9971

Sr. No. Of Question Paper : Roll No.: .....

Name of Course : B.Sc. (H) Electronics

Semester : VII

Name of the Paper : Neural Networks and Deep Learning

Unique Paper Code : 2513010018

Total Time : 3 Hours

Maximum Marks : 90

Instructions for candidates

1. **Attempt 5 questions in all.**
2. ***Question No. 1 is compulsory.***
3. ***All questions carry equal marks***

Q1

(a) How is Deep Learning different from conventional Machine Learning?

(b) What is a Hidden Layer in an Artificial Neural Network (ANN)? What is its main function?

(c) Explain the basic architecture of a RNN. How is it different from the Back Propagation Neural Network?

(d) Given the dataset :
Temperature = [22, 25, 27, 30, 33]
Humidity = [60, 55, 50, 45, 40]
Label (Rain) = [1, 1, 0, 0, 0]
Write the basic steps required to implement a Neural Network model to classify whether it may rain or not.

(e) What are hyperparameters in deep learning models? Identify the types of hyperparameters used during training and elaborate on any one of them.

(f) What are the key characteristics of ResNet? Where is it used?

(6 x 3=18)

Q2

(a) Explain the convolutional operation on the following grey scale image matrix

$$I = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 1 \\ 7 & 8 & 9 & 2 \\ 1 & 3 & 5 & 4 \end{bmatrix}$$

using a simple 2x2 filter

$$K = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and a stride s=2.

What will be the final feature map? (9)

(b) Without activation functions, neural networks act as linear models. Justify this statement. Provide the mathematical formulation of any three activation functions that help overcome this limitation (9)

Q3

(a) Backpropagation is widely used for training deep neural networks. Describe how it works and present the mathematical update rules used in Gradient Descent (6)

(b) Discuss how Radial Basis Neural Networks (RBNNs) use RBF kernels to map data into a higher-dimensional feature space. Explain the architecture and show how this results in nonlinear hyperplanes for curve fitting (6)

(c) RNNs extend BPNNs by introducing memory through recurrent connections. Explain the RNN architecture, contrast it with BPNNs, and describe how the BPTT algorithm enables learning from sequential data (6)

Q4

(a) Explain the PCA technique in detail. Discuss all mathematical operations involved—such as standardization, covariance matrix computation, eigenvalues, and eigenvectors (6)

(b) Write the mathematical connection between PCA eigenvectors and the left/right singular vectors obtained using SVD. Use a diagram or block representation (6)

(c) True target y=3.0. Three model predictions (from repeated training/runs) are [2.5, 3.5, 3.0]. Compute the squared bias, variance of predictions, and mean squared error (MSE). Verify decomposition MSE = bias² + variance (6)

Q5

(a) Describe the purpose of the pooling layer in a Convolutional Neural Network (CNN). Compare max pooling and average pooling in terms of feature extraction (6)

(b) Explain the concept of optimization in neural networks. Give a detailed comparison of two optimization algorithms and their role in CNN performance (6)

(c) A convolutional layer has a single weight w=0.5. The gradient of the loss with respect to this weight is $g=\frac{\partial L}{\partial w}$ =-0.8. The learning rate is η =0.1

Compute the updated weight using basic Stochastic Gradient Descent (SGD). Also, compute the updated weight using SGD with momentum (β=0.9) assuming initial velocity v_0=0. (6)

Q6

(a) What is meant by "perception" in self-driving cars? Discuss the types of sensors used for environmental perception in self-driving cars. (6)

(b) What are the key differences between retrieval-based and generative chatbots? Explain their pros, cons, and practical applications. (6)

(c) What role does prompt engineering play in using LLMs effectively? A Large Language Model (LLM) has a maximum context length of 2048 tokens. You provide a prompt containing 350 tokens and the model generates 500 tokens as output. How many tokens are remaining in the context window before hitting the limit? (6)

Q7

(a) What are the key differences between object detection and image classification in terms of input, output, and processing? (6)

(b) How do transformers and attention mechanisms address limitations of older NLP models? Explain their contributions to the field (6)

(c) A transformer has 2 attention heads. Each head projects input of dimension 4 to dimension 2. If the input token embedding is X=[1,2,3,4], what is the total dimension after concatenating the heads? (6)