U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

June 2025 Semester End Main Examinations

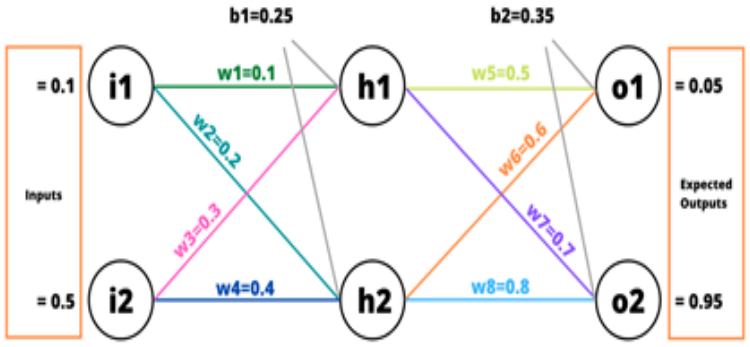
Programme: B.E. **Semester: V**

Branch: Artificial Intelligence & Machine Learning **Duration: 3 hrs.**

Course Code: 24AM5PCDEL **Max Marks: 100**

Course: DEEP LEARNING

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

		UNIT - I	CO	PO	Marks
1	a)	Apply perceptron learning algorithm to train a single layer perceptron for AND Gate with initial weights w1 = 1.2, w2 = 0.6, Threshold = 1 and Learning Rate n = 0.5. Refine the model's weights iteratively until convergence.	CO1	PO2	**10**
	b)	Explain the Various Benefits of Neural Networks. Provide an example.	CO1	PO1	**10**
		OR			
2	a)	Consider the Multilayer Perceptron with sigmoid function as activation function and learning rate = 0.6. Apply backpropagation algorithm and update the weight W5 and W6 in backward pass. 	CO1	PO2	**10**
	b)	Imagine an online learning model for real-time predictions in a stock trading application. To optimize the performance of the model, learning rate is to be adjusted during training. Given the nature of online learning, how optimal annealing strategy for the learning rate can be applied to ensure that the model converges effectively while adapting to new data?	CO1	PO2	**10**
		UNIT - II			
3	a)	A neural network is trained using Nesterov Accelerated Gradient Descent (NAG). The learning rate is 0.01, and the momentum coefficient β is =0.9. The velocity at time t−1 is Vt-1=0.5, and the	CO1	PO2	**10**

			gradient at time t−1 is ∇L(wt−1)=−0.3. Calculate the update for the weights at time step t using the NAG update rule.			
		b)	How does the vanishing gradient problem affect the training of deep neural networks, also describe the strategies that can be employed to mitigate its impact?	*CO1*	*PO2*	**10**
			OR			
	4	a)	Illustrate Nesterov accelerated gradient descent with relevant steps and equations.	*CO1*	*PO2*	**10**
		b)	Explain the Momentum based Gradient Descent in detail.	*CO1*	*PO2*	**10**
			UNIT - III			
	5	a)	How does Norm Penalties (NP) contribute in improving model performance and reducing overfitting? Explain L1 and L2 NPs.	*CO2*	*PO2*	**07**
		b)	A neural network has 200 parameters, and the L2 norm penalty is applied with a regularization factor λ = 0.1. The sum of squared weights before training is 15, and after training for 50 epochs, it reduces to 10. Calculate the change in the L2 norm penalty after 50 epochs.	*CO2*	*PO1*	**07**
		c)	Justify the advantages of early stopping in regularizing a model. Explain its working with an algorithm.	*CO2*	*PO2*	**06**
			OR			
	6	a)	How Tangent Prop technique can be used to improve the accuracy of classification?	*CO2*	*PO2*	**08**
		b)	Describe the role of Manifold Tangent Classifier in improving the accuracy of classification?	*CO2*	*PO1*	**06**
		c)	Differentiate between Adaboost and RMSProp algorithm.	*CO2*	*PO2*	**06**
			UNIT - IV			
	7	a)	Examine Padding from basic Structure of Convolutional Network.	*CO2*	*PO3*	**08**
		b)	How can strides be used to reduce the spatial footprint of the image? Explain.	*CO2*	*PO3*	**06**
		c)	Explain Pooling Operations of Convolutional Network.	*CO2*	*PO1*	**06**
			OR			
	8	a)	Consider a CNN architecture with the following specifications: Input image (I): $\begin{bmatrix} 4 & 3 & 2 & 1 & 0 & 6 \\ 7 & 8 & 9 & 2 & 3 & 5 \\ 2 & 3 & 1 & 6 & 5 & 4 \\ 1 & 4 & 7 & 8 & 6 & 2 \\ 5 & 3 & 2 & 1 & 4 & 7 \\ 6 & 8 & 9 & 2 & 3 & 1 \end{bmatrix}$ Filter (F): $\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$	*CO2*	*PO1*	**08**

			i. Perform convolution between I and F using stride 1 and padding 1 to generate the feature map. ii. Apply the ReLU activation function to the feature map. iii. Perform max-pooling on the updated feature map using a pooling size of 2×2 and stride 2. Represent the final matrix after max-pooling.			
		b)	Analyze the Backpropagation as Matrix Multiplications	*CO2*	*PO2*	**06**
		c)	Explain the working of Data Augmentation in CNN.	*CO2*	*PO1*	**06**
			UNIT - V			
	9	a)	Simplify the basic architecture of a recurrent Neural Networks.	*CO3*	*PO1*	**10**
		b)	Analyze the Challenges of Training Recurrent Neural Networks.	*CO3*	*PO2*	**10**
			OR			
	10	a)	Analyze the LSTM problems associated with Neural Networks.	*CO3*	*PO2*	**10**
		b)	Simplify the Gated Recurrent Units (GRUs) in Neural Networks.	*CO3*	*PO2*	**10**

****