Name: **Enrolment No:**	

UPES
End Semester Examination, May 2025

Course: Deep Learning — **Semester: 2**
Program: MCA — **Time: 03 hrs.**
Course Code: CSAI7020 — **Max. Marks: 100**

Instructions:

- **A calculator is allowed.**
- **Question 9 and 11 has an internal choice.**

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q1	Compare CNNs (Convolutional Neural Networks) and RNNs (Recurrent Neural Networks) in terms of their suitability for different types of tasks. • In which tasks would you prefer to use a CNN over an RNN and vice versa? • Justify your answer with specific examples, such as image classification (CNN) vs. time series forecasting (RNN).	**4**	**CO3**
Q2	Compare the Adam optimizer with Stochastic Gradient Descent (SGD) in terms of their performance in deep learning tasks.When would you use Adam instead of SGD, and what advantages does Adam offer in training deep networks?	**4**	**CO2**
Q3	Analyze the differences between feedforward neural networks (FNNs) and autoencoders. • What are the primary applications of each model? • Discuss how the architecture and the objective function differ between these models, especially in terms of training data and learning objectives.	**4**	**CO4**
Q4	You are designing a self-driving car system. Compare the use of Convolutional Neural Networks (CNNs) for object detection versus Recurrent Neural Networks (RNNs) for motion prediction in autonomous vehicles. • How could you combine both models to enhance performance? • Provide a high-level explanation of how the CNN and RNN can complement each other in such an application.	**4**	**CO1**

Q5	Consider a 2-layer feedforward neural network with the following parameters: Input: $\begin{bmatrix}2\\1\end{bmatrix}$, Weights of first layer: $W^{[1]} = \begin{bmatrix}0.5 & 0.1\\0.2 & 0.3\end{bmatrix}$, $b^{[1]} = \begin{bmatrix}0.2\\0.2\end{bmatrix}$ Weights of second layer: $W^{[2]} = [0.4 \quad 0.5]$, $b^{[2]} = 0.6$ Activation function is ReLU and Activation is Linear. Compute the output of the network for the given input.	**4**	**CO1**
SECTION B **(4Qx10M= 40 Marks)**			
Q6	Suppose the cost function is: $C = \frac{1}{2n}\sum_{i=1}^{n}(y_i - a_i)^2$, for $n = 1$ Let y=0.7, predicted a=0.9, and learning rate η=0.2. Perform a **gradient descent step** on weight w with a=σ(w·x+b) x=2.0, w=0.5, b=0	**10**	**CO2**
Q7	You design a 3-layer FFNN as follows: • Input layer: 784 units (for MNIST) • Hidden layer 1: 128 units • Hidden layer 2: 64 units • Output layer: 12 units (Softmax) Compute the total number of trainable parameters in the network.	**10**	**CO1**

Q8	Discuss the loss function of a Variational Autoencoder.
	• Derive the two main components of the VAE loss function:
	  ○ Reconstruction loss (how well the decoder reconstructs the input).
	  ○ KL divergence loss (how close the distribution of the latent space is to a standard normal distribution).
	• Why is it important for the latent space to follow a standard normal distribution in a VAE?
	• How does the VAE use these two terms to both reconstruct data and generate new data?
		10	**CO5**

Q9	a. Explain how the Transformer model processes the input sentence through the encoder-decoder framework. b. Discuss how the model's attention mechanism helps in handling words with different importance in the translation task. OR Consider a Convolutional Neural Network (CNN) layer where you are performing convolution on an input image. The image has the following properties:	**10**	**CO4/CO5**

	Input image size: 32 × 32 × 3 (Height × Width × Channels) Convolution filter size: 4 × 4 Number of filters: 5 Stride: 1 Padding: Same padding (output size should be the same as input size for each filter) Calculate the output size (Height × Width) for each filter in the convolutional layer. Assume the same padding is applied. Use the formula for output size in convolution: $$O = \frac{(I + 2P - F)}{S} + 1$$ where: • O = Output size (Height/Width) • I= Input size (Height/Width) • P = Padding • F= Filter size (Height/Width) • S= Stride		
SECTION-C **(2Qx20M=40 Marks)**			
Q10	You are working on a project for a media analytics company. The goal is to build a sentiment analysis model that classifies user reviews into positive or negative sentiments. The input is a sequence of word embeddings (pre-trained using Word2Vec), and you are experimenting with both RNN and LSTM to improve the model's performance. **Word** **Embedding** "yes" [0.1, -0.2, 0.3, 0.0] "this" [-0.3, 0.2, 0.1, 0.4] "is" [0.2, -0.4, 0.0, 0.1] "my" [0.0, 0.0, 0.0, 0.0] "book" [0.1, 0.1, 0.2, 0.2] (i). You decide to use sequences of 5 words, each represented as a 4-dimensional vector. What is the input shape to the RNN/LSTM model? (ii). If you use an LSTM layer with 4 input features and 10 hidden units, Calculate the total number of trainable parameters in this LSTM layer. *(Hint: LSTM parameters = 4 × [(input_dim + hidden_dim) × hidden_dim + hidden_dim])* (iii). During training, you notice that the RNN is unable to capture dependencies between distant words (e.g., "bad" and "movie"). (iv). Explain how the LSTM architecture overcomes this limitation. (v). After training, the LSTM outputs a value of 0.83 for a review, where 1 = positive, 0 = negative.	**20(each part is of 4 marks)**	**CO5**

<table>
<tr><td></td><td>If you apply a threshold of 0.5, what would be the final prediction class?
Justify whether this thresholding is suitable or if a better approach exists.</td><td></td><td></td></tr>
<tr><td>Q11</td><td>You are given a high-resolution image (1800×1580) containing multiple small and large objects (e.g., people, vehicles, street signs).
(i) Describe how R-CNN and Fast RCNN differ in object detection pipelines.
Include at least three core differences in terms of:
Region proposal generation
Feature extraction strategy
Prediction stage
(ii) Explain why FAST RCNN is significantly faster than R-CNN but often struggles with detecting small or overlapping objects.
(iii) Imagine a real-time surveillance system in an airport terminal.
You are tasked with choosing either R-CNN or Fast RCNN for this application.
Justify your choice using logical reasoning based on:
-Speed vs. accuracy trade-off
-Object size and density
-Deployment hardware (e.g., edge device vs. GPU server)
(iv). Compare Fast RCNN and R-CNN based on the handling of overlapping objects.

OR

Explain the architecture of a Generative Adversarial Network (GAN).<ol type="a">
Describe the roles of the Generator and Discriminator.
What is the objective of each network, and how do they interact during training?
Explain the concept of a minimax game in GANs.
Write the loss function used in standard GAN training and explain its components
You are tasked with using a GAN for generating realistic human face images (e.g., for a face dataset like CelebA).
How would you design the Generator and Discriminator networks?
What preprocessing steps are necessary for the image dataset?
Explain how you would evaluate the performance of your GAN (mention metrics like Inception Score or FID).
Suggest one ethical concern associated with using GANs in realistic image generation.</td><td>20(4+4+4+4+4)</td><td>CO5</td></tr>
</table>