

UNIT 3 DIMENSIONALITY REDUCTION, SUPERVISED LEARNING, BAYESIAN DECISION THEORY

1. Give any two reasons with examples why we need to do 'dimensionality reduction'. What are the techniques of dimensionality reduction? Explain **any one** of them with an example.
2. Describe the similarities & differences between following dimensionality reduction techniques:
(i) LDA and PCA (ii) PCA and ICA (iii) LDA and ICA
3. Illustrate Principal Component Analysis with an example
4. Explain the working of PCA as a dimensionality reduction technique using the feature vector

$$\begin{pmatrix} \mathbf{x_1} & \mathbf{x_2} \\ 2 & 4 \\ 1 & 3 \\ 0 & 1 \\ -1 & 0.5 \end{pmatrix}$$

(**Tip:** Make sure to clearly depict the values of Covariance matrix, Eigenvalues, Eigenvectors as well as the linear transformed feature vector).

5. What do you understand by the following terms w.r.t LDA:
(i) Fischer Ratio (ii) Scatter Within S_w (iii) Scatter Between S_B
6. Explain the working of LDA as a dimensionality reduction technique using the feature vectors

Samples for class ω_1 : $\mathbf{X}_1 = (x_1, x_2) = \{(4,2), (2,4), (2,3), (3,6), (4,4)\}$

Sample for class ω_2 : $\mathbf{X}_2 = (x_1, x_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$

(**Tip:** Make sure to clearly depict the values of S_w , S_B matrices, Eigenvalues, Eigenvectors as well as the linear transformed feature vector).

7. What are the limitations of PCA, LDA and ICA?
8. What is utility theory? How is it used in Bayesian Learning models?
9. How are probabilities estimated? What is the role of **m-estimate** in the estimation of probabilities?
10. What do you understand by 'Independent Component Analysis(ICA)'? Where is it used?
11. With example illustrations explain why LDA uses 'mean' as well as 'two types of scatter' while PCA uses only one type of 'scatter' (variance).
12. Take the covariance matrix

$$\begin{pmatrix} 1.667 & 2.08 \\ 2.08 & 2.727 \end{pmatrix}$$

and use PCA to arrive at the Eigenvalues, Eigenvectors. Which Eigenvalues will you consider for the linear transformation along PC1? What is the reason for your choice?

13. Take the scatter matrices (S_w , S_B) given below:

S_B	S_w
-------	-------

$$\begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} \quad \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}$$

and use LDA to arrive at the Eigenvalues, Eigenvectors. Which Eigenvalues will you consider for the linear transformation? What is the reason for your choice?

14. Exercises to Chapter 3 in Alpaydin *Bayesian Decision Theory*, all questions of Section 3.8.

15. Discuss how ‘Maximum Likelihood Hypothesis’ can be used in predicting probabilities.

16. With examples explain what do you understand by the terms:

(i) Probably Approximately Correct (PAC) Learning

(ii) VC (Vapnik-Chervonenkis) Dimension

17. “The loss for a high-risk applicant incorrectly accepted may be different from the potential gain for an erroneously rejected low-risk applicant”. Justify using algorithm of ‘Losses & Risks’.

18. How are losses & risks modeled in Bayesian Decision Theory? In a two-class, two-action problem, if the loss function is $\lambda_{11} = \lambda_{22} = 0$, $\lambda_{12} = 10$, $\lambda_{21} = 1$, write the optimal decision rule.

19. Explain with an example how ‘Association Rule’ is applied to Basket Analysis.

20. What is Minimum Description Length principle? How is it used in Bayesian methods?

21. What do you understand by the following terms w.r.t ‘Association Rule’:

(i) Support (ii) Confidence (iii) Lift (iv) Antecedent (v) Consequent

22. Discuss optimal classifier using Bayes method. Suppose h_1, h_2, h_3 are three hypothesis and their posterior probabilities are 0.4, 0.3, 0.3 respectively. Use optimal Bayes classifier to find the most probable classification of the new instance given the training data.

23.(a) Suppose the fraction of undergraduate students who smoke is 15% and the fraction of

graduate students who smoke is 23%. If one-fifth of the college students are graduate students and the rest are undergraduates, what is the probability that a student who smokes is a graduate student?

- (b)** Given the information in part (a), is a randomly chosen college student more likely to be a graduate or undergraduate student?
- (c)** Repeat part **(b)** assuming that the student is a smoker.
- (d)** Suppose 30% of the graduate students live in a dorm but only 10% of the undergraduate students live in a dorm. If a student smokes and lives in the dorm, is he or she more likely to be a graduate or undergraduate student? You can assume independence between students who live in a dorm and those who smoke.