

Department of Computer Science and Engineering

Indian Institute of Technology Jodhpur

CSL 7620: Machine Learning

Quiz 2

Total marks: 10

Set A

Time: 10 minutes

Instructions:

- During the exam, invigilators will not address any queries. If you encounter anything unclear or incorrect in a question, make a reasonable assumption and proceed
 - Each question may have multiple correct options. You have to mark all the correct options. Otherwise, no marks will be awarded.
 - Mark the correct choices with a pen on the OMR sheet. No correction is allowed in the marking.
-

[Each question carries 1 mark]

1. Which of the following best describes the main goal of Principal Component Analysis (PCA)?
 - A. Maximize class separation
 - B. **Reduce dimensionality by capturing variance**
 - C. None of the others
 - D. Cluster data into groups
2. K-Means clustering aims to minimize:
 - A. **Total within-cluster sum of squared distances**
 - B. Between-cluster variance
 - C. Euclidean distance between cluster centroids
 - D. Sum of distances between every pair of points
3. In DBSCAN, let ϵ (disc radius) = 2, MinPts (Minimum number of points for considering a point as a core point) = 3. Consider the following 1D points:

$$x = \{1, 1.5, 2, 8, 8.5, 9, 20\}$$

Which of the following statements are true (consider Euclidean distance)?

- The first and the third points (from left) in x will be included in the same cluster
 - The fourth and the sixth points (from left) in x will be included in the same cluster
 - The last point (from left) in x will be labeled as an outlier
 - There will be exactly 3 clusters
- In Linear Discriminant Analysis (LDA), what is optimized to separate classes?
 - Maximizing between-class variance**
 - Minimizing within-class variance
 - Maximizing the distance between clusters and the decision boundary
 - Finding orthogonal basis vectors for data decorrelation
 - Given the Covariance matrix of a dataset:

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The first principal component using PCA is.

- A. [0, 1]
 - B. [1, 0]**
 - C. [2, 0]
 - D. None of the others
6. Compared to K-Means, GMM clustering:
- A. Allows soft assignment of points to clusters.**
 - B. Automatically finds out the number of clusters
 - C. None of the others
 - D. Assumes all clusters have the same variance.
7. Which of the following is/are the possible termination criteria used in GMM?
- A. If μ_i and Σ_i for each cluster, do not change significantly over a few epochs.**
 - B. When the log likelihood of all data points does not change over a few epochs.**
 - C. When all the clusters are assigned with equal number of data points.
 - D. None of the others.
8. In Linear Discriminant Analysis (LDA), which of the following best explains the criterion used to find the optimal projection direction?
- A. The ratio of total scatter to within-class scatter is maximized.
 - B. The difference between total scatter and within-class scatter is minimized.
 - C. The ratio of between-class scatter to within-class scatter is maximized.**
 - D. None of the others
9. A binary classifier uses Naive Bayes to predict whether an email is *spam* (*S*) or *not spam* (*N*).

Given: $P(S)=0.4$, $P(N)=0.6$ and the word "offer" appears with

$$P(\text{"offer"} | S) = 0.7, P(\text{"offer"} | N) = 0.1$$

What is the class with the higher posterior probability for an email containing the word "offer"?

- A. Spam**
 - B. Not Spam
 - C. Both classes have equal probabilities
 - D. Cannot be determined
10. In a GMM with two 1D components:
- Mixing Coefficients: $\pi_1 = 0.3, \pi_2 = 0.7$,
 - Means: $\mu_1 = 2, \mu_2 = 5$,
 - Standard deviations: $\sigma_1 = \sigma_2 = 1$

Find the responsibility $\lambda_1(x = 3)$, i.e., the probability that the point $x = 3$ belongs to the first component. Choose the nearest value(s):

- A. 0.10
- B. None of the others
- C. 0.50
- D. 0.66**