

In [1]: #EXP-11

In [2]: #Aim: To perform and analysis of decision tree

In [3]:  
 # Name: Payal Devanand Manwar  
 # Roll no.: 37  
 # Sec:A  
 # Subject: ET1  
 # Date: 29/09/2025

In [4]:  
 import pandas as pd  
 import numpy as np

In [5]: import os

In [6]: os.getcwd()

Out[6]: 'C:\\\\Users\\\\USER'

In [7]: os.chdir("C:\\\\Users\\\\USER\\\\Desktop")

In [8]: data=pd.read\_csv("heart - heart.csv")

In [9]: data.head()

Out[9]:

	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	thal	target
<b>0</b>	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
<b>1</b>	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
<b>2</b>	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
<b>3</b>	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
<b>4</b>	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0

In [10]: data.tail()

Out[10]:

	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	thal	target
<b>1020</b>	59	1	1	140	221	0	1	164	1	0.0	2	0	2	1
<b>1021</b>	60	1	0	125	258	0	0	141	1	2.8	1	1	3	0
<b>1022</b>	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
<b>1023</b>	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
<b>1024</b>	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

# decision tree

```
In [11]: x=data.drop("target", axis=1)
y=data["target"]
```

```
In [12]: #splitting the data into training and testing data sets
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2 ,random_state=42)
```

```
In [14]: from sklearn.tree import DecisionTreeClassifier
```

```
In [18]: dt=DecisionTreeClassifier()
```

```
In [19]: dt.fit(x_train, y_train)
```

```
Out[19]: DecisionTreeClassifier()
```

```
In [20]: y_pred4=dt.predict(x_test)
```

```
In [21]: from sklearn.metrics import accuracy_score
```

```
In [22]: accuracy_score (y_test,y_pred4)
```

```
Out[22]: 0.9853658536585366
```

```
In [23]: x_train
```

	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	thal
835	49	1	2	118	149	0	0	126	0	0.8	2	3	2
137	64	0	0	180	325	0	1	154	1	0.0	2	0	2
534	54	0	2	108	267	0	0	167	0	0.0	2	0	2
495	59	1	0	135	234	0	1	161	0	0.5	1	0	3
244	51	1	2	125	245	1	0	166	0	2.4	1	0	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...
700	41	1	2	130	214	0	0	168	0	2.0	1	0	2
71	61	1	0	140	207	0	0	138	1	1.9	2	1	3
106	51	1	0	140	299	0	1	173	1	1.6	2	0	3
270	43	1	0	110	211	0	1	161	0	0.0	2	0	3
860	52	1	0	112	230	0	1	160	0	0.0	2	1	2

820 rows × 13 columns

In [24]: `x_test`

```
Out[24]:    age  sex  cp  trestbps  chol  fbs  restecg  thalach  exang  oldpeak  slope  ca  thal
      527   62   0   0     124   209   0     1     163     0     0.0     2   0     2
      359   53   0   2     128   216   0     0     115     0     0.0     2   0     0
      447   55   1   0     160   289   0     0     145     1     0.8     1   1     3
      31    50   0   1     120   244   0     1     162     0     1.1     2   0     2
      621   48   1   0     130   256   1     0     150     1     0.0     2   2     3
      ...
      832   68   1   2     118   277   0     1     151     0     1.0     2   1     3
      796   41   1   1     135   203   0     1     132     0     0.0     1   0     1
      644   44   1   2     120   226   0     1     169     0     0.0     2   0     2
      404   61   1   0     140   207   0     0     138     1     1.9     2   1     3
      842   58   1   2     112   230   0     0     165     0     2.5     1   1     3
```

205 rows × 13 columns

In [25]: `y_train`

```
Out[25]: 835    0
137    1
534    1
495    1
244    1
...
700    1
71     0
106    0
270    1
860    0
Name: target, Length: 820, dtype: int64
```

In [26]: `y_test`

```
Out[26]: 527    1
359    1
447    0
31     1
621    0
...
832    1
796    1
644    1
404    0
842    0
Name: target, Length: 205, dtype: int64
```

In [ ]: