```
In [2]:  1  import numpy as np
         2  import pandas as pd
         3  from sklearn.datasets import load_breast_cancer
         4  from sklearn.model_selection import train_test_split
         5  from sklearn.ensemble import RandomForestClassifier
         6  from sklearn.metrics import accuracy_score, classification_report
```

```
In [3]:  1  data = load_breast_cancer()
         2  X = data.data
         3  y = data.target
```

```
In [4]:  1  rf_classifier = RandomForestClassifier(n_estimators=100, random_state=42)
```

```
In [5]:  1  X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

```
In [6]:  1  rf_classifier.fit(X_train, y_train)
```

Out[6]:
```
  ▾              RandomForestClassifier

RandomForestClassifier(random_state=42)
```

```
In [7]:  1  y_pred = rf_classifier.predict(X_test)
         2  accuracy = accuracy_score(y_test, y_pred)
         3  print(accuracy)
         4  report = classification_report(y_test, y_pred)
```

```
0.9912280701754386
```

```
In [8]:  1  print(report)
         2
```

```
              precision    recall  f1-score   support

           0       1.00      0.97      0.99        40
           1       0.99      1.00      0.99        74

    accuracy                           0.99       114
   macro avg       0.99      0.99      0.99       114
weighted avg       0.99      0.99      0.99       114
```

knn

```
In [9]:    1  import numpy as np
           2  from sklearn.neighbors import KNeighborsClassifier
           3  from sklearn.model_selection import train_test_split
           4  from sklearn.metrics import accuracy_score
           5  from sklearn.cluster import KMeans
```

```
In [10]:   1  from sklearn.datasets import load_iris
           2  iris = load_iris()
           3  X = iris.data
           4  y = iris.target
           5
```

```
In [11]:   1  X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

```
In [12]:   1  k = 3 # Set the value of k
           2  knn_classifier = KNeighborsClassifier(n_neighbors=k)
           3  knn_classifier.fit(X_train, y_train)
           4
```

```
Out[12]:   ▼          KNeighborsClassifier

           KNeighborsClassifier(n_neighbors=3)
```

```
In [13]:   1  y_pred = knn_classifier.predict(X_test)
           2
```

```
In [14]:   1  accuracy = accuracy_score(y_test, y_pred)
           2  print(f"Accuracy: {accuracy * 100:.2f}%")
           3
```

```
Accuracy: 93.33%
```

kmean

```
In [18]:   1  data = pd.read_csv("C:/Users/HARISH BJ/Downloads/breast-cancer.csv")
```

```
In [19]:   1  selected_columns = ['id', 'diagnosis', 'radius_mean']
           2  X=data[selected_columns]
           3
```

```
In [20]:   1  X.isnull().sum()
```

```
Out[20]:   id             0
           diagnosis      0
           radius_mean    0
           dtype: int64
```

In [22]:
```python
k = 3
```

In [26]:
```python
kmeans = KMeans(n_clusters=K, random_state=0)
cluster_assignments = kmeans.fit_predict(X_scaled)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[26], line 2
      1 kmeans = KMeans(n_clusters=K, random_state=0)
----> 2 cluster_assignments = kmeans.fit_predict(X_scaled)

NameError: name 'X_scaled' is not defined
```

In [ ]:
```python

```