Printed Page:- Subject Code:- ACSML0602

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM: VI - THEORY EXAMINATION (20.. - 20....)

Subject: Deep Learning

Time: 3 Hours Max. Marks: 100

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*

1. This Question paper comprises of three Sections -A, B, & C. It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.

2. Maximum marks for each question are indicated on right -hand side of each question.

3. Illustrate your answers with neat sketches wherever necessary.

4. Assume suitable data if necessary.

5. Preferably, write the answers in sequential order.

6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.

SECTION-A 20

1. Attempt all parts:-

1-a. In Delta Rule for error minimization (CO1,K1) 1

(a) Weights are adjusted with respect to difference between desired output and actual output

(b) Weights are adjusted with respect to difference between input and output

(c) Weights are adjusted with respect to change in the output

(d) no-one

1-b. The complexity of ANN is dependent upon _____ (CO1,K2) 1

(a) Number of Neurons

(b) Number of Nodes

(c) Number of Anodes

(d) Number of Layers

1-c. Which is the following is an application of neural network? (CO2,K2) 1

(a) Sales forecasting

(b) Data validation

(c) Risk management

(d) All of the mentioned

1-d. Choose Common uses of RNNs. (CO2,K1) 1

(a) Detect fraudulent credit-card transaction

(b) Provide a caption for images

.

(c) Businesses Help securities traders to generate analytic reports

(d) All of the above

1-e. Full form of PAC is _____ (CO3,K1) 1

(a) Probably Approx Cost

(b) Probably Approximate Correct

(c) Probability Approx Communication

(d) Probably Approximate Computation

1-f. Among the following identify the one in which dimensionality reduction reduces. (CO3,K2) 1

(a) Performance

(b) Entropy

(c) Stochastic

(d) Collinearity

1-g. The following is FALSE about Weights and Bias. (CO4, K1) 1

(a) Biases are usually initialized to o (or close to o)

(b) Weights are usually initialized using Xavier technique

(c) Both weight and bias are hyperparameters

(d) None of the above

1-h. The following is NOT a generalization technique in neural networks.(CO4,K1) 1

(a) Xavier Weight Initialization

(b) Dropout

(c) Li and L2 Regularization

(d) Data Augmentation

1-i. Which of the following is correct about Denoising autoencoder? (CO5,K1) 1

(a) control the number of nodes in the hidden layers to increase its generalization capa

(b) add random noise to input to increase its generalization capabilities

(c) remove noise to input to increase its generalization capabilities

(d) All of the above

1-j. Which of the following techniques perform similar operations as dropout in a neural network? (CO5,K1) 1

(a) Bagging

(b) Stacking

(c) Non Stacking

(d) None of these

2. Attempt all parts:-

2.a. Define the **Curse of Dimensionality** and briefly explain how it affects model performance. (CO1, K2) 2

2.b. Define **computer vision** and mention one practical application of CNNs in this 2

.

field. (CO2, K2)

2.c. What is the purpose of **strided convolutions**? How do they affect the feature map size? (CO3, K2) 2

2.d. Briefly describe **Backpropagation Through Time (BPTT)** and its role in training RNNs.(CO4, K2) 2

2.e. Briefly explain the concept of **semi-supervised learning** and how stacked autoencoders can be used in this context. (CO5, K2) 2

SECTION-B 30

3. Answer any five of the following:-

3-a. Explain the **bias-variance trade-off** in the context of model performance. How does it relate to overfitting and underfitting? (CO1, K4) 6

3-b. Describe the **structure and components of an artificial neuron**. Explain how activation functions help in neural network learning.(CO1, K3) 6

3-c. Explain the working of a **convolutional neural network (CNN)** with a focus on its three core layers: convolutional, pooling, and fully connected layers. (CO2, K4) 6

3-d. What is **transfer learning** in CNNs? How can it help when training models on small datasets? Explain with an example. (CO2, K3) 6

3.e. Explain the architecture and motivation of **1x1 convolutions** and their role in **Networks-in-Networks** and **Inception networks**. (CO3, K4) 6

3.f. Compare **Gated Recurrent Unit (GRU)** and **LSTM** networks in terms of structure and performance. When would you prefer one over the other? (CO4, K3) 6

3.g. Explain the role of **batch normalization** in training deep neural networks. How does it improve model performance? (CO5, K4) 6

SECTION-C 50

4. Answer any one of the following:-

4-a. Explain the **Delta Learning Rule** (Widrow-Hoff rule) and its significance in training neural networks. How does it differ from the Perceptron Learning Rule? Consider a single neuron with two inputs and a linear activation function f(net)=net. 10

Initial weights: w1=0.4, w2=-0.6*, bias b=0.1*

Learning rate $\eta=0.2$

Given training input: x=[0.5,1.0]

Target output: t=0.8

Calculate: Net input, Output, Error, Weight updates using the Delta Learning Rule, Updated weights and bias

(CO1, K3)

4-b. Derive the **Backpropagation algorithm** used in training a multilayer perceptron. Explain how gradient descent and the delta rule contribute to weight updates. (CO1, K6) 10

5. Answer any one of the following:-

.

5-a. A convolutional layer takes an input image of size 32×32 with a single channel. The layer uses a 5×5 filter with a stride of 1 and **padding** applied such that the output size is the same as the input size. 10

a) Calculate the amount of padding needed on each side of the input.

b) What will be the size of the output feature map?

After this convolutional layer, a **max pooling** layer with a 2×2 window and stride 2 is applied.

c) Calculate the size of the output feature map after pooling.

(CO2, K3)

5-b. Compare **image classification** and **text classification**. Explain the role of **hyper-parameter tuning** in improving CNN performance.(CO2,K4) 10

6. Answer any <u>one</u> of the following:-

6-a. Explain YOLO algorithms with architecture. (CO3, K2) 10

6-b. Explain about Auto Encoder? Details about Encoder, Decoder and Bottleneck? (CO3,K2) 10

<u>OR</u>

Differentiate between 1x1 convolution and fully connected layer (CO3,K4) 10

7. Answer any <u>one</u> of the following:-

7-a. Discuss the challenges involved in training deep RNNs and describe how different RNN variant like LSTM address these issues. Include a discussion on sequence generation techniques. (CO4, K5) 10

7-b. Define the N-gram language model and describe how it predicts the next word in a text sequence. What are the common challenges faced by N-gram models, like data sparsity, and how can these challenges be overcome?(CO4,K2) 10

8. Answer any <u>one</u> of the following:-

8-a. Explain why using a bottleneck layer in the architecture of an autoencoder is crucial for learning a compressed representation of the input data(CO5,K2) 10

8-b. Name some of the Autoencoder Variations. Also, explain them (CO5,K2) 10

.