CSE 201

B.Tech. III[rd] SEMESTER EXAMINATION, 2024-25

B.TECH.

(Mathematics for Machine Learning)

(CSE)

AFFIX PRESCRIBED RUBBER STAMP

Date (तिथि) : 17/12/2024.

Paper ID
(To be filled in the OMR Sheet)

2109

अनुक्रमांक (अंकों में) :
Roll No. (In Figures) :

अनुक्रमांक (शब्दों में) : _____

Roll No. (In Words) : _____

Time : 1:30 Hrs.
समय : 1:30 घण्टे

Max. Marks : 75
अधिकतम अंक : 75

नोट : पुस्तिका में 50 प्रश्न दिये गये हैं, सभी प्रश्न करने होंगे। प्रत्येक प्रश्न 1.5 अंक का होगा।

Important Instructions :

1. The candidate will write his/her Roll Number only at the places provided for, i.e. on the cover page and on the OMR answer sheet at the end and nowhere else.
2. Immediately on receipt of the question booklet, the candidate should check up the booklet and ensure that it contains all the pages and that no question is missing. If the candidate finds any discrepancy in the question booklet, he/she should report the invigilator within 10 minutes of the issue of this booklet and a fresh question booklet without any discrepancy be obtained.

महत्वपूर्ण निर्देश :

1. अभ्यर्थी अपने अनुक्रमांक केवल उन्हीं स्थानों पर लिखेंगे जो इसके लिए दिये गये हैं, अर्थात् प्रश्न पुस्तिका के मुख्य पृष्ठ तथा साथ दिये गये ओ०एम०आर० उत्तर पत्र पर, तथा अन्यत्र कहीं नहीं लिखेंगे।
2. प्रश्न पुस्तिका मिलते ही अभ्यर्थी को जाँच करके सुनिश्चित कर लेना चाहिए कि इस पुस्तिका में पूरे पृष्ठ हैं और कोई प्रश्न छूटा तो नहीं है। यदि कोई विसंगति है तो प्रश्न पुस्तिका मिलने के 10 मिनट के भीतर ही कक्ष परिप्रेक्षक को सूचित करना चाहिए और बिना त्रुटि की दूसरी प्रश्न पुस्तिका प्राप्त कर लेना चाहिए।

1. How are input layer units connected to second layer in competitive learning networks ?

(A) Feed forward manner

(B) Feedback manner

(C) Feed Forward and feedback

(D) None

2. Which of the following is not a supervised machine learning Algorithm ?

(A) K-means

(B) Naive-Bayes

(C) SVM

(D) Decision Tree

3. An artificially intelligent car decreases its speed based on its distance from the car in front of it. Which algorithm is used ?

(A) Naive - Bayes

(B) Decision Tree

(C) Linear Regression

(D) Logistic Regression

4. The learner is trying to predict the cost of laptop based on its features/brand. The variable cost is -

(A) Independent variable

(B) Target variable

(C) Ranked variable

(D) Categorical variable

5. The learner is trying to predict the housing prices based on the size of the each house. The variable "Size" is -

(A) Dependent Variable

(B) Label Set Variable

(C) Target Variable

(D) Independent Variable

6. Among the following identify the one in which dimensionality reduction reduces.

(A) Performance

(B) Entropy

(C) Stochastics

(D) Collinearity

7. Which of the following is an unsupervised learning ?

(A) Naive Bayesian

(B) Linear Regression

(C) Decision Tree

(D) PCA

8. Identify the difficulties with the K – nearest neighbor algorithm .

(A) Cause of dimensionality

(B) Calculate the distance of the test case from all training cases

(C) Both (A) & (B)

(D) None

9. For the Binomial Variable, what is the variance in case of Beta Distribution ?

(A) $\frac{ab}{(a+b)(a+b+1)}$

(B) $\frac{ab}{(a+b)^2(a+b)}$

(C) $\frac{ab(a+b)}{(a+b+1)}$

(D) $\frac{ab}{(a+b)^2(a+b+1)}$

10. What will be the mean and variance in case of Gaussian distribution ?

(A) $\sum(x) = \mu$, $V(x) = \sigma^2$

(B) $\sum(x) = \mu^2$, $V(x) = \sigma$

(C) $\sum(x) = \sqrt{\mu}$, $V(x) = \sigma^2$

(D) $\sum(x) = \mu$, $V(x) = \sqrt{\sigma}$

11. Consider a variable having 8 possible states {a,b,c,d,e,f,g,h} for which the respective probabilities are given as $\left(\frac{1}{2},\frac{1}{4},\frac{1}{8},\frac{1}{16},\frac{1}{64},\frac{1}{64},\frac{1}{64},\frac{1}{64}\right)$ the find the entropy ?

(A) 2 bits

(B) 3 bits

(C) 4 bits

(D) 16 bits

12. For a multinomial variables, the normalized form of the distribution is given as $\frac{\Gamma(\alpha_o)}{\Gamma(\alpha_1).....\lceil\alpha_K}\prod_{K=1}^{K}\mu_K^{\alpha-1}$. Where $\Gamma(x)$ is called the gamma function this Distribution is called -

(A) Beta Distribution

(B) Gaussian Distribution

(C) Dirichlet Distribution

(D) Poisson Distribution

13. The probability distributions are specific example of a broad class of distributions called -

(A) Information Theory

(B) Decision Theory

(C) Bayesian Model

(D) The exponential Family

14. The exponential family of distributions over x_1 given parameters η is as -

(A) $P(x/\eta) = h(x)\, g(\eta) \exp\{\mu^T \eta(x)\}$

(B) $P(x/\eta) = h(x)\, g(\eta) \exp\{\eta^T \mu(x)\}$

(C) $P(x/\eta) = \frac{h(x)}{g(\eta)} \exp\{\eta^T \mu(x)\}$

(D) $P(x/\eta) = \frac{h(x)}{g(\eta)} \exp\{\mu^T \eta(x)\}$

15. For the linear models of Regression.

$$y(x, \omega) = \omega_0 + \sum_{y=1}^{m-1} \omega_j \, \emptyset_j(x)$$

Here the $\emptyset_j(x)$ is the useful class of functions by taking linear combinations of non linear functions of the input variables, known as -

(A) Dummy function

(B) Parametric function

(C) Basis function

(D) None

16. In case of Bias – Variance Decomposition the expected loss = _____.

(A) $(\text{Bias})^2 + (\text{Variance})^2 + \text{noise}$

(B) $(\text{Bias}) + (\text{Variance})^2 + \text{noise}$

(C) $(\text{Bias})^2 + (\text{Variance}) + \text{noise}$

(D) $(\text{Bias}) + (\text{Variance}) + \text{noise}$

17. Consider the problem, plot the regression line and predict the sales at 12th week.

Week	Sales in (thousand)
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

(A) 7.26

(B) 5.16

(C) 8.46

(D) None

18. For X = [1,1,1,1,0,1,1,1]

Y = [1,0,0,1,1,0,1,0]

Find the cosine similarity?

(A) 0.4743

(B) 0.5757

(C) 0.3733

(D) 0.9983

19. The sum of the absolute difference between the co-ordinates of the points are called -

(A) Euclidean distance

(B) Manhattan distance

(C) Hamming distance

(D) None

20. Evidence Approximation in machine learning literature is also know as -

(A) Empirical Bayes

(B) Type 2 maximum likelihood

(C) Generalized maximum likelihood

(D) All

21. In which category the linear regression belong to ?

(A) Unsupervised

(B) Reinforcement

(C) Supervised

(D) Both (A) & (B)

22. Using Gaussian Distribution find the value of σ if $\frac{1}{\sqrt{4\Pi}}\int_0^{\infty} \exp\left(\frac{-x^2}{8}\right) = 1$.

(A) 1

(B) 2

(C) Π

(D) 2Π

23. If X is uniformly distributed over (0,10) , calculate the probability if $3<x<8$.
(Using Uniform Distribution)

(A) $\frac{1}{10}$

(B) $\frac{3}{10}$

(C) $\frac{4}{10}$

(D) $\frac{1}{2}$

24. The probability density function is of the form $p(x) = Ke^{-\alpha|x|}$, $x\epsilon\ (-\infty, \infty)$ the value of K is -

(A) 0.5

(B) 1

(C) 0.5α

(D) 1α

25. With the help of General continuous probability Distribution, Find the standard deviation of the continuous random variable if probability density function is given as –

$$f(t) = \begin{cases} 1 + t \text{ for } -1 < t \leq 0 \\ 1 - t \text{ for } 0 \leq t \leq 1 \end{cases}$$

(A) $\frac{1}{\sqrt{3}}$

(B) $\frac{1}{\sqrt{6}}$

(C) $\frac{1}{3}$

(D) $\frac{1}{6}$

26. The PDF of a random variable X is $f(x) = \begin{cases} x^2 \text{ for } -1 \leq \times \leq 1 \text{ and} \\ 0 \text{ for other values} \end{cases}$ then the percentage probability $P\left(-\frac{1}{3} \leq \times \leq \frac{1}{3}\right)$ is -

(A) 0.247

(B) 2.47

(C) 24.7

(D) 247

27. For Binary variables , the formula for βeta distributions is -

(A) $\beta\left(\frac{\mu}{a,b}\right) = \frac{\sqrt{(a+b)}}{\sqrt{(a)}\sqrt{(b)}}\ \mu^{a-1}(1-\mu)^{b-1}$

(B) $\beta\left(\frac{\mu}{a,b}\right) = \frac{\sqrt{a}\ \sqrt{b}}{\sqrt{(a+b)}}\ \mu^{a-1}\ (1-\mu)^{\ b-1}$

(C) $\beta\left(\frac{\mu}{a,b}\right) = \frac{\sqrt{(a+b)}\ .\sqrt{a}}{\sqrt{b}}\ \mu^{a-1}\ (1-4)^{b-1}$

(D) All

28. A certain company sells tractors which fail at a rate of 1 out of 1000. If 500 tractors are purchased from this company what is the probability of 2 of them failing with in first year.

(A) 1011

(B) 1.011

(C) 0.1011

(D) 10.11

29. A certain airport receives on an average of 4 aircrafts per hour. What is the probability that no aircraft lands in a particular 2 hr period.

(A) e^{-7}

(B) e^{-8}

(C) e^{-9}

(D) e^{-10}

30. A box contains 20 defective items and 80 non-defective items. If two items are selected at random without replacement. What will be the probability that both items are defective ?

(A) $\frac{1}{5}$

(B) $\frac{1}{25}$

(C) $\frac{20}{99}$

(D) $\frac{19}{495}$

31. A fair win is tossed four time. What is the probability that two heads and two fails will result ?

(A) $\frac{3}{8}$

(B) $\frac{1}{2}$

(C) $\frac{5}{8}$

(D) $\frac{3}{4}$

32. 10 dices are thrown. What is the probability of getting exactly 2 sixes.

(A) 0.2907

(B) 0.2807

(C) 0.8207

(D) 0.9207

33. A fair coin is tossed 10 times. What is the probability that ONLY the first two tosses will yield heads ?

(A) $\left(\frac{1}{2}\right)^2$

(B) $10C_2\left(\frac{1}{2}\right)^2$

(C) $\left(\frac{1}{2}\right)^{10}$

(D) $10C_2\left(\frac{1}{2}\right)^{10}$

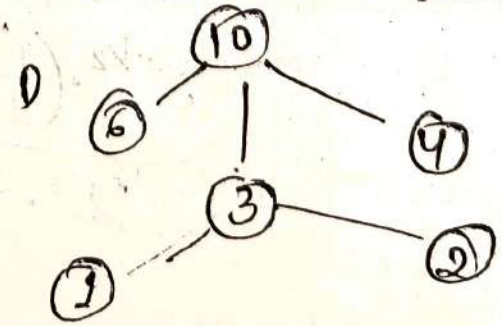
34. There are 10 markers on a table of which 6 are defective and 4 are not defective. If 3 are randomly taken from above it, what is the probability that exactly 1 of markers is defective ?

(A) 0.4

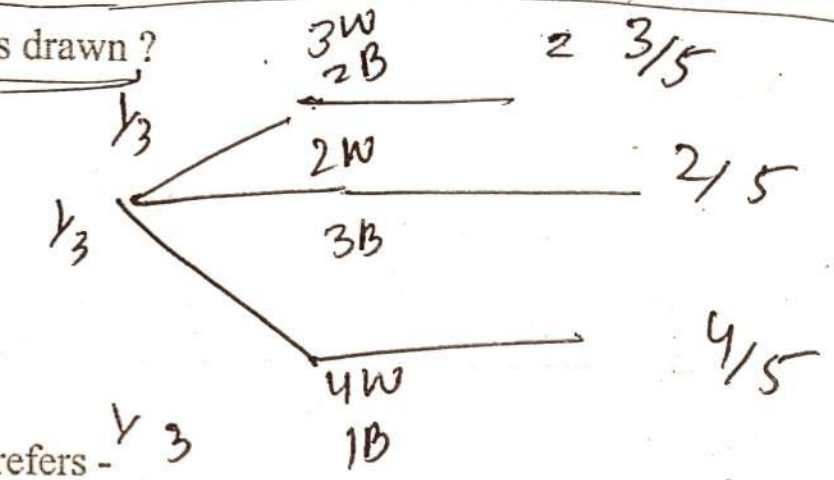
(B) 0.2

(C) 0.3

(D) None of these



35. There are three utensils containing (3 white and 2 black balls) , (2 white and 3 black balls) , and (1 Black and 4 white balls) . There is an equal probability of each utensils being choosen. If 1 ball is choosen at random, what is the probability that a white ball is drawn ?



(A) $\frac{3}{5}$

(B) $\frac{2}{5}$

(C) $\frac{1}{5}$

(D) None

36. The cause of dimensionality refers -

(A) All the problems that arise in higher dimension that did not exist in lower dimension

(B) All the problems that arise in lower dimension that did not exist in higher dimension

(C) All the problems that arise in both dimension but not exist later in lower dimension

(D) All of these

37. The curse of dimensionality relates to which problem of using operational data for BI reporting ?

(A) Inconsistent data

(B) Dirty data

(C) Non-integrated data

(D) Too much data

38. If we have D input variable, then a general polynomial with coefficients up to order 3 (for the curse of dimensionality) is $y(x_1 w) =$ _____

(A) $\omega_0 + \sum_{i=1}^{D} \omega_i x_i$

(B) $\omega_0 + \sum_{i=1}^{D} \sum_{j=1}^{D} \omega_{ij} \times i \times j$

(C) $\omega_0 + \sum_{i=1}^{D} \sum_{j=1}^{D} \omega_{ij} \times j \times i + \sum_{i=1}^{D} \sum_{j=1}^{D} \sum_{k=1}^{D} \omega_{ijk} \times i \times j \times k$

(D) $\omega_0 + \sum_{i=1}^{D} \omega_i \times i + \sum_{i=1}^{D} \sum_{j=1}^{D} \omega_{ij} \times i \times j + \sum_{i=1}^{D} \sum_{j=1}^{D} \sum_{K=1}^{D} \omega_{ijk} \times i \times j \times k$

39. The fundamental unit of network is -

(A) Brain

(B) Nucleus

(C) Neuron

(D) Axon

40. Where does the chemical reactions take place in neuron ?

(A) Dendrites

(B) Axon

(C) Synapses

(D) Nucleus

41. Function of Dendrites is -

(A) Transmitter

(B) Receptors

(C) Both (A) & (B)

(D) None

42. What is the purpose of AXON ?

(A) Receptors

(B) Transmitter & Process

(C) Transmission

(D) None

43. What was the name of the first model which can perform weighted sum of inputs?

(A) Marvin Minsky neuron model

(B) McCulloch –Pitts neuron model

(C) Hop field model of neuron

(D) None

44. Who proposed the first perceptron model in 1958 ?

(A) McCulloch -Pitts

(B) Marvin Minsky

(C) Hopfield

(D) Rosenblatt

45. A 4 input neuron has weight 1,2,3, and 4. The transfer function is linear with the constant of proportionality being equal to 2. The inputs are 4,10,5 and 20 respectively. The output will be -

(A) 238

(B) 76

(C) 119

(D) 123

46. What is back propagation?

(A) It is another name given to the curvy function in the perceptron

(B) It is the transmission of error back through the network to adjust the inputs

(C) It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn

(D) None

47. Consider a single perceptron with sign activation function. The perceptron is represented by weight vector $[0.4, -0.3, 0.1]^T$ and a bias $\theta=0$. If the input vector to the perceptron is X = [0.2 0.6 0.5] then the output of the perceptron is -

(A) 1

(B) 0

(C) -0.05

(D) -1

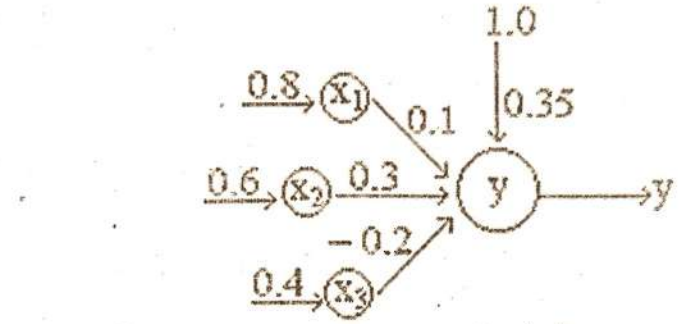
48. If input layer has 3 neurons as shown in figure with bias (b) then find the value of Y. choose the following option – [without using any sigmoid activation function]

(A) 0.625

(B) 0.259

(C) 0.56

(D) 0.53



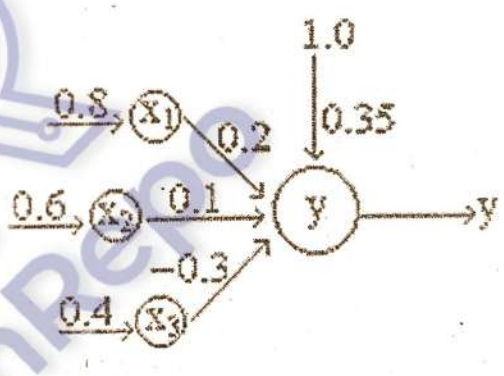
49. For the given figure, find the value of y using Binary Sigmoidal Activation Function.

(A) 0.625

(B) 0.259

(C) 0.56

(D) 0.53



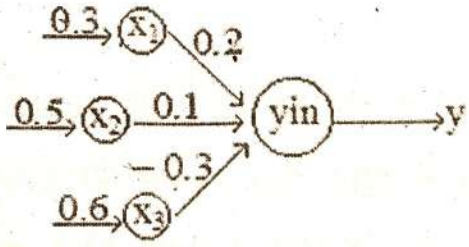
50. For the network shown, calculate the net input to the output neuron

(A) 0.29

(B) -0.07

(C) 0.07

(D) -0.29



*****