Handwritten Assignment-2
CSL2050 - Pattern Recognition and Machine Learning

NOTE:

1. This Handwritten Assignment contains ten problems. Total points for this assignment is 70. Please use A4 sheets and neatly solve all the problems in the proper order. Be precise, verbosity and messy writing will be penalized.
2. **(IMPORTANT)** Please write the following pledge in your handwriting with your signature on the first page of your assignment sheet (without this, the assignment will not be considered for evaluation):

   Honesty Pledge:

   "I affirm that this assignment is solely my work. I have not used unauthorized assistance, engaged in plagiarism, or violated ethical standards. Further, all references used and any discussion with anyone have been appropriately cited. Any breach may lead to disciplinary actions as per the course academic honesty policy discussed in Lecture-1."
3. **Deadline:** March 20, 2024, 9:05 AM before the class. Note please make the submission in handwritten sheet (hard copy).
4. **Late Submission Policy:** Late submissions will not be considered for this assignment.

---

1. **(4 points)** Compute mean vector and covariance matrix for the following data points: (1,2), (3,4), (4,5), (5,5), (5,6), (10,10).
2. **(PCA)** Given the following points: (-1,1), (-2,2), (-3,3).

   (a) **(2 points)** Compute zero mean data points

   (b) **(2 points)** Compute covariance matrix of zero mean data points.

   (c) **(2 points)** Find out the Eigen values and Eigen vectors of the covariance matrix.

(d) **(2 points)** Project all the zero mean data points to the first principle component.

(e) **(2 points)** Reconstruct the projected data points and compute reconstruction error.

3. **(6 points)** You are given a dataset of fruits that are either apples or oranges. Each fruit has two features: color (RED or ORANGE) and shape (ROUND or NOT ROUND). You are also provided with the following probabilities: $P(apple) = 0.6, P(orange) = 0.4$. $P(RED|Apple) = 0.8.P(Round|apple) = 0.7.P(RED|orange) = 0.6.P(Round|range) = 0.3$. (note: Capital letters ORANGE denotes color while small letters orange denotes fruit). Now, you have a fruit with the following characteristics: Color: RED, Shape: ROUND Using the Bayesian classification, determine whether the fruit is more likely to be an apple or an orange. Calculate the posterior probabilities for both classes and provide your classification decision.

4. **(5 points)** Compute the error probability of using Bayesian classification for the following two distributions: $P(X|+) : U[3, 6], P(X|-) : U[1, 4]$ and $P(+) = 0.8, P(-) = 0.2$. Here U denotes Uniform Distribution.

5. **(5 points)** Define Covariance Matrix and write down five Properties of Covariance Matrix.

6. **(6 points)** Prove that the zero mean data when projected to the space of Eigen Vectors of its covariance matrix, it remains zero mean.

7. **(8 points)** Suppose you are given the following set of data with three Boolean input variables a,b, and c, and a single Boolean output variable K.

1	0	1	1
1	1	1	1
0	1	1	0
1	1	0	0
1	0	1	0
0	0	0	1
0	0	0	1
0	0	1	0

(a) According to the naive Bayes classifier,what is $P(K = 1|a = 1 \wedge b = 1 \wedge c = 0)$? (b) According to the naive Bayes classifier, what is $P(K =$

$0|a = 1 \wedge b = 1)$? (c) According to the joint Bayes classifier,what is $P(K = 1|a = 1 \wedge b = 1 \wedge c = 0)$? (d) According to the joint Bayes classifier,what is $P(K = 0|a = 1 \wedge b = 1)$?

8. **(10 points)** Solve Problem 33 of Textbook of Chapter 4 (Duda, Hart and Stroke). (Also study how LDA can be extended to multi class cases, ref. Section 4.11 Multiple Discriminant Analysis)

9. **(10 points)** Solve Problems 1 and 2 of Chapter 3 (Duda, Hart, and Stroke).

10. **(6 points)** Let X be a discrete random variable with support $\{-2, -1, 0, 1, 2\}$. Further, suppose it follows the uniform distribution U for the aforementioned support. Compute expected value and variance for X.

End of Paper