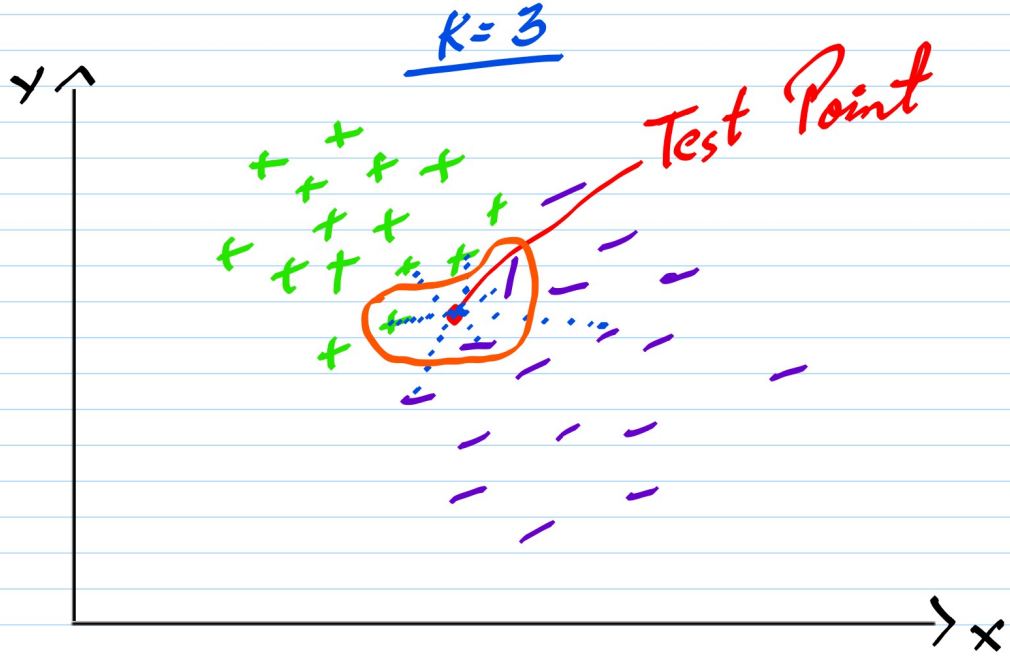
K-Nearest Neighbour.(KNN)

K-Means Clustering

Unsupervised ML used for grouping/Clustering

KNN :→ Supervised Machine Learning algorithm used for both Regression and Classification.

Logistic Reg → Classificatn
Liner Reg → Regression



K-means

K= number of Cluster

KNN

K = Nearest Neighbour to the Test Point.

If we are dealing with Classificat^n Problem,

→ Majority Voting is done

When K is odd

K=3 → 2(- -) / 1(+) → Majority Voting → (-) is the winner.

When K is Even

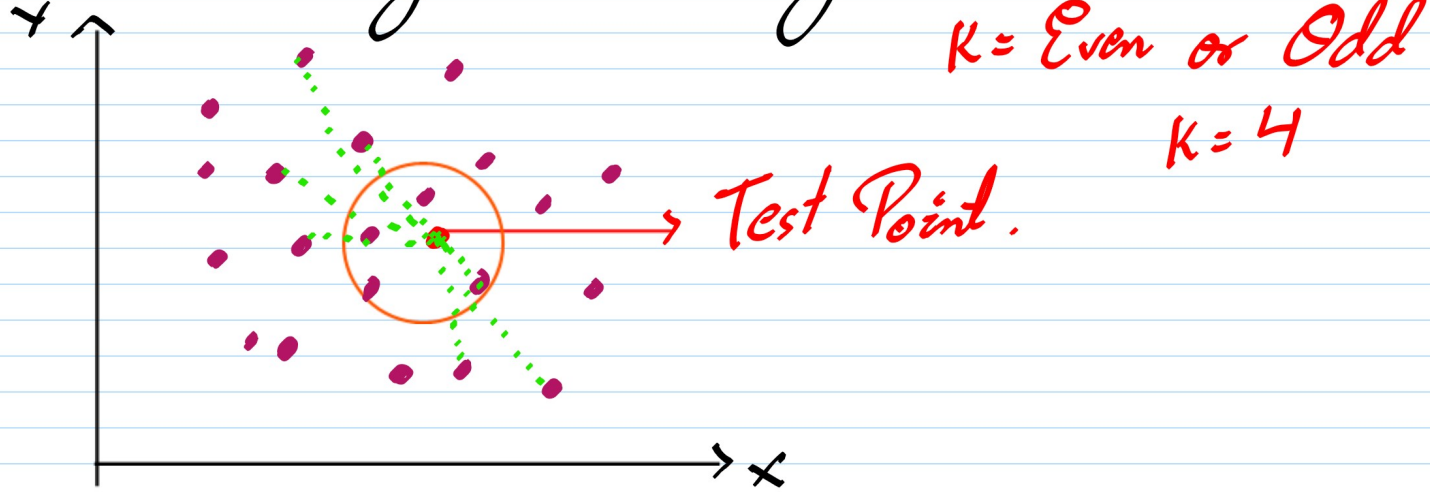
K=4 → 2(- -) / 2(+ +) → Majority Voting → Draw (creates Ambiguity)

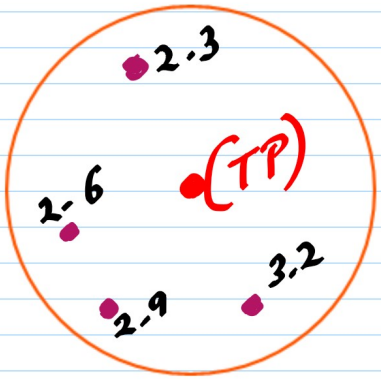
So always prefer K = odd value.

If we are dealing with Regression Problem,

K = Even or Odd

K = 4





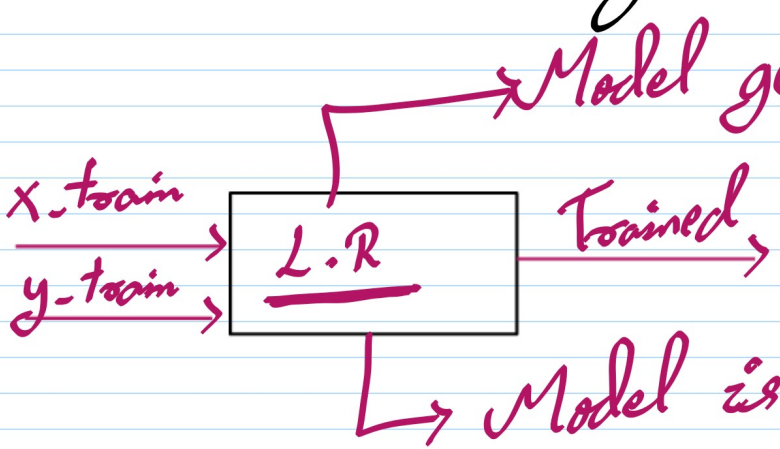
Average of my 4 datapoint, As we have taken K as 4.

$$= \frac{2.6 + 2.9 + 3.2 + 2.3}{4} = \frac{11}{4}$$

Test point = 2.75

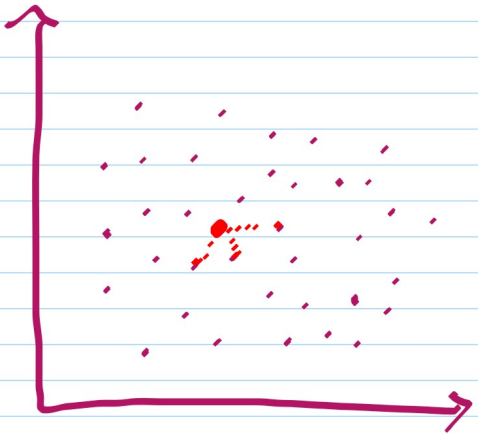
In Regression Problem, we take Average or Mean value.

Q Why KNN is called as Lazy Learner? (**)



Now, the model learned the data, and becomes generalized.

In KNN →



→ No Training (Training time is very very less).

→ Testing Data (It takes a lot of time as it has to do lot of computation.

As number of Dimensions increases, Your computation time also increases.

e.g. KNN is not used in Production.