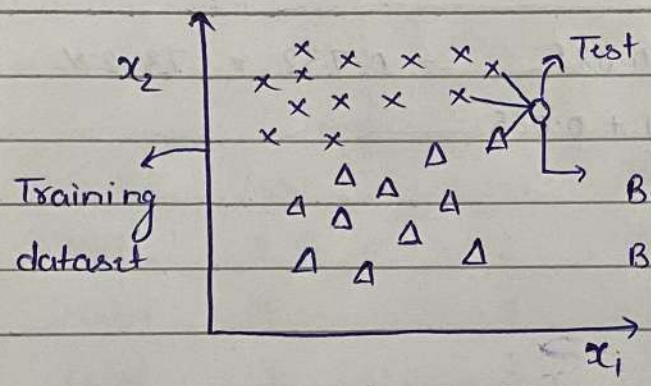
K-Nearest Neighbours

→ It is a supervised machine learning algorithm which is used to solve both classification and regression problem with the help of k neighbours datapoints.

- **classification :**



x_1	x_2	o/p
—	—	1
—	—	0

Belongs to x (1) class.
↓
Because more # of x(1) class neighbours

o/p → 1 → x
0 → Δ

Steps :

(i) Initialize the value of k.
Let k = 3 , k > 0

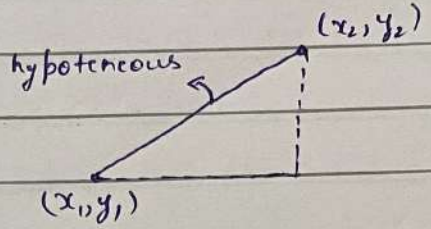
(ii) Find k-nearest datapoints by calculating distance.

(iii) Assign the test data to the majority class in the nearest neighbours.

→ Distance Metrics :

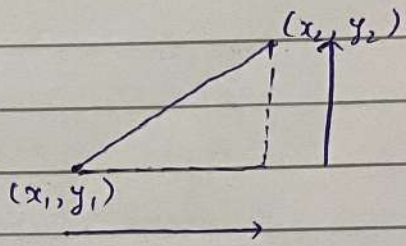
- Euclidean Distance -

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

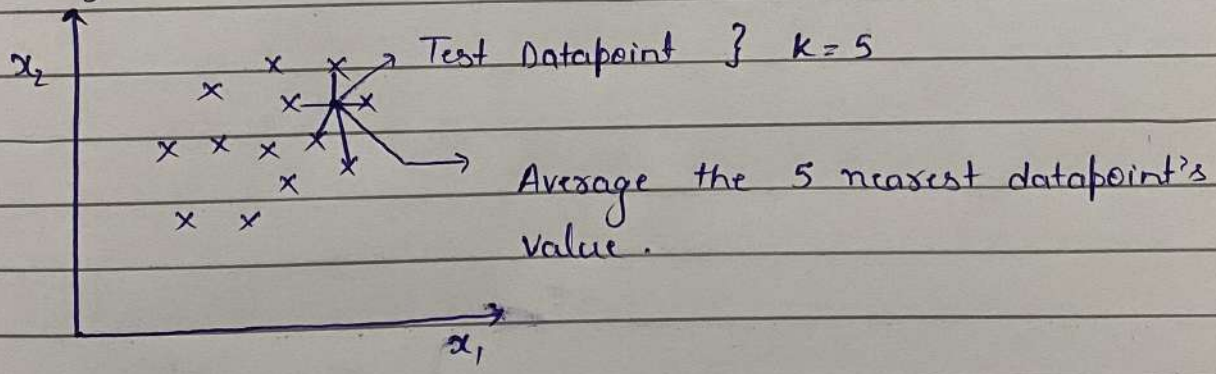


- Manhattan Distance -

$$M = |y_2 - y_1| + |x_2 - x_1|$$



- Regression :



* Limitations of KNearestNeighbours :

- computational expensive to huge dataset
- Sensitive to outliers
- Sensitive to missing value.