

RAMAKRISHNA MISSION VIDYAMANDIRA

(Residential Autonomous College affiliated to University of Calcutta)

M.A./M.Sc. THIRD SEMESTER EXAMINATION, JANUARY 2023

SECOND YEAR [BATCH 2021-23]

COMPUTER SCIENCE

Paper : CC11

Date : 07/01/2023

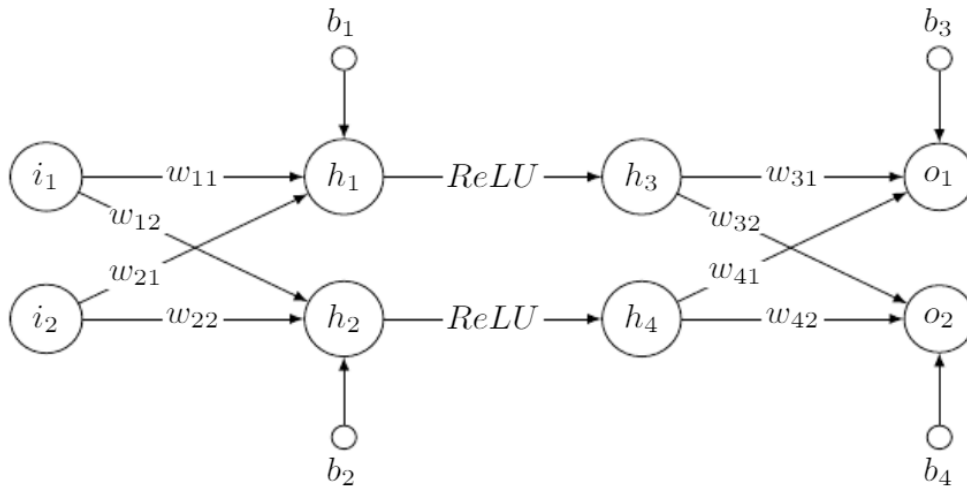
Time : 11.00 am – 1.00 pm

Full Marks : 50

Answer **any five** questions :

[5×10]

1. Given the following Neural Network with fully connection layer and ReLU activations, including two input units (i_1, i_2), four hidden units (h_1, h_2) and (h_3, h_4). The output units are indicated as (o_1, o_2) and their targets are indicated as (t_1, t_2). The weights and bias of fully connected layer are called w and b with specific sub-descriptors.



The values of variables are given in the following table:

Variable	i_1	i_2	w_{11}	w_{12}	w_{21}	w_{22}	w_{31}	w_{32}	w_{41}	w_{42}	b_1	b_2	b_3	b_4	t_1	t_2
Value	2.0	-1.0	1.0	-0.5	0.5	-1.0	0.5	-1.0	-0.5	1.0	0.5	-0.5	-1.0	0.5	1.0	0.5

- a) Compute the output (o_1, o_2) with the input (i_1, i_2) and network parameters as specified above. (Write down all calculations, including intermediate layer results.)
- b) Compute the mean squared error of the output (o_1, o_2) calculated above and the target (t_1, t_2).
- c) Update the weight w_{21} using gradient descent with learning rate 0.1 as well as the loss computed previously. (Please write down all your computations.) [3+3+4]
2. a) You have two datasets, D1 and D2, each of which include one million tagged training samples for classification. On dataset D2, your friend creates a model from scratch. After deciding to train on D1, you opt to train on D2 using transfer learning. State one problem your friend is likely to find with his approach. How does your approach address this problem? Explain it.
- b) What are the different layers of Autoencoders? Explain briefly.
- c) Given a Convolution Layer with 8 filters, a filter size of 6, a stride of 2, and a padding of 1. For an input feature map of $32 \times 32 \times 32$, what is the output dimensionality after applying the Convolution Layer to the input? [3+3+4]

3.
 - a) You need to find a classification strategy. 20 samples are used to train your network initially. Training converges, but there is a significant training loss. Afterward, you choose to train this network using 10,000 samples. Is your method for resolving the issue sound? If yes, explain the most likely results of training with 10,000 examples. If not, give a solution to this problem.
 - b) What is the importance of LSTM. Explain the architecture of LSTM.
 - c) After training a neural network, you observe a large gap between the training accuracy (100%) and the test accuracy (42%). Which method is commonly used to reduce this gap? [4+3+3]

4.
 - a) You train a neural network to automate the whole hiring process since you work for a tech startup that receives thousands of job applications every day. Your model categorises applicant resumes automatically, rejecting or contacting each applicant with a job offer as appropriate. Which of the following metrics has a higher priority for your model? Justify.

When implementing a neural network layer from scratch, we usually implement a ‘forward‘ and a ‘backward‘ function for each layer. Explain what these functions do, potential variables that they need to save, which arguments they take, and what they return.
 - b) What is the drawback of GAN? The process of scanning mis-classified examples to identify weaknesses of a model is known as which analysis? Justify
 - c) A large dataset is being used to train a neural network for picture categorization. Your mathematician buddy makes the following suggestion: "Your neural network would converge much quicker than using gradient descent if you would utilise Newton-Method for optimization!" Examine the veracity of this claim and the potential drawbacks of taking his advice. [4+3+(1.5+1.5)]

5.
 - a) Why is it necessary to introduce non-linearities in a neural network?
 - b) Describe two ways of dealing with the vanishing Gradient Problem in a neural network?
 - c) What is the difference between CNN and DNN? [3+5+2]

6.
 - a) How does splitting a dataset into train, dev and test sets help identify overfitting? Which regularization method leads to weight sparsity? Explain why.
 - b) In order to identify driver tiredness in vehicles, you are developing a deep learning system. Your model must be able to recognise tiredness in order to avoid any mishaps.
 - c) Which of the following assessment metrics is the best fit? Accuracy, Recall, Precision, and Value Loss. Justify your decision. [(2+2)+(2+1)+3]

7.
 - a) What are the main difficulties when training RNNs ? How can you handle them.
 - b) What is the difference between stateful RNN and stateless RNN? What are their pros and cons? [5+(2.5+2.5)]

_____ × _____