- K-means | Hierarchical | DBSCAN.

Clustering → Grouping → unsupervised.

KNN.

K - Nearest - Neighbor

Machine learning algo.

- Supervised [target column].
  - Regression
    - Linear reg.
    - Ridge reg.
    - Lasso reg.
    - Elastic net
    - SVM/SVR
    - DTR
    - RFR.
    - GBR
    - XBR.
  - Classification.
    - log reg
    - SVC/SVM.
    - DTC
    - RFC
    - GBC.
    - XBG.
    - NBC → classification → text data → NLP with ML
      - (1) text data.
      - (2) TF/IDF
      - (3) NB.
      - (4) evaluation

  → Practical.

  → (1) train. (2) test.

- unsupervised. [grouping based on dist]

KNN. → Distance based approach. → lazy learner

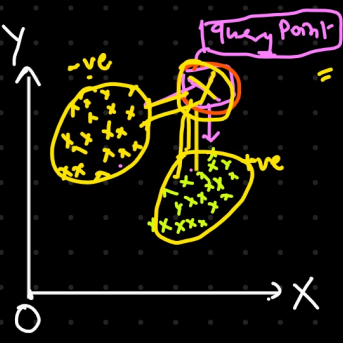
Machine learning. → learning. → fin → eq.

KNN.

K - Nearest - Neighbor

- Classification.
- Regression.

(1) classification.



= [Distance Based approach]

(1) Distance for each and every Point-

(2) K Point-

↓

hyperParameter

K = 2, 3, 5, 6, 7...

10 Point-

=

[New Point-] → E.D → 3NN → +ve

K=3

NP - 1 Point- = 3 → +ve = 3-Nearest Neighbor

NP - 2 Point- = 5

NP - 3 Point = 7

NP - 4 Point = 1 → +ve

NP - 5 Point = 0.1 → -ve

NP - 6 Point = 10

NP - 7 Point = 12

NP - 8 Point = 8

NP - 9 Point = 3

NP - 10 Point = 15

+ve
+ve
-ve → majority vote You gonna select query Point- will belong to which class.

my query Point would belong to +ve class.

5-NN

+ve
+ve
-ve
-ve
-ve

my query Point would belong to -ve class.

K-NN → K value at that time we don't consider

~~even no.~~

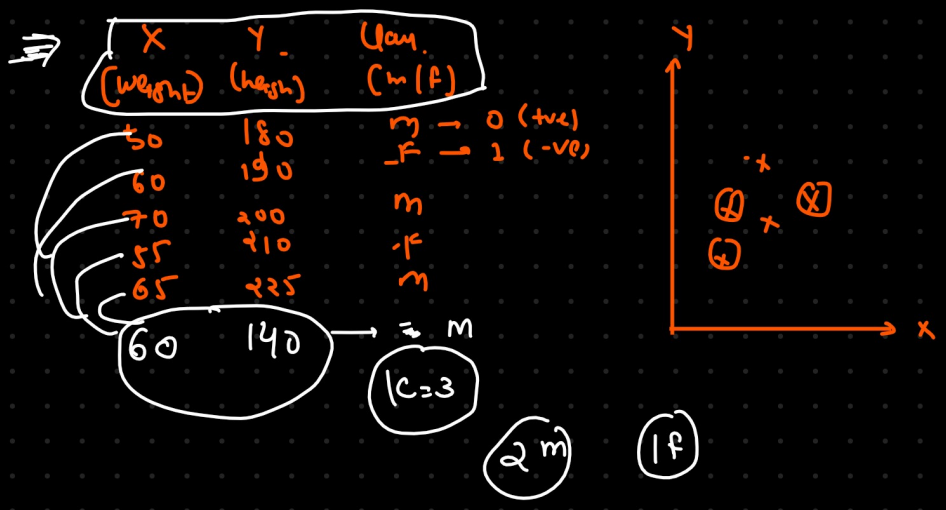
query Point-

~~K=4~~ = +ve, +ve, -ve, -ve → each Prob. $\frac{2}{4}, \frac{1}{4}, \frac{1}{4}$

$\frac{1}{2}$

always K = 3, 5, 7, 9, 11, ...N → Odd no.

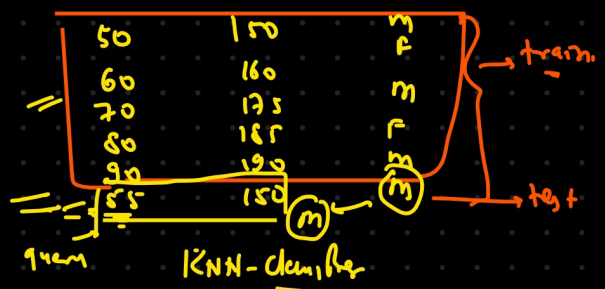
~~K = 2, 4, 6, 8~~



Classification - majority

Regression - mean of all the point

(1) classification.



(1) train (2) testing — Not ML algo — Lazy learner

KNN → (5)

(6) - (1) ⇒ 0.5 m
(6) - (2) ⇒ 0.2 f
(6) - (3) ⇒ 1 m
(6) - (4) ⇒ 2 f
(6) - (5) ⇒ 5 m

K = 3

(1) ball tree
(2) kd tree

(2) Regression.

weight	height
50	150
60	160
70	175
80	185
90	190
55	150

→ final

KNN Regressor

55 - =