Roll No											

PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING
END TERM EXAMINATION - MAY 2024

Semester : Semester VIII - B.Tech CSE - 2020
Course Code : CSE3010
Course Name : Sem VIII - CSE3010 - Deep Learning Techniques
Program : B.Tech. Computer Science and Engineering

Date : May 29, 2024
Time : 6:55 AM - 6:55 AM
Max Marks : 100
Weightage : 50%

Instructions:
(i) Read all questions carefully and answer accordingly.
(ii) Question paper consists of 3 parts.
(iii) Scientific and non-programmable calculator are permitted.
(iv) Do not write any information on the question paper other than Roll Number.

1.a) Based on the given details in the table, fill the remaining columns. (CO1)

Problem type	Output Type	Final activation function	Loss function
Regression	Numerical value		
Classification	Binary outcome		
Classification	Single label, multiple classes		
Classification	Multiple labels, multiple classes		

1.b) Compare and contrast shallow neural networks and deep neural networks in terms of their architecture, performance, and training. (CO1)

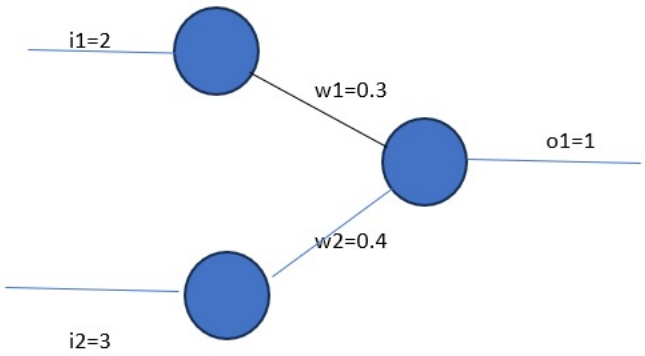
1.c) List and explain the various activation functions used in modeling of artificial neuron. Also explain their suitability with respect to applications. (CO1)

OR

2.a) Discuss any two ways to detect, whether your neural network is suffered with vanishing gradient Problem. (CO1)

2.b) Explain Backpropogation with its algorithm. (CO1)

Given a neural network, calculate updated weights for the layer using backpropogation. consider bias=0.5 and learning rate=0.01



2.c) (CO1)

3.a) Write the importance of Hyperparameter Tuning. (CO2)

3.b) Explain the concept of batch normalization and outline its benefits in the training phase of deep neural networks. (CO2)

3.c) Identify the differences between Parameter and Hyperparameter, explain in detail with an examples. (CO2)

OR

4.a) Explain why dropout can be considered a form of ensemble learning. (CO2)

4.b)

1. Consider a linear regression problem with the regularization term added to the cost function: Given m=3 data points and the following parameters X=[1,2,3], y=[2,4,6] ,λ=1. Calculate L(θ), L2 weight decay for θ1=0 and θ2=2 (i.e. k=0 to 1)
   y=θ1+θ2x

(CO2)

4.c) Compare and contrast dropout regularization with other regularization techniques such as L1 and L2 regularization. Highlight the unique advantages and limitations of dropout compared to these techniques. (CO2)

Given a simple RNN with single hidden layer and following parameters

- Input vector $x = [1, 2]$
- Weight matrix for input to hidden $W_{xh} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}$
- Weight matrix for hidden to hidden $W_{hh} = \begin{bmatrix} 0.5 & 0.6 \\ 0.7 & 0.8 \end{bmatrix}$
- Initial hidden state $h_0 = [0, 0]$
- Bias vector for hidden layer $b_h = [0.1, 0.2]$

5.a) Compute the hidden state at time step 1. (CO3)

5.b) Discuss the important criteria to consider when designing and evaluating a sequence model. (CO3)

Consider LSTM unit with the following parameters. Compute the updated cell state Ct and hidden state ht

- Input vector $x_t = [1, 2]$
- Previous hidden state $h_{t-1} = [0.5, 0.1]$
- Previous cell state $C_{t-1} = [0.4, 0.2]$

Weights and biases are:

- Input gate weights $W_i = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}, U_i = \begin{bmatrix} 0.5 & 0.6 \\ 0.7 & 0.8 \end{bmatrix}, b_i = [0.1, 0.2]$
- Forget gate weights $W_f = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}, U_f = \begin{bmatrix} 0.5 & 0.6 \\ 0.7 & 0.8 \end{bmatrix}, b_f = [0.1, 0.2]$
- Output gate weights $W_o = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}, U_o = \begin{bmatrix} 0.5 & 0.6 \\ 0.7 & 0.8 \end{bmatrix}, b_o = [0.1, 0.2]$
- Cell state weights $W_c = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}, U_c = \begin{bmatrix} 0.5 & 0.6 \\ 0.7 & 0.8 \end{bmatrix}, b_c = [0.1, 0.2]$

5.c) (CO3)

OR

6.a) What are the building blocks of convolutional neural network (CNN), how they are used to extract features from input data. (CO3)

6.b) Discuss the limitations of Recurrent Neural Networks (RNNs) and explain how Long Short-Term Memory (LSTM) networks can address these issues. (CO3)

6.c) Discuss the working principle of recurrent neural network, and explain how they are used to process sequential data. (CO3)

7.a) With the architecture explain training of Restricted Boltzmann Machines. (CO4)

7.b) Explain Restricted Boltzman Machine with the training and reconstruction phase. (CO4)

7.c) List the different types of Autoencoders. Explain Denoising autoencoder with their advantages and disadvantages. (CO4)

OR

8.a) Draw the architecture of Boltzmann Machine having two visible nodes and three hidden nodes. (CO4)

8.b) Explain Denoising autoencoder. (CO4)

8.c) Construct the Convolutional autoencoders. Give the advantages and Disadvantages. (CO4)

9.a) Define RNN as sequential data, and what is the primary reasons ANNs are not ideal for handling sequential data? (CO3)

9.b) Explain convolution operation. Given 5x5 input image, a 3x3 convolutional filter, a stride of 2, and padding of 1, calculate the dimensions of the resulting feature map. (CO3)

9.c)	What are convolutional layers and pooling layers in a CNN, and what roles do they play in feature extraction?	(CO3)

OR

10.a)	Write the differences between Boltzmann Machine and Restricted Boltzmann Machine.	(CO4)
10.b)	Justify the advantages of auto encoder over principal component analysis for dimensionality reduction.	(CO4)
10.c)	Explian in detail about Restricted Botzmann Machine.	(CO4)