INDIAN INSTITUTE OF INFORMATION TECHNOLOGY UNA [HP]

An Institute of National Importance under MoE

Saloh, Una (HP) – 177 209

Website: www.iiitu.ac.in

AY 2023-24

School of Computing

CURRICULUM: IIITUGCSE22

Cycle Test – I

16, Aug.'23

(02:00 PM – 03:00 PM)

Degree	B. Tech.	**Branch**	CSE/IT/ECE
Semester	VII		
Subject Code & Name	CSSE12: Deep Learning		
Time: 60 Minutes	**Answer All Questions**		**Maximum: 20 Marks**

S. No.	Question	Marks
1.a	What is overfitting?	(1)
1.b	Consider a neural network composed of McCulloch-Pitts units with binary thresholding activation functions. The network has three input units (X_1, X_2, X_3) and one output unit (Y). The thresholds for the units are: • $\theta(X_1) = 2$ • $\theta(X_2) = 3$ • $\theta(X_3) = 1$ • $\theta(Y) = 2$ The weights are: • $w(X_1 \rightarrow Y) = 1$ • $w(X_2 \rightarrow Y) = -2$ • $w(X_3 \rightarrow Y) = 3$ Determine the output (Y) of the neural network for an input pattern where $X_1 = 0$, $X_2 = 1$, and $X_3 = 1$. Show the step-by-step calculations and explain whether the output neuron fires or not based on the thresholding logic.	(2)

1.c	A dataset located within a two-dimensional feature space, containing two distinct categories are denoted as Class A and Class B. The provided information regarding the dataset is as follows: • Class A contains the points: (2, 5), (3, 4), (4, 3), (5, 2). • Class B contains the points: (1, 1), (2, 2), (3, 3), (4, 4) Compute the equation of a straight line that best separates the two classes, if possible. If the classes are not linearly separable, explain why.	(2)
2.a	What necessitates the utilization of the backpropagation technique?	(1)
2.b	Explain how the activation of a biological neuron is related to the idea of an activation function in artificial neural networks?	(2)
2.c	A linear regression model is being developed to predict housing prices based on various features, with the equation: $\textbf{Price} = \beta_0 + \beta_1 \times \textbf{Feature}_1 + \beta_2 \times \textbf{Feature}_2 + \beta_3 \times \textbf{Feature}_3 + \varepsilon$ L1 and L2 regularization techniques are being applied to the model, with the following regularization terms added to the loss function: • L1 Regularization Term: $\lambda \sum_{i=1}^{3} \lvert\beta_i\rvert$ • L2 Regularization Term: $\lambda \sum_{i=1}^{3} \beta_i^2$ Where λ represents the strength of regularization. Using $\lambda = 0.1$, compute the L1 regularization term and the L2 regularization term for the given model.	(2)
3.a	What are the different challenges associated with training deep neural networks?	(1)
3.b	Prove convergence theorem for the perceptron learning algorithm.	(2)
3.c	Discuss the concept of convergence in the context of the Linear Perceptron algorithm. How to determine if the algorithm has converged to a solution? Are there cases where the algorithm might not converge?	(2)
4.a	What is the role of the activation function in the Linear Perceptron algorithm? How does it help in making a decision about the class of a given data point?	(1)
4.b	A dataset comprises three data points: (1, 5), (2, 9), and (3, 14). The objective is to establish a linear regression model represented as $y = mx + b$ for this dataset through the utilization of gradient descent. The initial assumptions are: • $m_0 = 0.5$ • $b_0 = 2$ • $\alpha = 0.01$. Execute a single iteration of the gradient descent process to modify the values of m and b, employing the gradient derived from the mean squared error loss function.	(2)
4.c	Given a dataset with 80% of samples from Class A and 20% from Class B, what challenges might arise when applying ERM? How can these challenges affect the model's performance?	(2)