MACHINE LEARNING ASSIGNMENT – 6

Student ID – 700741281

NAME – VENKATARAMANA RAVURI

USERID – VXR12810

Question-1

Find out clustering representations & Dendrogram using Single, Complete and Average link Proximity function in Hierarchical clustering technique?

Point	X-coordinate	Y-coordinate
P_1	0.04005	0.5306
P_2	0.248	0.3854
P_3	0.3457	0.3156
P_4	0.2652	0.1875
P_5	0.789	0.4139
P_6	0.4548	0.3022

CamScanner

Single link proximity Distance between two clusters is the distance between the closet points Also

Distance matrix.

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	0	0.2357	0.2218	0.3688	0.3421	0.2347
P_2	0.2357	0	0.1483	0.2042	0.1388	0.254
P_3	0.2218	0.1483	0	0.1513	0.2843	0.11
P_4	0.3688	0.2042	0.1513	0	0.2932	0.2216
P_5	0.3421	0.1388	0.2843	0.2932	0	0.3921
P_6	0.2347	0.254	0.11	0.2216	0.3924	0

Smallest distance from above data is

	P_1	P_2	P_3	P_4	P_5
P_1	0	0.2357	0.2218	0.3688	0.3421
P_2	0.2357	0	0.1483	0.2042	0.1388
P_3	0.2218	0.1483	0	0.1513	0.2843
P_4	0.3688	0.2042	0.1513	0	0.2932
P_5	0.3421	0.1388	0.2843	0.2932	0

CamScanner

Smallest distance from above data is 0.1388

* P_2 & P_5 forms 2nd cluster

	P_1	P_{25}	P_{36}	P_4
P_1	0	0.2357	0.2218	0.3688
P_{25}	0.2357	0	0.1483	0.2042
P_{36}	0.2218	0.1483	0	0.1513
P_4	0.3688	0.2042	0.1513	0

Smallest distance from above data is 0.1483

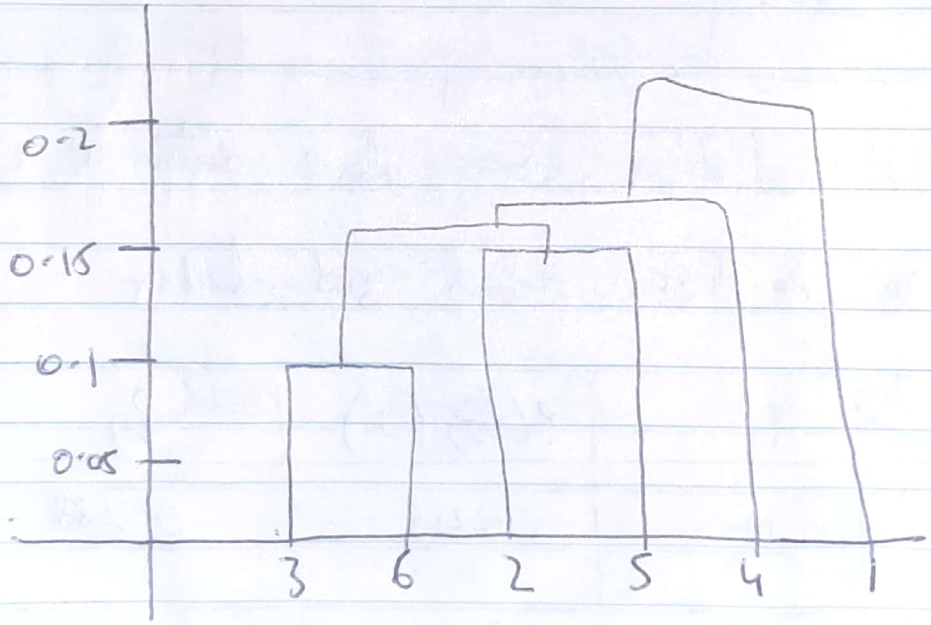
* P_{25} & P_{36} forms 3rd cluster

	P_1	$P_{(25)(36)}$	P_4
P_1	0	0.2218	0.3688
$P_{(25)(36)}$	0.2218	0	0.1513
P_4	0.3688	0.1513	0

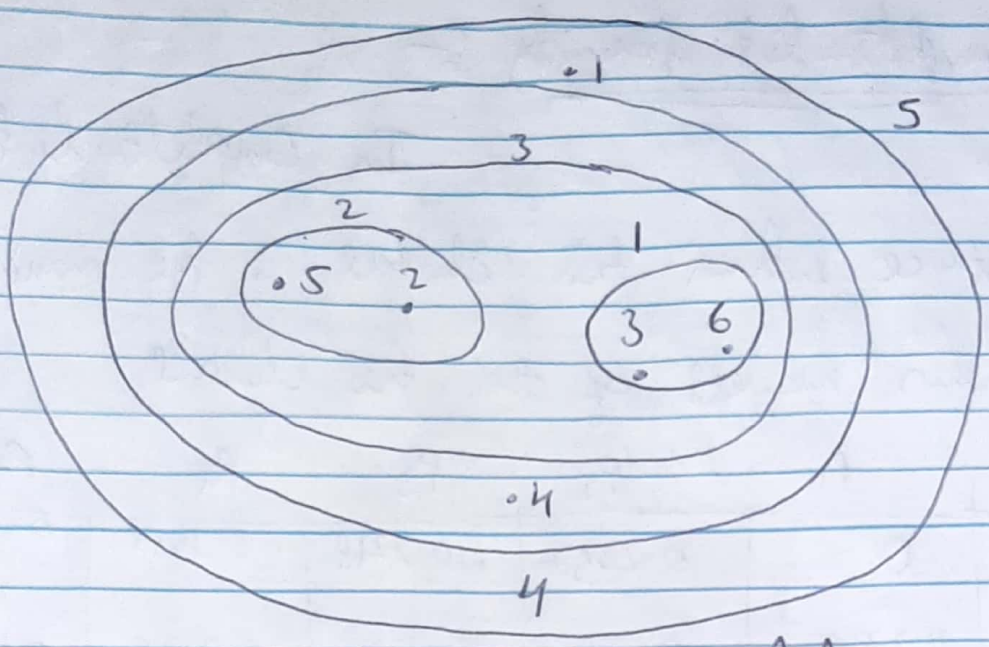
Smallest distance from above data is 0.153

* P(2,5) (3,6) & P4 forms 4th cluster

	P_1	$P_4(2,5)(3,6)$
P_1	0	0.2218
$P_4(2,5)(3,6)$	0.2218	0



Single link dendogram



Single link clustering Representation

Complete link proximity :-

In Complete linkage, the distance between two clusters is the maximum distance between members of the two clusters

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	0	0.2357	0.2218	0.3688	0.3421	0.2347
P_2	0.2357	0	0.1483	0.2042	0.1388	0.254
P_3	0.2218	0.1483	0	0.1513	0.2843	0.11
P_4	0.3688	0.2042	0.1513	0	0.2932	0.2216
P_5	0.3421	0.1388	0.2843	0.2937	0	0.3924
P_6	0.2347	0.254	0.11	0.2216	0.3924	0

Smallest distance from above data is 0.11

~~*~~ P_3 & P_6 form first cluster

	P_1	P_2	P_{36}	P_4	P_5
P_1	0	0.2357	0.2347	0.3688	0.3421
P_2	0.2357	0	0.254	0.2042	0.1388
P_{36}	0.2347	0.254	0	0.2216	0.3924

P_4	0.3688	0.2042	0.2216	0	0.2932
P_5	0.3421	0.1388	0.3921	0.2932	0

Smallest distance from above data is 0.1388

* P_2 & P_5 forms 2nd cluster

	P_1	P_{25}	P_{36}	P_4
P_1	0	0.3421	0.2342	0.3688
P_{25}	0.3421	0	0.3921	0.2932
P_{36}	0.2342	0.3921	0	0.2216
P_4	0.3688	0.2932	0.2216	0

Smallest distance from data is 0.2216

* P_{25} & P_{36} forms 3rd cluster

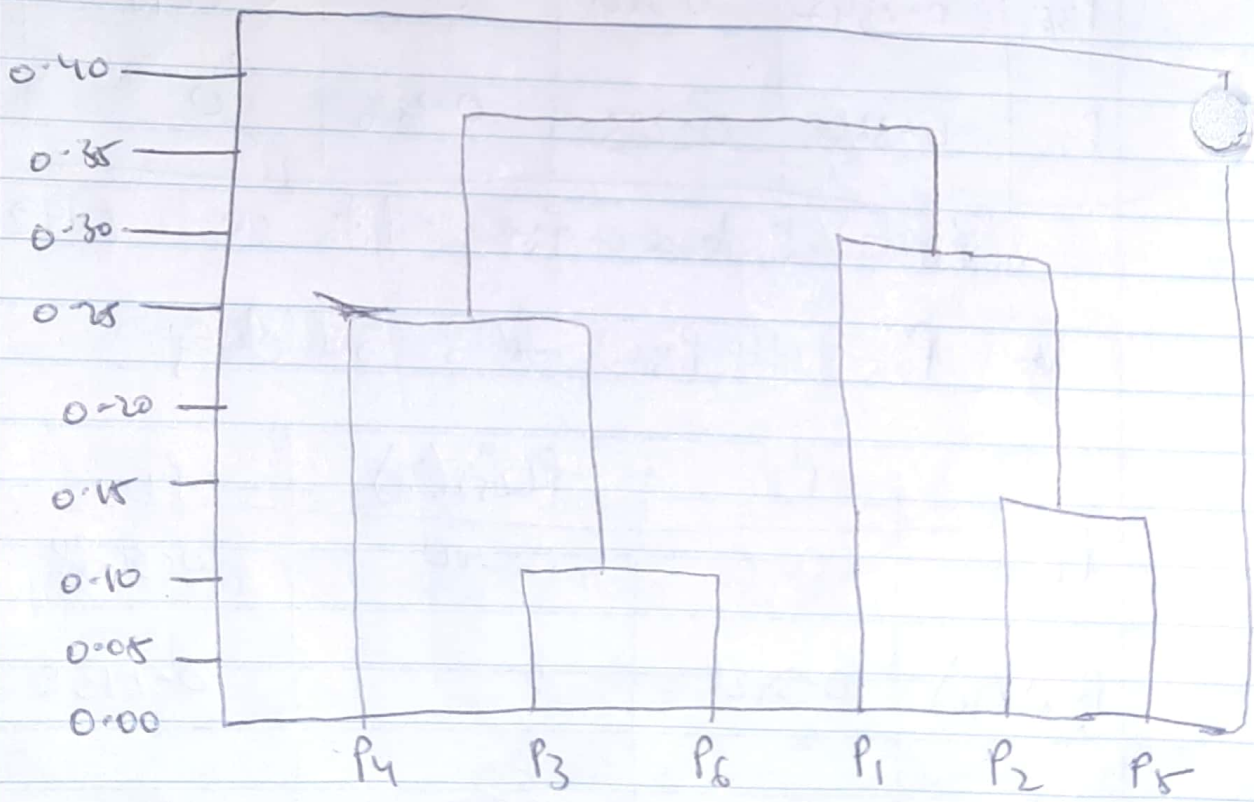
	P_1	$P_{(25)(36)}$	P_4
P_1	0	0.342	0.3688
$P_{(25)(36)}$	0.3421	0	0.2932
P_4	0.3688	0.2932	0

Smallest distance from above data is 0.2932

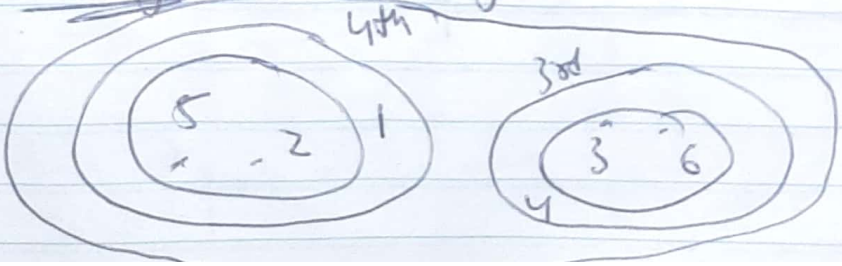
* P(25)(36) and P1 forms 4th cluster

	P1(25)(36)	P4
P1(25)(36)	0	0.1483
P4	0.3688	0

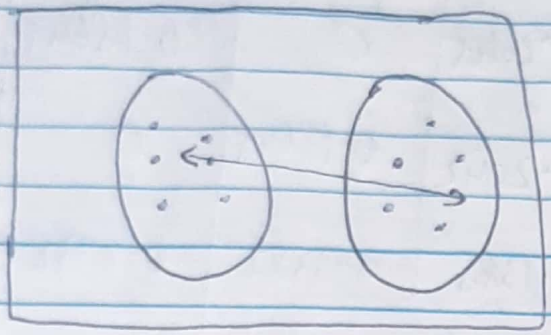
Complete linkage Dendrogram



Complete linkage clustering representation



Average link :- Distance between clusters is distance between the cluster centroids.



Distance matrix :-

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	0	0.2357	0.2218	0.3688	0.3421	0.2347
P_2	0.2357	0	0.1483	0.2042	0.1388	0.254
P_3	0.2218	0.1483	0	0.1513	0.2843	0.11
P_4	0.3688	0.2042	0.1513	0	0.2932	0.2216
P_5	0.3421	0.1388	0.2843	0.2932	0	0.3921
P_6	0.2347	0.254	0.11	0.2216	0.3921	0

Smallest distance from above data is 0.11

* P_3 & P_6 forms first cluster.

	P_1	P_2	P_{36}	P_4	P_5
P_1	0	0.2357	0.22825	0.3688	0.3421
P_2	0.2357	0	0.2015	0.2042	0.1388
P_{36}	0.22825	0.2015	0	0.18645	0.3382
P_4	0.3688	0.2042	0.18645	0	0.2932
P_5	0.3421	0.1388	0.3382	0.2932	0

Smallest distance from above data is 0.1388

* P_2 & P_5 forms 2nd cluster

	P_1	P_{25}	P_{36}	P_4
P_1	0	0.2889	0.2342	0.3688
P_{25}	0.2889	0	0.269675	0.2487
P_{36}	0.2342	0.269675	0	0.18645
P_4	0.3688	0.2487	0.18645	0

Smallest distance from above data is 0.18645

So P_{25} & P_{33} forms 3rd cluster

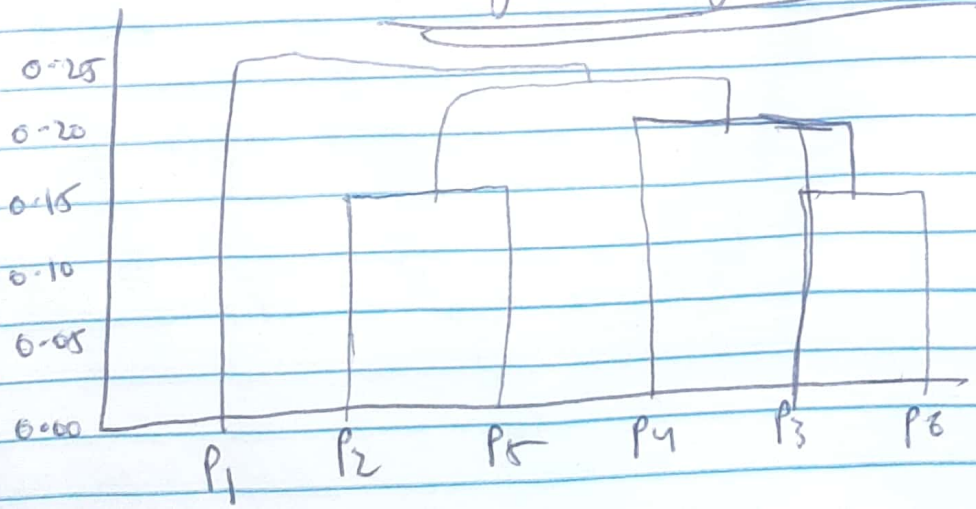
	P_1	$P(25)(36)$	P_4
P_1	0	0.2618	0.3688
$P_{25}(36)$	0.2618	0	0.217575
P_4	0.3688	0.217575	0

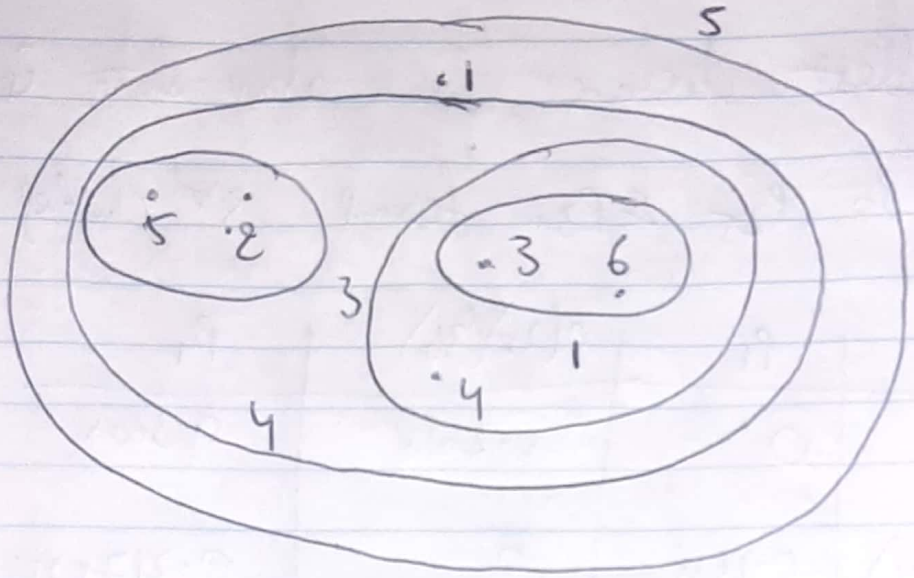
Table-4

The smallest distance from above data is 0.217575

	$P(25)(36)$	P_1
$P(25)(36)$	0	0.3153
P_1	0.3153	0

Average linkage Dendogram





Average clustering Representation