

KNN Algorithm

- It comes under supervised.
- KNN is used to solve both regression and classification problem.

KNN Regression [Regressor]

- KNN Regressor used when data has continuous target.

KNN classifier

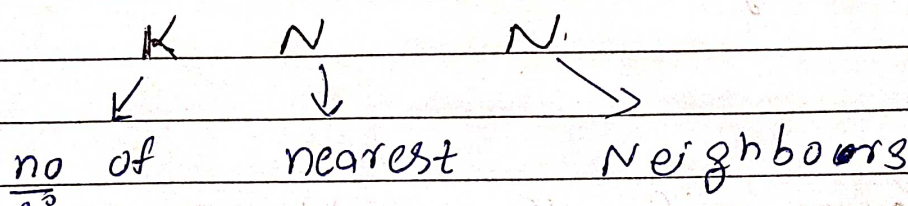
- KNN classifier used when data has categorical data

→ KNN algorithm is lazy algorithm because it will not learn anything during training.

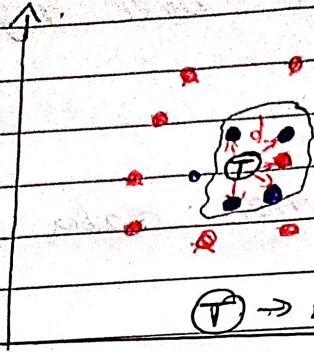
→ KNN assumes similar things exists in closer proximity.

→ KNN relies on similarities and distance measures to make predictions

→ KNN is called as distance based algorithm because it will use distance measures.



How KNN works?



• $\bullet \rightarrow$ diabetic

• $\bullet \rightarrow$ non diabetic

$\textcircled{T} \rightarrow$ Test point, Test data

$\textcircled{T} \rightarrow$ we can say non diabetic, ^{max} nearest neighbours are non diabetic

- plot data
- Take test data $\rightarrow \textcircled{T}$
- Find distance between test point and all other observation
- Define K , ex: $\{3, 5, \dots\}$ (only odd) \leftarrow
Let $K = 5$.
- Look for 5 nearest neighbors
- classify test data into one of the class based on "majority voting"
- In Knn Regressor prediction will be average of all the nearest neighbors
- Classification problem \rightarrow Majority voting
- Regression problem \rightarrow average