

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import confusion_matrix
from sklearn.metrics import f1_score
from sklearn.metrics import accuracy_score
```

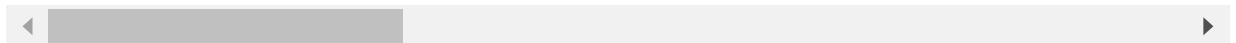
```
In [3]: df = pd.read_csv("C:\\Users\\Anjali\\OneDrive\\Machine Learning\\Breast_cancer.csv")
```

```
In [4]: df.head()
```

```
Out[4]:
```

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean
0	842302	M	17.99	10.38	122.80	1001.0	0.11840
1	842517	M	20.57	17.77	132.90	1326.0	0.08474
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960
3	84348301	M	11.42	20.38	77.58	386.1	0.14250
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030

5 rows × 33 columns



```
In [5]: df.shape
```

```
Out[5]: (569, 33)
```

```
In [6]: df.keys()
```

```
Out[6]: Index(['id', 'diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean',
              'area_mean', 'smoothness_mean', 'compactness_mean', 'concavity_mean',
              'concave points_mean', 'symmetry_mean', 'fractal_dimension_mean',
              'radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
              'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
              'fractal_dimension_se', 'radius_worst', 'texture_worst',
              'perimeter_worst', 'area_worst', 'smoothness_worst',
              'compactness_worst', 'concavity_worst', 'concave points_worst',
              'symmetry_worst', 'fractal_dimension_worst', 'Unnamed: 32'],
              dtype='object')
```

```
In [7]: df.isnull().sum()
```

```
Out[7]: id                                0
diagnosis                                0
radius_mean                             0
texture_mean                             0
perimeter_mean                           0
area_mean                                0
smoothness_mean                           0
compactness_mean                           0
```

```
concavity_mean      0
concave points_mean 0
symmetry_mean       0
fractal_dimension_mean 0
radius_se           0
texture_se          0
perimeter_se        0
area_se             0
smoothness_se       0
compactness_se      0
concavity_se        0
concave points_se   0
symmetry_se         0
fractal_dimension_se 0
radius_worst        0
texture_worst       0
perimeter_worst     0
area_worst          0
smoothness_worst    0
compactness_worst   0
concavity_worst     0
concave points_worst 0
symmetry_worst      0
fractal_dimension_worst 0
Unnamed: 32         569
dtype: int64
```

In [8]: `df = df.dropna(axis = 1)`

In [9]: `df.shape`

Out[9]: (569, 32)

In [10]: `df = df.drop(columns = 'id')`

In [11]: `df.shape`

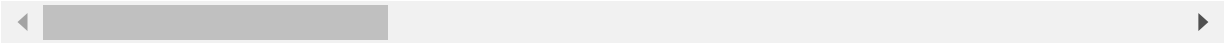
Out[11]: (569, 31)

In [12]: `df.head()`

Out[12]:

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactn
0	M	17.99	10.38	122.80	1001.0	0.11840	
1	M	20.57	17.77	132.90	1326.0	0.08474	
2	M	19.69	21.25	130.00	1203.0	0.10960	
3	M	11.42	20.38	77.58	386.1	0.14250	
4	M	20.29	14.34	135.10	1297.0	0.10030	

5 rows × 31 columns



In [13]: `df['diagnosis'] = df['diagnosis'].map({'B':0, 'M':1})`

Out[22]: 0.9

In [23]: `accuracy_score(predict, y_test)`

Out[23]: 0.9298245614035088

In [24]: `accuracy = model.score(X_train, y_train)`
`accuracy`

Out[24]: 0.9321608040201005

In []: