K-Nearest Neighbor

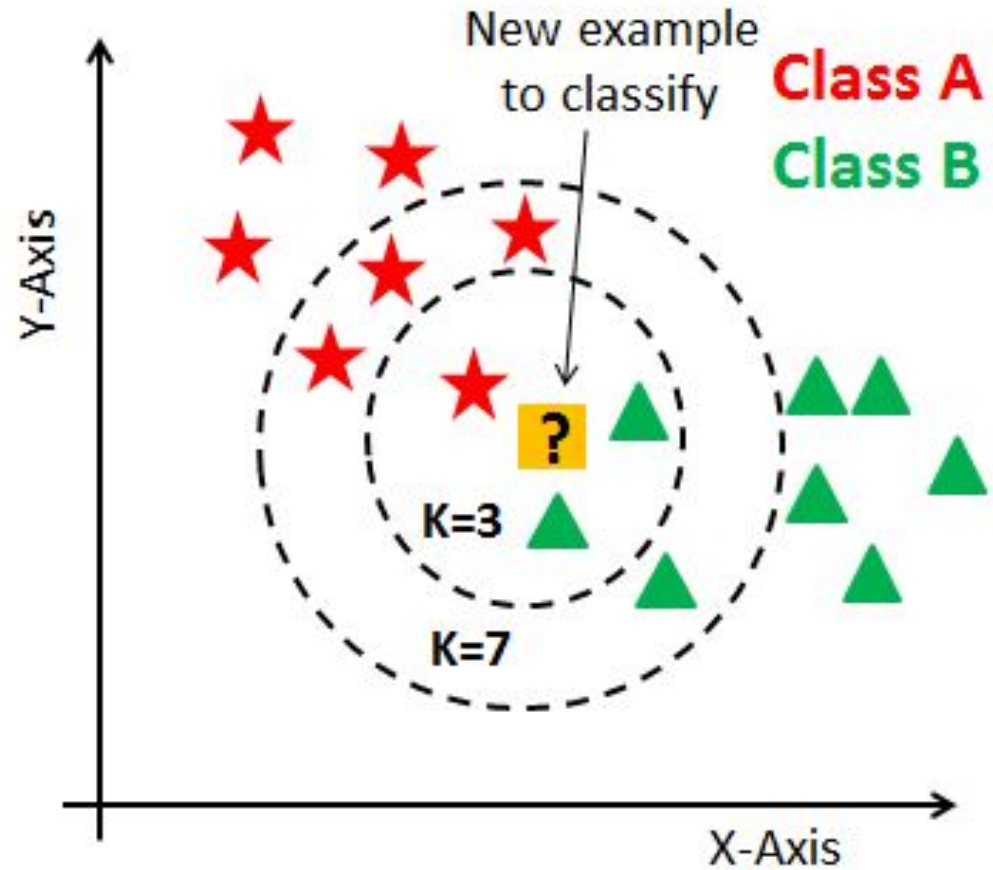
AMJ

It’s a Classification Algorithm

Example: Movie Genre Prediction

IMDB	Duration	Genre
8.0	160	Action
6.2	170	Action
7.2	168	Comedy
8.2	155	Comedy

Find the genre of Barbie, **IMDB-7.4**
Duration-114


New example
to classify
Class A
Class B
Y-Axis
?
K=3
K=7
X-Axis

Euclidean Distance

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Here,

- x_1, y_1: Barbie's attributes (IMDB, Duration).
- x_2, y_2: Attributes of each data point.

Now,

Step 1: Compute distances

$$d_1 = \sqrt{(7.4 - 8.0)^2 + (114 - 160)^2}$$

$$d_2 = \sqrt{(7.4 - 6.2)^2 + (114 - 170)^2}$$

$$d_3 = \sqrt{(7.4 - 7.2)^2 + (114 - 168)^2}$$

$$d_4 = \sqrt{(7.4 - 8.2)^2 + (114 - 155)^2}$$

Suppose K=3,

The distances to the 3 nearest neighbors are as follows:

1. **Comedy**: Distance = 41.01
2. **Action**: Distance = 46.00
3. **Comedy**: Distance = 54.00

Suppose K=3,

Genre Classification:

- Among the 3 nearest neighbors, **Comedy** appears 2 times, and **Action** appears 1 time.
- Therefore, the predicted genre for Barbie is **Comedy**.