410-Asif-Sayyed-ML-Case-Study-Question-3

In [7]:

```python
import numpy as np

# Define the dataset with numeric labels
dataset = np.array([
    [8.27, 5.59, 1],  # Label '+'
    [1.58, 5.87, 0],  # Label '-'
    [5.92, 5.87, 0],
    [9.44, 5.83, 1],
    [2.11, 5.57, 0],
    [4.71, 5.94, 1],
    [3.82, 5.84, 1],
    [6.98, 5.91, 0],
    [3.15, 5.42, 0],
    [8.9, 5.94, 0],
    [7.65, 5.77, 1],
    [9.83, 5.29, 0],
    [1.94, 5.36, 1],
    [7.13, 5.28, 0],
    [5.77, 5.47, 0],
    [4.36, 5.31, 1],
    [5.09, 5.65, 0],
    [3.42, 5.24, 1],
    [2.76, 5.71, 1],
    [9.6, 5.52, 0]
])

# Function to predict labels using KNN algorithm and return indices of closest neighbors
def predict_labels_with_indices(dataset, points, k):
    predicted_labels = []
    nearest_neighbors_indices = []
    for point in points:
        distances = np.sqrt(np.sum((dataset[:, :2] - point) ** 2, axis=1))  # Calculate distances
        nearest_indices = np.argsort(distances)[:k]  # Indices of k nearest neighbors
        nearest_neighbors_indices.append(nearest_indices)  # Store indices of nearest neighbors
        nearest_labels = dataset[nearest_indices, 2].astype(int)  # Convert labels to integers
        predicted_label = np.argmax(np.bincount(nearest_labels))  # Majority label
        predicted_labels.append(predicted_label)
    return predicted_labels, nearest_neighbors_indices

# Points to predict
points = np.array([
    [7.81, 5.33],
    [9.43, 5.29]
])

# Predict labels and find indices of closest neighbors for the points with K=3
predicted_labels, nearest_neighbors_indices = predict_labels_with_indices(dataset, points, 3)

# Map predicted label index to actual label
label_mapping = {1: '+', 0: '-'}  # Mapping 1 to '+' and 0 to '-'
predicted_labels = [label_mapping[label_index] for label_index in predicted_labels]

print("Predicted labels for point1 and point2 (with K=3):", predicted_labels)
print("Indices of closest neighbors for point1:", nearest_neighbors_indices[0])
print("Indices of closest neighbors for point2:", nearest_neighbors_indices[1])
```

```
Predicted labels for point1 and point2 (with K=3): ['+', '-']
Indices of closest neighbors for point1: [10  0 13]
Indices of closest neighbors for point2: [19 11  3]
```