Lecture - 17

* Condensed Nearest Neighbour.

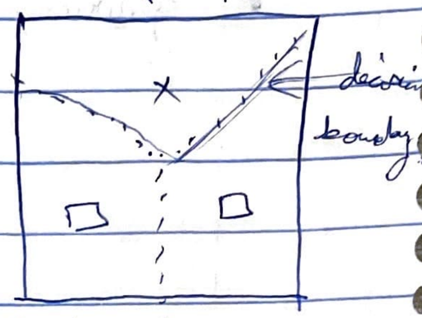
A greedy algorithm

$S \subseteq D$

$S \rightarrow$ condensed data set

$D \rightarrow$ data points.

K = 1.

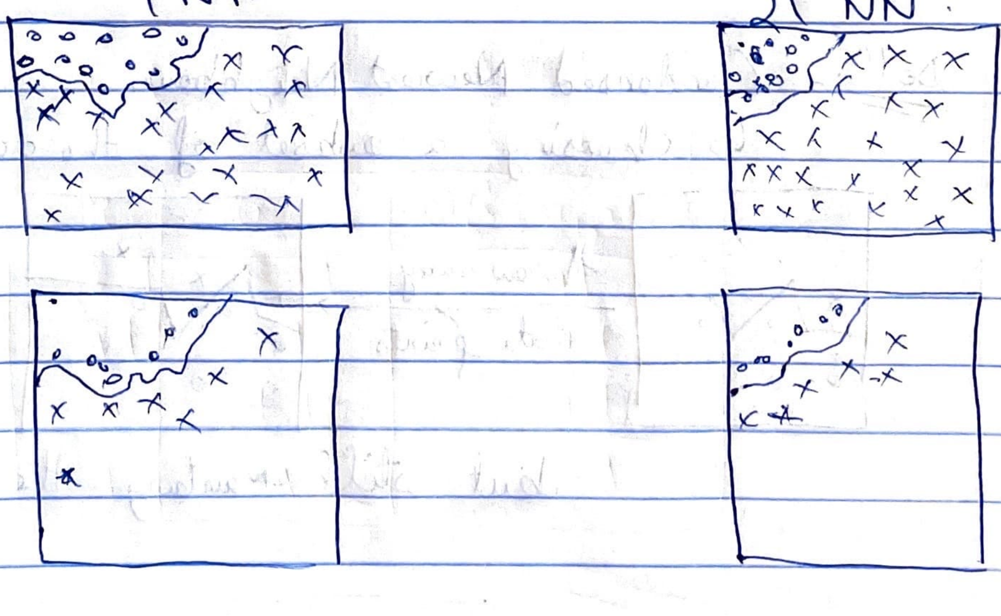


for $K = 3$ the whole region will be blue.
& the training Error would be $1/3^{rd} = 33\%$

* 1] Randomly select k-data point into S.
2] Classify all data accordingly to S.

* Condensing the data.

1 NN | 21 NN



★ Identifying K - representatives (centre).

Unsupervised K-means clustering.

partition input $X = (x_1, x_2, \ldots x_N)$
into K sets $S_1, S_2, \ldots S_K$ → clusters.
with centers $\mu_1, \mu_2, \mu_3 \ldots \mu_K$ → cluster centers.

Where μ_j "represents" every data point in cluster S_j

$$E_j = \sum_{x_n \in S} \sqrt{\| x_n - \mu_j \|}^2.$$

★ Finding clusters - Lloyd's Algorithm.
→ Pick random point as centre of cluster
→ Compute Voronoi regions as clusters.
↳ construct S_j to be all points closest to centre.
→ Update centre (μ_j) to equal the centroid of S_j.

$$\mu_j = \frac{1}{|S_j|} \sum_{x_n \in S_j} x_n$$

→ Repeate steps 2 & 3 until E in stops decreasing.

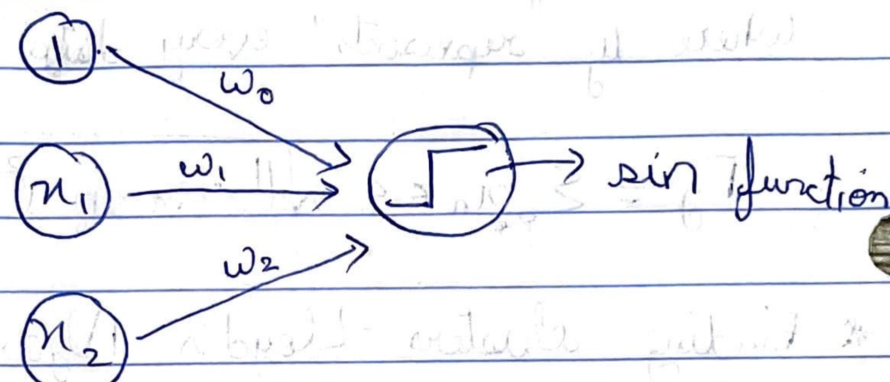
★

(Artificial) Neural Networks.

→ A generalization of the perceptron model.

Reminder of Perceptron model.

$$h = \text{sign}(w^T x) = \text{sign}(w_0 + w_1 x_1 + w_2 x_2)$$



$$AND(x_1, x_2) = \begin{cases} +1, & \text{if } x_1 = x_2 = +1 \\ -1, & \text{otherwise} \end{cases}$$

$$OR(x_1, x_2) = \begin{cases} +1 & \text{otherwise} \\ -1 & \text{if } x_1 = x_2 = -1. \end{cases}$$

x	x_2	$OR(x_1, x_2)$
+1	+1	+1
+1	-1	+1
-1	+1	+1
-1	-1	-1

x_1	x_2	$AND(x_1, x_2)$
+1	+1	+1
+1	-1	-1
-1	+1	-1
-1	-1	-1

OR

$w_0 + w_1 x_1 + w_2 x_2$.

let w_0, w_1 & $w_2 = 1$

$1 + w_1 x_1 + w_2 x_2$.

~~let w_1 & $w_2 = 1$~~

for $x_1, x_2 = +1, +1$.

$1 + 1 + 1 \Rightarrow +1$

for $x_1, x_2 \Rightarrow +1, -1$

$1 + 1 - 1 \Rightarrow +1$

for $x_1, x_2 \Rightarrow -1, +1$

$1 - 1 + 1 \Rightarrow +1$

for $x_1, x_2 \Rightarrow -1, -1$

$1 - 1 - 1 \Rightarrow -1$.

AND

let $w_0 = -1$

w_1 & $w_2 = 1$.