

Reg No.: _____



Jyothi Engineering College(Autonomous)
M.Tech Degree S2 (R) Examination, May 2026 (2025 Scheme)

25PADT201 - DEEP LEARNING



Total Mark: 60

PART A

Answer All Questions

1. Compare exploding gradients and vanishing gradient in deep neural networks. What are its causes and common techniques used to mitigate it? CO1 (5)
2. Consider a 4*4 matrix: CO2 (5)
$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 6 & 1 & 2 \\ 7 & 8 & 3 & 0 \\ 1 & 2 & 4 & 5 \end{bmatrix}$$

Apply 2×2 max pooling with stride = 2 and compute the output feature map.
3. Explain the architecture of a Recurrent Neural Network (RNN). Describe how the hidden state helps in maintaining memory of previous inputs. CO3 (5)
4. Analyze the difference between TF-IDF and Bag of Words. How does TF-IDF improve feature importance? CO4 (5)
5. Explain the concept of Bahdanau Attention. How does it improve sequence-to-sequence models compared to basic encoder-decoder architecture? CO5 (5)

PART B

Answer Any Five Question(s)

6. Consider a loss function $J(w) = (y - wx)^2 + \lambda w^2$ where $x=2$, $y=6$, $w=1$ and $\lambda=0.5$. Compute the value of the regularized loss function. Also, explain how L2 regularization affects the model parameters. CO1 (7)
7. Explain the architecture of AlexNet. Illustrate how convolution, ReLU, and pooling layers are arranged. Apply your understanding to explain how it improves performance over traditional CNNs. CO2 (7)
8. Analyze the working of Backpropagation Through Time (BPTT). How are gradients propagated across multiple time steps? Discuss the challenges involved in training deep RNNs using BPTT. CO3 (7)
9. Given the following documents: CO4 (7)
D1: "cat sat on mat"
D2: "cat sat"
D3: "dog sat on mat". Compute TF-IDF values for the word "cat" in all documents and Interpret the results.

10. Explain self-attention and multi-head attention mechanisms in Transformers. How do they improve representation learning? CO5 (7)
11. Explain the concept of large-scale pretraining with Transformers. Describe how pretraining and fine-tuning improve performance across tasks. CO5 (7)
12. Analyze the working of Word2Vec (CBOW and Skip-gram). Compare their approaches in learning word representations. Evaluate their advantages and limitations. CO4 (7)
