K-Nearest Neighbors (KNN) Classification Example

Given dataset:

Point A: (ProgrammingScore = 80, MathsScore = 90, EnglishScore = 75)
Point B: (ProgrammingScore = 65, MathsScore = 75, EnglishScore = 80)
Point C: (ProgrammingScore = 95, MathsScore = 85, EnglishScore = 70)
New Point X: (ProgrammingScore = 70, MathsScore = 80, EnglishScore = 85)

K-Nearest Neighbors for Different Values of k

For $k = 1$

Step 1: Choose $k = 1$. Step 2: Find the closest point to X based on the distances.

$$d_{AX} = \sqrt{(70 - 80)^2 + (80 - 90)^2 + (85 - 75)^2}$$
$$d_{BX} = \sqrt{(70 - 65)^2 + (80 - 75)^2 + (85 - 80)^2}$$
$$d_{CX} = \sqrt{(70 - 95)^2 + (80 - 85)^2 + (85 - 70)^2}$$

The smallest distance is d_{BX}, so the closest point to X is B. Step 3: Classify X based on the class of the closest point. Since Point B belongs to Class 1, X also belongs to Class 1.

For $k = 2$

Step 1: Choose $k = 2$. Step 2: Find the two closest points to X based on the distances.

$$d_{AX} = \sqrt{(70 - 80)^2 + (80 - 90)^2 + (85 - 75)^2}$$
$$d_{BX} = \sqrt{(70 - 65)^2 + (80 - 75)^2 + (85 - 80)^2}$$
$$d_{CX} = \sqrt{(70 - 95)^2 + (80 - 85)^2 + (85 - 70)^2}$$

The two smallest distances are d_{BX} and d_{AX}, so the closest points to X are B and A. Step 3: Classify X based on the classes of the two closest points. Both Point B and Point A belong to Class 1, so X also belongs to Class 1.

For $k = 3$

Step 1: Choose $k = 3$. Step 2: Find the three closest points to X based on the distances.

$$d_{AX} = \sqrt{(70-80)^2 + (80-90)^2 + (85-75)^2}$$
$$d_{BX} = \sqrt{(70-65)^2 + (80-75)^2 + (85-80)^2}$$
$$d_{CX} = \sqrt{(70-95)^2 + (80-85)^2 + (85-70)^2}$$

The three smallest distances are d_{BX}, d_{AX}, and d_{CX}, so the closest points to X are B, A, and C. Step 3: Classify X based on the classes of the three closest points. Two out of the three closest points (Point B and Point A) belong to Class 1, so X belongs to Class 1 as well.

1 Explanation on Class 1 and Class 0

In the k-nearest neighbors (KNN) algorithm, when we talk about "Class 1," we are referring to one of the classes to which the data points in the dataset belong. In the given example, we haven't explicitly defined the classes, so let's assume that Class 1 represents the "positive" or "approved" class, while Class 0 represents the "negative" or "rejected" class.

So, when we say that a point belongs to Class 1, it means that based on the features of that point (e.g., ProgrammingScore, MathsScore, EnglishScore), it is predicted to be in the positive or approved category. Conversely, if a point belongs to Class 0, it is predicted to be in the negative or rejected category.

In the provided example, we're classifying whether a student is likely to be approved (Class 1) or rejected (Class 0) based on their scores in programming, maths, and English.