KNN Algorithm

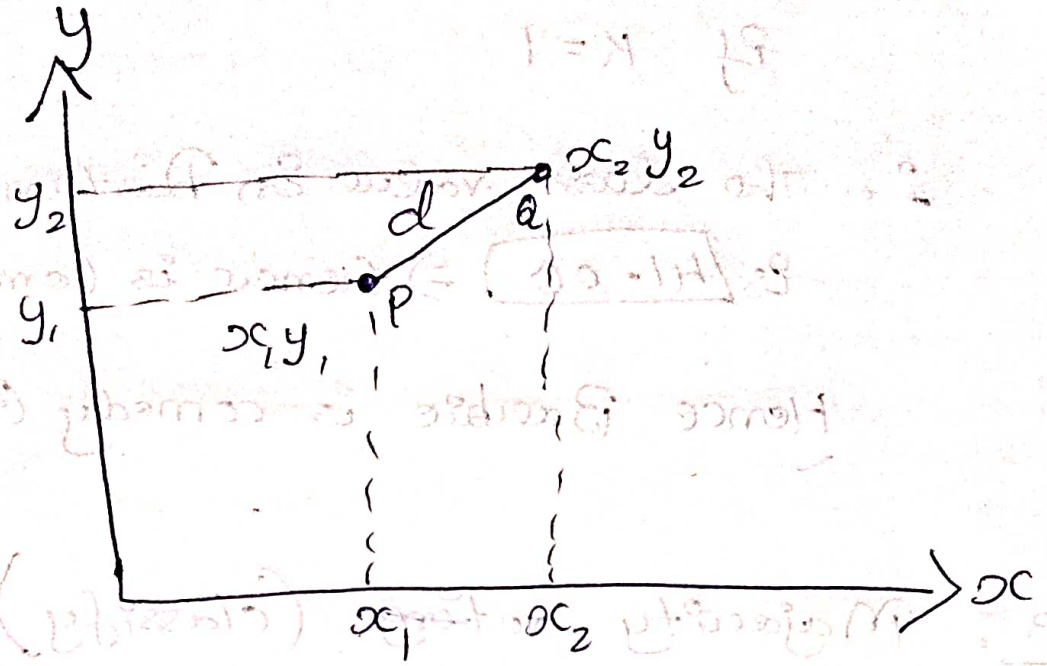
* KNN stands for K-Nearest Neighbours.
* classification - Predicting classes.

Eg: Predicting Movie Genres

IMDb Rating	Duration	Genres
8.0 (mission Impossible)	160	Action
6.2 (Gadar 2)	170	Action
7.2 (Rocky & Rani)	168	comedy
8.2 (OMG 2)	155	Comedy

Predict the Genres of 'Barbie' movie with IMDb 7.4 duration 114.

Step 1: Calculate Distances



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

$x_1 = 7.4$ $y_1 = 114$ ⇒ baabie.

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

(x_2 y_2)

$= \sqrt{0.36 + 2116}$

$= \underline{46.00}$

Distance (6.2, 16) = $\sqrt{((7.4 - 6.2))^2 + (114 - 170)^2}$

$= \sqrt{1.44 + 3136}$ $= \underline{56.01}$

Distance (7.2, 168) = $\underline{54.00}$

Distance (8.2, 155) = $\underline{41.00}$

Step 2: Select K Nearest Neighbours

If K = 1

∴ The least value in Distance is 41.00 ⇒ Genre is Comedy

Hence Barbie is comedy Genre

Step 3: Majority voting (classify)

If K = 3

take least 3 values from Distance

46.00 → Action
54.00 ⇒ comedy
41.00 → comedy } majority

∴ Barbie is comedy.