Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2080

Subject: **Deep Learning**

Course No: **MDS 655**

Level: **MDS /II Year /IV Semester**

Full Marks: 45

Pass Marks: 22.5

Time: 2 hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A   [3 ×5 = 15]

1. Differentiate between machine learning and deep learning.
2. Discuss mechanics of machine learning.
3. Discuss McCulloch Pits model of neuron.
4. Explain the concept of Underfitting and overfitting.
5. What is delta rule? Discuss briefly.

Group B   [5 × 6 = 30]

6. Derive weight update rule for backpropagation algorithm using binary cross entropy as loss function.

OR

Discuss gradient descent variations along with their pros and cons.

7. Why deep learning models are prone to overfitting? Discuss the roles of validation sets in building deep learning models.

8. Why regularization is used in deep learning? Discuss various regularization methods.

9. Consider an ANN having single neuron with two inputs and logistic activation function. Assume that learning rate is 0.1 and momentum hyperparameter is 0.9. Show weight updates for the training tuples {(0.7,0.5, 0.8), (1,0.4, 0.9)} using momentum. Assume that initial weights are (0.2,0.2).

OR

Consider a neural network having single neuron with two inputs and logistic activation function. Assume that learning rate is 0.2 and momentum hyperparameter is 0.9. Show weight updates for the training tuples {(0.9,0.3, 0.6), (0.3,1, 0.9)} using Adagrad. Assume that initial weights are (0.5,0.5).

10. Why the concept of learning rate annealing is important? Discuss the concept of second order methods used in training neural networks.

***

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
Second Assessment 2080

Subject: **Deep Learning**
Course No: **MDS 655**
Level: **MDS/II Year /IV Semester**

Full Marks: 45
Pass Marks: 22.5
Time: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt ALL questions.

Group A [5 × 3 = 15]

1. What is meant by model non-identifiablity? Explain briefly.
2. Why the concept of bias correction is important? Explain.
3. Discuss shortcomings of feature selection.
4. Why vanilla neural networks can't be used in image classification? Discuss briefly.
5. Discuss the working of PCA briefly.

Group A [5 × 6 = 30]

6. What is the concept behind learning rate adaptation? Explain various learning rate adaptation methods.

7. Discuss the working of convolution layer of CNN.

OR

Consider an 8-bit grayscale image having resolution 6x6 and filter of size 3x3. Show convolution operation on the image by assuming stride=1 and zero padding.

8. Why image preprocessing is important? List various image preprocessing operation. Also explain the concept of Batch normalization and its importance.

9. What is Auto encoder? Discuss the concept of sparse Auto encoder architecture in detail.

OR

What is context? When context is important? Explain with example. What are limitations of one-hot-encoding? How Auto encoders tackle with this problem? Explain.

10. Discuss the importance and working mechanism of Word2Vec framework.

***

Tribhuvan University
Institute of Science and Technology
2080
✿

Subject: **Deep Learning**
Course No: MDS 655
Level: MDS / II Year / IV Semester

Full Marks: 45
Pass Marks: 22.5
Time: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt **ALL** questions.

| Group A | [5 × 3 = 15] |

1. Discuss relationship between bias variance and MSE.

2. Discuss the concepts of Sigmoid, Tanh, and RELU Neurons.

3. Derive weight update rule for gradient descent by assuming logistic activation function.

4. Discuss the concept of momentum based optimization.

5. Discuss the concept of one hot encoding with example.

| Group B | [5 × 6 = 30] |

6. Derive backpropagation weight update rule using MSE as loss function

OR

What is delta rule? Describe stochastic and mini batch gradient descent.

7. Consider an ANN with single neuron and logistic activation function. Calculate weight updates for the training samples (0.7, 0.3, 0.8) and (0.2, 0.6, 0.4) using and RMSProp. Assume neuron with two inputs and one output and $\alpha = 0.2$ $\beta = 0.8$. Make your own assumptions as per requirements.

8. Discuss the working of convolution and pooling layers in CNN.

9. What is main objective of Autoencoders? Discuss working of SAE with suitable mathematical formulation and diagrams.

10. What is Neural Translation Network? Discuss its working in detail.

OR

What kinds of problems fall under sequence analysis? How Neural N-grams helps to tackle Seq2Seq problems.