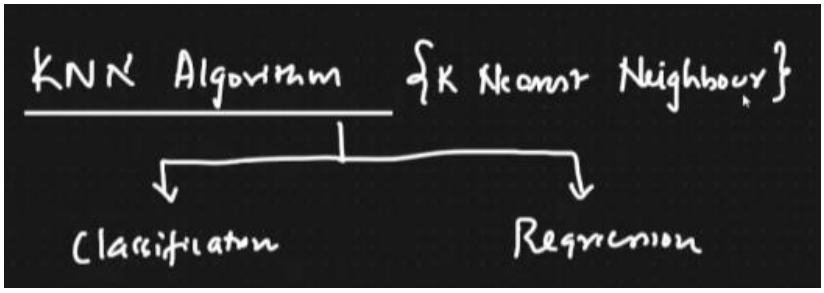
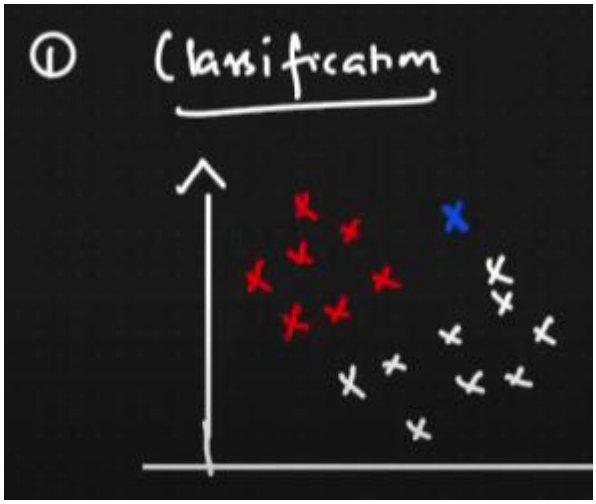
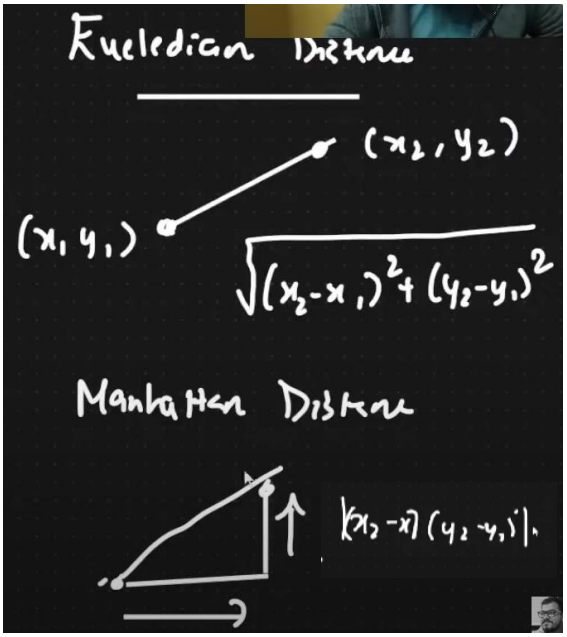




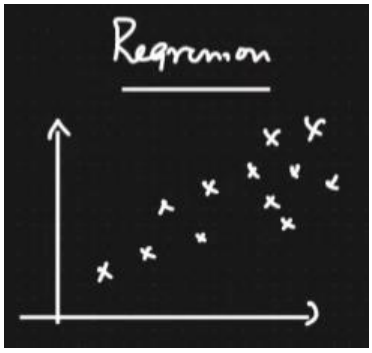
How do decide, whether the new data point (blue pt.) lies with the red part or the white part?

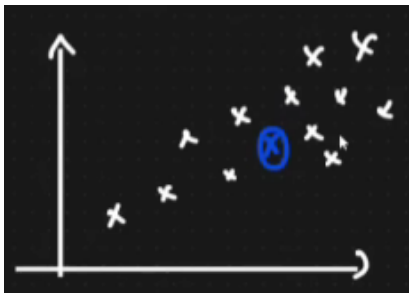
➔ by KNN!
Suppose, K=5 (hyperparameter);
The blue pt. is closest to 3 red pts. & 2 white pts.
Max pts. = Red;
Hence, blue pt. lies into the red part!

With the help of **Euclidian & Manhattan Distance**, we decide the future of a particular new data.

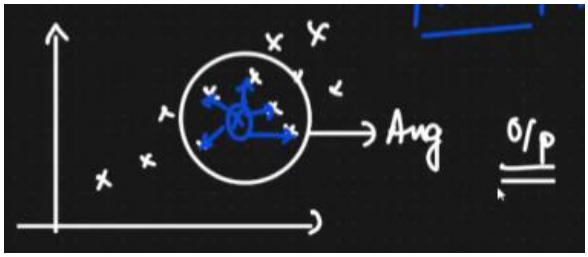


Similarly for Regression,





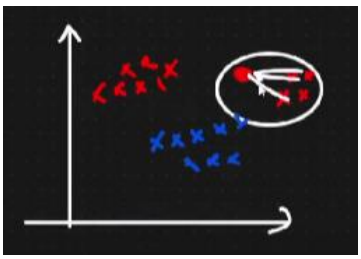
if, K=5 (a hyperparameter)



later, take the 5 pts. nearer to the blue & average it;
Hurrah, we got the answer!

In real life larger dataset, we try to calculate K from 1 to 50; later check the error rate, if the error rate is less than we select the model!

KNN doesn't works well with:
i) Outliers
ii) Imbalanced dataset
example,



over here, the red pt. is near towards the blue pts. but due to the outliers and imbalanced dataset, the red pt. goes towards the right-hand side red part!

Reference:
1) KNN Concept Explanation (from 1 hr 52 mins onwards)