

Single Option Correct Questions

- Which of the following is NOT true about multicollinearity?
 - It is represented by a high correlation coefficient between pairs of variables
 - It causes the parameter estimates to be biased
 - It leads to the feature matrix not being a full-rank matrix
 - It causes one or more features to be estimable from other features
- Which of the following is NOT a feature of PCA?
 - It creates uncorrelated features
 - It creates mutually orthogonal features
 - The features created are linear combinations of the older features
 - None of the above
- What is meant by “preserving maximum information” in context of PCA?
 - The sum of squares of residuals is maximized
 - The variance captured is maximized
 - The mean squared error is minimized
 - The variance captured is minimized
- Which one of the following is NOT a pair of basis vectors of the 2D space?
 - \hat{i}, \hat{j}
 - $-\hat{i}, -\hat{j}$
 - $\frac{1}{\sqrt{2}}(\hat{i} + \hat{j}), \frac{1}{\sqrt{2}}(\hat{i} - \hat{j})$
 - $2\hat{i}, 2\hat{j}$
- Which is NOT the fundamental property of basis vectors in a vector space?
 - Linear combinations of basis vectors are basis vectors
 - They have unit magnitude
 - They are orthogonal
 - They can uniquely describe each point in the vector space
- How do you determine the covariance matrix for a dataset?
 - By calculating pairwise correlation coefficients and using eigen-decomposition
 - By calculating pairwise correlation coefficients and normalizing the matrix
 - By calculating pairwise covariances and normalizing the matrix
 - By calculating pairwise covariances
- How are the directions of the principal components determined?
 - By eigen-decomposing the feature matrix
 - By eigen-decomposing the covariance matrix
 - By factorizing the covariance matrix
 - By factorizing the feature matrix
- Which of the following statements regarding PCA is NOT right?

- a. Eigenvalue corresponding to a particular eigenvector represent the percentage variance explained by that eigenvector
 - b. Eigenvalues when performing PCA can be sometimes negative
 - c. Eigenvectors represent the directions of principal components
 - d. Eigenvectors represent linear transformation of the original basis vectors
9. Arrange the following steps for the PCA algorithm in the correct order
- i. Represent all data in a covariance matrix after normalizing them
 - ii. Eigen-decompose the covariance matrix to get set of eigenvectors and eigenvalues
 - iii. New basis vectors are created from eigenvectors and assigned principal components based on decreasing magnitude of its corresponding eigenvalues
 - iv. Points are assigned coordinates based on the new basis vectors
- a. I -> II -> III -> IV
 - b. II -> I -> III -> IV
 - c. III -> II -> I -> IV
 - d. IV -> I -> II -> III
10. How do you choose the value of k in KMeans?
- a. Kink in the elbow plot
 - b. Minimizing silhouette score
 - c. Maximizing area under the elbow curve
 - d. All of the above
11. What pre-processing step do you need to perform before fitting the data to KMeans?
- a. Standard Scaling
 - b. Label encoding of categorical variables
 - c. One hot encoding of categorical variables
 - d. All of the above
12. What is NOT one of the drawbacks of the KMeans method?
- a. The value of k has to be set beforehand
 - b. The clusters are spherical in shape
 - c. The loss function is the within cluster sum squared value
 - d. Categorical variables need to be dropped before
13. To solve which hindrance of the KMeans method was the KMeans++ method devised?
- a. KMeans suffers from the problem of the loss function being composed of Euclidean distances
 - b. KMeans clustering is susceptible to cluster centroid initialization
 - c. KMeans requires input of k, the number of clusters
 - d. KMeans cannot return non-spherical clusters
14. What is/are the advantage(s) of hierarchical clustering over KMeans?
- a. The number of clusters do not need to be passed

- b. The distance measure can be non-euclidean
- c. We can select arbitrary number of clusters based on threshold values
- d. All of the above

15. What does DBSCAN stand for?

- a. Density Based Statistical Clustering with Amplified Noise
- b. Density Based Spatial Clustering of Applications with Noise
- c. Density Based Statistical Clustering of Applications with Noise
- d. Density Based Spatial Clustering with Amplified Noise

16. What is the complexity vs interpretability tradeoff?

17. What is the bias vs variance tradeoff?

18. What is NOT true in the context of overfitting?

- a. It occurs when the model has learnt the test data really well, but cannot generalize to train data
- b. It is undesirable
- c. It is usually associated with more complex models
- d. We can reduce overfitting by creating bootstrapped training datasets

19. Why do we need cross-validation?

20. What are the types of regularization present in Ridge and Lasso and ElasticNet?

- a. L1 for Ridge, L2 for Lasso, L1 & L2 for ElasticNet
- b. L1 for Lasso, L2 for Ridge, L1 & L2 for ElasticNet
- c. L1 for ElasticNet, L2 for Ridge, L1 & L2 for Lasso
- d. L1 for ElasticNet, L2 for Lasso, L1 & L2 for Ridge