No. of Printed Pages : 6

MST-026

M. SC. (APPLIED STATISTICS)

(MSCAST)

Term-End Examination

December, 2025

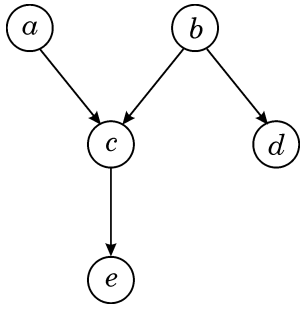
MST-026 : INTRODUCTION TO MACHINE LEARNING

Time : 3 Hours *Maximum Marks : 50*

Note : *Question No. **1** is compulsory. Attempt any **four** questions from question numbers 2 to 6. Use of scientific (non-programmable) calculator is allowed. Symbols have their usual meanings.*

1. (a) Consider the following graphical model, mark which of the following pairs of random variables are independent

given no evidence ? Give reason in support of your answer : 2



(i) a, d

(ii) e, d

(iii) c, e

(iv) a, b

(b) If $g(z)$ is the sigmoid function, then its derivative with respect of z may be written in term of $g(z)$. Which one of the following options is correct ? Give reason in support of your answer : 3

(i) $g(z)\,(g(z) - 1)$

(ii) $g(z)\,(1 + g(z))$

(iii) $-\,g(z)\,(1 + g(z))$

(iv) $g(z)\,(1 - g(z))$

(c) Assume, you want to cluster 7 observations into 3 clusters using k-means clustering algorithm. After iteration, the clusters : C1, C2, C3 have the following observations : 5

C1 : {(1, 1), (4, 4), (7, 7)}

C2 : {(0, 4), (4, 0)}

C3 : {(5, 5), (9, 9)}

Find the cluster centroids after the first iteration.

2. What is the role of an activation function in neural networks ? Give example of *two* commonly used activation functions. 10

3. Solve the network to approximate the function $g(x) = 1 + \sin\left(\frac{\pi x}{4}\right)$ for $-2 \leq x \leq 2$, choosing initial weights and bias as the random numbers. 10

4. (a) Consider the following data for 800 instances of home, 900 instances of office and 1050 instances of factory type buildings :

Building	Balcony	Multi-storeyed	Power backup	Total
Home	400	200	100	700
Office	300	150	450	900
Factory	150	450	450	1050
Total	850	800	1000	2650

Suppose a building has a balcony and power-backup, but is not multistoreyed. According to the Naive Bayes algorithm, it is of type (i) Home, (ii) Office, (iii) Factory. Give detailed explanation. 6

(b) Consider the problem of learning a function $X \rightarrow Y$, where Y is boolean, X is an input vector (X_1, X_2), where X_1 is categorical and takes 3 values, and X_2

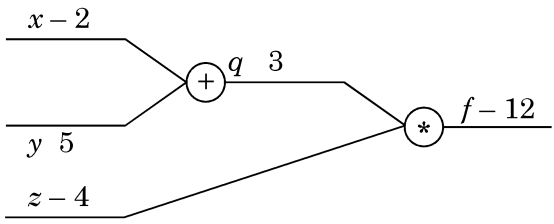
is a continuous variable (normally distributed). What would be the minimum number of parameters required to define a Naïve Bayes model for this function ? 4

5. (a) Suppose you have inputs as x, y and z with values – 2, 5 and – 4, respectively. You have a neuron 'q' and neuron 'f ' with functions :

$$q = x + y$$

$$f = q * z$$

Graphical representation of the functions is as follows :



What is the gradient of f with respect to x, y and z ? 5

(b) Suppose you have run Adaboost on a training set for three boosting iterations. The results are classified as h_1, h_2, h_3 with coefficients $\alpha_1 = 0.2$, $\alpha_2 = -0.3$ and $\alpha_3 = -0.2$. For a given test input x, you find that the classified results are $h_1(x) = 1$, $h_2(x) = 1$ and $h_3(x) = -1$. What is the class returned by the Adaboost ensemble classifier H on test example x ? 5

6. Write any *five* applications of self-organising map. 10

× × × × ×