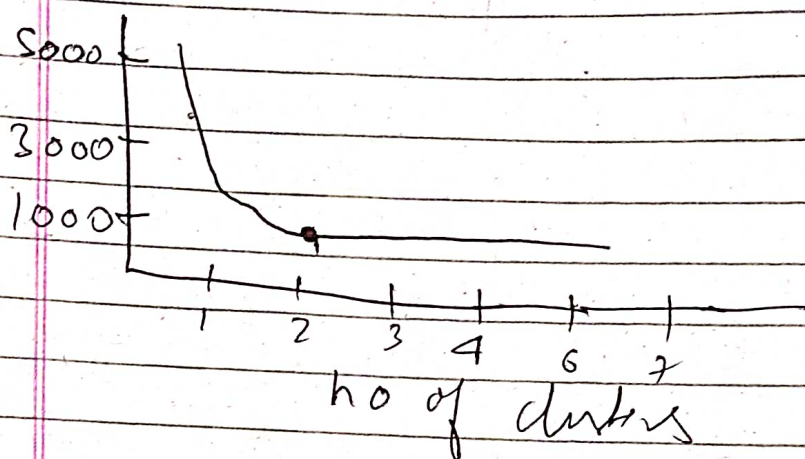


# Machine Learning Assignment

Answer 1

Elbow Method ; In this method, we plot the within cluster sum of square (WCSS) against the no of clusters. The WCSS measures the sum of squared distance between each point and assigned centroid.



Answer 2

K-means clustering algorithm

Iteration - 1

1) Calculate the distance of each point from each of the centre of the cluster.

2) Distance is calculated by formula

$$P(a,b) = |x_2 - x_1| + |y_2 - y_1|$$



Distance b/w  $A_1(2, 10)$  and  $C_1(2, 10)$

$$\begin{aligned}
 P(A_1, C_1) &= |x_2 - x_1| + |y_2 - y_1| \\
 &= |2 - 2| + |10 - 10| \\
 &= 0
 \end{aligned}$$

Points	Cluster 1	Cluster 2	Cluster 3	Cluster 4
$A_1(2, 10)$	(2, 10)	(5, 8)	(1, 2)	
$A_2(2, 10)$	0	5	9	$C_1$
$A_2(2, 5)$	5	6	4	$C_3$
$A_3(8, 4)$	12	7	9	$C_2$
$A_4(5, 8)$	5	0	10	$C_2$
$A_5(7, 5)$	16	5	9	$C_2$
$A_6(6, 4)$	10	5	7	$C_2$
$A_7(1, 2)$	9	10	0	$C_3$
$A_8(4, 9)$	3	2	10	$C_2$

Cluster - 1 contains :  $A_1(2, 10)$

Cluster - 2 contains :  $A_3(8, 4), A_4(5, 8), A_5(7, 5)$   
 $A_6(6, 4), A_8(4, 9)$

Cluster - 3 contains :  $A_2(2, 5), A_7(1, 2)$

Re-computing new clusters by taking means of all points

Cluster - 1 :  $(2, 10)$

Cluster - 2 :  $\frac{8+5+7+6+4}{5}, \frac{4+8+5+4+9}{5} = (6, 6)$



Cluster 3:  $\frac{2+1}{2}, \frac{5+1}{2}$  (1.5, 3.5)

So, the clusters formed after the Iteration are (3, 9.5), (6.5, 5.25) and (1.5, 3.5)

Ques 3 Divide the clustering?

Answer: 3 The value of no of clusters  $K=2$ , let two centroid of these two clusters

Iteration 1:  $C_1 (185, 72)$   
 $C_2 (170, 56)$

Now calculating distance of each point from centroids

Distance Formula  $\rightarrow \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

~~Ques~~

Height and weight Distance from clusters  
 $C_1 (185, 72), C_2 (170, 56)$

		$C_1$	$C_2$	
185	72	0	481	$C_1$
170	56	481	20	$C_2$
168	66	433	20	$C_1$
179	63	52	225	$C_1$
182	72	161	85	$C_2$
188	74	34	765	$C_1$
180	71	26	323	$C_1$
180	70	425	116	$C_1$
183	80	118	953	$C_1$



Cluster-1 contains

$A_1 (185, 72)$ ,  $A_4 (179, 68)$

$A_6 (188, 77)$ ,  $A_7 (180, 71)$

$A_9 (183, 84)$

Cluster-2 contains :

$A_2 (170, 56)$ ,  $A_5 (168, 60)$

$A_8 (177, 62)$ ,  $A_3 (186, 70)$

Now new centroid of clusters

$C_1$  : Cluster 1 :  $(183, 79.4)$

$C_2$  Cluster :  $(173.2, 57.5)$

Iteration 2

follows the steps of Iteration-1

Height	weight	Distance from $C_1$ $(183, 79.4)$	Distance $C_2 (173.2, 57.5)$	cluster
--------	--------	--------------------------------------	---------------------------------	---------

185	72	19.7	338.81	$C_1$
170	56	507.5	16.31	$C_2$
168	60	563.5	39.3	$C_2$
179	68	56.9	137.8	$C_1$
182	62	159.7	30.8	$C_2$
188	77	31.76	583.3	$C_1$
180	71	20.5	221.3	$C_1$
180	52	510.1	69.3	$C_2$
183	84	92.16	787.1	$C_1$

Cluster - 1 Centres :  $A_1 (9.7, 336.8)$   
 $A_4 (56.9, 132.8)$   
 $A_6 (31.7, 583.3)$ ,  $A_7 (20.5, 221.3)$   
 $A_9 (92.16, 782.8)$

Cluster - 2 Centres  $A_2 (50.75, 16.31)$ ,  $A_3 (86.3, 5.39)$   
 $A_5 (151.7, 36.8)$ ,  $A_8 (510.7, 69.3)$

new Centres after iteration 2 are

$$C_1 (\text{Cluster 1}) : \frac{9.7 + 56.9 + 31.7 + 20.5 + 92.16}{5}$$

$$\frac{782.8 + 336.8 + 132.8 + 883.3}{4}$$

$$C_1 = (182.19, 78.9)$$

$$C_2 (\text{Cluster 2}) = (172.33, 61.33)$$