About Dataset

This dataset provides a comprehensive view of students enrolled in various undergraduate degrees offered at a higher education institution. It includes demographic data, social-economic factors and academic performance information.

Link : https://www.kaggle.com/datasets/thedevastator/higher-education-predictors-of-student-retention?resource=download

Data Loading

```
import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import seaborn as sns


import sklearn
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import LabelEncoder
```

```
df = pd.read_csv('/content/studentdatasetclass.csv')
```

```
df.head()
```

	Marital status	Application mode	Application order	Course	Daytime/evening attendance	Previous qualification	Nacionality
0	1	8	5	2	1	1	1
1	1	6	1	11	1	1	1
2	1	1	5	5	1	1	1
3	1	8	2	15	1	1	1
4	2	12	1	3	0	1	1

5 rows × 35 columns

```
df.tail()
```

	Marital status	Application mode	Application order	Course	Daytime/evening attendance	Previous qualification	Nacionali
4419	1	1	6	15	1	1	
4420	1	1	2	15	1	1	
4421	1	1	1	12	1	1	
4422	1	1	1	9	1	1	
4423	1	5	1	15	1	1	

5 rows × 35 columns

```
df.describe()
```

	Marital status	Application mode	Application order	Course	Daytime/evening attendance	Previous qualification
count	4424.000000	4424.000000	4424.000000	4424.000000	4424.000000	4424.000000
mean	1.178571	6.886980	1.727848	9.899186	0.890823	2.531420
std	0.605747	5.298964	1.313793	4.331792	0.311897	3.963707
min	1.000000	1.000000	0.000000	1.000000	0.000000	1.000000
25%	1.000000	1.000000	1.000000	6.000000	1.000000	1.000000
50%	1.000000	8.000000	1.000000	10.000000	1.000000	1.000000
75%	1.000000	12.000000	2.000000	13.000000	1.000000	1.000000

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4424 entries, 0 to 4423
Data columns (total 35 columns):
 #   Column                                          Non-Null Count  Dtype
---  ------                                          --------------  -----
 0   Marital status                                  4424 non-null   int64
 1   Application mode                                4424 non-null   int64
 2   Application order                               4424 non-null   int64
 3   Course                                          4424 non-null   int64
 4   Daytime/evening attendance                      4424 non-null   int64
 5   Previous qualification                          4424 non-null   int64
 6   Nacionality                                     4424 non-null   int64
 7   Mother's qualification                          4424 non-null   int64
 8   Father's qualification                          4424 non-null   int64
 9   Mother's occupation                             4424 non-null   int64
 10  Father's occupation                             4424 non-null   int64
 11  Displaced                                       4424 non-null   int64
 12  Educational special needs                       4424 non-null   int64
 13  Debtor                                          4424 non-null   int64
 14  Tuition fees up to date                         4424 non-null   int64
 15  Gender                                          4424 non-null   int64
 16  Scholarship holder                              4424 non-null   int64
 17  Age at enrollment                               4424 non-null   int64
 18  International                                   4424 non-null   int64
 19  Curricular units 1st sem (credited)             4424 non-null   int64
 20  Curricular units 1st sem (enrolled)             4424 non-null   int64
 21  Curricular units 1st sem (evaluations)          4424 non-null   int64
 22  Curricular units 1st sem (approved)             4424 non-null   int64
 23  Curricular units 1st sem (grade)                4424 non-null   float64
 24  Curricular units 1st sem (without evaluations)  4424 non-null   int64
 25  Curricular units 2nd sem (credited)             4424 non-null   int64
 26  Curricular units 2nd sem (enrolled)             4424 non-null   int64
 27  Curricular units 2nd sem (evaluations)          4424 non-null   int64
 28  Curricular units 2nd sem (approved)             4424 non-null   int64
 29  Curricular units 2nd sem (grade)                4424 non-null   float64
 30  Curricular units 2nd sem (without evaluations)  4424 non-null   int64
 31  Unemployment rate                               4424 non-null   float64
 32  Inflation rate                                  4424 non-null   float64
 33  GDP                                             4424 non-null   float64
 34  Target                                          4424 non-null   int64
dtypes: float64(5), int64(30)
memory usage: 1.2 MB
```

```
df.columns
```

```
Index(['Marital status', 'Application mode', 'Application order', 'Course',
       'Daytime/evening attendance', 'Previous qualification', 'Nacionality',
       'Mother's qualification', 'Father's qualification',
       'Mother's occupation', 'Father's occupation', 'Displaced',
       'Educational special needs', 'Debtor', 'Tuition fees up to date',
       'Gender', 'Scholarship holder', 'Age at enrollment', 'International',
       'Curricular units 1st sem (credited)',
       'Curricular units 1st sem (enrolled)',
       'Curricular units 1st sem (evaluations)',
       'Curricular units 1st sem (approved)',
       'Curricular units 1st sem (grade)',
       'Curricular units 1st sem (without evaluations)',
       'Curricular units 2nd sem (credited)',
       'Curricular units 2nd sem (enrolled)',
       'Curricular units 2nd sem (evaluations)',
       'Curricular units 2nd sem (approved)',
       'Curricular units 2nd sem (grade)',
       'Curricular units 2nd sem (without evaluations)', 'Unemployment rate',
```

```
              'Inflation rate', 'GDP', 'Target'],
          dtype='object')
```

Types of Data Errors:

Missing Data: Occurs when certain data values are not recorded or unavailable.

Duplicate Data: Refers to the presence of identical or replicated data entries in the dataset.

Inconsistent Data: Arises when there are discrepancies or contradictions in data values across different sources or attributes.

Outliers: Represent extreme or abnormal data points that deviate significantly from the overall pattern or distribution of the data.

Missing Data

```
missing_values = df.isnull().sum()
print("Missing Values:\n", missing_values)
```

```
    Missing Values:
     Marital status                                   0
    Application mode                                  0
    Application order                                 0
    Course                                            0
    Daytime/evening attendance                        0
    Previous qualification                            0
    Nacionality                                       0
    Mother's qualification                            0
    Father's qualification                            0
    Mother's occupation                               0
    Father's occupation                               0
    Displaced                                         0
    Educational special needs                         0
    Debtor                                            0
    Tuition fees up to date                           0
    Gender                                            0
    Scholarship holder                                0
    Age at enrollment                                 0
    International                                      0
    Curricular units 1st sem (credited)               0
    Curricular units 1st sem (enrolled)               0
    Curricular units 1st sem (evaluations)            0
    Curricular units 1st sem (approved)               0
    Curricular units 1st sem (grade)                  0
    Curricular units 1st sem (without evaluations)    0
    Curricular units 2nd sem (credited)               0
    Curricular units 2nd sem (enrolled)               0
    Curricular units 2nd sem (evaluations)            0
    Curricular units 2nd sem (approved)               0
    Curricular units 2nd sem (grade)                  0
    Curricular units 2nd sem (without evaluations)    0
    Unemployment rate                                 0
    Inflation rate                                    0
    GDP                                               0
    Target                                            0
    dtype: int64
```

No Missing values exist in the student performance database

Duplicate Data

```
duplicate_records = df[df.duplicated()]
print("Duplicate Records:\n", duplicate_records)
```

```
    Duplicate Records:
     Empty DataFrame
    Columns: [Marital status, Application mode, Application order, Course, Daytime/evening attendance, Previous qualification, Nacionality,
    Index: []

    [0 rows x 35 columns]
```

No Duplicated values exist in the student performance database

## Inconsitent Data

```
data_types = df.dtypes
print("Data Types:\n", data_types)
```

```
Data Types:
 Marital status                                    int64
Application mode                                   int64
Application order                                  int64
Course                                             int64
Daytime/evening attendance                         int64
Previous qualification                             int64
Nacionality                                        int64
Mother's qualification                             int64
Father's qualification                             int64
Mother's occupation                                int64
Father's occupation                                int64
Displaced                                          int64
Educational special needs                          int64
Debtor                                             int64
Tuition fees up to date                            int64
Gender                                             int64
Scholarship holder                                 int64
Age at enrollment                                  int64
International                                       int64
Curricular units 1st sem (credited)                int64
Curricular units 1st sem (enrolled)                int64
Curricular units 1st sem (evaluations)             int64
Curricular units 1st sem (approved)                int64
Curricular units 1st sem (grade)                 float64
Curricular units 1st sem (without evaluations)     int64
Curricular units 2nd sem (credited)                int64
Curricular units 2nd sem (enrolled)                int64
Curricular units 2nd sem (evaluations)             int64
Curricular units 2nd sem (approved)                int64
Curricular units 2nd sem (grade)                 float64
Curricular units 2nd sem (without evaluations)     int64
Unemployment rate                                float64
Inflation rate                                   float64
GDP                                              float64
Target                                            object
dtype: object
```

Outlier

```
num_col = ['Inflation rate','Unemployment rate','GDP']
for i in num_col:
  sns.catplot(x= i, y ="Target", data =df, kind = "box", aspect = 2)
  plt.title("Boxplot for Target vs Population")
  plt.show()
```

Boxplot for Target vs Population (Inflation rate)

Boxplot for Target vs Population (Unemployment rate)


Boxplot for Target vs Population (GDP)
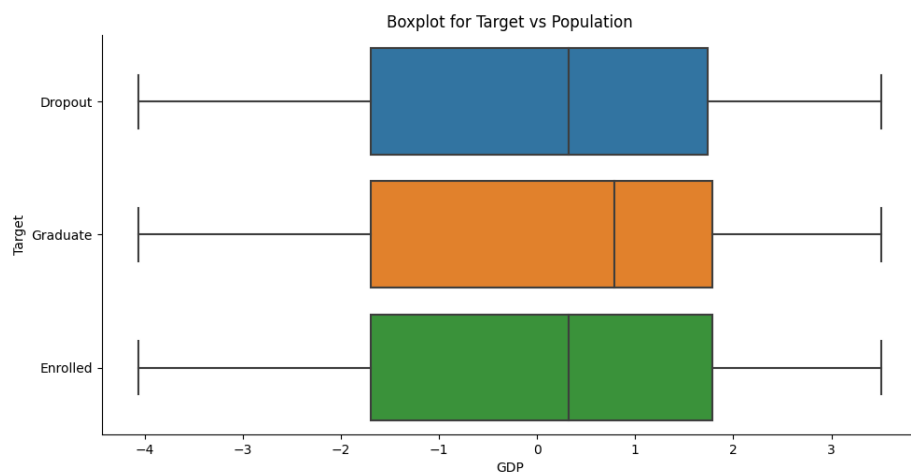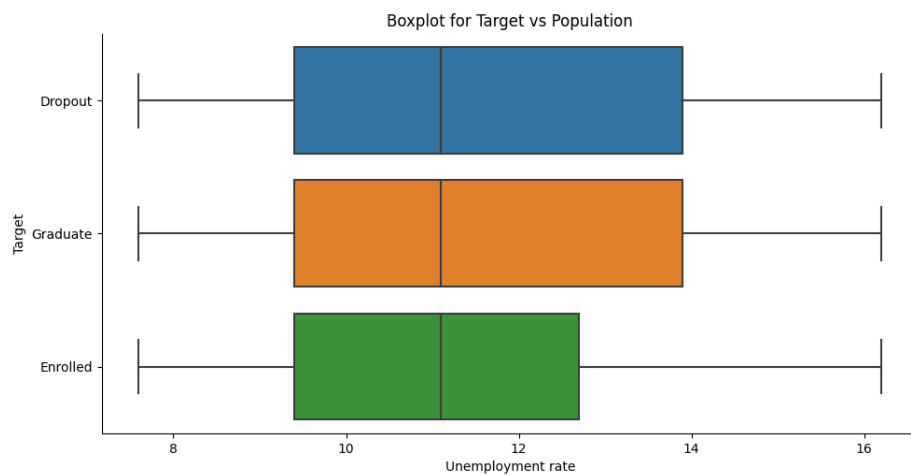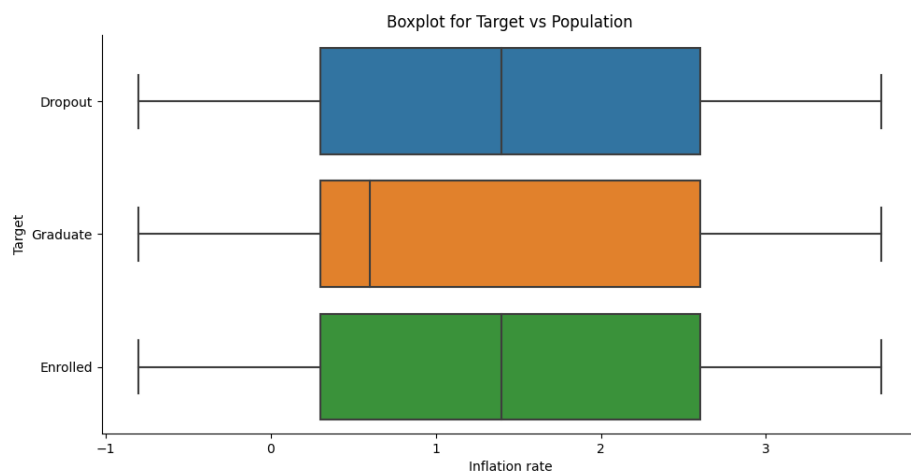
Visualization

```
plt.figure(figsize = (20,5))
sns.countplot(x='Inflation rate', data=df)
plt.xlabel('Inflation Rate')
plt.ylabel('Count')
plt.title('Distribution of Inflation Rate')
plt.show()
```

Distribution of Inflation Rate
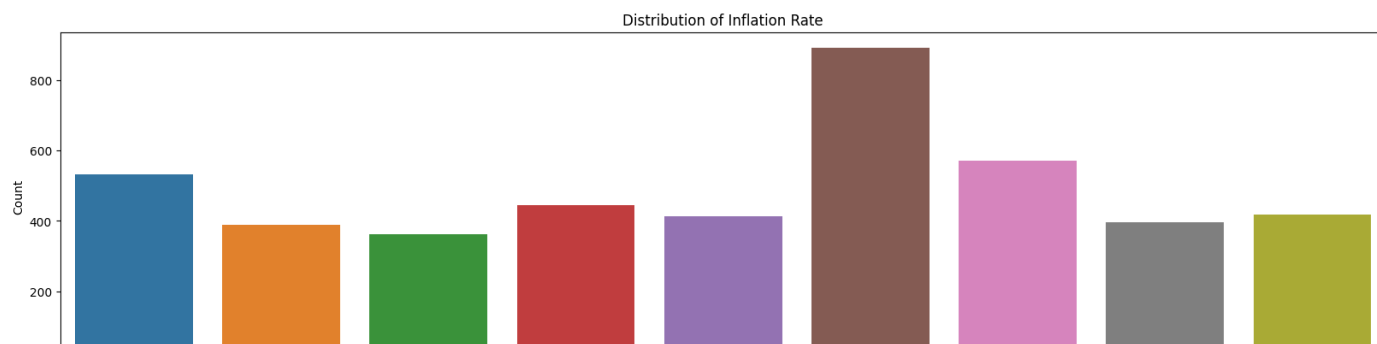

```
plt.figure(figsize = (20,5))
sns.countplot(x='Unemployment rate', data=df)
plt.xlabel('Unemployment rate')
plt.ylabel('Count')
plt.title('Distribution of Unemployment rate')
plt.show()
```

Distribution of Unemployment rate
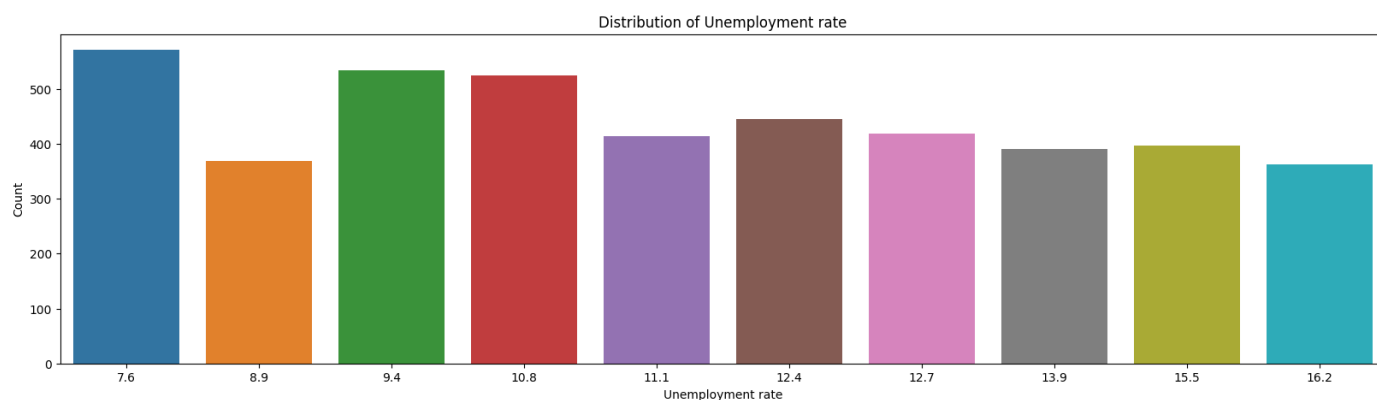

```
plt.figure(figsize = (20,5))
sns.countplot(x='GDP', data=df)
plt.xlabel('GDP')
plt.ylabel('Count')
plt.title('Distribution of GDP')
plt.show()
```

Distribution of GDP
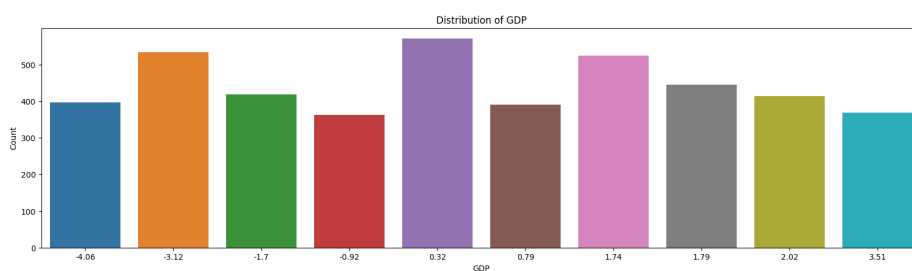

```
plt.figure(figsize = (20,5))
sns.countplot(x='Target', data=df)
plt.xlabel('Target')
plt.ylabel('Count')
plt.title('Distribution of Target')
plt.show()
```

Distribution of Target
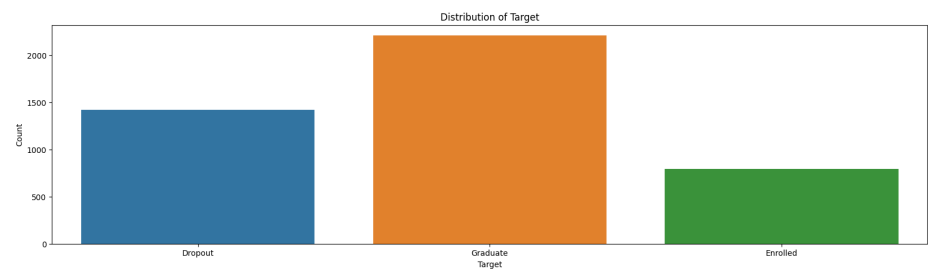
df.corr()

```
<ipython-input-16-2f6f6606aa2c>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version
  df.corr()
```

	Marital status	Application mode	Application order	Course	Daytime/evening attendance	Previous qualification	Nacionality	Mother's qualification	Father's qualification
Marital status	1.000000	0.224855	-0.125854	0.018925	-0.274939	0.120925	-0.020722	0.185522	0.128326
Application mode	0.224855	1.000000	-0.246497	-0.085116	-0.268616	0.433028	-0.001360	0.092867	0.072798
Application order	-0.125854	-0.246497	1.000000	0.118928	0.158657	-0.199029	-0.029385	-0.061719	-0.049936
Course	0.018925	-0.085116	0.118928	1.000000	-0.070232	-0.158382	-0.004761	0.058909	0.045659
Daytime/evening attendance	-0.274939	-0.268616	0.158657	-0.070232	1.000000	-0.103022	0.024433	-0.195346	-0.137769
Previous qualification	0.120925	0.433028	-0.199029	-0.158382	-0.103022	1.000000	-0.038997	0.018868	0.013152
Nacionality	-0.020722	-0.001360	-0.029385	-0.004761	0.024433	-0.038997	1.000000	-0.043847	-0.088892
Mother's qualification	0.185522	0.092867	-0.061719	0.058909	-0.195346	0.018868	-0.043847	1.000000	0.524529
Father's qualification	0.128326	0.072798	-0.049936	0.045659	-0.137769	0.013152	-0.088892	0.524529	1.000000
Mother's occupation	0.069734	0.033489	-0.046591	0.029672	-0.037986	0.006190	0.044123	0.295178	0.207067
Father's occupation	0.024351	0.001253	-0.029754	0.016489	0.000845	0.005381	0.024538	0.115989	0.184001
Displaced	-0.234886	-0.263079	0.332362	0.006142	0.251767	-0.149356	-0.010774	-0.075864	-0.055007
Educational special needs	-0.028343	-0.030868	0.025597	-0.001886	0.031017	-0.015015	-0.002399	-0.019808	0.000917
Debtor	0.034304	0.114348	-0.072151	-0.053149	0.006658	0.117447	0.070860	0.018776	-0.006125

```
correlation_matrix = df.corr()
plt.figure(figsize = (30,20))
sns.heatmap(correlation_matrix, annot=True, cmap= 'Pastel1')
plt.title('Correlation Matrix of Iris Dataset')
plt.show()
```

```
<ipython-input-17-f177a3e7da49>:1: FutureWarning: The default value of numeric_only in Dat
  correlation_matrix = df.corr()
```


Correlation Matrix of Iris Dataset

```
df['Target']
```

```
0          Dropout
1         Graduate
2          Dropout
3         Graduate
4         Graduate
           ...
4419      Graduate
4420       Dropout
4421       Dropout
4422      Graduate
4423      Graduate
Name: Target, Length: 4424, dtype: object
```

```
le = LabelEncoder()
df['Target'] = le.fit_transform(df['Target'])
df['Target']
```

```
0        0
1        2
2        0
3        2
4        2
        ..
4419     2
4420     0
4421     0
4422     2
4423     2
Name: Target, Length: 4424, dtype: int64
```

```
X = df.drop(columns= ['Target'])
Y = df['Target']
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.20)
```

MODEL

1. Logistic Regression | Accuracy : 77.40
2. Decision Tree | Accuracy : 72.88
3. Support Vector Machine | Accuracy : 74.6
4. Random Forest | Accuracy : 77.06

Logistic Regression

```
from sklearn.linear_model import LogisticRegression
model = LogisticRegression(verbose = 40)
```

```
model.fit(X_train, Y_train)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/linear_model/_logistic.py:458: Convergence
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
[Parallel(n_jobs=1)]: Done    1 tasks      | elapsed:    1.5s
[Parallel(n_jobs=1)]: Done    1 tasks      | elapsed:    1.5s
```

```
  ▾        LogisticRegression
LogisticRegression(verbose=40)
```

```python
print("Accuracy: ", model.score(X_test, Y_test) * 100)
```

```
Accuracy:  77.40112994350282
```

```python
y_pred = model.predict(X_test)
score = accuracy_score(Y_test, y_pred)
accuracy = score*100
print(accuracy)
```

```
77.40112994350282
```

```python
from sklearn.metrics import confusion_matrix

# PRINT THE CONFUSION MATRIX
print("Confusion Matrix")
cm = confusion_matrix(Y_test, y_pred)
print(cm)
```

```
Confusion Matrix
[[224  18  48]
 [ 40  54  62]
 [ 16  16 407]]
```

```python
plt.figure(figsize = (6, 4))
sns.heatmap(cm, annot = True, fmt = 'd', cmap = 'Blues', cbar = False, annot_kws = {'size' : 14})
plt.xlabel('Predicted Labels', fontsize = 14)
plt.ylabel('True Labels', fontsize = 14)
plt.title('Confusion Matrix', fontsize = 16)
plt.show()
```
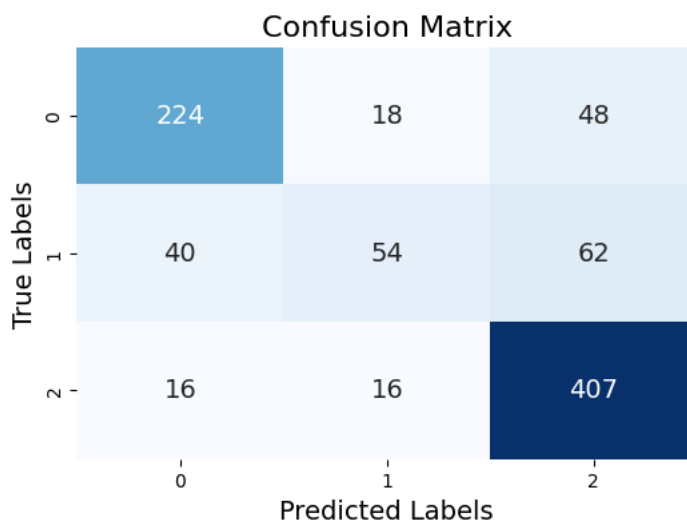

▾ Decision Tree

```
from sklearn import tree
clf = tree.DecisionTreeClassifier(random_state=42, max_depth=10, min_samples_split = 2)
clf = clf.fit(X_train, Y_train)


y_pred = clf.predict(X_test)
score = accuracy_score(Y_test, y_pred)
accuracy = score*100
print(accuracy)
```

```
    72.88135593220339
```

```
tree.plot_tree(clf)
```

⊐→

```
from sklearn import tree
clf = tree.DecisionTreeClassifier(random_state=42, max_depth=10, min_samples_split = 2)
clf = clf.fit(X_train, Y_train)
```

```
[Text(0.41114672364672367, 0.9545454545454546, 'x[28] <= 4.5\ngini = 0.615\nsamples = 3539\nvalue = [1131, 638, 1770]'),
 Text(0.1777065527065527, 0.8636363636363636, 'x[28] <= 1.5\ngini = 0.534\nsamples = 1494\nvalue = [932, 364, 198]'),
 Text(0.08404558404558404, 0.7727272727272727, 'x[20] <= 0.5\ngini = 0.309\nsamples = 778\nvalue = [639, 79, 60]'),
 Text(0.038461538461538464, 0.6818181818181818, 'x[14] <= 0.5\ngini = 0.616\nsamples = 141\nvalue = [60, 21, 60]'),
 Text(0.009971509971509971, 0.5909090909090909, 'x[32] <= 3.25\ngini = 0.204\nsamples = 18\nvalue = [16, 1, 1]'),
 Text(0.007122507122507123, 0.5, 'x[8] <= 27.5\ngini = 0.111\nsamples = 17\nvalue = [16, 1, 0]'),
 Text(0.004273504273504274, 0.4090909090909091, 'gini = 0.0\nsamples = 14\nvalue = [14, 0, 0]'),
 Text(0.009971509971509971, 0.4090909090909091, 'x[33] <= -2.02\ngini = 0.444\nsamples = 3\nvalue = [2, 1, 0]'),
 Text(0.007122507122507123, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
 Text(0.01282051282051282, 0.3181818181818182, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
 Text(0.01282051282051282, 0.5, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
 Text(0.06695156695156695, 0.5909090909090909, 'x[17] <= 25.5\ngini = 0.616\nsamples = 123\nvalue = [44, 20, 59]'),
 Text(0.04843304843304843, 0.5, 'x[31] <= 12.55\ngini = 0.612\nsamples = 115\nvalue = [37, 20, 58]'),
 Text(0.0292022792022792, 0.4090909090909091, 'x[2] <= 3.5\ngini = 0.557\nsamples = 68\nvalue = [15, 12, 41]'),
 Text(0.018518518518518517, 0.3181818181818182, 'x[13] <= 0.5\ngini = 0.526\nsamples = 58\nvalue = [9, 12, 37]'),
 Text(0.011396011396011397, 0.22727272727272727, 'x[9] <= 9.5\ngini = 0.494\nsamples = 54\nvalue = [6, 12, 36]'),
 Text(0.005698005698005698, 0.13636363636363635, 'x[17] <= 18.5\ngini = 0.594\nsamples = 35\nvalue = [6, 10, 19]'),
 Text(0.002849002849002849, 0.045454545454545456, 'gini = 0.497\nsamples = 13\nvalue = [0, 7, 6]'),
 Text(0.008547008547008548, 0.045454545454545456, 'gini = 0.558\nsamples = 22\nvalue = [6, 3, 13]'),
 Text(0.017094017094017096, 0.13636363636363635, 'x[15] <= 0.5\ngini = 0.188\nsamples = 19\nvalue = [0, 2, 17]'),
 Text(0.014245014245014245, 0.045454545454545456, 'gini = 0.0\nsamples = 8\nvalue = [0, 0, 8]'),
 Text(0.019943019943019943, 0.045454545454545456, 'gini = 0.298\nsamples = 11\nvalue = [0, 2, 9]'),
 Text(0.02564102564102564, 0.22727272727272727, 'x[1] <= 8.5\ngini = 0.375\nsamples = 4\nvalue = [3, 0, 1]'),
 Text(0.022792022792022793, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
 Text(0.02849002849002849, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
 Text(0.039886039886039885, 0.3181818181818182, 'x[32] <= 1.0\ngini = 0.48\nsamples = 10\nvalue = [6, 0, 4]'),
 Text(0.037037037037037035, 0.22727272727272727, 'x[8] <= 9.0\ngini = 0.444\nsamples = 6\nvalue = [2, 0, 4]'),
 Text(0.03418803418803419, 0.13636363636363635, 'x[32] <= -0.1\ngini = 0.444\nsamples = 3\nvalue = [2, 0, 1]'),
 Text(0.03133903133903134, 0.045454545454545456, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
 Text(0.037037037037037035, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
 Text(0.039886039886039885, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [0, 0, 3]'),
 Text(0.042735042735042736, 0.22727272727272727, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
 Text(0.06766381766381767, 0.4090909090909091, 'x[2] <= 1.5\ngini = 0.621\nsamples = 47\nvalue = [22, 8, 17]'),
 Text(0.05555555555555555, 0.3181818181818182, 'x[8] <= 2.0\ngini = 0.583\nsamples = 37\nvalue = [21, 7, 9]'),
 Text(0.04843304843304843, 0.22727272727272727, 'x[1] <= 14.5\ngini = 0.219\nsamples = 8\nvalue = [7, 1, 0]'),
 Text(0.045584045584045586, 0.13636363636363635, 'gini = 0.0\nsamples = 7\nvalue = [7, 0, 0]'),
 Text(0.05128205128205128, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
 Text(0.06267806267806268, 0.22727272727272727, 'x[7] <= 9.0\ngini = 0.628\nsamples = 29\nvalue = [14, 6, 9]'),
 Text(0.05698005698005698, 0.13636363636363635, 'x[9] <= 2.5\ngini = 0.604\nsamples = 13\nvalue = [3, 3, 7]'),
 Text(0.05413105413105413, 0.045454545454545456, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
 Text(0.05982905982905983, 0.045454545454545456, 'gini = 0.512\nsamples = 11\nvalue = [3, 1, 7]'),
 Text(0.06837606837606838, 0.13636363636363635, 'x[8] <= 20.5\ngini = 0.477\nsamples = 16\nvalue = [11, 3, 2]'),
 Text(0.06552706552706553, 0.045454545454545456, 'gini = 0.642\nsamples = 9\nvalue = [4, 3, 2]'),
 Text(0.07122507122507123, 0.045454545454545456, 'gini = 0.0\nsamples = 7\nvalue = [7, 0, 0]'),
 Text(0.07977207977207977, 0.3181818181818182, 'x[2] <= 4.5\ngini = 0.34\nsamples = 10\nvalue = [1, 1, 8]'),
 Text(0.07692307692307693, 0.22727272727272727, 'x[8] <= 20.5\ngini = 0.198\nsamples = 9\nvalue = [0, 1, 8]'),
 Text(0.07407407407407407, 0.13636363636363635, 'gini = 0.0\nsamples = 8\nvalue = [0, 0, 8]'),
 Text(0.07977207977207977, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
 Text(0.08262108262108261, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
 Text(0.08547008547008547, 0.5, 'x[5] <= 6.0\ngini = 0.219\nsamples = 8\nvalue = [7, 0, 1]'),
 Text(0.08262108262108261, 0.4090909090909091, 'gini = 0.0\nsamples = 6\nvalue = [6, 0, 0]'),
 Text(0.08831908831908832, 0.4090909090909091, 'x[33] <= 1.03\ngini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
 Text(0.08547008547008547, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
 Text(0.09116809116809117, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
 Text(0.12962962962962962, 0.6818181818181818, 'x[28] <= 0.5\ngini = 0.166\nsamples = 637\nvalue = [579, 58, 0]'),
 Text(0.10826210826210826, 0.5909090909090909, 'x[24] <= 7.5\ngini = 0.117\nsamples = 544\nvalue = [510, 34, 0]'),
 Text(0.10541310541310542, 0.5, 'x[27] <= 4.5\ngini = 0.111\nsamples = 542\nvalue = [510, 32, 0]'),
 Text(0.10256410256410256, 0.4090909090909091, 'gini = 0.0\nsamples = 190\nvalue = [190, 0, 0]'),
 Text(0.10826210826210826, 0.4090909090909091, 'x[17] <= 18.5\ngini = 0.165\nsamples = 352\nvalue = [320, 32, 0]'),
 Text(0.09686609686609686, 0.3181818181818182, 'x[23] <= 11.5\ngini = 0.453\nsamples = 26\nvalue = [17, 9, 0]'),
 Text(0.09401709401709402, 0.22727272727272727, 'x[11] <= 0.5\ngini = 0.492\nsamples = 16\nvalue = [7, 9, 0]'),
 Text(0.08831908831908832, 0.13636363636363635, 'x[7] <= 22.5\ngini = 0.375\nsamples = 8\nvalue = [6, 2, 0]'),
 Text(0.08547008547008547, 0.045454545454545456, 'gini = 0.0\nsamples = 5\nvalue = [5, 0, 0]'),
 Text(0.09116809116809117, 0.045454545454545456, 'gini = 0.444\nsamples = 3\nvalue = [1, 2, 0]'),
 Text(0.09971509971509972, 0.13636363636363635, 'x[21] <= 13.5\ngini = 0.219\nsamples = 8\nvalue = [1, 7, 0]'),
 Text(0.09686609686609686, 0.045454545454545456, 'gini = 0.0\nsamples = 7\nvalue = [0, 7, 0]'),
 Text(0.10256410256410256, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
 Text(0.09971509971509972, 0.22727272727272727, 'gini = 0.0\nsamples = 10\nvalue = [10, 0, 0]'),
 Text(0.11965811965811966, 0.3181818181818182, 'x[9] <= 16.0\ngini = 0.131\nsamples = 326\nvalue = [303, 23, 0]'),
 Text(0.1168091168091168, 0.22727272727272727, 'x[8] <= 1.5\ngini = 0.121\nsamples = 324\nvalue = [303, 21, 0]'),
 Text(0.1111111111111111, 0.13636363636363635, 'x[32] <= 0.0\ngini = 0.259\nsamples = 59\nvalue = [50, 9, 0]'),
 Text(0.10826210826210826, 0.045454545454545456, 'gini = 0.494\nsamples = 9\nvalue = [5, 4, 0]'),
 Text(0.11396011396011396, 0.045454545454545456, 'gini = 0.18\nsamples = 50\nvalue = [45, 5, 0]'),
 Text(0.1225071225071225, 0.13636363636363635, 'x[14] <= 0.5\ngini = 0.086\nsamples = 265\nvalue = [253, 12, 0]'),
 Text(0.11965811965811966, 0.045454545454545456, 'gini = 0.017\nsamples = 117\nvalue = [116, 1, 0]'),
 Text(0.12535612535612536, 0.045454545454545456, 'gini = 0.138\nsamples = 148\nvalue = [137, 11, 0]'),
 Text(0.1225071225071225, 0.22727272727272727, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
 Text(0.1111111111111111, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
 Text(0.150997150997151, 0.5909090909090909, 'x[7] <= 16.0\ngini = 0.383\nsamples = 93\nvalue = [69, 24, 0]'),
 Text(0.14245014245014245, 0.5, 'x[14] <= 0.5\ngini = 0.483\nsamples = 49\nvalue = [29, 20, 0]'),
 Text(0.13960113960113960, 0.4090909090909091, 'gini = 0.0\nsamples = 8\nvalue = [8, 0, 0]'),
 Text(0.14529914529914531, 0.4090909090909091, 'x[31] <= 14.7\ngini = 0.5\nsamples = 41\nvalue = [21, 20, 0]'),
 Text(0.14245014245014245, 0.3181818181818182, 'x[22] <= 1.5\ngini = 0.494\nsamples = 36\nvalue = [16, 20, 0]'),
 Text(0.13675213675213677, 0.22727272727272727, 'x[27] <= 14.5\ngini = 0.26\nsamples = 13\nvalue = [2, 11, 0]'),
```

```
Text(0.1339031339031339, 0.13636363636363635, 'x[10] <= 10.5\ngini = 0.153\nsamples = 12\nvalue = [1, 11, 0]'),
Text(0.13105413105413105, 0.045454545454545456, 'gini = 0.0\nsamples = 11\nvalue = [0, 11, 0]'),
Text(0.13675213675213677, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.1396011396011396, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.14814814814814814, 0.22727272727272727, 'x[3] <= 10.0\ngini = 0.476\nsamples = 23\nvalue = [14, 9, 0]'),
Text(0.1452991452991453, 0.13636363636363635, 'x[27] <= 7.5\ngini = 0.498\nsamples = 17\nvalue = [8, 9, 0]'),
Text(0.14245014245014245, 0.045454545454545456, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
Text(0.14814814814814814, 0.045454545454545456, 'gini = 0.426\nsamples = 13\nvalue = [4, 9, 0]'),
Text(0.150997150997151, 0.13636363636363635, 'gini = 0.0\nsamples = 6\nvalue = [6, 0, 0]'),
Text(0.14814814814814814, 0.3181818181818182, 'gini = 0.0\nsamples = 5\nvalue = [5, 0, 0]'),
Text(0.15954415954415954, 0.5, 'x[27] <= 15.5\ngini = 0.165\nsamples = 44\nvalue = [40, 4, 0]'),
Text(0.15669515669515668, 0.4090909090909091, 'x[29] <= 15.5\ngini = 0.091\nsamples = 42\nvalue = [40, 2, 0]'),
Text(0.15384615384615385, 0.3181818181818182, 'gini = 0.0\nsamples = 39\nvalue = [39, 0, 0]'),
Text(0.15954415954415954, 0.3181818181818182, 'x[17] <= 29.5\ngini = 0.444\nsamples = 3\nvalue = [1, 2, 0]'),
Text(0.15669515669515668, 0.22727272727272727, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
Text(0.1623931623931624, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.1623931623931624, 0.4090909090909091, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
Text(0.27136752136752135, 0.7727272727272727, 'x[14] <= 0.5\ngini = 0.637\nsamples = 716\nvalue = [293, 285, 138]'),
Text(0.18803418803418803, 0.6818181818181818, 'x[33] <= 1.765\ngini = 0.26\nsamples = 113\nvalue = [96, 15, 2]'),
Text(0.17663817663817663, 0.5909090909090909, 'x[6] <= 8.5\ngini = 0.099\nsamples = 78\nvalue = [74, 2, 2]'),
Text(0.1737891737891738, 0.5, 'x[22] <= 7.5\ngini = 0.051\nsamples = 76\nvalue = [74, 2, 0]'),
Text(0.16809116809116809, 0.4090909090909091, 'x[8] <= 27.5\ngini = 0.027\nsamples = 74\nvalue = [73, 1, 0]'),
Text(0.16524216524216523, 0.3181818181818182, 'gini = 0.0\nsamples = 62\nvalue = [62, 0, 0]'),
Text(0.17094017094017094, 0.3181818181818182, 'x[32] <= 0.0\ngini = 0.153\nsamples = 12\nvalue = [11, 1, 0]'),
Text(0.16809116809116809, 0.22727272727272727, 'x[22] <= 4.0\ngini = 0.5\nsamples = 2\nvalue = [1, 1, 0]'),
Text(0.16524216524216523, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.17094017094017094, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.1737891737891738, 0.22727272727272727, 'gini = 0.0\nsamples = 10\nvalue = [10, 0, 0]'),
Text(0.1794871794871795, 0.4090909090909091, 'x[9] <= 4.5\ngini = 0.5\nsamples = 2\nvalue = [1, 1, 0]'),
Text(0.17663817663817663, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.18233618233618235, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.1794871794871795, 0.5, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
Text(0.19943019943019943, 0.5909090909090909, 'x[16] <= 0.5\ngini = 0.467\nsamples = 35\nvalue = [22, 13, 0]'),
Text(0.19658119658119658, 0.5, 'x[17] <= 25.0\ngini = 0.412\nsamples = 31\nvalue = [22, 9, 0]'),
Text(0.19373219373219372, 0.4090909090909091, 'x[21] <= 9.5\ngini = 0.495\nsamples = 20\nvalue = [11, 9, 0]'),
Text(0.18803418803418803, 0.3181818181818182, 'x[20] <= 5.5\ngini = 0.219\nsamples = 8\nvalue = [7, 1, 0]'),
Text(0.18518518518518517, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.1908831908831909, 0.22727272727272727, 'gini = 0.0\nsamples = 7\nvalue = [7, 0, 0]'),
Text(0.19943019943019943, 0.3181818181818182, 'x[21] <= 11.5\ngini = 0.444\nsamples = 12\nvalue = [4, 8, 0]'),
Text(0.19658119658119658, 0.22727272727272727, 'gini = 0.0\nsamples = 8\nvalue = [0, 8, 0]'),
Text(0.2022792022792023, 0.22727272727272727, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
Text(0.19943019943019943, 0.4090909090909091, 'gini = 0.0\nsamples = 11\nvalue = [11, 0, 0]'),
Text(0.2022792022792023, 0.5, 'gini = 0.0\nsamples = 4\nvalue = [0, 4, 0]'),
Text(0.3547008547008547, 0.6818181818181818, 'x[28] <= 3.5\ngini = 0.642\nsamples = 603\nvalue = [197, 270, 136]'),
Text(0.28383190883190884, 0.5909090909090909, 'x[17] <= 22.5\ngini = 0.587\nsamples = 323\nvalue = [136, 153, 34]'),
Text(0.24643874643874644, 0.5, 'x[11] <= 0.5\ngini = 0.565\nsamples = 228\nvalue = [75, 128, 25]'),
Text(0.21794871794871795, 0.4090909090909091, 'x[30] <= 2.5\ngini = 0.445\nsamples = 76\nvalue = [16, 54, 6]'),
Text(0.2150997150997151, 0.3181818181818182, 'x[9] <= 2.5\ngini = 0.425\nsamples = 74\nvalue = [14, 54, 6]'),
Text(0.20797720797720798, 0.22727272727272727, 'x[10] <= 3.0\ngini = 0.375\nsamples = 4\nvalue = [3, 1, 0]'),
Text(0.20512820512820512, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.21082621082621084, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.2222222222222222, 0.22727272727272727, 'x[27] <= 17.5\ngini = 0.395\nsamples = 70\nvalue = [11, 53, 6]'),
Text(0.21652421652421652, 0.13636363636363635, 'x[3] <= 14.5\ngini = 0.365\nsamples = 68\nvalue = [10, 53, 5]'),
Text(0.21367521367521367, 0.045454545454545456, 'gini = 0.311\nsamples = 61\nvalue = [7, 50, 4]'),
Text(0.21937321937321938, 0.045454545454545456, 'gini = 0.612\nsamples = 7\nvalue = [3, 3, 1]'),
Text(0.22792022792022792, 0.13636363636363635, 'x[9] <= 19.0\ngini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.22507122507122507, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.23076923076923078, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.2207977207977208, 0.3181818181818182, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.27492877492877493, 0.4090909090909091, 'x[22] <= 5.5\ngini = 0.597\nsamples = 152\nvalue = [59, 74, 19]'),
Text(0.2564102564102564, 0.3181818181818182, 'x[8] <= 27.5\ngini = 0.565\nsamples = 136\nvalue = [55, 70, 11]'),
Text(0.245014245014245, 0.22727272727272727, 'x[15] <= 0.5\ngini = 0.584\nsamples = 111\nvalue = [49, 51, 11]'),
Text(0.23931623931623933, 0.13636363636363635, 'x[3] <= 15.5\ngini = 0.587\nsamples = 65\nvalue = [21, 35, 9]'),
Text(0.23646723646723647, 0.045454545454545456, 'gini = 0.564\nsamples = 56\nvalue = [14, 33, 9]'),
Text(0.24216524216524216, 0.045454545454545456, 'gini = 0.346\nsamples = 9\nvalue = [7, 2, 0]'),
Text(0.25071225071225073, 0.13636363636363635, 'x[8] <= 1.5\ngini = 0.507\nsamples = 46\nvalue = [28, 16, 2]'),
Text(0.24786324786324787, 0.045454545454545456, 'gini = 0.521\nsamples = 13\nvalue = [4, 8, 1]'),
Text(0.2535612535612536, 0.045454545454545456, 'gini = 0.411\nsamples = 33\nvalue = [24, 8, 1]'),
Text(0.2678062678062678, 0.22727272727272727, 'x[7] <= 17.5\ngini = 0.365\nsamples = 25\nvalue = [6, 19, 0]'),
Text(0.2621082621082621, 0.13636363636363635, 'x[10] <= 5.0\ngini = 0.198\nsamples = 18\nvalue = [2, 16, 0]'),
Text(0.25925925925925924, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.26495726495726496, 0.045454545454545456, 'gini = 0.111\nsamples = 17\nvalue = [1, 16, 0]'),
Text(0.27350427350427353, 0.13636363636363635, 'x[2] <= 2.5\ngini = 0.49\nsamples = 7\nvalue = [4, 3, 0]'),
Text(0.2706552706552707, 0.045454545454545456, 'gini = 0.375\nsamples = 4\nvalue = [1, 3, 0]'),
Text(0.27635327635327633, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.2934472934472934, 0.3181818181818182, 'x[17] <= 19.5\ngini = 0.625\nsamples = 16\nvalue = [4, 4, 8]'),
Text(0.28774928774928776, 0.22727272727272727, 'x[2] <= 1.5\ngini = 0.245\nsamples = 7\nvalue = [0, 1, 6]'),
Text(0.2849002849002849, 0.13636363636363635, 'x[10] <= 4.5\ngini = 0.5\nsamples = 2\nvalue = [0, 1, 1]'),
Text(0.28205128205128205, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.28774928774928776, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.2905982905982906, 0.13636363636363635, 'gini = 0.0\nsamples = 5\nvalue = [0, 0, 5]'),
Text(0.29914529914529914, 0.22727272727272727, 'x[10] <= 7.5\ngini = 0.642\nsamples = 9\nvalue = [4, 3, 2]'),
Text(0.2962962962962963, 0.13636363636363635, 'x[23] <= 12.333\ngini = 0.611\nsamples = 6\nvalue = [1, 3, 2]'),
Text(0.2934472934472934, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.29914529914529914, 0.045454545454545456, 'gini = 0.444\nsamples = 3\nvalue = [1, 0, 2]'),
```

```
Text(0.301994301994302, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.3212250712250712, 0.5, 'x[10] <= 15.5\ngini = 0.509\nsamples = 95\nvalue = [61, 25, 9]'),
Text(0.3183760683760684, 0.4090909090909091, 'x[21] <= 6.5\ngini = 0.481\nsamples = 90\nvalue = [61, 20, 9]'),
Text(0.3076923076923077, 0.3181818181818182, 'x[26] <= 5.5\ngini = 0.56\nsamples = 10\nvalue = [2, 6, 2]'),
Text(0.30484330484330485, 0.22727272727272727, 'gini = 0.0\nsamples = 6\nvalue = [0, 6, 0]'),
Text(0.31054131054131057, 0.22727272727272727, 'x[17] <= 24.5\ngini = 0.5\nsamples = 4\nvalue = [2, 0, 2]'),
Text(0.3076923076923077, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
Text(0.31339031339031337, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.3290598290598291, 0.3181818181818182, 'x[28] <= 2.5\ngini = 0.418\nsamples = 80\nvalue = [59, 14, 7]'),
Text(0.32193732193732194, 0.22727272727272727, 'x[33] <= 1.905\ngini = 0.16\nsamples = 35\nvalue = [32, 1, 2]'),
Text(0.3190883190883191, 0.13636363636363635, 'x[27] <= 21.0\ngini = 0.111\nsamples = 34\nvalue = [32, 0, 2]'),
Text(0.3162393162393162, 0.045454545454545456, 'gini = 0.061\nsamples = 32\nvalue = [31, 0, 1]'),
Text(0.32193732193732194, 0.045454545454545456, 'gini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.3247863247863248, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.33618233618233617, 0.22727272727272727, 'x[22] <= 2.5\ngini = 0.544\nsamples = 45\nvalue = [27, 13, 5]'),
Text(0.33048433048433046, 0.13636363636363635, 'x[17] <= 31.5\ngini = 0.541\nsamples = 14\nvalue = [5, 8, 1]'),
Text(0.32763532763532766, 0.045454545454545456, 'gini = 0.469\nsamples = 8\nvalue = [5, 3, 0]'),
Text(0.3333333333333333, 0.045454545454545456, 'gini = 0.278\nsamples = 6\nvalue = [0, 5, 1]'),
Text(0.3418803418803419, 0.13636363636363635, 'x[31] <= 12.55\ngini = 0.454\nsamples = 31\nvalue = [22, 5, 4]'),
Text(0.33903133903133903, 0.045454545454545456, 'gini = 0.609\nsamples = 19\nvalue = [10, 5, 4]'),
Text(0.34472934472934474, 0.045454545454545456, 'gini = 0.0\nsamples = 12\nvalue = [12, 0, 0]'),
Text(0.32407407407407407, 0.4090909090909091, 'gini = 0.0\nsamples = 5\nvalue = [0, 5, 0]'),
Text(0.4255698005698006, 0.5909090909090909, 'x[22] <= 4.5\ngini = 0.645\nsamples = 280\nvalue = [61, 117, 102]'),
Text(0.38746438746438744, 0.5, 'x[26] <= 6.5\ngini = 0.634\nsamples = 152\nvalue = [39, 73, 40]'),
Text(0.38461538461538464, 0.4090909090909091, 'x[29] <= 13.625\ngini = 0.617\nsamples = 143\nvalue = [30, 73, 40]'),
Text(0.37037037037037035, 0.3181818181818182, 'x[3] <= 16.5\ngini = 0.604\nsamples = 132\nvalue = [29, 71, 32]'),
Text(0.358974358974359, 0.22727272727272727, 'x[1] <= 13.0\ngini = 0.628\nsamples = 114\nvalue = [26, 56, 32]'),
Text(0.35327635327635326, 0.13636363636363635, 'x[22] <= 3.5\ngini = 0.641\nsamples = 98\nvalue = [22, 44, 32]'),
Text(0.3504273504273504, 0.045454545454545456, 'gini = 0.601\nsamples = 50\nvalue = [13, 27, 10]'),
Text(0.3561253561253561, 0.045454545454545456, 'gini = 0.629\nsamples = 48\nvalue = [9, 17, 22]'),
Text(0.3646723646723647, 0.13636363636363635, 'x[31] <= 8.25\ngini = 0.375\nsamples = 16\nvalue = [4, 12, 0]'),
Text(0.36182336182336183, 0.045454545454545456, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.36752136752136755, 0.045454545454545456, 'gini = 0.245\nsamples = 14\nvalue = [2, 12, 0]'),
Text(0.3817663817663818, 0.22727272727272727, 'x[9] <= 5.5\ngini = 0.278\nsamples = 18\nvalue = [3, 15, 0]'),
Text(0.37606837606837606, 0.13636363636363635, 'x[7] <= 3.0\ngini = 0.444\nsamples = 3\nvalue = [2, 1, 0]'),
Text(0.3732193732193732, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.3789173789173789, 0.045454545454545456, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.38746438746438744, 0.13636363636363635, 'x[31] <= 15.85\ngini = 0.124\nsamples = 15\nvalue = [1, 14, 0]'),
Text(0.38461538461538464, 0.045454545454545456, 'gini = 0.0\nsamples = 14\nvalue = [0, 14, 0]'),
Text(0.3903133903133903, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.39886039886039887, 0.3181818181818182, 'x[11] <= 0.43\ngini = 0.43\nsamples = 11\nvalue = [1, 2, 8]'),
Text(0.396011396011396, 0.22727272727272727, 'gini = 0.0\nsamples = 7\nvalue = [0, 0, 7]'),
Text(0.4017094017094017, 0.22727272727272727, 'x[8] <= 7.5\ngini = 0.625\nsamples = 4\nvalue = [1, 2, 1]'),
Text(0.39886039886039887, 0.13636363636363635, 'x[29] <= 14.625\ngini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.396011396011396, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.4017094017094017, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.4045584045584046, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
Text(0.3903133903133903, 0.4090909090909091, 'gini = 0.0\nsamples = 9\nvalue = [9, 0, 0]'),
Text(0.4636752136752137, 0.5, 'x[16] <= 0.5\ngini = 0.618\nsamples = 128\nvalue = [22, 44, 62]'),
Text(0.4458689458689459, 0.4090909090909091, 'x[2] <= 1.5\ngini = 0.637\nsamples = 103\nvalue = [20, 40, 43]'),
Text(0.42735042735042733, 0.3181818181818182, 'x[10] <= 9.5\ngini = 0.627\nsamples = 68\nvalue = [14, 33, 21]'),
Text(0.41595441595441596, 0.22727272727272727, 'x[3] <= 12.5\ngini = 0.584\nsamples = 43\nvalue = [12, 24, 7]'),
Text(0.41025641025641024, 0.13636363636363635, 'x[27] <= 10.5\ngini = 0.487\nsamples = 31\nvalue = [6, 21, 4]'),
Text(0.4074074074074074, 0.045454545454545456, 'gini = 0.568\nsamples = 18\nvalue = [6, 10, 2]'),
Text(0.4131054131054131, 0.045454545454545456, 'gini = 0.26\nsamples = 13\nvalue = [0, 11, 2]'),
Text(0.42165242165242167, 0.13636363636363635, 'x[1] <= 7.0\ngini = 0.625\nsamples = 12\nvalue = [6, 3, 3]'),
Text(0.4188034188034188, 0.045454545454545456, 'gini = 0.656\nsamples = 8\nvalue = [2, 3, 3]'),
Text(0.42450142450142453, 0.045454545454545456, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
Text(0.43874643874643876, 0.22727272727272727, 'x[31] <= 14.7\ngini = 0.55\nsamples = 25\nvalue = [2, 9, 14]'),
Text(0.43304843304843305, 0.13636363636363635, 'x[3] <= 14.5\ngini = 0.505\nsamples = 20\nvalue = [2, 5, 13]'),
Text(0.4301994301994302, 0.045454545454545456, 'gini = 0.406\nsamples = 16\nvalue = [2, 2, 12]'),
Text(0.4358974358974359, 0.045454545454545456, 'gini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
Text(0.4444444444444444, 0.13636363636363635, 'x[23] <= 12.217\ngini = 0.32\nsamples = 5\nvalue = [0, 4, 1]'),
Text(0.4415954415954416, 0.045454545454545456, 'gini = 0.0\nsamples = 4\nvalue = [0, 4, 0]'),
Text(0.4472934472934473, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.46438746438746437, 0.3181818181818182, 'x[13] <= 0.5\ngini = 0.536\nsamples = 35\nvalue = [6, 7, 22]'),
Text(0.46153846153846156, 0.22727272727272727, 'x[1] <= 10.5\ngini = 0.496\nsamples = 33\nvalue = [4, 7, 22]'),
Text(0.45584045584045585, 0.13636363636363635, 'x[17] <= 18.5\ngini = 0.416\nsamples = 29\nvalue = [1, 7, 21]'),
Text(0.452991452991453, 0.045454545454545456, 'gini = 0.0\nsamples = 9\nvalue = [0, 0, 9]'),
Text(0.45868945868945687, 0.045454545454545456, 'gini = 0.515\nsamples = 20\nvalue = [1, 7, 12]'),
Text(0.4672364672364672, 0.13636363636363635, 'x[33] <= 0.41\ngini = 0.375\nsamples = 4\nvalue = [3, 0, 1]'),
Text(0.46438746438746437, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.4700854700854701, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.4672364672364672, 0.22727272727272727, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.48148148148148145, 0.4090909090909091, 'x[29] <= 11.225\ngini = 0.39\nsamples = 25\nvalue = [2, 4, 19]'),
Text(0.47578347578347758, 0.3181818181818182, 'x[33] <= -1.87\ngini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
Text(0.47293447293447294, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.47863247863247865, 0.22727272727272727, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.48717948717948717, 0.3181818181818182, 'x[3] <= 15.5\ngini = 0.254\nsamples = 21\nvalue = [2, 1, 18]'),
Text(0.4843304843304843, 0.22727272727272727, 'x[23] <= 14.833\ngini = 0.185\nsamples = 20\nvalue = [1, 1, 18]'),
Text(0.48148148148148145, 0.13636363636363635, 'x[17] <= 23.0\ngini = 0.1\nsamples = 19\nvalue = [0, 1, 18]'),
Text(0.47863247863247865, 0.045454545454545456, 'gini = 0.0\nsamples = 16\nvalue = [0, 0, 16]'),
Text(0.4843304843304843, 0.045454545454545456, 'gini = 0.444\nsamples = 3\nvalue = [0, 1, 2]'),
Text(0.48717948717948717, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.49002849002849, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
```

```
Text(0.6445868945868946, 0.8636363636363636, 'x[14] <= 0.5\ngini = 0.382\nsamples = 2045\nvalue = [199, 274, 1572]'),
Text(0.5185185185185185, 0.7727272727272727, 'x[33] <= 1.905\ngini = 0.572\nsamples = 89\nvalue = [52, 17, 20]'),
Text(0.5071225071225072, 0.6818181818181818, 'x[26] <= 0.5\ngini = 0.38\nsamples = 55\nvalue = [41, 0, 14]'),
Text(0.5042735042735043, 0.5909090909090909, 'x[28] <= 5.5\ngini = 0.484\nsamples = 34\nvalue = [20, 0, 14]'),
Text(0.4985754985754986, 0.5, 'x[18] <= 0.5\ngini = 0.255\nsamples = 20\nvalue = [17, 0, 3]'),
Text(0.49572649572649574, 0.4090909090909091, 'x[16] <= 0.5\ngini = 0.105\nsamples = 18\nvalue = [17, 0, 1]'),
Text(0.4928774928774929, 0.3181818181818182, 'gini = 0.0\nsamples = 15\nvalue = [15, 0, 0]'),
Text(0.4985754985754986, 0.3181818181818182, 'x[27] <= 8.0\ngini = 0.444\nsamples = 3\nvalue = [2, 0, 1]'),
Text(0.49572649572649574, 0.22727272727272727, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.5014245014245015, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5014245014245015, 0.4090909090909091, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
Text(0.50997150997151, 0.5, 'x[8] <= 2.0\ngini = 0.337\nsamples = 14\nvalue = [3, 0, 11]'),
Text(0.5071225071225072, 0.4090909090909091, 'x[7] <= 2.0\ngini = 0.375\nsamples = 4\nvalue = [3, 0, 1]'),
Text(0.5042735042735043, 0.3181818181818182, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.50997150997151, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5128205128205128, 0.4090909090909091, 'gini = 0.0\nsamples = 10\nvalue = [0, 0, 10]'),
Text(0.50997150997151, 0.5909090909090909, 'gini = 0.0\nsamples = 21\nvalue = [21, 0, 0]'),
Text(0.5299145299145299, 0.6818181818181818, 'x[13] <= 0.5\ngini = 0.614\nsamples = 34\nvalue = [11, 17, 6]'),
Text(0.5242165242165242, 0.5909090909090909, 'x[26] <= 10.0\ngini = 0.475\nsamples = 18\nvalue = [1, 12, 5]'),
Text(0.5213675213675214, 0.5, 'x[0] <= 1.5\ngini = 0.338\nsamples = 15\nvalue = [1, 12, 2]'),
Text(0.5185185185185185, 0.4090909090909091, 'x[6] <= 11.5\ngini = 0.245\nsamples = 14\nvalue = [0, 12, 2]'),
Text(0.5156695156695157, 0.3181818181818182, 'x[29] <= 14.243\ngini = 0.142\nsamples = 13\nvalue = [0, 12, 1]'),
Text(0.5128205128205128, 0.22727272727272727, 'gini = 0.0\nsamples = 12\nvalue = [0, 12, 0]'),
Text(0.5185185185185185, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5213675213675214, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5242165242165242, 0.4090909090909091, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.5270655270655271, 0.5, 'gini = 0.0\nsamples = 3\nvalue = [0, 0, 3]'),
Text(0.5356125356125356, 0.5909090909090909, 'x[29] <= 12.633\ngini = 0.508\nsamples = 16\nvalue = [10, 5, 1]'),
Text(0.5327635327635327, 0.5, 'gini = 0.0\nsamples = 7\nvalue = [7, 0, 0]'),
Text(0.5384615384615384, 0.5, 'x[8] <= 20.5\ngini = 0.568\nsamples = 9\nvalue = [3, 5, 1]'),
Text(0.5327635327635327, 0.4090909090909091, 'x[27] <= 7.0\ngini = 0.32\nsamples = 5\nvalue = [0, 4, 1]'),
Text(0.5299145299145299, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5356125356125356, 0.3181818181818182, 'gini = 0.0\nsamples = 4\nvalue = [0, 4, 0]'),
Text(0.5441595441595442, 0.4090909090909091, 'x[21] <= 16.5\ngini = 0.375\nsamples = 4\nvalue = [3, 1, 0]'),
Text(0.5413105413105413, 0.3181818181818182, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.5470085470085471, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7706552706552706, 0.7727272727272727, 'x[21] <= 8.5\ngini = 0.348\nsamples = 1956\nvalue = [147, 257, 1552]'),
Text(0.6597222222222222, 0.6818181818181818, 'x[28] <= 5.5\ngini = 0.247\nsamples = 1271\nvalue = [73, 102, 1096]'),
Text(0.583511396011396, 0.5909090909090909, 'x[26] <= 5.5\ngini = 0.39\nsamples = 340\nvalue = [31, 50, 259]'),
Text(0.5612535612535613, 0.5, 'x[13] <= 0.5\ngini = 0.22\nsamples = 182\nvalue = [11, 11, 160]'),
Text(0.5555555555555556, 0.4090909090909091, 'x[2] <= 7.5\ngini = 0.184\nsamples = 171\nvalue = [8, 9, 154]'),
Text(0.5527065527065527, 0.3181818181818182, 'x[7] <= 26.5\ngini = 0.175\nsamples = 170\nvalue = [8, 8, 154]'),
Text(0.5498575498575499, 0.22727272727272727, 'x[23] <= 10.35\ngini = 0.166\nsamples = 169\nvalue = [8, 7, 154]'),
Text(0.5470085470085471, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.5527065527065527, 0.13636363636363635, 'x[8] <= 2.5\ngini = 0.156\nsamples = 168\nvalue = [8, 6, 154]'),
Text(0.5498575498575499, 0.045454545454545456, 'gini = 0.416\nsamples = 31\nvalue = [4, 4, 23]'),
Text(0.5555555555555556, 0.045454545454545456, 'gini = 0.085\nsamples = 137\nvalue = [4, 2, 131]'),
Text(0.5555555555555556, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.5584045584045584, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.5669515669515669, 0.4090909090909091, 'x[3] <= 11.5\ngini = 0.595\nsamples = 11\nvalue = [3, 2, 6]'),
Text(0.5641025641025641, 0.3181818181818182, 'x[27] <= 6.5\ngini = 0.494\nsamples = 9\nvalue = [1, 2, 6]'),
Text(0.5612535612535613, 0.22727272727272727, 'gini = 0.0\nsamples = 5\nvalue = [0, 0, 5]'),
Text(0.5669515669515669, 0.22727272727272727, 'x[29] <= 12.517\ngini = 0.625\nsamples = 4\nvalue = [1, 2, 1]'),
Text(0.5641025641025641, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
Text(0.5698005698005698, 0.13636363636363635, 'x[7] <= 8.0\ngini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.5669515669515669, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5726495726495726, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.5698005698005698, 0.3181818181818182, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.6057692307692307, 0.5, 'x[31] <= 8.25\ngini = 0.53\nsamples = 158\nvalue = [20, 39, 99]'),
Text(0.5811965811965812, 0.4090909090909091, 'x[21] <= 8.5\ngini = 0.48\nsamples = 15\nvalue = [1, 10, 4]'),
Text(0.5754985754985755, 0.3181818181818182, 'x[29] <= 11.5\ngini = 0.375\nsamples = 4\nvalue = [0, 1, 3]'),
Text(0.5726495726495726, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.5783475783475783, 0.22727272727272727, 'gini = 0.0\nsamples = 3\nvalue = [0, 0, 3]'),
Text(0.5868945868945868, 0.3181818181818182, 'x[9] <= 2.0\ngini = 0.314\nsamples = 11\nvalue = [1, 9, 1]'),
Text(0.5840045584045584, 0.22727272727272727, 'x[7] <= 21.0\ngini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.5811965811965812, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.5868945868945868, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.5897435897435898, 0.22727272727272727, 'gini = 0.0\nsamples = 9\nvalue = [0, 9, 0]'),
Text(0.6303418803418803, 0.4090909090909091, 'x[29] <= 12.387\ngini = 0.5\nsamples = 143\nvalue = [19, 29, 95]'),
Text(0.6096866096866097, 0.3181818181818182, 'x[7] <= 8.5\ngini = 0.639\nsamples = 47\nvalue = [12, 13, 22]'),
Text(0.5982905982905983, 0.22727272727272727, 'x[8] <= 20.5\ngini = 0.406\nsamples = 16\nvalue = [2, 2, 12]'),
Text(0.5925925925925926, 0.13636363636363635, 'x[19] <= 1.0\ngini = 0.26\nsamples = 13\nvalue = [2, 0, 11]'),
Text(0.5897435897435898, 0.045454545454545456, 'gini = 0.153\nsamples = 12\nvalue = [1, 0, 11]'),
Text(0.5954415954415955, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.603988603988604, 0.13636363636363635, 'x[23] <= 12.633\ngini = 0.444\nsamples = 3\nvalue = [0, 2, 1]'),
Text(0.6011396011396012, 0.045454545454545456, 'gini = 0.0\nsamples = 2\nvalue = [0, 2, 0]'),
Text(0.6068376068376068, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.6210826210826211, 0.22727272727272727, 'x[23] <= 11.733\ngini = 0.666\nsamples = 31\nvalue = [10, 11, 10]'),
Text(0.6153846153846154, 0.13636363636363635, 'x[15] <= 0.5\ngini = 0.562\nsamples = 11\nvalue = [4, 1, 6]'),
Text(0.6125356125356125, 0.045454545454545456, 'gini = 0.406\nsamples = 8\nvalue = [1, 1, 6]'),
Text(0.6182336182336182, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [3, 0, 0]'),
Text(0.6267806267806267, 0.13636363636363635, 'x[9] <= 8.5\ngini = 0.62\nsamples = 20\nvalue = [6, 10, 4]'),
Text(0.6239316239316239, 0.045454545454545456, 'gini = 0.346\nsamples = 9\nvalue = [0, 7, 2]'),
Text(0.6296296296296297, 0.045454545454545456, 'gini = 0.595\nsamples = 11\nvalue = [6, 3, 2]'),
```

```
Text(0.6509971509971509, 0.3181818181818182, 'x[27] <= 7.5\ngini = 0.389\nsamples = 96\nvalue = [7, 16, 73]'),
Text(0.6438746438746439, 0.22727272727272727, 'x[7] <= 8.0\ngini = 0.579\nsamples = 31\nvalue = [4, 10, 17]'),
Text(0.6381766381766382, 0.13636363636363635, 'x[22] <= 5.5\ngini = 0.542\nsamples = 12\nvalue = [4, 1, 7]'),
Text(0.6353276353276354, 0.045454545454545456, 'gini = 0.0\nsamples = 6\nvalue = [0, 0, 6]'),
Text(0.6410256410256411, 0.045454545454545456, 'gini = 0.5\nsamples = 6\nvalue = [4, 1, 1]'),
Text(0.6495726495726496, 0.13636363636363635, 'x[22] <= 5.5\ngini = 0.499\nsamples = 19\nvalue = [0, 9, 10]'),
Text(0.6467236467236467, 0.045454545454545456, 'gini = 0.278\nsamples = 6\nvalue = [0, 5, 1]'),
Text(0.6524216524216524, 0.045454545454545456, 'gini = 0.426\nsamples = 13\nvalue = [0, 4, 9]'),
Text(0.6581196581196581, 0.22727272727272727, 'x[21] <= 5.5\ngini = 0.247\nsamples = 65\nvalue = [3, 6, 56]'),
Text(0.6552706552706553, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.6609686609686609, 0.13636363636363635, 'x[22] <= 3.5\ngini = 0.226\nsamples = 64\nvalue = [3, 5, 56]'),
Text(0.6581196581196581, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.6638176638176638, 0.045454545454545456, 'gini = 0.204\nsamples = 63\nvalue = [3, 4, 56]'),
Text(0.7359330484330484, 0.5909090909090909, 'x[30] <= 0.5\ngini = 0.187\nsamples = 931\nvalue = [42, 52, 837]'),
Text(0.7197293447293447, 0.5, 'x[23] <= 12.041\ngini = 0.178\nsamples = 920\nvalue = [39, 49, 832]'),
Text(0.7015665515669516, 0.4090909090909091, 'x[3] <= 15.5\ngini = 0.363\nsamples = 101\nvalue = [8, 14, 79]'),
Text(0.688034188034188, 0.3181818181818182, 'x[1] <= 13.5\ngini = 0.311\nsamples = 94\nvalue = [6, 11, 77]'),
Text(0.6780626780626781, 0.22727272727272727, 'x[17] <= 42.5\ngini = 0.252\nsamples = 85\nvalue = [5, 7, 73]'),
Text(0.6723646723646723, 0.13636363636363635, 'x[9] <= 4.5\ngini = 0.219\nsamples = 83\nvalue = [4, 6, 73]'),
Text(0.6695156695156695, 0.045454545454545456, 'gini = 0.595\nsamples = 11\nvalue = [3, 2, 6]'),
Text(0.6752136752136753, 0.045454545454545456, 'gini = 0.131\nsamples = 72\nvalue = [1, 4, 67]'),
Text(0.6837606837606838, 0.13636363636363635, 'x[29] <= 11.651\ngini = 0.5\nsamples = 2\nvalue = [1, 1, 0]'),
Text(0.6809116809116809, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.6866096866096866, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.698005698005698, 0.22727272727272727, 'x[21] <= 7.5\ngini = 0.593\nsamples = 9\nvalue = [1, 4, 4]'),
Text(0.6951566951566952, 0.13636363636363635, 'x[26] <= 8.5\ngini = 0.32\nsamples = 5\nvalue = [1, 0, 4]'),
Text(0.6923076923076923, 0.045454545454545456, 'gini = 0.0\nsamples = 4\nvalue = [0, 0, 4]'),
Text(0.698005698005698, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.7008547008547008, 0.13636363636363635, 'gini = 0.0\nsamples = 4\nvalue = [0, 4, 0]'),
Text(0.7150997150997151, 0.3181818181818182, 'x[33] <= 1.03\ngini = 0.653\nsamples = 7\nvalue = [2, 3, 2]'),
Text(0.7094017094017094, 0.22727272727272727, 'x[29] <= 12.417\ngini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
Text(0.7065527065527065, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.7122507122507122, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.7207977207977208, 0.22727272727272727, 'x[27] <= 9.5\ngini = 0.444\nsamples = 3\nvalue = [2, 0, 1]'),
Text(0.717948717948718, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.7236467236467237, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.7378917378917379, 0.4090909090909091, 'x[19] <= 4.5\ngini = 0.151\nsamples = 819\nvalue = [31, 35, 753]'),
Text(0.7350427350427351, 0.3181818181818182, 'x[9] <= 31.5\ngini = 0.149\nsamples = 818\nvalue = [31, 34, 753]'),
Text(0.7321937321937322, 0.22727272727272727, 'x[20] <= 5.5\ngini = 0.147\nsamples = 817\nvalue = [31, 33, 753]'),
Text(0.7293447293447294, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7350427350427351, 0.13636363636363635, 'x[22] <= 5.5\ngini = 0.145\nsamples = 816\nvalue = [31, 32, 753]'),
Text(0.7321937321937322, 0.045454545454545456, 'gini = 0.395\nsamples = 29\nvalue = [3, 4, 22]'),
Text(0.7378917378917379, 0.045454545454545456, 'gini = 0.135\nsamples = 787\nvalue = [28, 28, 731]'),
Text(0.7378917378917379, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7407407407407407, 0.3181818181818182, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7521367521367521, 0.5, 'x[31] <= 14.45\ngini = 0.645\nsamples = 11\nvalue = [3, 3, 5]'),
Text(0.7492877492877493, 0.4090909090909091, 'x[23] <= 12.592\ngini = 0.612\nsamples = 7\nvalue = [3, 3, 1]'),
Text(0.7464387464387464, 0.3181818181818182, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
Text(0.7521367521367521, 0.3181818181818182, 'x[9] <= 2.5\ngini = 0.56\nsamples = 5\nvalue = [1, 3, 1]'),
Text(0.7492877492877493, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
Text(0.7549857549857549, 0.22727272727272727, 'x[7] <= 22.5\ngini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
Text(0.7521367521367521, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.7578347578347578, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
Text(0.7549857549857549, 0.4090909090909091, 'gini = 0.0\nsamples = 4\nvalue = [0, 0, 4]'),
Text(0.8815883190883191, 0.6818181818181818, 'x[22] <= 5.5\ngini = 0.494\nsamples = 685\nvalue = [74, 155, 456]'),
Text(0.8222934472934473, 0.5909090909090909, 'x[26] <= 6.5\ngini = 0.597\nsamples = 229\nvalue = [31, 82, 116]'),
Text(0.7905982905982906, 0.5, 'x[3] <= 10.5\ngini = 0.509\nsamples = 172\nvalue = [15, 47, 110]'),
Text(0.7763532763532763, 0.4090909090909091, 'x[29] <= 14.618\ngini = 0.424\nsamples = 113\nvalue = [7, 24, 82]'),
Text(0.7735042735042735, 0.3181818181818182, 'x[21] <= 10.5\ngini = 0.404\nsamples = 110\nvalue = [7, 21, 82]'),
Text(0.7663817663817664, 0.22727272727272727, 'x[0] <= 4.5\ngini = 0.246\nsamples = 51\nvalue = [3, 4, 44]'),
Text(0.7635327635327636, 0.13636363636363635, 'x[23] <= 14.208\ngini = 0.218\nsamples = 50\nvalue = [3, 3, 44]'),
Text(0.7606837606837606, 0.045454545454545456, 'gini = 0.188\nsamples = 49\nvalue = [3, 2, 44]'),
Text(0.7663817663817664, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7692307692307693, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
Text(0.7806267806267806, 0.22727272727272727, 'x[23] <= 10.775\ngini = 0.498\nsamples = 59\nvalue = [4, 17, 38]'),
Text(0.7749287749287749, 0.13636363636363635, 'x[11] <= 0.5\ngini = 0.56\nsamples = 5\nvalue = [1, 3, 1]'),
Text(0.7720797720797721, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.7777777777777778, 0.045454545454545456, 'gini = 0.5\nsamples = 2\nvalue = [1, 0, 1]'),
Text(0.7863247863247863, 0.13636363636363635, 'x[15] <= 0.5\ngini = 0.46\nsamples = 54\nvalue = [3, 14, 37]'),
Text(0.7834757834757835, 0.045454545454545456, 'gini = 0.373\nsamples = 39\nvalue = [2, 7, 30]'),
Text(0.7891737891737892, 0.045454545454545456, 'gini = 0.56\nsamples = 15\nvalue = [1, 7, 7]'),
Text(0.7792022792022792, 0.3181818181818182, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
Text(0.8048433048433048, 0.4090909090909091, 'x[9] <= 3.5\ngini = 0.604\nsamples = 59\nvalue = [8, 23, 28]'),
Text(0.801994301994302, 0.3181818181818182, 'gini = 0.0\nsamples = 5\nvalue = [0, 5, 0]'),
Text(0.8076923076923077, 0.3181818181818182, 'x[29] <= 12.55\ngini = 0.598\nsamples = 54\nvalue = [8, 18, 28]'),
Text(0.8005698005698005, 0.22727272727272727, 'x[29] <= 12.367\ngini = 0.631\nsamples = 39\nvalue = [7, 16, 16]'),
Text(0.7977207977207977, 0.13636363636363635, 'x[23] <= 11.775\ngini = 0.633\nsamples = 35\nvalue = [7, 12, 16]'),
Text(0.7948717948717948, 0.045454545454545456, 'gini = 0.643\nsamples = 14\nvalue = [5, 6, 3]'),
Text(0.8005698005698005, 0.045454545454545456, 'gini = 0.526\nsamples = 21\nvalue = [2, 6, 13]'),
Text(0.8034188034188035, 0.13636363636363635, 'gini = 0.0\nsamples = 4\nvalue = [0, 4, 0]'),
Text(0.8148148148148148, 0.22727272727272727, 'x[1] <= 12.5\ngini = 0.338\nsamples = 15\nvalue = [1, 2, 12]'),
Text(0.8091168091168092, 0.13636363636363635, 'x[27] <= 11.0\ngini = 0.142\nsamples = 13\nvalue = [0, 1, 12]'),
Text(0.8062678062678063, 0.045454545454545456, 'gini = 0.0\nsamples = 11\nvalue = [0, 0, 11]'),
Text(0.8119658119658112, 0.045454545454545456, 'gini = 0.5\nsamples = 2\nvalue = [0, 1, 1]'),
Text(0.8205128205128205, 0.13636363636363635, 'x[17] <= 30.5\ngini = 0.5\nsamples = 2\nvalue = [1, 1, 0]'),
```

```
    Text(0.8176638176638177, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
    Text(0.8233618233618234, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
    Text(0.853988603988604, 0.5, 'x[27] <= 9.5\ngini = 0.533\nsamples = 57\nvalue = [16, 35, 6]'),
    Text(0.8361823361823362, 0.4090909090909091, 'x[33] <= 1.905\ngini = 0.291\nsamples = 30\nvalue = [2, 25, 3]'),
    Text(0.8262108262108262, 0.3181818181818182, 'x[17] <= 33.5\ngini = 0.091\nsamples = 21\nvalue = [1, 20, 0]'),
    Text(0.8233618233618234, 0.22727272727272727, 'gini = 0.0\nsamples = 19\nvalue = [0, 19, 0]'),
    Text(0.8290598290598291, 0.22727272727272727, 'x[22] <= 4.5\ngini = 0.5\nsamples = 2\nvalue = [1, 1, 0]'),
    Text(0.8262108262108262, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
    Text(0.8319088319088319, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
    Text(0.8461538461538461, 0.3181818181818182, 'x[23] <= 13.324\ngini = 0.568\nsamples = 9\nvalue = [1, 5, 3]'),
    Text(0.8404558404558404, 0.22727272727272727, 'x[31] <= 10.0\ngini = 0.278\nsamples = 6\nvalue = [0, 5, 1]'),
    Text(0.8376068376068376, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
    Text(0.8433048433048433, 0.13636363636363635, 'gini = 0.0\nsamples = 5\nvalue = [0, 5, 0]'),
    Text(0.8518518518518519, 0.22727272727272727, 'x[26] <= 7.5\ngini = 0.444\nsamples = 3\nvalue = [1, 0, 2]'),
    Text(0.8490028490028491, 0.13636363636363635, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
    Text(0.8547008547008547, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
    Text(0.8717948717948718, 0.4090909090909091, 'x[23] <= 12.344\ngini = 0.582\nsamples = 27\nvalue = [14, 10, 3]'),
    Text(0.8660968660968661, 0.3181818181818182, 'x[23] <= 11.85\ngini = 0.398\nsamples = 16\nvalue = [12, 3, 1]'),
    Text(0.8632478632478633, 0.22727272727272727, 'x[33] <= -1.31\ngini = 0.594\nsamples = 8\nvalue = [4, 3, 1]'),
    Text(0.8603988603988604, 0.13636363636363635, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
    Text(0.8660968660968661, 0.13636363636363635, 'x[1] <= 13.0\ngini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
    Text(0.8632478632478633, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
    Text(0.8689458689458689, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
    Text(0.8689458689458689, 0.22727272727272727, 'gini = 0.0\nsamples = 8\nvalue = [8, 0, 0]'),
    Text(0.8774928774928775, 0.3181818181818182, 'x[32] <= 2.1\ngini = 0.529\nsamples = 11\nvalue = [2, 7, 2]'),
    Text(0.8746438746438746, 0.22727272727272727, 'x[5] <= 7.5\ngini = 0.346\nsamples = 9\nvalue = [2, 7, 0]'),
    Text(0.8717948717948718, 0.13636363636363635, 'gini = 0.0\nsamples = 7\nvalue = [0, 7, 0]'),
    Text(0.8774928774928775, 0.13636363636363635, 'gini = 0.0\nsamples = 2\nvalue = [2, 0, 0]'),
    Text(0.8803418803418803, 0.22727272727272727, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
    Text(0.9408831908831908, 0.5909090909090909, 'x[29] <= 11.942\ngini = 0.41\nsamples = 456\nvalue = [43, 73, 340]'),
    Text(0.9123931623931624, 0.5, 'x[3] <= 14.0\ngini = 0.591\nsamples = 101\nvalue = [19, 26, 56]'),
    Text(0.896011396011396, 0.4090909090909091, 'x[21] <= 9.5\ngini = 0.53\nsamples = 84\nvalue = [9, 23, 52]'),
    Text(0.8888888888888888, 0.3181818181818182, 'x[2] <= 4.5\ngini = 0.265\nsamples = 20\nvalue = [1, 2, 17]'),
    Text(0.886039886039886, 0.22727272727272727, 'x[29] <= 11.155\ngini = 0.194\nsamples = 19\nvalue = [1, 1, 17]'),
    Text(0.8831908831908832, 0.13636363636363635, 'x[29] <= 11.071\ngini = 0.625\nsamples = 4\nvalue = [1, 1, 2]'),
    Text(0.8803418803418803, 0.045454545454545456, 'gini = 0.444\nsamples = 3\nvalue = [0, 1, 2]'),
    Text(0.886039886039886, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [1, 0, 0]'),
    Text(0.8888888888888888, 0.13636363636363635, 'gini = 0.0\nsamples = 15\nvalue = [0, 0, 15]'),
    Text(0.8917378917378918, 0.22727272727272727, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
    Text(0.9031339031339032, 0.3181818181818182, 'x[33] <= 1.265\ngini = 0.578\nsamples = 64\nvalue = [8, 21, 35]'),
    Text(0.8974358974358975, 0.22727272727272727, 'x[29] <= 10.845\ngini = 0.489\nsamples = 40\nvalue = [5, 8, 27]'),
    Text(0.8945868945868946, 0.13636363636363635, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
    Text(0.9002849002849003, 0.13636363636363635, 'x[28] <= 6.5\ngini = 0.431\nsamples = 37\nvalue = [5, 5, 27]'),
    Text(0.8974358974358975, 0.045454545454545456, 'gini = 0.64\nsamples = 10\nvalue = [2, 4, 4]'),
    Text(0.9031339031339032, 0.045454545454545456, 'gini = 0.261\nsamples = 27\nvalue = [3, 1, 23]'),
    Text(0.9088319088319088, 0.22727272727272727, 'x[10] <= 4.5\ngini = 0.58\nsamples = 24\nvalue = [3, 13, 8]'),
    Text(0.905982905982906, 0.13636363636363635, 'gini = 0.0\nsamples = 4\nvalue = [0, 0, 4]'),
    Text(0.9116809116809117, 0.13636363636363635, 'x[29] <= 11.422\ngini = 0.515\nsamples = 20\nvalue = [3, 13, 4]'),
    Text(0.9088319088319088, 0.045454545454545456, 'gini = 0.667\nsamples = 9\nvalue = [3, 3, 3]'),
    Text(0.9145299145299145, 0.045454545454545456, 'gini = 0.165\nsamples = 11\nvalue = [0, 10, 1]'),
    Text(0.9287749287749287, 0.4090909090909091, 'x[25] <= 3.0\ngini = 0.567\nsamples = 17\nvalue = [10, 3, 4]'),
    Text(0.9259259259259259, 0.3181818181818182, 'gini = 0.0\nsamples = 2\nvalue = [0, 0, 2]'),
    Text(0.9316239316239316, 0.3181818181818182, 'x[29] <= 11.469\ngini = 0.498\nsamples = 15\nvalue = [10, 3, 2]'),
    Text(0.9287749287749287, 0.22727272727272727, 'x[17] <= 28.5\ngini = 0.642\nsamples = 9\nvalue = [4, 3, 2]'),
    Text(0.9230769230769231, 0.13636363636363635, 'x[22] <= 14.5\ngini = 0.32\nsamples = 5\nvalue = [4, 0, 1]'),
    Text(0.9202279202279202, 0.045454545454545456, 'gini = 0.0\nsamples = 4\nvalue = [4, 0, 0]'),
    Text(0.9259259259259259, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
    Text(0.9344729344729344, 0.13636363636363635, 'x[31] <= 11.05\ngini = 0.375\nsamples = 4\nvalue = [0, 3, 1]'),
    Text(0.9316239316239316, 0.045454545454545456, 'gini = 0.0\nsamples = 3\nvalue = [0, 3, 0]'),
    Text(0.9373219373219374, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 0, 1]'),
    Text(0.9344729344729344, 0.22727272727272727, 'gini = 0.0\nsamples = 6\nvalue = [6, 0, 0]'),
    Text(0.9693732193732194, 0.5, 'x[28] <= 5.5\ngini = 0.338\nsamples = 355\nvalue = [24, 47, 284]'),
    Text(0.9572649572649573, 0.4090909090909091, 'x[25] <= 4.5\ngini = 0.615\nsamples = 39\nvalue = [11, 8, 20]'),
    Text(0.9544159544159544, 0.3181818181818182, 'x[17] <= 47.0\ngini = 0.567\nsamples = 34\nvalue = [6, 8, 20]'),
    Text(0.9515669515669516, 0.22727272727272727, 'x[16] <= 0.5\ngini = 0.531\nsamples = 32\nvalue = [4, 8, 20]'),
    Text(0.9458689458689459, 0.13636363636363635, 'x[27] <= 11.5\ngini = 0.616\nsamples = 22\nvalue = [4, 7, 11]'),
    Text(0.9430199430199431, 0.045454545454545456, 'gini = 0.551\nsamples = 15\nvalue = [4, 2, 9]'),
    Text(0.9487179487179487, 0.045454545454545456, 'gini = 0.408\nsamples = 7\nvalue = [0, 5, 2]'),
    Text(0.9572649572649573, 0.13636363636363635, 'x[33] <= 0.555\ngini = 0.18\nsamples = 10\nvalue = [0, 1, 9]'),
    Text(0.9544159544159544, 0.045454545454545456, 'gini = 0.0\nsamples = 1\nvalue = [0, 1, 0]'),
    Text(0.9601139601139601, 0.045454545454545456, 'gini = 0.0\nsamples = 9\nvalue = [0, 0, 9]'),
```

```python
from sklearn.metrics import confusion_matrix

# PRINT THE CONFUSION MATRIX
print("Confusion Matrix")
cm = confusion_matrix(Y_test, y_pred)
print(cm)
```
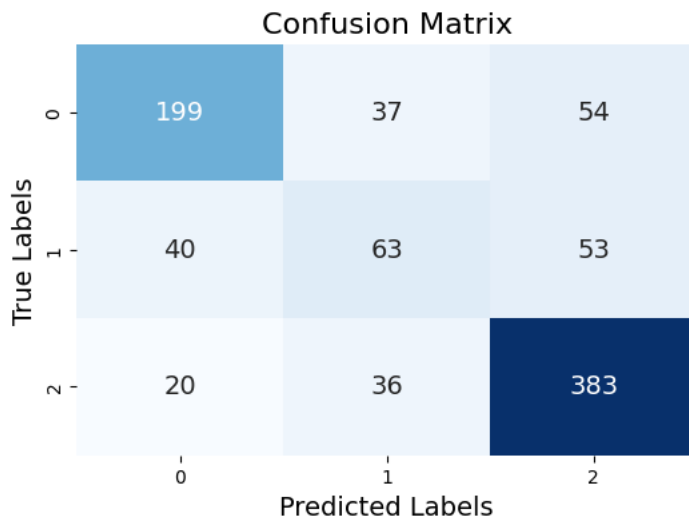
```
    Confusion Matrix
    [[199  37  54]
     [ 40  63  53]
     [ 20  36 383]]
```

```
plt.figure(figsize = (6, 4))
sns.heatmap(cm, annot = True, fmt = 'd', cmap = 'Blues', cbar = False, annot_kws = {'size' : 14})
plt.xlabel('Predicted Labels', fontsize = 14)
plt.ylabel('True Labels', fontsize = 14)
plt.title('Confusion Matrix', fontsize = 16)
plt.show()
```

Confusion Matrix

Support Vector Machine

```
from sklearn import svm
clf = svm.SVC()
clf.fit(X_train, Y_train)
```

```
  ▼ SVC
  SVC()
```

```
y_pred = clf.predict(X_test)
score = accuracy_score(Y_test, y_pred)
accuracy = score*100
print(accuracy)
```
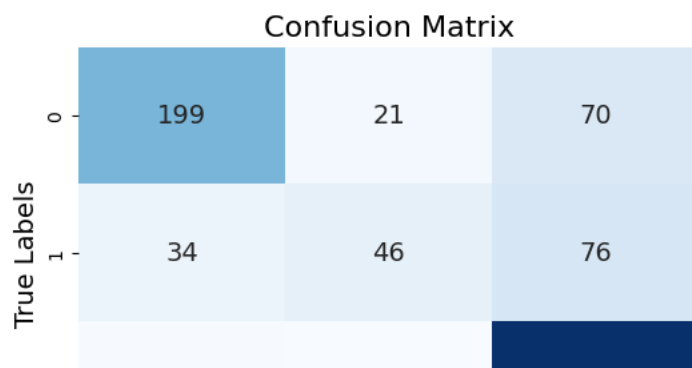
```
    74.46327683615819
```

```
from sklearn.metrics import confusion_matrix

# PRINT THE CONFUSION MATRIX
print("Confusion Matrix")
cm = confusion_matrix(Y_test, y_pred)
print(cm)
```

```
    Confusion Matrix
    [[199  21  70]
     [ 34  46  76]
     [ 14  11 414]]
```

```
plt.figure(figsize = (6, 4))
sns.heatmap(cm, annot = True, fmt = 'd', cmap = 'Blues', cbar = False, annot_kws = {'size' : 14})
plt.xlabel('Predicted Labels', fontsize = 14)
plt.ylabel('True Labels', fontsize = 14)
plt.title('Confusion Matrix', fontsize = 16)
plt.show()
```

Confusion Matrix

True Labels			
0	199	21	70
1	34	46	76

Random Forest

```
from sklearn.ensemble import RandomForestClassifier
clf = RandomForestClassifier(max_depth=20, random_state=0, n_estimators = 124)
clf.fit(X_train, Y_train)
```

```
              ▾              RandomForestClassifier
    RandomForestClassifier(max_depth=20, n_estimators=124, random_state=0)
```

```
y_pred = clf.predict(X_test)
score = accuracy_score(Y_test, y_pred)
accuracy = score*100
print(accuracy)
```
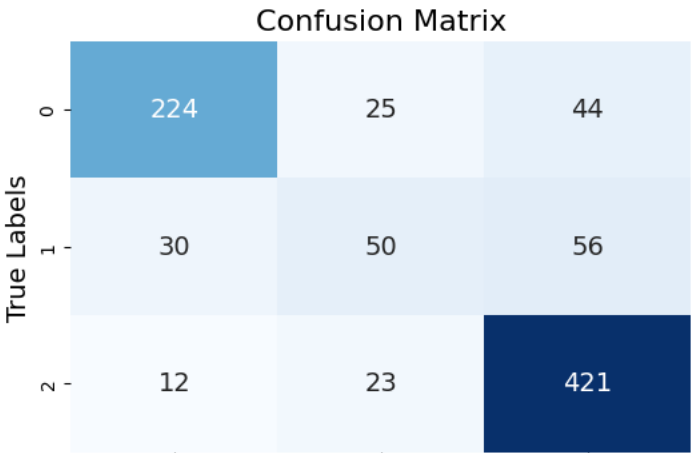
```
    77.06214689265536
```

```
from sklearn.metrics import confusion_matrix

# PRINT THE CONFUSION MATRIX
print("Confusion Matrix")
cm = confusion_matrix(Y_test, y_pred)
print(cm)
```

```
    Confusion Matrix
    [[213  17  60]
     [ 30  64  62]
     [ 15  19 405]]
```

```
plt.figure(figsize = (6, 4))
sns.heatmap(cm, annot = True, fmt = 'd', cmap = 'Blues', cbar = False, annot_kws = {'size' : 14})
plt.xlabel('Predicted Labels', fontsize = 14)
plt.ylabel('True Labels', fontsize = 14)
plt.title('Confusion Matrix', fontsize = 16)
plt.show()
```

Confusion Matrix

	0	1	2
0	224	25	44
1	30	50	56
2	12	23	421

True Labels

✓ 0s   completed at 3:58 AM