Machine Learning BITS F464	04-11-2022

Sem 1 2022-23 Mid Sem Exam, BITS Pilani KK Birla Goa Campus

Instructor in Charge: Harikrishnan N B   *Student:*

READ THE FOLLOWING CAREFULLY:

- **Honour Code for Students:** I shall be honest in my efforts and will make my parents proud. Write the oath and sign it on the answer sheet.

- **Duration: 9:30 AM - 11:00 AM**

- **Each question carries six marks (Max. Marks = 30).**

- **Students can use calculator. Mobile phones are strictly prohibited.**

- **Answer all the questions**.

1. Solve the following problem using ideas from **Linear Algebra**: Three ponds and three temples, stand interspersed, one after the other : P1 T1 P2 T2 P3 T3. You're bringing some flowers with you to the first pond (P1). Here's what you do at each pond:

    - You take a dip in the pond and the number of flowers gets doubled.
    - You keep some flowers at the temple next to the pond, and move on to the next pond with the remaining flowers.

   You should keep the same number of flowers at each temple. After you visit the last temple, you shouldn't have any flowers left. Answer the following questions.

   (a) How many flowers should you carry with you in the beginning?
   (b) For the same problem, find a generalized solution for n ponds and n temples (where after you visit the n-th temple, you shouldn't have any flowers left).

2. Answer the following questions.

   (a) Write any three assumptions of Principal Component Analysis (PCA)?
   (b) What is the assumption in k-Nearest Neighbour classifier?
   (c) What is the assumption in perceptron algorithm for classification?

3. Consider the singular value decomposition (SVD) of the matrix $A = U\Sigma V^T$. Answer the following.

   (a) The columns of U are the eigenvectors of _____ (Explain).
   (b) The columns of V are the eigenvectors of _____ (Explain).

4. Definition: Given random variables X, Y, and Z. X is conditionally independent of Y given Z, denoted by $X \perp Y|Z$, if and only if:

$$P(X = x_i|Y = y_i, Z = z_k) = P(X = x_i|Z = z_k)\forall i, j, k$$

   Given this definition, answer the following questions:

(a) Given $X \perp Y|Z$, can we say $P(X, Y|Z) = P(X|Z)P(Y|Z)$? Explain.

(b) Given $X \perp Y|Z$, can we say $P(X, Y) = P(X)P(Y)$? Explain.

(c) What are the assumptions in Gaussian Naive Bayes Classifier?

(d) Alice and Bob loves playing tennis. As a BITS F464 Machine Learning student help Alice and Bob to develop a Naive Bayes Classifier to decide whether they should play tennis or not given the attributes: Outlook = Sunny, Temperature = Cool, Humidity = High, and Windy = Strong. The data recordings are provided in Table 0.1.

Table 0.1: Weather Data Set (all nominal).

Item No.	Outlook	Temperature	Humidity	Windy	Play
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rainy	Mild	High	Weak	Yes
5	Rainy	Cool	Normal	Weak	Yes
6	Rainy	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rainy	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rainy	Mild	High	Strong	No

5. Answer the following questions.

   (a) Show that Logistic Regression is a linear classifier?

   (b) An experiment consists of tossing 2 coins simultaneously. One of the coins is unbiased, but the other coin has HEADS on both sides. List the sample space and find the self-information of all possible outcomes of this random experiment. What is the Shannon Entropy of this experiment?