END SEM , ME-781, November 23, 2021

Max Marks: 150, Total time: 2 Hr

No explanation for any question would be provided. Please make any assumptions as you see fit and solve the questions.

This is an open-notes exam. You are not allowed to use a computer or calculator. You are not required to numerically solve an arithmetic expression.

5
10
10

1.
   a. In standardization, we transform data to have a mean of 0 and a standard deviation of 1. For a simple linear regression with standardized data what would be the value of β_0.
   b. For a generalized multiple linear regression find the dot product of $(\boldsymbol{y_i} - \hat{\boldsymbol{y}}_{\boldsymbol{i}})$ and $(\beta_0 + \beta_1 x_i)$.

5
5

2. In a supervised learning model, briefly explain how to handle:
   a. when the predictor variable is of nominal data type
   b. when the output (the response) variable is of nominal data type

5
5
5

3. An experiment of throwing two unbiased dice (each having numbers from 1 to 6) is performed. The result of the experiment is the product of the numbers obtained in the two dice.
   a. What is the sample space (set of all possible outcomes) for this experiment?
   b. What is the smallest Sigma algebra for this experiment?
   c. Give an example of a Sigma algebra that is greater than the smallest Sigma Algebra and smaller than the power set for this experiment.

5

4. For a random household with two children, let the event that the first child is a boy be ***A***, and the event that both the children are boys be ***B***. Using conditional probability, find out whether these two events are independent or not.

5
5

5. Eigenvalue and eigenvector for matrix

$\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$ is $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 & -\frac{1}{2} \\ 1 & 1 \end{bmatrix}$

Then find the eigenvalues and eigenvectors for the following matrix:

$$\begin{bmatrix} 18 & 7 \\ 14 & 11 \end{bmatrix}$$

2+2
2+4

6. Two data points in a three-dimensional space are (7,11,13) and (13,7, 11). Find their Euclidean norm, Frobenius norm, and Diagonal norm (with the diagonals being -1,10,1) . Also, find the Cosine similarity for these today points. You need not numerically solve this arithmetic expression of your answer.

5
5

7. If the underlying relation between a predictor and response variable is linear (with an added random noise of zero mean). Then, briefly explain how would the Precision and Accuracy of a linear regression model change with
    a. a decrease in the number of data points for training and
    b. an increase in the variance of random noise.

5
5
5
5

8. <u>Seven</u> random points are sampled from a population of independent and identically distributed random variables (of mean 0 and variance 3). What would be the expected value of:
    a. Mean value of the sample
    b. The variance of the sum of is samples
    c. The variance of the mean value of the samples
    d. If the population (of independent and identically distributed random variables) is normally distributed, then how would the mean value of the sample be distributed?

5
5

9. For which of the following two cases of multiple linear regression is the null hypothesis more likely to be true:
    - F=1.1 for the F-statistic and number of data points, n= 50, and the number of predictor variables, P =45
    - F=1.1 for the F-statistic and number of data points, n= 50, and the number of predictor variables, P =25

5
5

10.
    a. The leave-one-out cross-validation process is an extreme case of K-fold cross-validation. If the total size of the data set is n, how much more computation has to be done to perform leave-one-out cross-validation instead of K fold cross-validation?
    b. A measure to assess the goodness of a machine learning model is to compare the mean squared error of testing. An estimate of the mean square error for testing for least square models is Cp and BIC. Show that for any reasonably sized dataset, BIC is a more stringent criterion. What could be the size of a reasonably sized dataset?

5
5

11. In a 2D space, one can define a reflection transformation about y-axis as (x,y) transforming to (-x,y). Similarly, one can define reflection about the origin as (x,y) transforming to (-x,-y).
    a. Write the transformation Matrix for these two transformations.
    b. Are these transformations orthogonal?

5

12. A classification model is developed to predict defective parts (classify the product in one of the two classes: Defective, Non-defective) in a production process. The model is able to predict 99% of the time the correct class. Is this a good classification model? Briefly explain and provide an alternative methodology to assess the goodness of the model.

5

13. In K-Means Clustering Algorithm, there could be many different possible cluster formations depending upon the initial random assignment. Could you suggest a possible solution to take care of this problem?

5

14. The Projection Pursuit Regressions were the precursors to neural networks. For a two-dimensional predictor variable space (X_1, X_2), suggest possible ω_m, p-vectors, such that projection along these when used in the projection Pursuit regression can provide $X_1 X_2$ term.