In [27]:
```
import pandas as pd
import numpy as np

heart_disease = pd.read_csv('heart.csv')
```

In [20]:
```
heart_disease
```

Out[20]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca
0	52	1	0	125	212	0	1	168	0	1.0	2	2
1	53	1	0	140	203	1	0	155	1	3.1	0	0
2	70	1	0	145	174	0	1	125	1	2.6	0	0
3	61	1	0	148	203	0	1	161	0	0.0	2	1
4	62	0	0	138	294	1	1	106	0	1.9	1	3
...	...	...	...	...	...	...	...	...	...	...	...	...
1020	59	1	1	140	221	0	1	164	1	0.0	2	0
1021	60	1	0	125	258	0	0	141	1	2.8	1	1
1022	47	1	0	110	275	0	0	118	1	1.0	1	1
1023	50	0	0	110	254	0	0	159	0	0.0	2	0
1024	54	1	0	120	188	0	1	113	0	1.4	1	1

1025 rows × 14 columns

In [21]:
```
X = heart_disease.drop("target", axis=1)
y = heart_disease["target"]
```

In [22]:
```
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier()

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)

model.fit(X_train, y_train)
model.score(X_test, y_test)
```

Out[22]: 1.0

In [26]:
```
X_test
```

Out[26]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca
685	63	0	0	150	407	0	0	154	0	4.0	1	3
136	55	0	1	132	342	0	1	166	0	1.2	2	0
721	45	1	0	115	260	0	0	185	0	0.0	2	0
252	55	1	0	132	353	0	1	132	1	1.2	1	1
321	48	0	2	130	275	0	1	139	0	0.2	2	0
...	...	...	...	...	...	...	...	...	...	...	...	...
520	59	1	0	140	177	0	1	162	1	0.0	2	1
702	71	0	1	160	302	0	1	162	0	0.4	2	2
1005	55	0	0	128	205	0	2	130	1	2.0	1	1
277	44	1	1	130	219	0	0	188	0	0.0	2	0
894	51	1	0	140	299	0	1	173	1	1.6	2	0

205 rows × 13 columns

In [23]:
```
y_preds = model.predict(X_test)
```

In [24]:
```
y_test
```

Out[24]:
```
685    0
136    1
721    1
252    0
321    1
      ..
520    0
702    1
1005   0
277    1
894    0
Name: target, Length: 205, dtype: int64
```

In [25]:
```
from sklearn.metrics import accuracy_score

accuracy_score(y_test, y_preds)
```

Out[25]:
```
1.0
```

In [40]:
```
try_data = np.array([[50,1,2,129,196,0,1,163,0,0,2,0,2]])
try_data_df = pd.DataFrame(try_data)
try_data_df
```

Out[40]:

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	50	1	2	129	196	0	1	163	0	0	2	0	2

In [41]:
```
heart_disease_predict = model.predict(try_data)
heart_disease_predict
```

```
/home/jovyan/.local/lib/python3.10/site-packages/sklearn/base.py:439: UserWarnin
g: X does not have valid feature names, but RandomForestClassifier was fitted wit
h feature names
  warnings.warn(
```

Out[41]:
```
array([1])
```

In [43]:
```
import pickle

pickle.dump(model, open("random_forest_model_1.pkl", "wb"))
```

In [ ]:
```

```