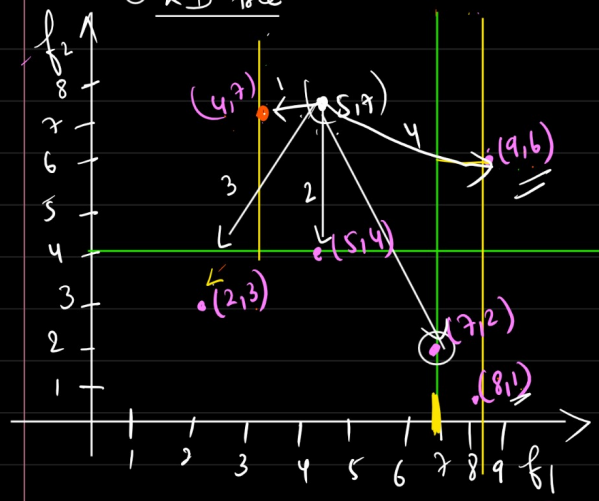
Variants of KNN

① KD Tree (K-dimensional tree)

② Ball Tree

① KD tree



f_1	f_2
7	2
5	4
9	6
2	3
4	7
8	1

→ Distance calculation from each of dp of the test data is to be Calculated. ⟹ Computationally expensive.

* KD tree

↳ Partition the data in a binary tree.

Step-1 Sort the feature f_1 & f_2

Step-2 Calculate the median of f_1 & f_2

Step-3 – Partition the data based on median

Step-4 – Partition all the data recursively in each of sub partition.

① f_1 – 2, 4, 5, 7, 8, 9 | f_2 = 1, 2, 3, 4, 6, 7

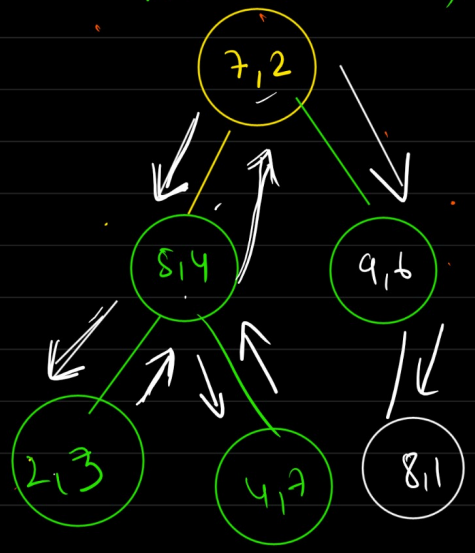
② median → $\frac{5+7}{2}$ ~~6~~ 5 → take 7 as median. | take 4 as median.

③ Partition of data based on median.

④ recursively → (9, 6), (8, 1)

f_1 = 8, 9 → 9

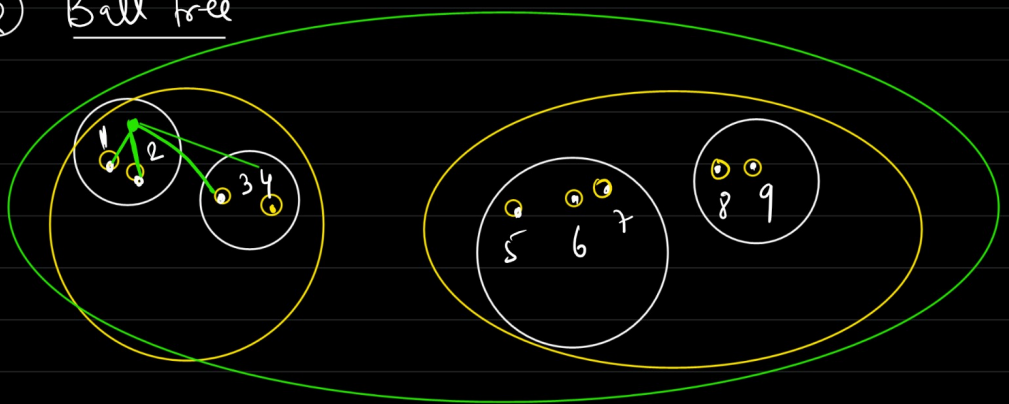
KD-tree (Binary tree)



Advantage

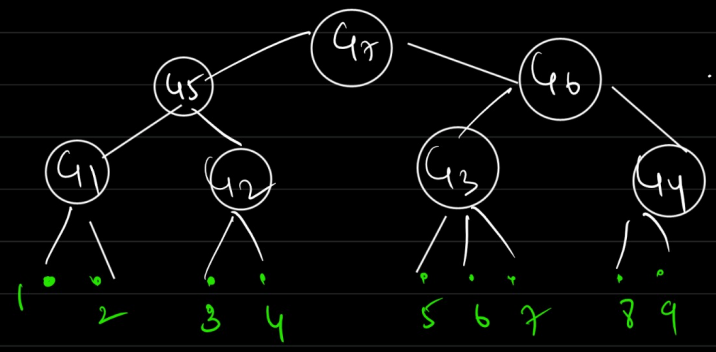
→ you don't need to calculate all the distance.

② Ball tree



Advantage

→ only need calculate distance of the nearest group elements


G7
G5
G6
G1
G2
G3
G4
1
2
3
4
5
6
7
8
9