

Part1:

Supervised Pattern Recognition:

Supervised Pattern recognition makes use of data with a label associated with each data. These labels are used to design of classifier.

There are 2 phases in Supervised pattern recognition , a pre-processing and post processing phase.

Output of the Supervised is dependent on input provided. Training and testing phases are the 2 phases in Supervised Pattern Recognition.

End result in supervised pattern recognition is decision.

UnSupervised Pattern Recognition:

UnSupervised Pattern Recognition does not use labels. They don't have pre and post processing phase.

There are no training and testing phases. The algorithm changes according to the situation based on a measure.

End result in unsupervised pattern recognition is clusters.

Disadvantages of K-means Clustering:

1. Determining K value (number of clusters) is difficult.
2. The output is dependent on initial clustering. Different portioning results in different clusters.
3. Choosing clusters randomly may not yield accurate result.
4. The number of centres should be specified before the algorithm is started, which may not be the total number of partition at all.
5. Algorithm will fail for overlapping (close valued) data. It may also fail for non linear data.

Step 1: Dissimilarity Matrix

	#1	#2	#3	#4	#5
#1	0	3.16	4.24	2.23	1
#2		0	2	3.6	3.6
#3			0	5.38	5
#4				0	1.41
#5					0

The distance between #1 and #5 is the shortest. So, clustering both we get the matrix below.

Step 2: Dissimilarity matrix

	C1(#1,#5)	#2	#3	#4
C1(#1,#5)	0	3.16	4.24	1.41
#2		0	2	3.6
#3			0	5.38

#4				0
----	--	--	--	---

The distance between C1 and #4 is the shortest. So, clustering both we get the matrix below.

Step 3: Dissimilarity matrix

	$C2(\#1, \#5, \#4)$	#2	#3
$C2(\#1, \#5, \#4)$	0	3.16	4.24
#2		0	2
#3			0

The distance between #2 and #3 is the shortest. So, clustering both we get the matrix below.

Step 4: Dissimilarity matrix

	$C2(\#1, \#5, \#4)$	$C3(\#2, \#3)$
$C2(\#1, \#5, \#4)$	0	3.16
$C3(\#2, \#3)$		0

Step 5: Final matrix

	$C4(\#1, \#2, \#3, \#4, \#5)$
$C4(\#1, \#2, \#3, \#4, \#5)$	0

Dendrogram :

