      0.93       249
           1       0.69      0.21      0.32        43

    accuracy                           0.87       292
   macro avg       0.79      0.60      0.62       292
```

```
weighted avg       0.85      0.87      0.84       292
```

[150]:
```
# precision = TP/(TP+FP)
9/(9+4)
```

[150]: 0.6923076923076923

[151]:
```
# Recall = TP/(FN+TP)
9/(34+9)
```

[151]: 0.20930232558139536

[152]:
```
# F1 score
# 2*precision*Recall/(Precision+Recall)
2*0.69*0.20/(0.69+0.20)
```

[152]: 0.3101123595505618

Decision Tree

[153]:
```
from sklearn.tree import DecisionTreeClassifier
dtc=DecisionTreeClassifier()
```

[154]:
```
dtc.fit(x_train,y_train)
```

[154]: DecisionTreeClassifier()

[155]:
```
pred=dtc.predict(x_test)
```

[156]:
```
pred
```

[156]:
```
array([0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1,
       0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1,
       0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0,
       0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1,
       0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1,
       1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1,
       1, 0, 0, 0, 0, 0])
```

[157]:
```
y_test
```

[157]:
```
517     0
973     0
236     1
697     0
35      0
       ..
315     0
1440    0
41      0
1253    0
76      0
Name: Attrition, Length: 292, dtype: int64
```

[158]:
```
df
```

[158]:

	Attrition	BusinessTravel	Department	EducationField	Gender	JobRole	\
0	1	2	2	1	0	7	
1	0	1	1	1	1	6	
2	1	2	1	4	1	2	
3	0	1	1	1	0	6	
4	0	2	1	3	1	2	
…	…	…	…	…	…	…	
1465	0	1	1	3	1	2	
1466	0	2	1	3	1	0	
1467	0	2	1	1	1	4	
1468	0	1	2	3	1	7	
1469	0	2	1	3	1	2	

	MaritalStatus	MonthlyIncome	Over18	OverTime	TotalWorkingYears	\
0	2	5993	0	1	8	
1	1	5130	0	0	10	
2	2	2090	0	1	7	
3	1	2909	0	1	8	
4	1	3468	0	0	6	
…	…	…	…	…	…	
1465	1	2571	0	0	17	
1466	1	9991	0	0	9	
1467	1	6142	0	1	6	
1468	1	5390	0	0	17	
1469	1	4404	0	0	6	

	YearsAtCompany	YearsInCurrentRole	YearsSinceLastPromotion	\
0	6	4	0	
1	10	7	1	
2	0	0	0	
3	8	7	3	
4	2	2	2	

```
…             …                 …                       …
1465           5                 2                        0
1466           7                 7                        1
1467           6                 2                        0
1468           9                 6                        0
1469           4                 3                        1

      YearsWithCurrManager
0                        5
1                        7
2                        0
3                        0
4                        2
…                    …
1465                     3
1466                     7
1467                     3
1468                     8
1469                     2

[1456 rows x 15 columns]
```

[159]:
```
dtc.predict(ms.transform([[2,2,1,0,7,2,5993,0,1,8,6,4,0,5]]))
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but MinMaxScaler was fitted with feature names
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but DecisionTreeClassifier was fitted with feature
names
  warnings.warn(
```

[159]:
```
array([1])
```

[160]:
```
accuracy_score(y_test,pred)
```

[160]:
```
0.791095890410959
```

[161]:
```
confusion_matrix(y_test,pred)
```

[161]:
```
array([[213,  36],
       [ 25,  18]])
```

[162]:
```
#predicted no     predicted yes
#Actual No      213=TN                 36=FP
#Actual yes     25=FN                  18=TP
```

[187]:
```
#Accuracy
(213+18)/292
```

[187]:
```
0.791095890410959
```

[164]:
```
print(classification_report(y_test,pred))
```

```
              precision    recall  f1-score   support

           0       0.89      0.86      0.87       249
           1       0.33      0.42      0.37        43

    accuracy                           0.79       292
   macro avg       0.61      0.64      0.62       292
weighted avg       0.81      0.79      0.80       292
```

[188]:
```
# precision = TP/(TP+FP)
18/(54)
```

[188]:
```
0.3333333333333333
```

[189]:
```
# Recall = TP/(FN+TP)
18/(25+18)
```

[189]:
```
0.4186046511627907
```

[190]:
```
#F1 score
# 2*precision*Recall/(Precision+Recall)
2*0.33*0.41/(0.33+0.41)
```

[190]:
```
0.3656756756756757
```

ROC-AUC Curve

[168]:
```
probability=dtc.predict_proba(x_test)[:,1]
```

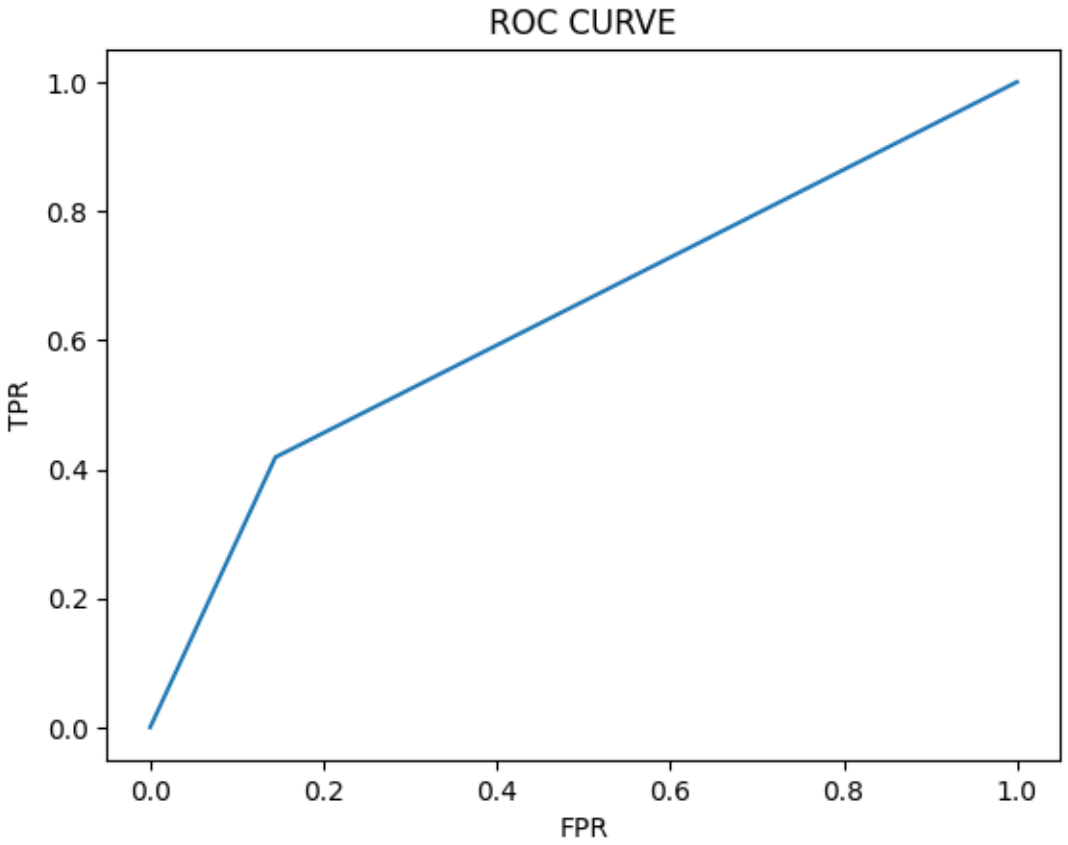
[169]:
```
probability
```

[169]:
```
array([0., 0., 0., 1., 0., 0., 0., 0., 1., 0., 0., 0., 0., 0., 0., 0., 0.,
       1., 0., 0., 0., 1., 0., 0., 1., 1., 0., 0., 0., 1., 0., 1., 0., 0.,
       0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 1., 0., 1.,
       0., 0., 1., 0., 0., 0., 0., 0., 0., 1., 0., 1., 0., 0., 1., 0., 0.,
       0., 0., 1., 0., 0., 0., 1., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 0., 1., 0., 0., 0., 1., 1., 0., 0., 0., 1., 0., 0., 0., 0.,
       0., 0., 0., 0., 0., 1., 1., 0., 0., 1., 0., 0., 0., 0., 0., 0., 0.,
       0., 0., 1., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 1., 0., 0.,
       0., 1., 0., 0., 0., 1., 0., 0., 1., 1., 0., 1., 0., 0., 0., 0., 0.,
```

```
1., 0., 0., 0., 0., 1., 0., 0., 1., 1., 0., 1., 0., 0., 0., 0., 0.,
0., 0., 0., 0., 0., 0., 0., 1., 0., 0., 0., 0., 1., 0., 0., 0., 0.,
0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0.,
1., 0., 1., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 1., 0., 1., 1.,
1., 0., 0., 0., 1., 0., 0., 0., 1., 0., 0., 0., 1., 0., 0., 0., 0.,
0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 0., 1., 0., 0., 0.,
0., 0., 0., 0., 0., 1., 0., 0., 0., 0., 0., 0., 0., 1., 0., 0., 1.,
0., 0., 0., 0., 0., 0., 0., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0.,
0., 0., 0.])
```

[170]:
```
fpr,tpr,threshsholds = roc_curve(y_test,probability)
```

[171]:
```
plt.plot(fpr,tpr)
plt.xlabel('FPR')
plt.ylabel('TPR')
plt.title('ROC CURVE')
plt.show()
```



Hyper parameter tuning

[172]:
```
from sklearn import tree
plt.figure(figsize=(25,15))
tree.plot_tree(dtc,filled=True)
```

[172]:
```
[Text(0.5312069163602942, 0.9736842105263158, 'x[8] <= 0.5\ngini =
0.278\nsamples = 1164\nvalue = [970, 194]'),
 Text(0.3007869944852941, 0.9210526315789473, 'x[9] <= 0.063\ngini =
0.189\nsamples = 832\nvalue = [744, 88]'),
 Text(0.060294117647058824, 0.868421052631579, 'x[5] <= 0.75\ngini =
0.46\nsamples = 64\nvalue = [41, 23]'),
 Text(0.023529411764705882, 0.8157894736842105, 'x[2] <= 0.1\ngini =
0.327\nsamples = 34\nvalue = [27, 7]'),
 Text(0.01764705882352941, 0.7631578947368421, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.029411764705882353, 0.7631578947368421, 'x[12] <= 0.033\ngini =
0.225\nsamples = 31\nvalue = [27, 4]'),
 Text(0.011764705882352941, 0.7105263157894737, 'x[6] <= 0.083\ngini =
0.095\nsamples = 20\nvalue = [19, 1]'),
 Text(0.0058823529411764705, 0.6578947368421053, 'gini = 0.0\nsamples =
13\nvalue = [13, 0]'),
 Text(0.01764705882352941, 0.6578947368421053, 'x[6] <= 0.086\ngini =
0.245\nsamples = 7\nvalue = [6, 1]'),
 Text(0.011764705882352941, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.023529411764705882, 0.6052631578947368, 'gini = 0.0\nsamples = 6\nvalue
= [6, 0]'),
 Text(0.047058823529411764, 0.7105263157894737, 'x[2] <= 0.9\ngini =
0.397\nsamples = 11\nvalue = [8, 3]'),
 Text(0.041176470588235294, 0.6578947368421053, 'x[6] <= 0.047\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.03529411764705882, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.047058823529411764, 0.6052631578947368, 'x[0] <= 0.75\ngini =
0.198\nsamples = 9\nvalue = [8, 1]'),
 Text(0.041176470588235294, 0.5526315789473685, 'x[3] <= 0.5\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.03529411764705882, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.047058823529411764, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.052941176470588235, 0.5526315789473685, 'gini = 0.0\nsamples = 6\nvalue
= [6, 0]'),
 Text(0.052941176470588235, 0.6578947368421053, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.09705882352941177, 0.8157894736842105, 'x[11] <= 0.028\ngini =
0.498\nsamples = 30\nvalue = [14, 16]'),
 Text(0.08235294117647059, 0.7631578947368421, 'x[9] <= 0.013\ngini =
0.42\nsamples = 20\nvalue = [6, 14]'),
 Text(0.07058823529411765, 0.7105263157894737, 'x[4] <= 0.5\ngini =
```

```
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.06470588235294118, 0.6578947368421053, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.07647058823529412, 0.6578947368421053, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.09411764705882353, 0.7105263157894737, 'x[6] <= 0.023\ngini =
0.305\nsamples = 16\nvalue = [3, 13]'),
 Text(0.08823529411764706, 0.6578947368421053, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.1, 0.6578947368421053, 'x[6] <= 0.067\ngini = 0.42\nsamples = 10\nvalue
= [3, 7]'),
 Text(0.09411764705882353, 0.6052631578947368, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.10588235294117647, 0.6052631578947368, 'x[6] <= 0.096\ngini =
0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.1, 0.5526315789473685, 'gini = 0.0\nsamples = 6\nvalue = [0, 6]'),
 Text(0.11176470588235295, 0.5526315789473685, 'x[6] <= 0.103\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.10588235294117647, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.11764705882352941, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.11176470588235295, 0.7631578947368421, 'x[6] <= 0.056\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.10588235294117647, 0.7105263157894737, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.11764705882352941, 0.7105263157894737, 'x[6] <= 0.091\ngini =
0.198\nsamples = 9\nvalue = [8, 1]'),
 Text(0.11176470588235295, 0.6578947368421053, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.12352941176470589, 0.6578947368421053, 'x[2] <= 0.5\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.11764705882352941, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.12941176470588237, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5412798713235294, 0.868421052631579, 'x[4] <= 0.938\ngini =
0.155\nsamples = 768\nvalue = [703, 65]'),
 Text(0.5002068014705883, 0.8157894736842105, 'x[12] <= 0.967\ngini =
0.144\nsamples = 743\nvalue = [685, 58]'),
 Text(0.4474724264705882, 0.7631578947368421, 'x[6] <= 0.51\ngini =
0.14\nsamples = 738\nvalue = [682, 56]'),
 Text(0.3772977941176471, 0.7105263157894737, 'x[6] <= 0.475\ngini =
0.162\nsamples = 607\nvalue = [553, 54]'),
 Text(0.28400735294117646, 0.6578947368421053, 'x[13] <= 0.107\ngini =
0.148\nsamples = 585\nvalue = [538, 47]'),
 Text(0.18235294117647058, 0.6052631578947368, 'x[2] <= 0.9\ngini =
0.266\nsamples = 95\nvalue = [80, 15]'),
 Text(0.1588235294117647, 0.5526315789473685, 'x[6] <= 0.298\ngini =
```

```
0.247\nsamples = 90\nvalue = [77, 13]'),
 Text(0.12941176470588237, 0.5, 'x[9] <= 0.138\ngini = 0.209\nsamples =
76\nvalue = [67, 9]'),
 Text(0.10588235294117647, 0.4473684210526316, 'x[9] <= 0.113\ngini =
0.351\nsamples = 22\nvalue = [17, 5]'),
 Text(0.1, 0.39473684210526316, 'gini = 0.0\nsamples = 9\nvalue = [9, 0]'),
 Text(0.11176470588235295, 0.39473684210526316, 'x[5] <= 0.75\ngini =
0.473\nsamples = 13\nvalue = [8, 5]'),
 Text(0.1, 0.34210526315789475, 'x[0] <= 0.25\ngini = 0.346\nsamples = 9\nvalue
= [7, 2]'),
 Text(0.09411764705882353, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.10588235294117647, 0.2894736842105263, 'x[6] <= 0.174\ngini =
0.219\nsamples = 8\nvalue = [7, 1]'),
 Text(0.1, 0.23684210526315788, 'x[6] <= 0.14\ngini = 0.375\nsamples = 4\nvalue
= [3, 1]'),
 Text(0.09411764705882353, 0.18421052631578946, 'gini = 0.0\nsamples = 3\nvalue
= [3, 0]'),
 Text(0.10588235294117647, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.11176470588235295, 0.23684210526315788, 'gini = 0.0\nsamples = 4\nvalue
= [4, 0]'),
 Text(0.12352941176470589, 0.34210526315789475, 'x[1] <= 0.75\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.11764705882352941, 0.2894736842105263, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.12941176470588237, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.15294117647058825, 0.4473684210526316, 'x[10] <= 0.237\ngini =
0.137\nsamples = 54\nvalue = [50, 4]'),
 Text(0.1411764705882353, 0.39473684210526316, 'x[2] <= 0.5\ngini =
0.109\nsamples = 52\nvalue = [49, 3]'),
 Text(0.13529411764705881, 0.34210526315789475, 'gini = 0.0\nsamples = 30\nvalue
= [30, 0]'),
 Text(0.14705882352941177, 0.34210526315789475, 'x[10] <= 0.038\ngini =
0.236\nsamples = 22\nvalue = [19, 3]'),
 Text(0.1411764705882353, 0.2894736842105263, 'x[2] <= 0.7\ngini =
0.397\nsamples = 11\nvalue = [8, 3]'),
 Text(0.13529411764705881, 0.23684210526315788, 'x[0] <= 0.75\ngini =
0.32\nsamples = 10\nvalue = [8, 2]'),
 Text(0.12352941176470589, 0.18421052631578946, 'x[6] <= 0.148\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.11764705882352941, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.12941176470588237, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.14705882352941177, 0.18421052631578946, 'x[4] <= 0.125\ngini =
```

```
0.219\nsamples = 8\nvalue = [7, 1]'),
 Text(0.1411764705882353, 0.13157894736842105, 'x[9] <= 0.312\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.13529411764705881, 0.07894736842105263, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.14705882352941177, 0.07894736842105263, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.15294117647058825, 0.13157894736842105, 'gini = 0.0\nsamples = 6\nvalue
= [6, 0]'),
 Text(0.14705882352941177, 0.23684210526315788, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.15294117647058825, 0.2894736842105263, 'gini = 0.0\nsamples = 11\nvalue
= [11, 0]'),
 Text(0.16470588235294117, 0.39473684210526316, 'x[4] <= 0.562\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.1588235294117647, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.17058823529411765, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.18823529411764706, 0.5, 'x[6] <= 0.303\ngini = 0.408\nsamples =
14\nvalue = [10, 4]'),
 Text(0.18235294117647058, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.19411764705882353, 0.4473684210526316, 'x[4] <= 0.25\ngini =
0.165\nsamples = 11\nvalue = [10, 1]'),
 Text(0.18823529411764706, 0.39473684210526316, 'x[0] <= 0.75\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.18235294117647058, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.19411764705882353, 0.34210526315789475, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.2, 0.39473684210526316, 'gini = 0.0\nsamples = 8\nvalue = [8, 0]'),
 Text(0.20588235294117646, 0.5526315789473685, 'x[4] <= 0.625\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.2, 0.5, 'gini = 0.0\nsamples = 3\nvalue = [3, 0]'),
 Text(0.21176470588235294, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.38566176470588237, 0.6052631578947368, 'x[2] <= 0.9\ngini =
0.122\nsamples = 490\nvalue = [458, 32]'),
 Text(0.3360294117647059, 0.5526315789473685, 'x[6] <= 0.332\ngini =
0.104\nsamples = 438\nvalue = [414, 24]'),
 Text(0.2897058823529412, 0.5, 'x[11] <= 0.139\ngini = 0.084\nsamples =
365\nvalue = [349, 16]'),
 Text(0.25, 0.4473684210526316, 'x[12] <= 0.367\ngini = 0.153\nsamples =
156\nvalue = [143, 13]'),
 Text(0.22941176470588234, 0.39473684210526316, 'x[9] <= 0.138\ngini =
0.135\nsamples = 151\nvalue = [140, 11]'),
 Text(0.20588235294117646, 0.34210526315789475, 'x[6] <= 0.076\ngini =
```

```
0.038\nsamples = 51\nvalue = [50, 1]'),
 Text(0.2, 0.2894736842105263, 'x[6] <= 0.072\ngini = 0.219\nsamples = 8\nvalue
= [7, 1]'),
 Text(0.19411764705882353, 0.23684210526315788, 'gini = 0.0\nsamples = 7\nvalue
= [7, 0]'),
 Text(0.20588235294117646, 0.23684210526315788, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.21176470588235294, 0.2894736842105263, 'gini = 0.0\nsamples = 43\nvalue
= [43, 0]'),
 Text(0.2529411764705882, 0.34210526315789475, 'x[9] <= 0.188\ngini =
0.18\nsamples = 100\nvalue = [90, 10]'),
 Text(0.22941176470588234, 0.2894736842105263, 'x[6] <= 0.152\ngini =
0.334\nsamples = 33\nvalue = [26, 7]'),
 Text(0.21764705882352942, 0.23684210526315788, 'x[4] <= 0.188\ngini =
0.105\nsamples = 18\nvalue = [17, 1]'),
 Text(0.21176470588235294, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.2235294117647059, 0.18421052631578946, 'gini = 0.0\nsamples = 17\nvalue
= [17, 0]'),
 Text(0.2411764705882353, 0.23684210526315788, 'x[6] <= 0.232\ngini =
0.48\nsamples = 15\nvalue = [9, 6]'),
 Text(0.23529411764705882, 0.18421052631578946, 'x[5] <= 0.25\ngini =
0.375\nsamples = 8\nvalue = [2, 6]'),
 Text(0.22941176470588234, 0.13157894736842105, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.2411764705882353, 0.13157894736842105, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.24705882352941178, 0.18421052631578946, 'gini = 0.0\nsamples = 7\nvalue
= [7, 0]'),
 Text(0.27647058823529413, 0.2894736842105263, 'x[6] <= 0.108\ngini =
0.086\nsamples = 67\nvalue = [64, 3]'),
 Text(0.2647058823529412, 0.23684210526315788, 'x[6] <= 0.099\ngini =
0.18\nsamples = 20\nvalue = [18, 2]'),
 Text(0.25882352941176473, 0.18421052631578946, 'x[9] <= 0.237\ngini =
0.1\nsamples = 19\nvalue = [18, 1]'),
 Text(0.2529411764705882, 0.13157894736842105, 'gini = 0.0\nsamples = 12\nvalue
= [12, 0]'),
 Text(0.2647058823529412, 0.13157894736842105, 'x[6] <= 0.069\ngini =
0.245\nsamples = 7\nvalue = [6, 1]'),
 Text(0.25882352941176473, 0.07894736842105263, 'gini = 0.0\nsamples = 4\nvalue
= [4, 0]'),
 Text(0.27058823529411763, 0.07894736842105263, 'x[6] <= 0.074\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.2647058823529412, 0.02631578947368421, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.27647058823529413, 0.02631578947368421, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
```

```
 Text(0.27058823529411763, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.28823529411764703, 0.23684210526315788, 'x[13] <= 0.464\ngini =
0.042\nsamples = 47\nvalue = [46, 1]'),
 Text(0.2823529411764706, 0.18421052631578946, 'gini = 0.0\nsamples = 41\nvalue
= [41, 0]'),
 Text(0.29411764705882354, 0.18421052631578946, 'x[12] <= 0.033\ngini =
0.278\nsamples = 6\nvalue = [5, 1]'),
 Text(0.28823529411764703, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.3, 0.13157894736842105, 'gini = 0.0\nsamples = 5\nvalue = [5, 0]'),
 Text(0.27058823529411763, 0.39473684210526316, 'x[13] <= 0.536\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.2647058823529412, 0.34210526315789475, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.27647058823529413, 0.34210526315789475, 'gini = 0.0\nsamples = 3\nvalue
= [3, 0]'),
 Text(0.32941176470588235, 0.4473684210526316, 'x[2] <= 0.5\ngini =
0.028\nsamples = 209\nvalue = [206, 3]'),
 Text(0.3235294117647059, 0.39473684210526316, 'gini = 0.0\nsamples = 126\nvalue
= [126, 0]'),
 Text(0.3352941176470588, 0.39473684210526316, 'x[6] <= 0.191\ngini =
0.07\nsamples = 83\nvalue = [80, 3]'),
 Text(0.32941176470588235, 0.34210526315789475, 'x[6] <= 0.183\ngini =
0.105\nsamples = 54\nvalue = [51, 3]'),
 Text(0.3235294117647059, 0.2894736842105263, 'x[10] <= 0.113\ngini =
0.073\nsamples = 53\nvalue = [51, 2]'),
 Text(0.31176470588235294, 0.23684210526315788, 'x[9] <= 0.175\ngini =
0.245\nsamples = 7\nvalue = [6, 1]'),
 Text(0.3058823529411765, 0.18421052631578946, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.3176470588235294, 0.18421052631578946, 'x[6] <= 0.115\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.31176470588235294, 0.13157894736842105, 'gini = 0.0\nsamples = 2\nvalue
= [2, 0]'),
 Text(0.3235294117647059, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3352941176470588, 0.23684210526315788, 'x[6] <= 0.164\ngini =
0.043\nsamples = 46\nvalue = [45, 1]'),
 Text(0.32941176470588235, 0.18421052631578946, 'gini = 0.0\nsamples = 36\nvalue
= [36, 0]'),
 Text(0.3411764705882353, 0.18421052631578946, 'x[6] <= 0.169\ngini =
0.18\nsamples = 10\nvalue = [9, 1]'),
 Text(0.3352941176470588, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.34705882352941175, 0.13157894736842105, 'gini = 0.0\nsamples = 9\nvalue
= [9, 0]'),
```

```
 Text(0.3352941176470588, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.3411764705882353, 0.34210526315789475, 'gini = 0.0\nsamples = 29\nvalue
= [29, 0]'),
 Text(0.38235294117647056, 0.5, 'x[6] <= 0.335\ngini = 0.195\nsamples =
73\nvalue = [65, 8]'),
 Text(0.3764705882352941, 0.4473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.38823529411764707, 0.4473684210526316, 'x[6] <= 0.367\ngini =
0.155\nsamples = 71\nvalue = [65, 6]'),
 Text(0.3764705882352941, 0.39473684210526316, 'x[6] <= 0.364\ngini =
0.32\nsamples = 20\nvalue = [16, 4]'),
 Text(0.37058823529411766, 0.34210526315789475, 'x[9] <= 0.275\ngini =
0.266\nsamples = 19\nvalue = [16, 3]'),
 Text(0.36470588235294116, 0.2894736842105263, 'x[12] <= 0.233\ngini =
0.42\nsamples = 10\nvalue = [7, 3]'),
 Text(0.3588235294117647, 0.23684210526315788, 'x[10] <= 0.237\ngini =
0.219\nsamples = 8\nvalue = [7, 1]'),
 Text(0.35294117647058826, 0.18421052631578946, 'gini = 0.0\nsamples = 6\nvalue
= [6, 0]'),
 Text(0.36470588235294116, 0.18421052631578946, 'x[3] <= 0.5\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.3588235294117647, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.37058823529411766, 0.13157894736842105, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.37058823529411766, 0.23684210526315788, 'gini = 0.0\nsamples = 2\nvalue
= [0, 2]'),
 Text(0.3764705882352941, 0.2894736842105263, 'gini = 0.0\nsamples = 9\nvalue =
[9, 0]'),
 Text(0.38235294117647056, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.4, 0.39473684210526316, 'x[6] <= 0.453\ngini = 0.075\nsamples =
51\nvalue = [49, 2]'),
 Text(0.3941176470588235, 0.34210526315789475, 'gini = 0.0\nsamples = 36\nvalue
= [36, 0]'),
 Text(0.40588235294117647, 0.34210526315789475, 'x[6] <= 0.454\ngini =
0.231\nsamples = 15\nvalue = [13, 2]'),
 Text(0.4, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.4117647058823529, 0.2894736842105263, 'x[12] <= 0.267\ngini =
0.133\nsamples = 14\nvalue = [13, 1]'),
 Text(0.40588235294117647, 0.23684210526315788, 'gini = 0.0\nsamples = 11\nvalue
= [11, 0]'),
 Text(0.4176470588235294, 0.23684210526315788, 'x[6] <= 0.464\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.4117647058823529, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
```

```
 Text(0.4235294117647059, 0.18421052631578946, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.43529411764705883, 0.5526315789473685, 'x[6] <= 0.073\ngini =
0.26\nsamples = 52\nvalue = [44, 8]'),
 Text(0.4235294117647059, 0.5, 'x[13] <= 0.321\ngini = 0.5\nsamples = 8\nvalue =
[4, 4]'),
 Text(0.4176470588235294, 0.4473684210526316, 'x[6] <= 0.066\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.4117647058823529, 0.39473684210526316, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.4235294117647059, 0.39473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4294117647058823, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.4470588235294118, 0.5, 'x[5] <= 0.75\ngini = 0.165\nsamples = 44\nvalue
= [40, 4]'),
 Text(0.4411764705882353, 0.4473684210526316, 'gini = 0.0\nsamples = 28\nvalue =
[28, 0]'),
 Text(0.45294117647058824, 0.4473684210526316, 'x[10] <= 0.113\ngini =
0.375\nsamples = 16\nvalue = [12, 4]'),
 Text(0.4411764705882353, 0.39473684210526316, 'x[6] <= 0.114\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.43529411764705883, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue
= [1, 0]'),
 Text(0.4470588235294118, 0.34210526315789475, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.4647058823529412, 0.39473684210526316, 'x[6] <= 0.39\ngini =
0.26\nsamples = 13\nvalue = [11, 2]'),
 Text(0.4588235294117647, 0.34210526315789475, 'x[6] <= 0.121\ngini =
0.153\nsamples = 12\nvalue = [11, 1]'),
 Text(0.45294117647058824, 0.2894736842105263, 'x[11] <= 0.306\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.4470588235294118, 0.23684210526315788, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.4588235294117647, 0.23684210526315788, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.4647058823529412, 0.2894736842105263, 'gini = 0.0\nsamples = 10\nvalue =
[10, 0]'),
 Text(0.47058823529411764, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue
= [0, 1]'),
 Text(0.47058823529411764, 0.6578947368421053, 'x[3] <= 0.5\ngini =
0.434\nsamples = 22\nvalue = [15, 7]'),
 Text(0.4647058823529412, 0.6052631578947368, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.4764705882352941, 0.6052631578947368, 'x[6] <= 0.482\ngini =
0.484\nsamples = 17\nvalue = [10, 7]'),
 Text(0.47058823529411764, 0.5526315789473685, 'gini = 0.0\nsamples = 2\nvalue =
```

```
[0, 2]'),
 Text(0.4823529411764706, 0.5526315789473685, 'x[10] <= 0.3\ngini =
0.444\nsamples = 15\nvalue = [10, 5]'),
 Text(0.47058823529411764, 0.5, 'x[6] <= 0.506\ngini = 0.198\nsamples = 9\nvalue
= [8, 1]'),
 Text(0.4647058823529412, 0.4473684210526316, 'gini = 0.0\nsamples = 8\nvalue =
[8, 0]'),
 Text(0.4764705882352941, 0.4473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.49411764705882355, 0.5, 'x[5] <= 0.25\ngini = 0.444\nsamples = 6\nvalue
= [2, 4]'),
 Text(0.48823529411764705, 0.4473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.5, 0.4473684210526316, 'gini = 0.0\nsamples = 4\nvalue = [0, 4]'),
 Text(0.5176470588235295, 0.7105263157894737, 'x[6] <= 0.992\ngini =
0.03\nsamples = 131\nvalue = [129, 2]'),
 Text(0.5058823529411764, 0.6578947368421053, 'x[4] <= 0.062\ngini =
0.016\nsamples = 126\nvalue = [125, 1]'),
 Text(0.5, 0.6052631578947368, 'x[5] <= 0.25\ngini = 0.278\nsamples = 6\nvalue =
[5, 1]'),
 Text(0.49411764705882355, 0.5526315789473685, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5058823529411764, 0.5526315789473685, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.5117647058823529, 0.6052631578947368, 'gini = 0.0\nsamples = 120\nvalue
= [120, 0]'),
 Text(0.5294117647058824, 0.6578947368421053, 'x[6] <= 0.994\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.5235294117647059, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5352941176470588, 0.6052631578947368, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.5529411764705883, 0.7631578947368421, 'x[1] <= 0.75\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.5470588235294118, 0.7105263157894737, 'x[9] <= 0.637\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.5411764705882353, 0.6578947368421053, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.5529411764705883, 0.6578947368421053, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5588235294117647, 0.7105263157894737, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.5823529411764706, 0.8157894736842105, 'x[11] <= 0.194\ngini =
0.403\nsamples = 25\nvalue = [18, 7]'),
 Text(0.5764705882352941, 0.7631578947368421, 'x[5] <= 0.25\ngini =
0.475\nsamples = 18\nvalue = [11, 7]'),
 Text(0.5705882352941176, 0.7105263157894737, 'gini = 0.0\nsamples = 2\nvalue =
```

```
[0, 2]'),
 Text(0.5823529411764706, 0.7105263157894737, 'x[9] <= 0.2\ngini = 0.43\nsamples
= 16\nvalue = [11, 5]'),
 Text(0.5764705882352941, 0.6578947368421053, 'x[6] <= 0.165\ngini =
0.337\nsamples = 14\nvalue = [11, 3]'),
 Text(0.5705882352941176, 0.6052631578947368, 'x[0] <= 0.75\ngini =
0.26\nsamples = 13\nvalue = [11, 2]'),
 Text(0.5647058823529412, 0.5526315789473685, 'x[2] <= 0.8\ngini =
0.444\nsamples = 6\nvalue = [4, 2]'),
 Text(0.5588235294117647, 0.5, 'x[12] <= 0.033\ngini = 0.32\nsamples = 5\nvalue
= [4, 1]'),
 Text(0.5529411764705883, 0.4473684210526316, 'x[9] <= 0.1\ngini = 0.5\nsamples
= 2\nvalue = [1, 1]'),
 Text(0.5470588235294118, 0.39473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5588235294117647, 0.39473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5647058823529412, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.5705882352941176, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.5764705882352941, 0.5526315789473685, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.5823529411764706, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.5882352941176471, 0.6578947368421053, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.5882352941176471, 0.7631578947368421, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.7616268382352941, 0.9210526315789473, 'x[6] <= 0.144\ngini =
0.435\nsamples = 332\nvalue = [226, 106]'),
 Text(0.6647058823529411, 0.868421052631579, 'x[5] <= 0.75\ngini =
0.497\nsamples = 119\nvalue = [55, 64]'),
 Text(0.6411764705882353, 0.8157894736842105, 'x[6] <= 0.08\ngini =
0.491\nsamples = 76\nvalue = [43, 33]'),
 Text(0.6176470588235294, 0.7631578947368421, 'x[9] <= 0.212\ngini =
0.475\nsamples = 31\nvalue = [12, 19]'),
 Text(0.6058823529411764, 0.7105263157894737, 'x[6] <= 0.071\ngini =
0.403\nsamples = 25\nvalue = [7, 18]'),
 Text(0.6, 0.6578947368421053, 'x[12] <= 0.1\ngini = 0.465\nsamples = 19\nvalue
= [7, 12]'),
 Text(0.5941176470588235, 0.6052631578947368, 'x[6] <= 0.069\ngini =
0.492\nsamples = 16\nvalue = [7, 9]'),
 Text(0.5882352941176471, 0.5526315789473685, 'x[6] <= 0.019\ngini =
0.459\nsamples = 14\nvalue = [5, 9]'),
 Text(0.5823529411764706, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.5941176470588235, 0.5, 'x[4] <= 0.5\ngini = 0.426\nsamples = 13\nvalue =
[4, 9]'),
```

```
 Text(0.5823529411764706, 0.4473684210526316, 'x[0] <= 0.75\ngini =
0.245\nsamples = 7\nvalue = [1, 6]'),
 Text(0.5764705882352941, 0.39473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.5882352941176471, 0.39473684210526316, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.6058823529411764, 0.4473684210526316, 'x[13] <= 0.179\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.6, 0.39473684210526316, 'x[10] <= 0.05\ngini = 0.375\nsamples = 4\nvalue
= [1, 3]'),
 Text(0.5941176470588235, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6058823529411764, 0.34210526315789475, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.611764705882353, 0.39473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6, 0.5526315789473685, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.6058823529411764, 0.6052631578947368, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.611764705882353, 0.6578947368421053, 'gini = 0.0\nsamples = 6\nvalue =
[0, 6]'),
 Text(0.6294117647058823, 0.7105263157894737, 'x[12] <= 0.433\ngini =
0.278\nsamples = 6\nvalue = [5, 1]'),
 Text(0.6235294117647059, 0.6578947368421053, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.6352941176470588, 0.6578947368421053, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6647058823529411, 0.7631578947368421, 'x[6] <= 0.14\ngini =
0.429\nsamples = 45\nvalue = [31, 14]'),
 Text(0.6588235294117647, 0.7105263157894737, 'x[9] <= 0.45\ngini =
0.387\nsamples = 42\nvalue = [31, 11]'),
 Text(0.6529411764705882, 0.6578947368421053, 'x[13] <= 0.179\ngini =
0.349\nsamples = 40\nvalue = [31, 9]'),
 Text(0.6470588235294118, 0.6052631578947368, 'x[4] <= 0.875\ngini =
0.444\nsamples = 27\nvalue = [18, 9]'),
 Text(0.6411764705882353, 0.5526315789473685, 'x[11] <= 0.139\ngini =
0.403\nsamples = 25\nvalue = [18, 7]'),
 Text(0.6294117647058823, 0.5, 'x[10] <= 0.013\ngini = 0.32\nsamples = 20\nvalue
= [16, 4]'),
 Text(0.6235294117647059, 0.4473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6352941176470588, 0.4473684210526316, 'x[6] <= 0.124\ngini =
0.266\nsamples = 19\nvalue = [16, 3]'),
 Text(0.6235294117647059, 0.39473684210526316, 'x[0] <= 0.75\ngini =
0.133\nsamples = 14\nvalue = [13, 1]'),
 Text(0.6176470588235294, 0.34210526315789475, 'x[6] <= 0.097\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
```

```
 Text(0.611764705882353, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6235294117647059, 0.2894736842105263, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6294117647058823, 0.34210526315789475, 'gini = 0.0\nsamples = 11\nvalue
= [11, 0]'),
 Text(0.6470588235294118, 0.39473684210526316, 'x[2] <= 0.4\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.6411764705882353, 0.34210526315789475, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6529411764705882, 0.34210526315789475, 'x[6] <= 0.133\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.6470588235294118, 0.2894736842105263, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.6588235294117647, 0.2894736842105263, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6529411764705882, 0.5, 'x[4] <= 0.5\ngini = 0.48\nsamples = 5\nvalue =
[2, 3]'),
 Text(0.6470588235294118, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.6588235294117647, 0.4473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.6529411764705882, 0.5526315789473685, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.6588235294117647, 0.6052631578947368, 'gini = 0.0\nsamples = 13\nvalue =
[13, 0]'),
 Text(0.6647058823529411, 0.6578947368421053, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.6705882352941176, 0.7105263157894737, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.6882352941176471, 0.8157894736842105, 'x[13] <= 0.036\ngini =
0.402\nsamples = 43\nvalue = [12, 31]'),
 Text(0.6823529411764706, 0.7631578947368421, 'gini = 0.0\nsamples = 13\nvalue =
[0, 13]'),
 Text(0.6941176470588235, 0.7631578947368421, 'x[6] <= 0.07\ngini =
0.48\nsamples = 30\nvalue = [12, 18]'),
 Text(0.6823529411764706, 0.7105263157894737, 'x[4] <= 0.5\ngini =
0.198\nsamples = 9\nvalue = [1, 8]'),
 Text(0.6764705882352942, 0.6578947368421053, 'x[9] <= 0.188\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.6705882352941176, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6823529411764706, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.6882352941176471, 0.6578947368421053, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.7058823529411765, 0.7105263157894737, 'x[13] <= 0.25\ngini =
```

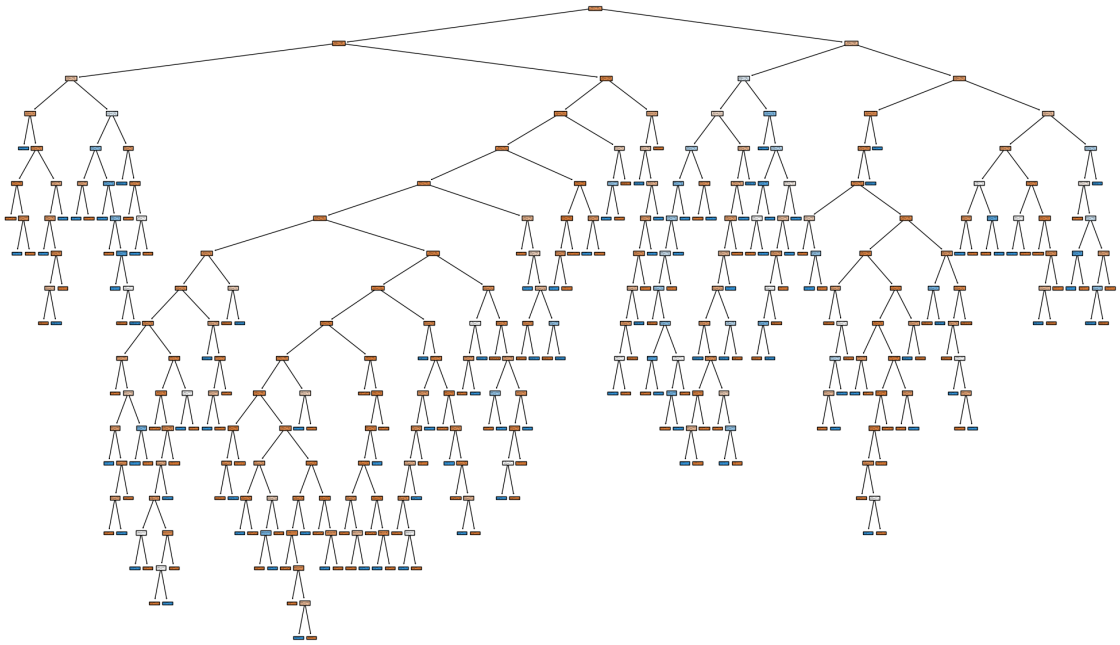
```
0.499\nsamples = 21\nvalue = [11, 10]'),
 Text(0.7, 0.6578947368421053, 'x[6] <= 0.108\ngini = 0.457\nsamples = 17\nvalue
= [11, 6]'),
 Text(0.6941176470588235, 0.6052631578947368, 'x[12] <= 0.033\ngini =
0.337\nsamples = 14\nvalue = [11, 3]'),
 Text(0.6882352941176471, 0.5526315789473685, 'x[6] <= 0.094\ngini =
0.5\nsamples = 6\nvalue = [3, 3]'),
 Text(0.6823529411764706, 0.5, 'x[6] <= 0.072\ngini = 0.375\nsamples = 4\nvalue
= [1, 3]'),
 Text(0.6764705882352942, 0.4473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.6882352941176471, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.6941176470588235, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [2, 0]'),
 Text(0.7, 0.5526315789473685, 'gini = 0.0\nsamples = 8\nvalue = [8, 0]'),
 Text(0.7058823529411765, 0.6052631578947368, 'gini = 0.0\nsamples = 3\nvalue =
[0, 3]'),
 Text(0.711764705882353, 0.6578947368421053, 'gini = 0.0\nsamples = 4\nvalue =
[0, 4]'),
 Text(0.858547794117647, 0.868421052631579, 'x[4] <= 0.812\ngini =
0.317\nsamples = 213\nvalue = [171, 42]'),
 Text(0.778860294117647, 0.8157894736842105, 'x[6] <= 0.988\ngini =
0.212\nsamples = 141\nvalue = [124, 17]'),
 Text(0.7729779411764706, 0.7631578947368421, 'x[9] <= 0.963\ngini =
0.202\nsamples = 140\nvalue = [124, 16]'),
 Text(0.7670955882352941, 0.7105263157894737, 'x[9] <= 0.125\ngini =
0.193\nsamples = 139\nvalue = [124, 15]'),
 Text(0.7235294117647059, 0.6578947368421053, 'x[6] <= 0.163\ngini =
0.48\nsamples = 5\nvalue = [3, 2]'),
 Text(0.7176470588235294, 0.6052631578947368, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.7294117647058823, 0.6052631578947368, 'x[2] <= 0.7\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.7235294117647059, 0.5526315789473685, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.7352941176470589, 0.5526315789473685, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8106617647058824, 0.6578947368421053, 'x[12] <= 0.3\ngini =
0.175\nsamples = 134\nvalue = [121, 13]'),
 Text(0.774264705882353, 0.6052631578947368, 'x[10] <= 0.038\ngini =
0.127\nsamples = 103\nvalue = [96, 7]'),
 Text(0.7470588235294118, 0.5526315789473685, 'x[3] <= 0.5\ngini =
0.375\nsamples = 12\nvalue = [9, 3]'),
 Text(0.7411764705882353, 0.5, 'gini = 0.0\nsamples = 6\nvalue = [6, 0]'),
 Text(0.7529411764705882, 0.5, 'x[4] <= 0.562\ngini = 0.5\nsamples = 6\nvalue =
[3, 3]'),
 Text(0.7470588235294118, 0.4473684210526316, 'x[6] <= 0.256\ngini =
```

```
0.48\nsamples = 5\nvalue = [2, 3]'),
 Text(0.7411764705882353, 0.39473684210526316, 'x[9] <= 0.512\ngini =
0.444\nsamples = 3\nvalue = [2, 1]'),
 Text(0.7352941176470589, 0.34210526315789475, 'gini = 0.0\nsamples = 2\nvalue =
[2, 0]'),
 Text(0.7470588235294118, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7529411764705882, 0.39473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.7588235294117647, 0.4473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8014705882352942, 0.5526315789473685, 'x[13] <= 0.75\ngini =
0.084\nsamples = 91\nvalue = [87, 4]'),
 Text(0.7852941176470588, 0.5, 'x[2] <= 0.1\ngini = 0.067\nsamples = 87\nvalue =
[84, 3]'),
 Text(0.7705882352941177, 0.4473684210526316, 'x[5] <= 0.25\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.7647058823529411, 0.39473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7764705882352941, 0.39473684210526316, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.8, 0.4473684210526316, 'x[12] <= 0.233\ngini = 0.048\nsamples =
82\nvalue = [80, 2]'),
 Text(0.788235294117647, 0.39473684210526316, 'x[4] <= 0.312\ngini =
0.026\nsamples = 77\nvalue = [76, 1]'),
 Text(0.7823529411764706, 0.34210526315789475, 'x[0] <= 0.75\ngini =
0.087\nsamples = 22\nvalue = [21, 1]'),
 Text(0.7764705882352941, 0.2894736842105263, 'x[13] <= 0.286\ngini =
0.245\nsamples = 7\nvalue = [6, 1]'),
 Text(0.7705882352941177, 0.23684210526315788, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.7823529411764706, 0.23684210526315788, 'x[6] <= 0.286\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.7764705882352941, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.788235294117647, 0.18421052631578946, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.788235294117647, 0.2894736842105263, 'gini = 0.0\nsamples = 15\nvalue =
[15, 0]'),
 Text(0.7941176470588235, 0.34210526315789475, 'gini = 0.0\nsamples = 55\nvalue
= [55, 0]'),
 Text(0.8117647058823529, 0.39473684210526316, 'x[4] <= 0.625\ngini =
0.32\nsamples = 5\nvalue = [4, 1]'),
 Text(0.8058823529411765, 0.34210526315789475, 'gini = 0.0\nsamples = 4\nvalue =
[4, 0]'),
 Text(0.8176470588235294, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
```

```
 Text(0.8176470588235294, 0.5, 'x[12] <= 0.167\ngini = 0.375\nsamples = 4\nvalue
= [3, 1]'),
 Text(0.8117647058823529, 0.4473684210526316, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8235294117647058, 0.4473684210526316, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.8470588235294118, 0.6052631578947368, 'x[11] <= 0.361\ngini =
0.312\nsamples = 31\nvalue = [25, 6]'),
 Text(0.8352941176470589, 0.5526315789473685, 'x[11] <= 0.083\ngini =
0.375\nsamples = 4\nvalue = [1, 3]'),
 Text(0.8294117647058824, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.8411764705882353, 0.5, 'gini = 0.0\nsamples = 3\nvalue = [0, 3]'),
 Text(0.8588235294117647, 0.5526315789473685, 'x[10] <= 0.275\ngini =
0.198\nsamples = 27\nvalue = [24, 3]'),
 Text(0.8529411764705882, 0.5, 'x[6] <= 0.316\ngini = 0.375\nsamples = 12\nvalue
= [9, 3]'),
 Text(0.8470588235294118, 0.4473684210526316, 'gini = 0.0\nsamples = 6\nvalue =
[6, 0]'),
 Text(0.8588235294117647, 0.4473684210526316, 'x[0] <= 0.75\ngini = 0.5\nsamples
= 6\nvalue = [3, 3]'),
 Text(0.8529411764705882, 0.39473684210526316, 'gini = 0.0\nsamples = 2\nvalue =
[0, 2]'),
 Text(0.8647058823529412, 0.39473684210526316, 'x[9] <= 0.737\ngini =
0.375\nsamples = 4\nvalue = [3, 1]'),
 Text(0.8588235294117647, 0.34210526315789475, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.8705882352941177, 0.34210526315789475, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8647058823529412, 0.5, 'gini = 0.0\nsamples = 15\nvalue = [15, 0]'),
 Text(0.778860294117647, 0.7105263157894737, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.7847426470588236, 0.7631578947368421, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9382352941176471, 0.8157894736842105, 'x[5] <= 0.75\ngini =
0.453\nsamples = 72\nvalue = [47, 25]'),
 Text(0.9, 0.7631578947368421, 'x[10] <= 0.113\ngini = 0.315\nsamples =
46\nvalue = [37, 9]'),
 Text(0.8764705882352941, 0.7105263157894737, 'x[12] <= 0.1\ngini = 0.5\nsamples
= 12\nvalue = [6, 6]'),
 Text(0.8647058823529412, 0.6578947368421053, 'x[0] <= 0.75\ngini =
0.278\nsamples = 6\nvalue = [5, 1]'),
 Text(0.8588235294117647, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.8705882352941177, 0.6052631578947368, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.888235294117647, 0.6578947368421053, 'x[0] <= 0.25\ngini =
0.278\nsamples = 6\nvalue = [1, 5]'),
```

```
 Text(0.8823529411764706, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.8941176470588236, 0.6052631578947368, 'gini = 0.0\nsamples = 5\nvalue =
[0, 5]'),
 Text(0.9235294117647059, 0.7105263157894737, 'x[13] <= 0.071\ngini =
0.161\nsamples = 34\nvalue = [31, 3]'),
 Text(0.9117647058823529, 0.6578947368421053, 'x[9] <= 0.212\ngini =
0.5\nsamples = 2\nvalue = [1, 1]'),
 Text(0.9058823529411765, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[0, 1]'),
 Text(0.9176470588235294, 0.6052631578947368, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9352941176470588, 0.6578947368421053, 'x[10] <= 0.237\ngini =
0.117\nsamples = 32\nvalue = [30, 2]'),
 Text(0.9294117647058824, 0.6052631578947368, 'gini = 0.0\nsamples = 19\nvalue =
[19, 0]'),
 Text(0.9411764705882353, 0.6052631578947368, 'x[13] <= 0.607\ngini =
0.26\nsamples = 13\nvalue = [11, 2]'),
 Text(0.9352941176470588, 0.5526315789473685, 'x[11] <= 0.417\ngini =
0.444\nsamples = 6\nvalue = [4, 2]'),
 Text(0.9294117647058824, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.9411764705882353, 0.5, 'gini = 0.0\nsamples = 4\nvalue = [4, 0]'),
 Text(0.9470588235294117, 0.5526315789473685, 'gini = 0.0\nsamples = 7\nvalue =
[7, 0]'),
 Text(0.9764705882352941, 0.7631578947368421, 'x[12] <= 0.1\ngini =
0.473\nsamples = 26\nvalue = [10, 16]'),
 Text(0.9705882352941176, 0.7105263157894737, 'x[6] <= 0.191\ngini =
0.499\nsamples = 19\nvalue = [10, 9]'),
 Text(0.9647058823529412, 0.6578947368421053, 'gini = 0.0\nsamples = 3\nvalue =
[3, 0]'),
 Text(0.9764705882352941, 0.6578947368421053, 'x[9] <= 0.237\ngini =
0.492\nsamples = 16\nvalue = [7, 9]'),
 Text(0.9647058823529412, 0.6052631578947368, 'x[2] <= 0.9\ngini =
0.219\nsamples = 8\nvalue = [1, 7]'),
 Text(0.9588235294117647, 0.5526315789473685, 'gini = 0.0\nsamples = 7\nvalue =
[0, 7]'),
 Text(0.9705882352941176, 0.5526315789473685, 'gini = 0.0\nsamples = 1\nvalue =
[1, 0]'),
 Text(0.9882352941176471, 0.6052631578947368, 'x[10] <= 0.038\ngini =
0.375\nsamples = 8\nvalue = [6, 2]'),
 Text(0.9823529411764705, 0.5526315789473685, 'x[3] <= 0.5\ngini =
0.444\nsamples = 3\nvalue = [1, 2]'),
 Text(0.9764705882352941, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [0, 2]'),
 Text(0.9882352941176471, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.9941176470588236, 0.5526315789473685, 'gini = 0.0\nsamples = 5\nvalue =
[5, 0]'),
 Text(0.9823529411764705, 0.7105263157894737, 'gini = 0.0\nsamples = 7\nvalue =
```

```
[0, 7]')]
```



[173]:
```
from sklearn.model_selection import GridSearchCV
parameter={
 'criterion':['gini','entropy'],
  'splitter':['best','random'],
  'max_depth':[1,2,3,4,5],
  'max_features':['auto', 'sqrt', 'log2']

}
```

[174]:
```
grid_search=GridSearchCV(estimator=dtc,param_grid=parameter,cv=5,scoring="accuracy")
```

[175]:
```
grid_search.fit(x_train,y_train)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
```

```
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
```

```
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
```

```
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
```

```
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
```

```
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
```

```
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
```

```
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
```

```
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
```

```
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/tree/_classes.py:269:
FutureWarning: `max_features='auto'` has been deprecated in 1.1 and will be
removed in 1.3. To keep the past behaviour, explicitly set
`max_features='sqrt'`.
  warnings.warn(
```

[175]:
```
GridSearchCV(cv=5, estimator=DecisionTreeClassifier(),
             param_grid={'criterion': ['gini', 'entropy'],
                         'max_depth': [1, 2, 3, 4, 5],
                         'max_features': ['auto', 'sqrt', 'log2'],
                         'splitter': ['best', 'random']},
             scoring='accuracy')
```

[176]:
```
grid_search.best_params_
```

[176]:
```
{'criterion': 'gini',
 'max_depth': 5,
 'max_features': 'sqrt',
 'splitter': 'random'}
```

[177]:
```
dtc_cv=DecisionTreeClassifier(criterion= 'entropy',
 max_depth=3,
 max_features='sqrt',
 splitter='best')
dtc_cv.fit(x_train,y_train)
```

[177]:
```
DecisionTreeClassifier(criterion='entropy', max_depth=3, max_features='sqrt')
```

[178]:
```
pred=dtc_cv.predict(x_test)
```

[179]:
```
print(classification_report(y_test,pred))
```

```
              precision    recall  f1-score   support

           0       0.88      0.92      0.90       249
           1       0.39      0.28      0.32        43

    accuracy                           0.83       292
   macro avg       0.63      0.60      0.61       292
weighted avg       0.81      0.83      0.82       292
```

Accuracy for this model is 83%

Random Forest

[180]:
```
from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier()
```

[181]:
```
forest_params = [{'max_depth': list(range(10, 15)), 'max_features':␣
 ↪list(range(0,14))}]
```

[182]:
```
rfc_cv= GridSearchCV(rfc,param_grid=forest_params,cv=10,scoring="accuracy")
```

[183]:
```
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 {'auto' (deprecated), 'log2', 'sqrt'}
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
0.8505084  0.85222517 0.85137784 0.84877689 0.85310934
 0.84019304 0.84791482 0.84448128 0.84189508 0.84703802 0.84704539
 0.84448128 0.84104775        nan 0.85307987 0.85051577 0.8479443
 0.85394194 0.85309461 0.84619805 0.8496537  0.84706012 0.84876953
```

```
0.84448128 0.84707486 0.84448865 0.84620542        nan 0.84706012
0.84963896 0.85223254 0.84793693 0.84878426 0.84620542 0.84533599
0.84359711 0.84190245 0.84534335 0.84190981 0.84705276 0.84189508
       nan 0.84707486 0.85137784 0.8479443  0.85051577 0.84448128
0.84963896 0.84790009 0.84446655 0.84446655 0.84619805 0.84535072
0.84103301 0.83930887        nan 0.84964633 0.8530725  0.85308724
0.8470896  0.84792956 0.84963159 0.84106248 0.84532862 0.8384468
0.84102564 0.84016357 0.84447392 0.84019304]
 warnings.warn(
```

[183]:
```
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')
```

[184]:
```
pred=rfc_cv.predict(x_test)
```

[191]:
```
confusion_matrix(y_test,pred)
```

[191]:
```
array([[241,   8],
       [ 32,  11]])
```

[192]:
```
#accuracy
(241+11)/292
```

[192]:
```
0.863013698630137
```

[185]:
```
print(classification_report(y_test,pred))
```

```
              precision    recall  f1-score   support

           0       0.88      0.97      0.92       249
           1       0.58      0.26      0.35        43

    accuracy                           0.86       292
   macro avg       0.73      0.61      0.64       292
weighted avg       0.84      0.86      0.84       292
```

Accuracy for this model is- 86%

[193]:
```
#precision = TP/(TP+FP)
11/(19)
```

[193]:
```
0.5789473684210527
```

[194]:
```
# Recall = TP/(FN+TP)
11/(32+11)
```

[194]: 0.2558139534883721

[195]:
```
#F1 score
# 2*precision*Recall/(Precision+Recall)
2*0.57*0.25/(0.57+0.25)
```

[195]: 0.3475609756097561

[186]:
```
rfc_cv.best_params_
```

[186]: {'max_depth': 11, 'max_features': 4}