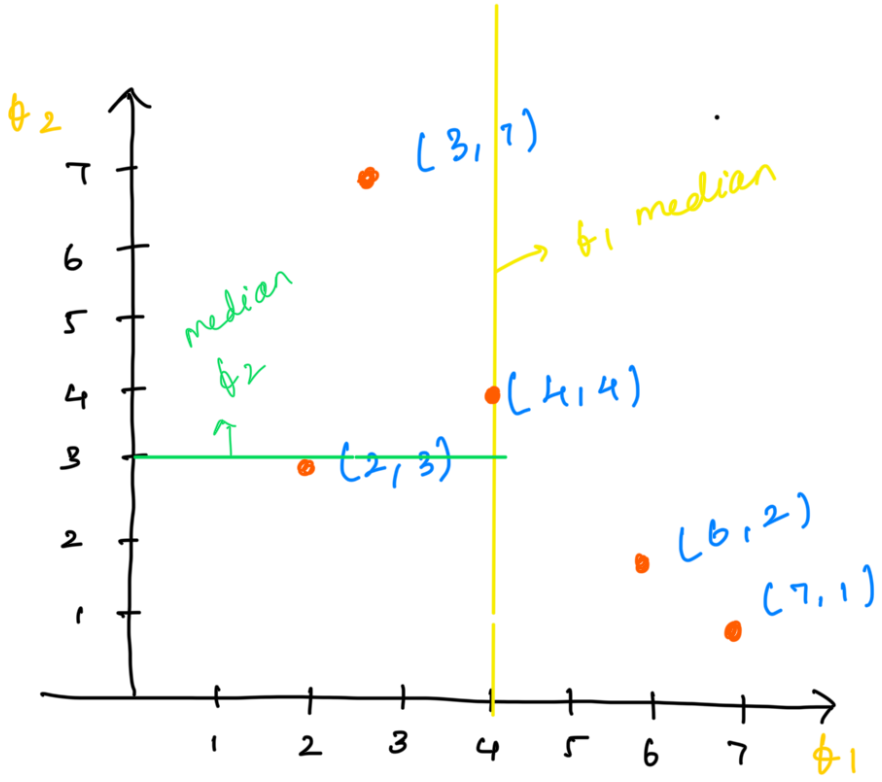
Variants of KNN

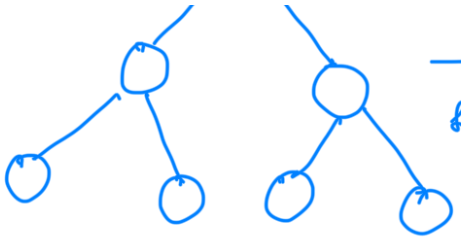
KD Tree



ϕ_1	ϕ_2
7	1
6	2
4	4
3	7
2	3

data plotted on the above plot

using KD tree will compute the binary tree



→ In this way search will be limited to one particular node.

how do we derive this?

In this case we will have to find out the median of ϕ_1 and ϕ_2

median of ϕ_1

2, 3, [4], 6, 7

= 4

median of ϕ_2

1, 2, [3], 4, 7

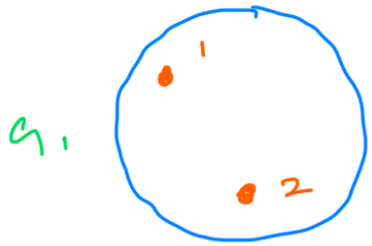
= 3

Actually, by this way, we are creating a spliting in the form of a tree, if you look at above plot, we have splited and this would in turn behave as tree.

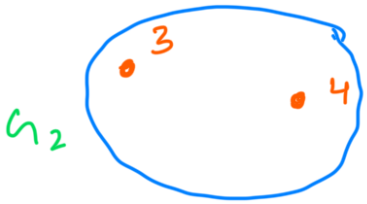
So, now if we get any new data points, it will be easier to traverse to the nearest node. In KD tree back tracking also possible to find the next nearest point. KD tree will be useful to reduce the time complexity.

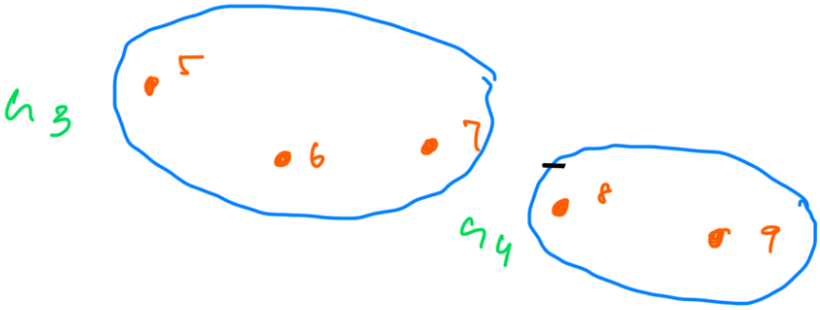
② Ball Tree

This will the better approach when compared to KD tree, because in KD we have to back tracking. in ball tree we dont have to do that.

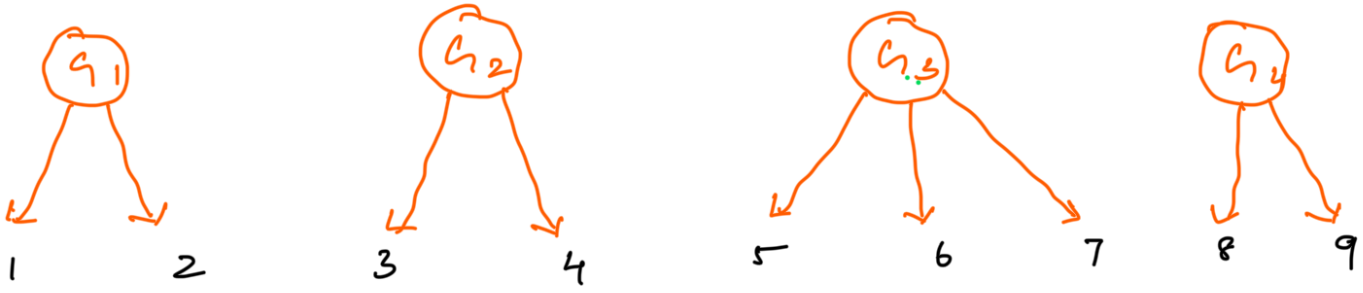


As the name suggests, we will group the nearest data points together



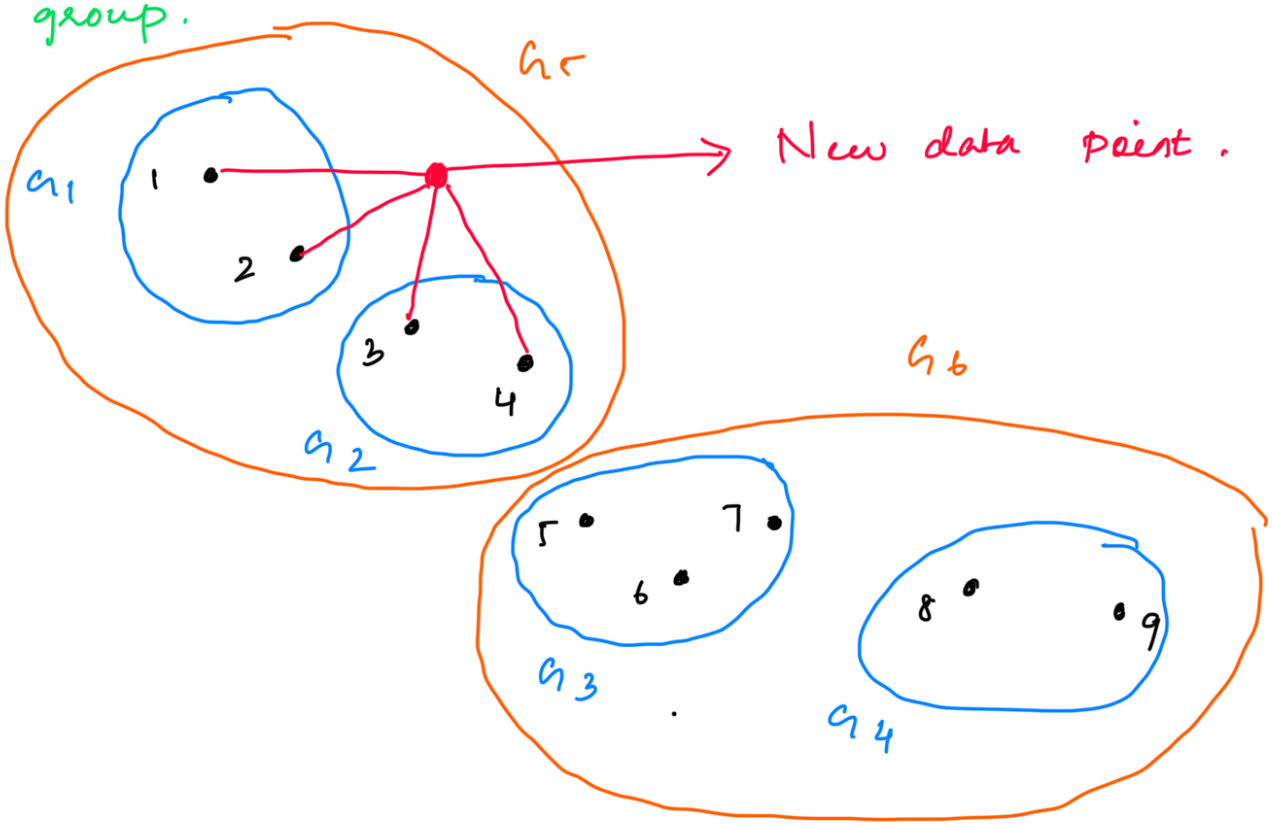


so basically, it can be discribed as

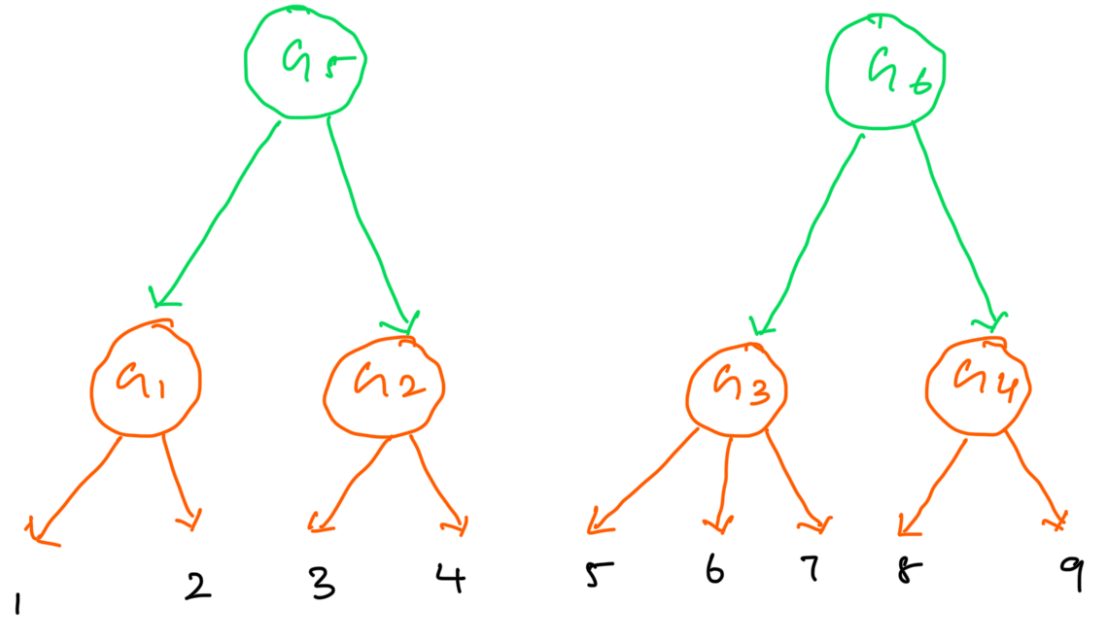


Next step will be combining the nearest

group.



So now my grouping will be.



Further, I can again combine entire group to new group, in our case it will be G7. If any new data point comes in, we can easily calculate the distance

instead of calculating the distance of each node.

With both KD tree and Ball tree we are basically, constructing a binary tree.

Why we are preparing the binary tree?

TO reduce the time complexity.