← Go Back to Unsupervised Learning

Course Content

Weekly Quiz - Hierarchical Clustering and PCA

Type	:	Graded Quiz
Attempts	:	1/1
Questions	:	10
Time	:	30m
Due Date	:	Feb 26, 1:30 AM CET
Your Marks	:	15/15

Instructions

Attempt History

Attempt #1
Feb 19, 6:28 PM

Marks: 15

Q No: 1 Correct Answer

Marks: 2/2

Which of the following statement is/are true about the difference between PCA and Hierarchical Clustering?

Cluster analysis groups observations while PCA is used for dimensionality reduction.
PCA groups observations while cluster analysis is used for dimensionality reduction.
PCA can be used to reduce the number of variables in the data whereas cluster analysis cannot.
Clustering analysis can be used to reduce the number of variables in the data whereas PCA cannot.

1 and 2

3 and 4

1 and 3 You Selected

2 and 4

PCA extracts principal components which capture the highest variance in the data, while clustering forms clusters to maximize homogeneity within the clusters and heterogeneity between the clusters. PCA works column-wise whereas clustering works row-wise.

Q No: 2 Correct Answer

Marks: 1/1

PCA is used for

detecting outliers

reducing dimensions You Selected

detecting missing values

scaling the data for better model performance

PCA is a dimensionality reduction technique. From a given set of variables, we compute principal components that indicate the captured variance in the data. By choosing relevant principal components, we reduce the no. of dimensions in the data.

Q No: 3 Correct Answer

Marks: 2/2

In the case of a dataset with multiple numeric variables with different units of measurement, which of the below two statements hold true?

I. It is necessary to scale data before applying PCA

II. It is necessary to scale data before applying Hierarchical clustering

- ○ II only
- ○ I only
- ○ Both are false
- ◉ Both are true — You Selected

Since PCA and hierarchical clustering involve distance calculations, we need to scale the data to avoid the influence of the units of measurement.

Q No: 4 Correct Answer

Marks: 1/1

Covariance matrix is a mathematical representation of

- ◉ Variance of individual dimensions and covariance between pairs of dimensions — You Selected
- ○ Variance of individual dimensions only
- ○ Covariance between pairs of dimensions only
- ○ None of the above

In a covariance matrix, the diagonal values represent the variances of individual attributes, and the off diagonal values represent the covariance of the attributes corresponding to the respective row and column.

Q No: 5 Correct Answer

Marks: 2/2

If we have 4 components in PCA and the percentage of variance explained by each of them are 10%, 15%, 25%, and 50%, what percentage of variance will be explained by the first principal component?

25%

50% You Selected

10%

15%

The magnitude of the eigen value corresponding to a principal component determines the percentage of variance explained in the data. The principal components are chosen in the descending order of their magnitude. Hence, the first principal component has the highest eigen value and correspondingly explains the highest amount of variation in the data.

Q No: 6 Correct Answer

Marks: 1/1

Feature elimination techniques reduce dimensionality by creating few new variables using the original variables.

True

False You Selected

Feature extraction techniques reduce dimensionality by creating few new variables using the original variables, while feature elimination techniques involve dropping one or more of the original variables.

Q No: 7 Correct Answer

Marks: 2/2

What does measuring the distance between clusters A and B mean in the case of complete linkage?

Minimum Distance between pair of records in cluster A and B respectively

Maximum Distance between pair of records in cluster A and B respectively You Selected

Distance between centroids of cluster A and B

Average of distances between pair of records in cluster A and B

In the case of complete linkage, the distance between 2 clusters is measured as the maximum possible distance between the points in two different clusters.

Q No: 8 Correct Answer

Marks: 1/1

Which of the following linkage methods involves analysis of variance of clusters while combining clusters using the agglomerative approach of clustering?

Single Linkage

Average Linkage

Ward Linkage You Selected

Complete Linkage

The Ward linkage analyzes the variance of clusters. It measures how much the within-cluster sum of squares (WCSS) will increase when one cluster is merged with another and merges those two clusters such that the increase in WCSS is minimum.

Q No: 9 Correct Answer

Marks: 1/1

The angle between any two symmetric eigenvectors for a given matrix is

45 degrees

90 degrees You Selected

180 degrees

360 degrees

Symmetric eigen vectors are orthogonal to each other. So, the angle between them is 90 degrees.

Q No: 10 Correct Answer

Marks: 2/2

For a data matrix X with n rows and p columns, the number of eigenvalues possible for the covariance matrix of X is ___.

p-1

p+1

p^2

Previous Next