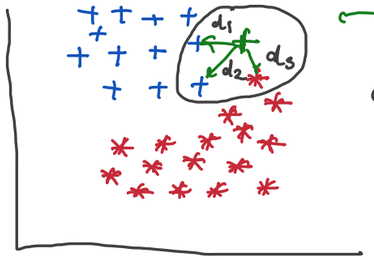
04 May 2022 11:33

k-Nearest Neighbour :



← test when k=3

d_1, d_2 and d_3 are minimum distances from k points

Now the circle contain 2 + and 1 * Class so KNN choose frequently occured Class.

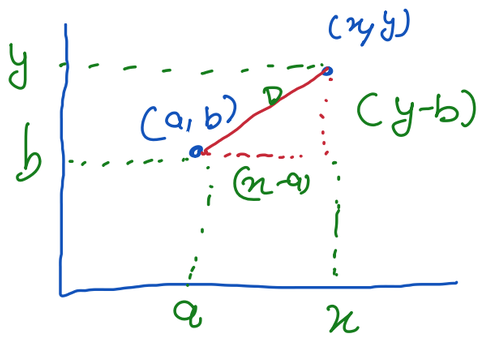
So, our answer will be (+) class.

Here we got that we have to work on distances.

Let us see how many ways we have to calc. distance.

i] Euclidean
ii] Manhattan
iii] Hamming

i] Euclidean distance :



$$h^2 = p^2 + b^2$$

$$D^2 = (y-b)^2 + (x-a)^2$$

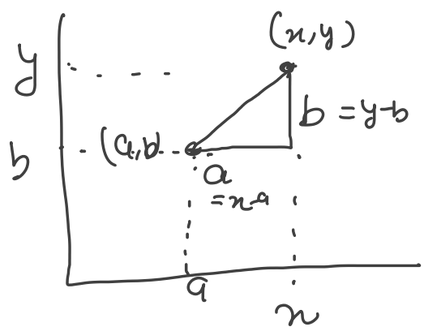
$$= \sqrt{(y-b)^2 + (x-a)^2}$$

Now this is for 2-Dimension let us see for d dimension:

$$= \left(\sum_{i=1}^{d} (x_i - a_i)^2 \right)^{1/2}$$

ii) Manhattan distance :

It is the total length covered between two points.



$M = a+b = (x-a)+(y-b)$

for d dimensions:

$$\left(\sum_{i=0}^{d} (x_i - a_i) \right)^{1}$$

Generalized form :-

$$\left(\sum_{i=1}^{d} (x_i - a_i)^p \right)^{1/p}$$

Hamming Distance :

This tells the uncommon elements.

a = | 0 | 1 | 1 | 0 | 1 | 0 |

b = | 0 | 1 | 1 | 0 | 1 | 0 |

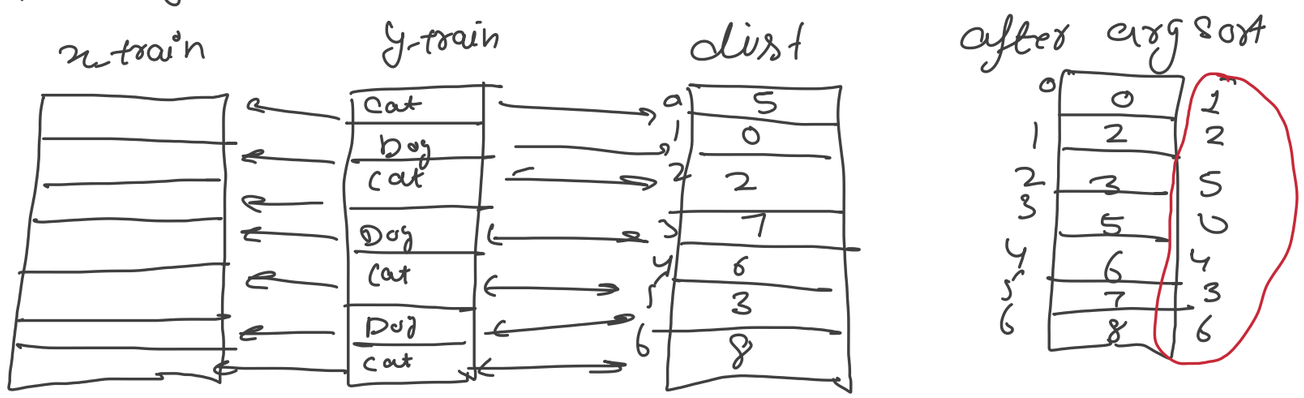
$h(a,b) = \underline{0}$ (all common)

Now The Scratch implementation:

In Scratch implementation one thing need to know is that we are only working when we are testing on data.

See code.

Steps followed:



for arg sort array take out k min.

$k<=3$:- $arg[:k] = [1,2,5]$

Now check for labels : Cat Cat Dog

Check frequently occurred class - (in this case)

= Cat