Aug 2021: End Semester Assessment (ESA) B Tech VI Sem
This is a closed book exam. All questions are compulsory
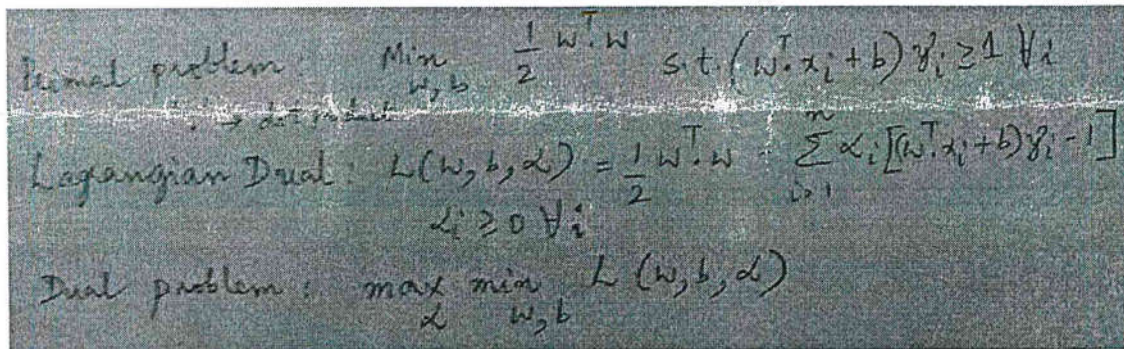Provide **concise (short and sharp)** answers. Mention clearly any assumptions made.

UE18CS343: Topics in Deep Learning

Time: 3 Hrs	Answer All Questions	Max Marks: 100

1. a. Explain AdaGrad and Adam techniques of optimization   8 marks (4+4)

   b. Explain Temporal Difference Learning with appropriate equations   5 marks

   c. Sketch the graphs of ReLU and tanh activation functions. What are the advantages of Leaky ReLU over ReLU   3 marks (1+1+1)

   d. Explain the concept of Dropout. How is Stochastic Gradient Descent different from Batch Gradient Descent?   4 marks (2+2)

2. a. Derive the dual formulation of the SVM classifier (linearly separable case) 4 marks

Discuss the sparsity of the Dual solution (why only a few alphas are non-zero) using the KKT complementary slackness condition)   3 marks
Discuss how the solution for 'b' is found   1 mark

   b. A 2D Graph is given to you containing 2 points x_1(+) at (0,0) and x_2(-) at (4,4)
      1. Find the parameters of the **maximum margin** SVM Boundary (Preferably formulate a **General** solution as any arbitrary multiple of a boundary eqn is also a solution)   2 marks

      2. Solve for the **alphas** (the dual variables) (α_i)   2 marks

      3. We move x_2(-) to (k,k)? Find the relationship between **alphas** and **k**   2 marks

      4. Express **alphas** in terms of the **margin width** and interpret the result   2 marks

   c. Explain the general idea of the Sequential Minimal Optimization (SMO) algo 4 marks

3. a. Derive the RNN Loss Gradient of L3 with respect to W ,V for the following network.

8 marks
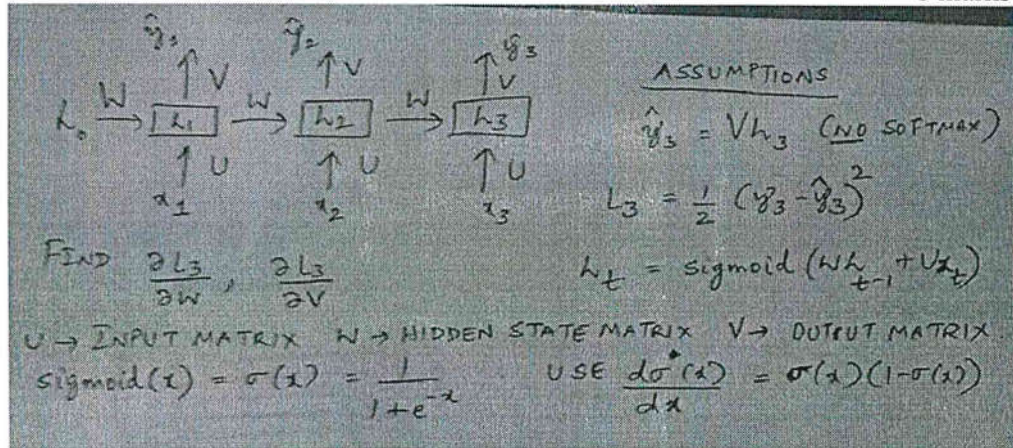
b. Explain the architecture of a Variational Autoencoder. How is a VaE different from an Autoencoder?   6 marks (4 + 2)

c. Explain seq2seq model with Attention   6 marks

4. a. Explain the architecture of a Generative Adversarial Network(GAN). Describe the Training process of a GAN   10 marks (5 + 5)

b. What are the operations performed within capsules? Explain the working of Dynamic Routing in Capsule Networks.   7 marks (3+4)

c. An input image matrix of size 24 X 24 is operated on by a filter of size 3 X 3 with a Stride of 2 and a padding of 1. Determine the size of the convoluted matrix.   3 marks

5. a. List the two fundamental operations forming the 'blueprint' of a Graph neural Network(GNN)? Explain the working of a GNN using Neural Message Passing as the basis, depicting the usage of the 'blueprint' for State updates   12 marks (2+ 10)

b. 'Convolutional Neural Networks cannot be directly applied on Graphs'. Provide 2 reasons to substantiate the above statement?   2 marks

c. What is the conceptual difference between the initial representation of a graph node(input to GNN) and the output representation of a graph node(output of GNN)? 2 marks

d. State whether the following 2 statements are TRUE or FALSE with relevant reasons:
     'The Aggregate procedure in a Graph Neural Network does NOT need to be Permutation Invariant' (Hint: Use the Neural Message passing idea)   2 marks
        'Constructing a Deep Graph Neural Network poses challenges to performance and imposes penalties'   2 marks