Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

B.Tech Degree S8 (R,S) Examinations April 2026 (2019 Scheme)

Course Code: CDT402
Course Name: DEEP LEARNING FOR DATA SCIENCE

Max. Marks: 100 **Duration: 3 Hours**

PART A

Answer all questions, each carries 3 marks. Marks

1		How can overfitting affect model generalization?	(3)
2		Specify the advantages of ReLU over sigmoid activation function.	(3)
3		Explain how L2 regularization improves the performance of deep feed forward neural networks.	(3)
4		How does dataset augmentation help when working with limited training data?	(3)
5		Illustrate the strengths and weaknesses of convolutional neural networks.	(3)
6		With the help of an example, illustrate convolution and pooling operations.	(3)
7		How does a recursive neural network work?	(3)
8		Explain how LSTM addresses the long-term dependency issues in language models.	(3)
9		Explain Denoising Autoencoders.	(3)
10		Discuss the applications of GAN.	(3)

PART B

Answer any one full question from each module, each carries 14 marks.

Module I

11	a)	You are training a fully connected neural network with 4 hidden layers, each containing 9 hidden units. The input dimension is 15, and the output is a scalar. The network includes bias terms for each neuron in both the hidden and output layers. Calculate the total number of trainable parameters in the network.	(7)
	b)	Consider the case of the XOR function in which the two points {(0, 0), (1, 1)} belong to one class, and the other two points {(1, 0), (0, 1)} belong to the	(7)

other class. Design a multilayer perceptron for this binary classification problem**.**

OR

12 a) With an example of classification problem, explain the following terms: (8)

i) Hyper parameters ii) Training set iii) Validation sets iv) Bias.

b) Illustrate the limitation of a single layer perceptron with an example. (6)

Module II

13 a) Differentiate gradient descent with and without momentum. Give equations for weight updation in GD with and without momentum. Illustrate plateaus, saddle points and slowly varying gradients. (14)

OR

14 a) Explain how L1 and L2 regularization techniques help in preventing overfitting in neural networks. (8)

b) Write weight updating rule in gradient descent when the error function is a) mean squared error b) cross entropy. (6)

Module III

15 a) Explain the architecture of ResNet model. (8)

b) Give two benefits of using convolutional layers instead of fully connected ones for visual tasks. (6)

OR

16 a) Consider an input image with dimensions of 28 x 28 pixels. You apply a convolutional operation with a kernel (filter) size of 3x3, a padding of 0, and a stride of 2. Calculate the dimensions of the output feature map. Also, calculate the padding value if we need the output to have the same size as the input with a stride of 1. (8)

b) Explain the following functions (i) kernel flipping, (ii) down sampling, (iii) pooling, (iv) strides. (6)

Module IV

17 a) Differentiate between Vanilla RNN and Gated RNN. (5)

b) Illustrate the working of RNN-based model used in text summarization applications where a long input document is converted into a shorter version. (9)

OR

18	a)	Explain the architecture of LSTM.	(7)
	b)	The vanishing gradient problem is more pronounced in RNN than in traditional neural networks. Give reasons. Discuss a solution for the problem.	(7)

Module V

19	a)	With the help of a diagram, explain the training process of Generative Adversarial Networks (GANs) and the adversarial relationship between Generator and Discriminator.	(8)
	b)	Explain the concept of Variational Autoencoders.	(6)

OR

20	a)	Explain the Deep belief Networks and their significance in the field of deep learning.	(6)
	b)	Compare Boltzmann machine and traditional neural network.	(8)

****