K Nearest Neighbour — KNN is supervised machine learning algorithm used for classification and regression

working — It finds the k nearest data point (neighbour) to a query point and predict the output based on majority vote (classification) or average (Regression)

Distance metrics — commonly uses Euclidean distance Manhattan.

Lazy learner — No training phase, just store data → prediction happen at query time.

Key Hyperparameter — K (number of neighbours)
Small k = noisy . large k = smother

Ex — Classification

f_1	f_2	f_3	O/P
—	—	✗	0
—	—	✗	1

y — fixed number of categories.

1 ⇒ 2
0 ⇒ 3

Step.

k hyperparameter.

① K=5

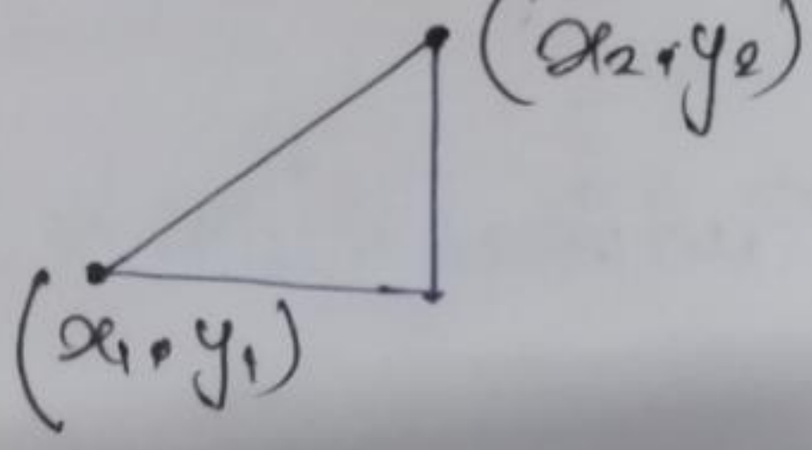
② Now finding k nearest data point

③ Now finding the category of nearest data point

④ How find the Nearest data point ⇒ through Calculate the distance of data points.

① Euclidian Distance

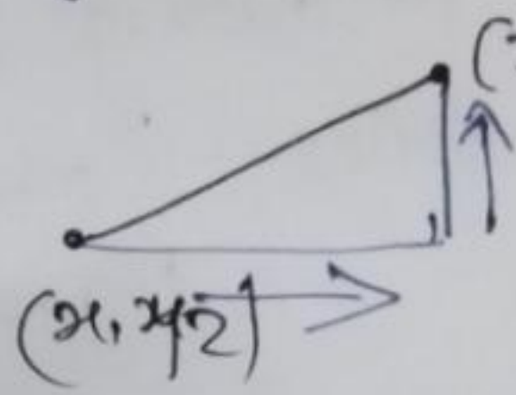
The straight-line (as the crow flies) distance between two points in space.



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

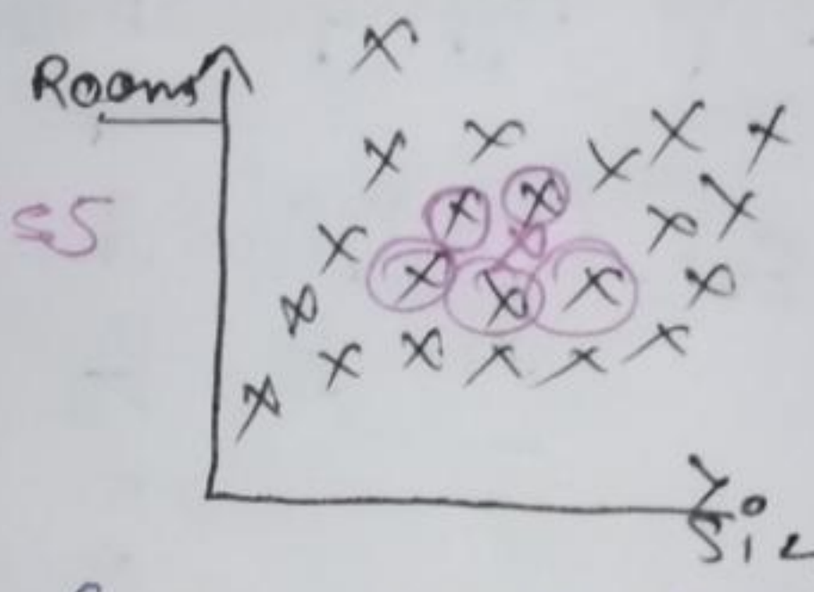
② Manhattan Distance

The distance measured along axes at right angles (like moving in city blocks)



Ex - Regression (House pricing)

Size	Room	price
—	—	Continuous
—	—	"
—	—	



o/p ⇒ Avg (nearest data point)

Limitation ⇒

1. In Huge dataset it creates a problem.
2. Outliers - Sensitive to Sensitive to Missingvalue.