HYPERPARAMETERS

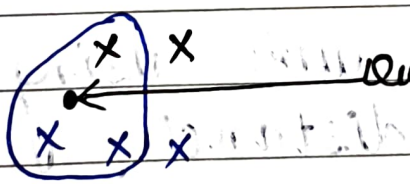
1) Weight → 'uniform' or 'distance'

- uniform → each vote has same value. Mukesh ambani & Ujwal vote are considered same.

- distance : The neighbors which are close will be assigned more weight
- generally → $w = \frac{1}{\text{distance}}$.

Ex1 →

- When weight is uniform



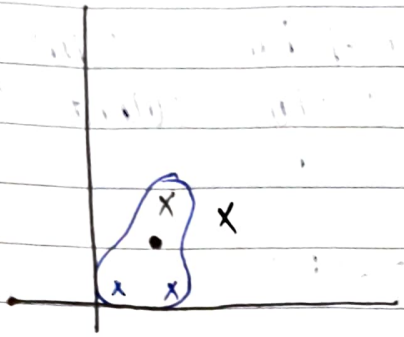
query, and K=3

majority count.

no of blue	2
no of black	1

- Since blue count is more for K=3, new query will be assigned to blue.

Ex2 When w1 = 'distance'



pt	distance	w
X	0.1	1/0.1
X	0.9	1/0.9
X	0.8	1/0.9

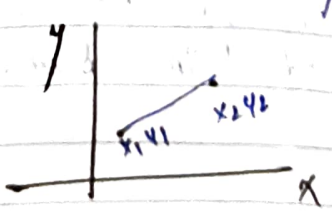
$\left(\frac{1}{0.1} > \frac{1}{0.9} + \frac{1}{0.9}\right)$

- Since Black weightage is more than of blue pts the query pt will be assigned to black.

(2) Types of Distances

- Distance we were using was <u>euclidient</u> distance.
- There are other types of distances too:

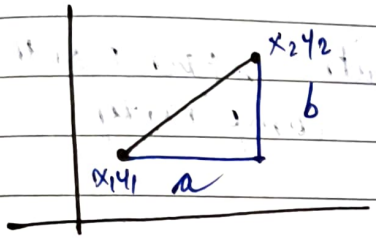
A) <u>Euclidient distance</u>: shortest distance between 2 points



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

- for 3 dimensions → $\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$
- for n dimension → $\left(\sum_{i=1}^{d} (x_{2i} - x_{1i})^2 \right)^{\frac{1}{2}}$
  - d → no of dimensions
- Also called L2 norm.

4) Manhatten distance



- a + b = manhatten distance.
- also called taxi cab distance
- L1 Norm

- $m = \sum_{i=1}^{d} | x_{2i} - x_{1i} |$ → for n dimensions
- for 2 ~~pts~~ → dim ⇒ $|(x_2 - x_1) + (y_2 - y_1)|$

5) Minkowski → custom Norm.

p=2 ⇒ Euc → $\sum_{i=1}^{d} \left((x_{2i} - x_{1i})^2 \right)^{\frac{1}{2}}$

p=1 ⇒ Manhatten → $\sum_{i=1}^{d} \left(|x_{2i} - x_{1i}|^1 \right)^{\frac{1}{1}}$

p=3 ⇒ L3 Norm ⇒ $\sum_{i=1}^{d} \left(|x_{2i} - x_{1i}|^3 \right)^{\frac{1}{3}}$

p ⇒ p $\quad \sum_{i}^{d} \left(|x_{2i} - x_{1i}|^p \right)^{\frac{1}{p}}$