Birla Institute of Technology and Science Pilani
Dept. of Computer Science & Information Systems
Deep Learning (CS F425)

Fri May 13, 2022
8:00–11:00 AM

Time: 3 Hour | *Comprehensive Exam (Open Book)* | Maximum Marks: 70 (35%)

INSTRUCTIONS:

1. There are **8** questions in this question paper, spread over **2** pages, make sure you have all the printed pages.
2. Attempt **ALL** questions in any order you like.
3. Use any printed material/books/handwritten notes you brought with you but, do not share/exchange with others. Use of calculator is **allowed** in this exam.
4. You are free to make reasonable assumption that you may need to logically answer the question (mention them clearly in the beginning of your answer).

1. Consider the data points having two attributes as given below. A perceptron is trained using Perceptron Training Rule with learning rate $\eta = 0.01$. [15]

x_1	x_2	y
1	9	green
10	9	green
4	7	green
4	5	red
5	3	red
8	9	green
4	2	red
2	5	red
7	1	red
2	10	green
8	5	green
1	2	red
8	2	red

During the training let the weights at some particular point of time be as below.

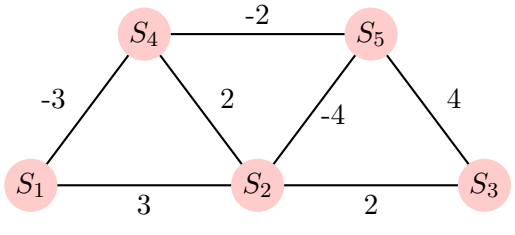
Let classification of y is obtained using $\sum w_i \times x_i \geq 0$ signifying to green or positive.

w0=-0.600, w1=-0.100, w2=0.280	err=2
w0=?, w1=?, w2=?	err=?

Determine the weights after immediate next epoch and the number of misclassified points.

2. Address the following queries in their context
   (a) State and prove universal approximation theorem. [3]
   (b) Describe regularized autoencoder and one of its use case. [3]
   (c) Write steps of the ADAM optimizer. Why does the bias correction is required? [3]
   (d) Why the L_2 loss does not enforces sharp images generation. [2]

3. Consider a Hopfield net given below where links are symmetric and states can be 0 or 1. Determine the most stable configuration for the networks in terms of states assignment (report state values when network is in minimum energy state). Also determine the minimum energy of the network. Assume bias to be zero. [10]



4. Suggest a suitable deep learning architecture for converting a dog images to lion. Describe the architecture, training strategy, and provide justifications why the proposed architecture would work? [8]

5. Recall CosFace and ArchFace loss to regularize deep feature space. On similar line of thought what next could be proposed. Justify merits of your proposal? [5]

6. What is the philosophical difference between DenseNet and ResNet. What architectural difference it brings and how that helps in a better classification. [5]

7. Consider two random variables X and Y, taking value in the range [0,7] and having the probability distribution as below [5+1]

	0	1	2	3	4	5	6	7
X		0.35	0.05	0.05	0.1	0.2	0.05	0.2
Y	0.2	0.15	0.1		0.1	0.05	0.3	0.1

Compute JS-Divergence between the two distributions. Without any computation, comment on the possible value of KL-divergence between the distribution of X and Y.

8. (a) What technique can significantly boost accuracy by data augmentation for robust training? [1]
   (b) What is poison in training with adversarial examples. [2]
   (c) Formally define list of desirable properties for a good performance measure for classification. [2]
   (d) Why Do Better Loss Functions Lead to Less Transferable Features? [2]
   (e) Explain how CoFrNet brings intrepretability in NN architecture. Highlight its limitation. [3]
