wajveo9cb

April 18, 2024

Name :- Arun M Register No :- 23122110

```python
[129]: import pandas as pd
       from sklearn import svm
       import matplotlib.pyplot as plt
       from sklearn import tree
       from sklearn.tree import DecisionTreeClassifier
       from sklearn.svm import SVC
       from sklearn.metrics import confusion_matrix
```

```python
[ ]:
```

```python
[130]: import pandas as pd
       df=pd.read_csv("nasav3.csv")
       print(df)
```

```
      Absolute Magnitude  Est Dia in KM(min)  Est Dia in KM(max)  \
0                 21.600            0.127220            0.284472
1                 21.300            0.146068            0.326618
2                 20.300            0.231502            0.517654
3                 27.400            0.008801            0.019681
4                 21.600            0.127220            0.284472
...                  ...                 ...                 ...
4682              23.900            0.044112            0.098637
4683              28.200            0.006089            0.013616
4684              22.700            0.076658            0.171412
4685              21.800            0.116026            0.259442
4686              19.109            0.400641            0.895860

      Epoch Date Close Approach  Relative Velocity km per hr  Miles per hour  \
0                   7.890000e+11                  22017.00380    13680.509940
1                   7.890000e+11                  65210.34609    40519.173110
2                   7.900000e+11                  27326.56018    16979.661800
3                   7.900000e+11                  40225.94819    24994.839860
4                   7.900000e+11                  35426.99179    22012.954980
...                          ...                          ...             ...
4682                1.470000e+12                  79755.35427    49556.875550
4683                1.470000e+12                  11610.53958     7214.337772
```

1

```
4684                  1.470000e+12                         25889.91063     16086.983630
4685                  1.470000e+12                         40867.52231     25393.489070
4686                  1.470000e+12                        129408.66630     80409.512650


      Miss Dist.(kilometers)  Orbit Uncertainity  Minimum Orbit Intersection  \
0               6.275369e+07                   5                    0.025282
1               5.729815e+07                   3                    0.186935
2               7.622912e+06                   0                    0.043058
3               4.268362e+07                   6                    0.005512
4               6.101082e+07                   1                    0.034798
...                      ...                 ...                         ...
4682            6.187511e+06                   8                    0.019777
4683            9.677324e+05                   6                    0.006451
4684            9.126775e+06                   6                    0.059972
4685            3.900908e+07                   5                    0.177510
4686            6.916986e+07                   6                    0.051777


      Jupiter Tisserand Invariant  …  Inclination  Asc Node Longitude  \
0                           4.634  …     6.025981          314.373913
1                           5.457  …    28.412996          136.717242
2                           4.557  …     4.237961          259.475979
3                           5.093  …     7.905894           57.173266
4                           5.154  …    16.793382           84.629307
...                           …  …         …                   …
4682                        5.156  …    39.880491          164.183305
4683                        5.742  …     5.360249          345.225230
4684                        4.410  …     4.405467           37.026468
4685                        4.477  …    21.080244          163.802909
4686                        4.108  …    53.574922          187.642183


      Orbital Period  Perihelion Distance  Perihelion Arg  Aphelion Dist  \
0         609.599786             0.808259       57.257470       2.005764
1         425.869294             0.718200      313.091975       1.497352
2         643.580228             0.950791      248.415038       1.966857
3         514.082140             0.983902       18.707701       1.527904
4         495.597821             0.967687      158.263596       1.483543
...              …                    …               …               …
4682      457.179984             0.741558      276.395697       1.581299
4683      407.185767             0.996434       42.111064       1.153835
4684      690.054279             0.965760      274.692712       2.090708
4685      662.048343             1.185467      180.346090       1.787733
4686      653.679098             0.876110      222.436688       2.071980


      Perihelion Time  Mean Anomaly  Mean Motion  Hazardous
0         2458161.642    264.837533     0.590551          1
1         2457794.969    173.741112     0.845330          0
2         2458120.468    292.893654     0.559371          1
3         2457902.337     68.741007     0.700277          0
```

```
4            2457814.455      135.142133        0.726395             1
...               ...             ...              ...              ...
4682         2457708.228      304.306024        0.787436             0
4683         2458087.617      282.978786        0.884117             0
4684         2458300.480      203.501147        0.521698             0
4685         2458288.261      203.524965        0.543767             0
4686         2458318.587      184.820424        0.550729             0

[4687 rows x 23 columns]
```

[131]:
```python
print(df.columns)
```

```
Index(['Absolute Magnitude', 'Est Dia in KM(min)', 'Est Dia in KM(max)',
       'Epoch Date Close Approach', 'Relative Velocity km per hr',
       'Miles per hour', 'Miss Dist.(kilometers)', 'Orbit Uncertainity',
       'Minimum Orbit Intersection', 'Jupiter Tisserand Invariant',
       'Epoch Osculation', 'Eccentricity', 'Semi Major Axis', 'Inclination',
       'Asc Node Longitude', 'Orbital Period', 'Perihelion Distance',
       'Perihelion Arg', 'Aphelion Dist', 'Perihelion Time', 'Mean Anomaly',
       'Mean Motion', 'Hazardous'],
      dtype='object')
```

[132]:
```python
features = ['Absolute Magnitude', 'Est Dia in KM(min)', 'Est Dia in KM(max)',
       'Epoch Date Close Approach', 'Relative Velocity km per hr',
       'Miles per hour', 'Miss Dist.(kilometers)', 'Orbit Uncertainity',
       'Minimum Orbit Intersection', 'Jupiter Tisserand Invariant',
       'Epoch Osculation', 'Eccentricity', 'Semi Major Axis', 'Inclination',
       'Asc Node Longitude', 'Orbital Period', 'Perihelion Distance',
       'Perihelion Arg', 'Aphelion Dist', 'Perihelion Time', 'Mean Anomaly',
       'Mean Motion']

X = df[features]
y = df['Hazardous']

print(X)
print(y)
```

```
       Absolute Magnitude   Est Dia in KM(min)   Est Dia in KM(max)  \
0                  21.600             0.127220             0.284472
1                  21.300             0.146068             0.326618
2                  20.300             0.231502             0.517654
3                  27.400             0.008801             0.019681
4                  21.600             0.127220             0.284472
...                   ...                  ...                  ...
4682               23.900             0.044112             0.098637
4683               28.200             0.006089             0.013616
4684               22.700             0.076658             0.171412
4685               21.800             0.116026             0.259442
```

		19.109	0.400641	0.895860
4686				

	Epoch Date Close Approach	Relative Velocity km per hr	Miles per hour \
0	7.890000e+11	22017.00380	13680.509940
1	7.890000e+11	65210.34609	40519.173110
2	7.900000e+11	27326.56018	16979.661800
3	7.900000e+11	40225.94819	24994.839860
4	7.900000e+11	35426.99179	22012.954980
…	…	…	…
4682	1.470000e+12	79755.35427	49556.875550
4683	1.470000e+12	11610.53958	7214.337772
4684	1.470000e+12	25889.91063	16086.983630
4685	1.470000e+12	40867.52231	25393.489070
4686	1.470000e+12	129408.66630	80409.512650

	Miss Dist.(kilometers)	Orbit Uncertainity	Minimum Orbit Intersection \
0	6.275369e+07	5	0.025282
1	5.729815e+07	3	0.186935
2	7.622912e+06	0	0.043058
3	4.268362e+07	6	0.005512
4	6.101082e+07	1	0.034798
…	…	…	…
4682	6.187511e+06	8	0.019777
4683	9.677324e+05	6	0.006451
4684	9.126775e+06	6	0.059972
4685	3.900908e+07	5	0.177510
4686	6.916986e+07	6	0.051777

	Jupiter Tisserand Invariant	…	Semi Major Axis	Inclination \
0	4.634	…	1.407011	6.025981
1	5.457	…	1.107776	28.412996
2	4.557	…	1.458824	4.237961
3	5.093	…	1.255903	7.905894
4	5.154	…	1.225615	16.793382
…	…	… …	…	…
4682	5.156	…	1.161429	39.880491
4683	5.742	…	1.075134	5.360249
4684	4.410	…	1.528234	4.405467
4685	4.477	…	1.486600	21.080244
4686	4.108	…	1.474045	53.574922

	Asc Node Longitude	Orbital Period	Perihelion Distance	Perihelion Arg \
0	314.373913	609.599786	0.808259	57.257470
1	136.717242	425.869294	0.718200	313.091975
2	259.475979	643.580228	0.950791	248.415038
3	57.173266	514.082140	0.983902	18.707701
4	84.629307	495.597821	0.967687	158.263596
…	…	…	…	…

```
4682          164.183305          457.179984                    0.741558          276.395697
4683          345.225230          407.185767                    0.996434           42.111064
4684           37.026468          690.054279                    0.965760          274.692712
4685          163.802909          662.048343                    1.185467          180.346090
4686          187.642183          653.679098                    0.876110          222.436688

        Aphelion Dist  Perihelion Time  Mean Anomaly  Mean Motion
0            2.005764      2458161.642     264.837533      0.590551
1            1.497352      2457794.969     173.741112      0.845330
2            1.966857      2458120.468     292.893654      0.559371
3            1.527904      2457902.337      68.741007      0.700277
4            1.483543      2457814.455     135.142133      0.726395
...               ...              ...           ...           ...
4682         1.581299      2457708.228     304.306024      0.787436
4683         1.153835      2458087.617     282.978786      0.884117
4684         2.090708      2458300.480     203.501147      0.521698
4685         1.787733      2458288.261     203.524965      0.543767
4686         2.071980      2458318.587     184.820424      0.550729

[4687 rows x 22 columns]
0       1
1       0
2       1
3       0
4       1
       ..
4682    0
4683    0
4684    0
4685    0
4686    0
Name: Hazardous, Length: 4687, dtype: int64
```

```python
import numpy as np
X=np.array(df)
y=X[:,22]
X=X[:,:21]

print(X)
print(y)
```

```
[[2.16000000e+01 1.27219879e-01 2.84472297e-01 … 2.00576367e+00
  2.45816164e+06 2.64837533e+02]
 [2.13000000e+01 1.46067964e-01 3.26617897e-01 … 1.49735229e+00
  2.45779497e+06 1.73741112e+02]
 [2.03000000e+01 2.31502122e-01 5.17654482e-01 … 1.96685667e+00
  2.45812047e+06 2.92893654e+02]
 …
```

```
[2.27000000e+01 7.66575570e-02 1.71411509e-01 … 2.09070784e+00
  2.45830048e+06 2.03501147e+02]
[2.18000000e+01 1.16025908e-01 2.59441818e-01 … 1.78773308e+00
  2.45828826e+06 2.03524965e+02]
[1.91090000e+01 4.00640618e-01 8.95859655e-01 … 2.07198006e+00
  2.45831859e+06 1.84820424e+02]]
[1. 0. 1. … 0. 0. 0.]
```

```
[134]: from sklearn.preprocessing import StandardScaler
       scaler=StandardScaler()
       scaler.fit(X)
       X=scaler.transform(X)
```

```
[135]: from sklearn.model_selection import train_test_split
       X_train, X_test, y_train, y_test = train_test_split(X,y ,
                                          random_state=104,
                                          test_size=0.25,
                                          shuffle=True)
```

```
[136]: dtree = DecisionTreeClassifier(criterion='gini',max_depth=3)
       dtree = dtree.fit(X_train, y_train)
```

```
[137]: tree.plot_tree(dtree, feature_names=features,fontsize=7)
       print(X_train.shape)
```

```
(3515, 21)
```
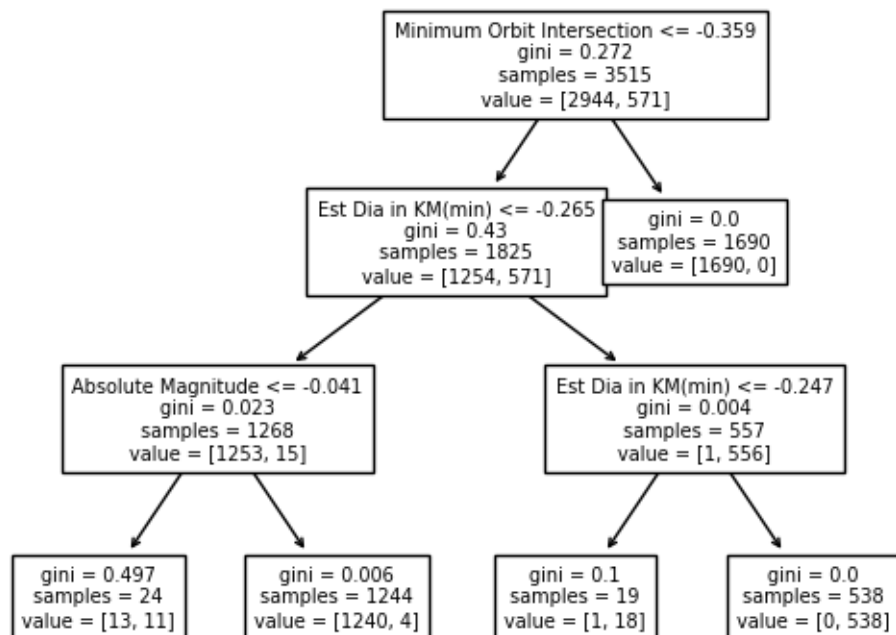
```python
[138]: from sklearn.model_selection import train_test_split
       from sklearn.metrics import accuracy_score

       # For Decesion TReee
       # Predict labels for test set
       y_pred = dtree.predict(X_test)
       # Calculate accuracy
       accuracy = accuracy_score(y_test, y_pred)
       print("Accuracy:", accuracy)
```

Accuracy: 0.9948805460750854

```python
[ ]:
```

```python
[139]: Sclf = SVC(kernel='sigmoid')
       Sclf = Sclf.fit(X_train, y_train)
```

```python
[140]: # For For support vector classifier
       # Predict labels for test set
       y_pred = Sclf.predict(X_test)
       # Calculate accuracy
       accuracy = accuracy_score(y_test, y_pred)
       print("Accuracy:", accuracy)
```

Accuracy: 0.8651877133105802

```python
[141]: cm=confusion_matrix(y_test,y_pred)
       print("confusion_matrix")
       print(cm)
```

confusion_matrix
[[907 81]
 [77 107]]