

MIT ACADEMY OF ENGINEERING**COURSE CODE: CS422T****13 DECEMBER 2019****BTECH SEMESTER - VII 2019-2020 EXAMINATION****DEPARTMENT OF COMPUTER ENGINEERING****END SEMESTER EXAMINATION****DEEP LEARNING****TIME : 3 HOURS****MAX MARKS : 100 MARKS****TOTAL NO OF QUESTIONS: 5****TOTAL NO OF PRINTED PAGES:2****INSTRUCTIONS TO CANDIDATES:**

1. Assume suitable data wherever necessary
2. Non programmable scientific calculators are allowed
3. Black figures to the right indicate full marks

- 1 a) "Normal distribution is a probability distribution that is [05] CO1 L2 symmetric about the mean and median. In graph form, normal distribution will appear as a bell curve." Illustrate the statement with suitable example.
(Example - 2M,Justification - 3M)
- 1 b) Two standard dice with 6 sides are thrown and the faces [05] CO1 L1 are recorded. Given that the sum of the two faces equals to 10, what is the probability that the first throw equals to 5?
- 1 c) Why does regularization reduce Over fitting? Make use of [10] CO2 L3 two forms of regularization to describe it.
(Reason - 4M,Two forms - 6M)
- 2 a) Analyze the process to control the size of output image in [06] CO3 L4 each CNN layer using hyper parameter like depth, stride and zero-padding.
(2M each hyper parameter)
- 2 b) Identify the steps involved in the working of softmax [04] CO3 L3 classification layer of a neural network and describe them?
Softmax Explanation(1M) with mathematical model (1M) and suitable example (2M)

- 2 c) Compare back propagation in a recurrent neural network with a feed-forward neural network? [06] CO4 L4
(Comparison discussion with suitable example)
- 2 d) Explain how the traditional Recurrent Neural Network (RNN) units suffer from the vanishing gradient problem? (RNN-2m, Vanishing gradient -2M) [04] CO4 L2
- 3 a) Apply LSTM model and specify how it overcome short term memory challenges in RNN model? [10] CO4 L3
(Short term memory problem in RNN-2M, Explanation LSTM model steps - 6M, Diagram of LSTM model - 2M)
- 3 b) Make use of output gate and cell state to operate on previous hidden state information and current input? [05] CO4 L3
(0.5M - diagram of each gate, explanation - 2M each)
- 3 c) Interpret how RNN train sequential data set. [10] CO5 L5
(RNN purpose-2m, Example with a suitable example -8M)
- 4 a) Determine the procedure to automate Ulcer Detection in Wireless Capsule Endoscopy Images using CNN. [10] CO5 L5
(CNN purpose-2m, Example with an example-5M and diagram -3M)
- 4 b) Determine difference between hierarchical and multiagent reinforcement learning with suitable example? [06] CO6 L5
(3 comparisons - 3M, example of each- 1.5M)
- 4 c) Explain how robot will learn to walk in room using reinforcement learning? What parameters affects the performance of robot's action. [06] CO6 L5
(Explanation - 3M, Parameters - 3M)
- 5 a) Q-Learning algorithm is used to maximize its reward in the long run. Justify your answer. [10] CO6 L5
(Q learning algorithm with explanation - 5M, Example - 5M)
- 5 b) What is role of policy gradient in deep reinforcement learning? [03] CO6 L1
(Explanation 3M)

MIT ACADEMY OF ENGINEERING**COURSE CODE: CS422T****4 OCTOBER 2019****BTECH SEMESTER - VII 2019-2020 EXAMINATION****DEPARTMENT OF COMPUTER ENGINEERING****IN SEMESTER EXAMINATION****DEEP LEARNING****TIME : 2 HOURS****MAX MARKS : 50 MARKS****TOTAL NO OF QUESTIONS: 5****TOTAL NO OF PRINTED PAGES:03****INSTRUCTIONS TO CANDIDATES:**

1. Assume suitable data wherever necessary
2. Non programmable scientific calculators are allowed
3. Black figures to the right indicate full marks

- 1
 - a) The owner of the Ches Tahoe restaurant is interested in how much people spend at the restaurant. He examines 10 randomly selected receipts for parties and writes down the following data.
 44, 50, 38, 96, 42, 47, 40, 39, 46, 50
 Calculate Mean, Mode, Median, Standard Deviation and Variance of selected samples. **[05] CO1 L2**
 - b) You toss a fair coin three times: **[03] CO1 L2**
 What is the probability of three heads, HHH?(1.5M)
 What is the probability that you observe exactly one heads?(1.5M)
 - c) Let X be a continuous random variable with PDF **[02] CO1 L2**
 $f(x) = 4x^3 \quad 0 < x \leq 1$
 $= 0 \quad \text{otherwise}$
 Find Probability density function $P(X \leq 2/3 \mid X > 1/3)$
- 2
 - a) Compare the discrete and continuous random variables in population with example(any four comparisons(4M) and one example of each(1M)) **[05] CO1 L2**
 - b) How bias and variance measured and balanced for analysis of outcomes of linear regression model? **[05] CO2 L2**
- 3
 - a) What is data augmentation? and State the significance of data augmentation in real time machine learning applications. Explain with suitable examples **[05] CO2 L3**

- b) Consider linear regression algorithm to find the predicted salary of employees based on their experience in company and calculate Sum of Squared Errors (SSE). (2M) Assume the initial values of Intercept $c=10.15$ and Slope $m=5.80$ to reduce SSE. And find updated values of m and c to reduce SSE using gradient descent (4M) and new predicted salary values of every data points (4M). Assume learning rate = 0.001

Years of Experience	Salary in 1000\$
2	15
3	28
5	42
13	64
8	50
16	90
11	58
1	8
9	54

- 4 a) How to reduce the overfitting using K-fold Cross validation explain with suitable example. (Explanation 3M, Example 2M) [05] CO2 L3
- b) There is huge and high dimensional dataset of cat, dog, lion images. Mr. Clarry want to classify this dataset into Cat, Dog and Lion classes. How can you help him to classify the selected dataset using convolutional neural network, explain with suitable diagram. (Layers in CNN models 7M, diagram -3M) [10] CO3 L4
- 5 a) Consider following CNN model for classification of images of size $1025 \times 1025 \times 3$, contain two convolutional layer and two fully connected networks. Assume following hyper parameter for model. Calculate the size of output tensor (image) in conv layer1, conv layer2, pooling layer1, pooling layer2 and parameters in fully connected layer 1 and 2 along with diagram of CNN model with calculated parameters and size of tensors. (1M each parameters in two layers (2M), 1M tensor size in each layer (4M), Diagram -4M) [10] CO3 L4

Layer	Filter size	No of Filters	Stride	Padding	Number of Neurons
Conv Layer 1	5*5	130	2	0	-
Max-Pool layer 1	3*3	-	1	-	-
Conv Layer 2	3*3	230	2	1	-
Max-Pool layer 2	3*3	-	2	-	-
FC layer 1	-	-	-	-	3900
FC layer 2	-	-	-	-	10

MIT ACADEMY OF ENGINEERING**COURSE CODE: ET414T****11 DECEMBER 2019****BTECH SEMESTER - VII 2019-2020 EXAMINATION****DEPARTMENT OF ELECTRONICS AND TELECOMMUNICATION ENGINEERING****END SEMESTER EXAMINATION****MACHINE LEARNING****TIME : 3 HOURS****MAX MARKS : 100 MARKS****TOTAL NO OF QUESTIONS: 5****TOTAL NO OF PRINTED PAGES:2****INSTRUCTIONS TO CANDIDATES:**

1. Assume suitable data wherever necessary
2. Non programmable scientific calculators are allowed
3. Black figures to the right indicate full marks

1 a) Describe Supervised, Unsupervised and Reinforcement Learning with suitable example. [6] CO2 L2
[Each 2 Marks]

b) How does regularization prevent Overfitting? [2 Marks] [6] CO1 L1
What is L1 and L2 regularization? [2 Marks Each]

c) Explain Cost Function of Logistic Regression. [4Marks] [4] CO3 L3

2 a) Consider the given Dataset, apply Naïve Bay's algorithm & predict that if a fruit has the following properties then which type of fruit it is. [8] CO4 L4
Fruit Type = { Yellow, Sweet, Long) = ?

Fruit	Yellow	Sweet	Long	Total	[8 Marks]
Mango	350	450	0	650	
Banana	400	300	350	400	
Others	50	100	50	150	
Total	800	850	400	1200	

b) Explain Random Forest Tree Algorithm with suitable block diagram [6 Marks]. List advantages of Random Forest Tree [2 Marks] [8] CO4 L3

- 3 a) Explain why SVM called as Large Margin Classifier [10] CO3 L4
with diagram.[4 marks]
State and proof Margin width of SVM $2/(||w||)$ with proper diagram.[6 Marks]
- b) What's the "kernel trick" and how is it useful in face recognition using SVM? [5 Marks] [8] CO3 L3
List different kernels functions in SVM. [3 Marks]
- 4 a) Forecasting is required extensively in everyday business decisions in economic and monetary policy, in finance and stock market. More often, forecasting problems are complex, for example, predicting stock prices is a complex problem with many underlying factors .Traditional-forecasting models throw up limitations in terms of taking into account these complex, non-linear relationships.
So to handle above case explain how to train Multilayer Neural Network with back propagation algorithm. [Draw Multilayer diagram [5 Marks] with mathematical equation, Explain learning Algorithm with back propagation [5 Marks]] [10] CO4 L4
- b) Explain any three activation function with the help of mathematical equation and graph. [2 Marks Each] [6] CO1 L2
- c) There is huge and high dimensional dataset of vehicles Car, Bus and Trucks images. Mr. Jhone wants to classify this dataset into Car, Bus and Trucks classes. How can you help him to classify the selected dataset using convolution neural network, explain with suitable diagram. [Layers in CNN models 7Marks, diagram - 3Marks] [10] CO4 L5
- 5 a) Explain how Ensemble Learning handles low bias and high variance. [6 Marks] [6] CO1 L2
- b) Describe ADBOOST Ensemble Learning Algorithm in detail. [10 Marks] [10] CO4 L4
- c) Explain Clustering approach in Machine Learning with example.[4 Marks] [8] CO5 L4
How is the k-nearest neighbour algorithm different from k-means clustering? [Marks 4]

MIT ACADEMY OF ENGINEERING

COURSE CODE: ET414T

3 OCTOBER 2019

BTECH SEMESTER - VII 2019-2020 EXAMINATION

DEPARTMENT OF ELECTRONICS AND TELECOMMUNICATION ENGINEERING

IN SEMESTER EXAMINATION

MACHINE LEARNING

TIME : 2 HOURS

MAX MARKS : 50 MARKS

TOTAL NO OF QUESTIONS: 6

TOTAL NO OF PRINTED PAGES: 3

INSTRUCTIONS TO CANDIDATES:

1. Assume suitable data wherever necessary
2. Non programmable scientific calculators are allowed
3. Black figures to the right indicate full marks

Q1	I). Classify following for Machine Learning problem is of Classification or Regression [1 Marks Each]	[6]	CO1 CO2	L2
-----------	--	------------	--------------------	-----------

a) Let's say there are linguistics researchers studying grammar structures in languages. They have a bunch of text files of transcribed speeches that they want to analyze. They want to teach a computer to recognize parts of speech, such as adjectives, subject, and verbs, in the sentence.

b)Kinematics of a robot arm: predict workspace location from angles.

c) Let's say you are trying to write a machine learning program that will be able to detect cancerous tumors in lungs. It takes in images of lung x-rays as input and determines if there is a tumor.

**II). Classify following for types of Machine Learning
[1 Marks Each]**

a) Sentiment analysis of tweets classifying them as positive, negative or neutral.

b) Market research - for differentiating groups of customers based on some attributes.

c) Attendance monitoring system based on Face Image.

Q2	<p>Explain Gradient Descent Algorithm for Linear Regression Machine Learning Model.</p> <p>1. Write down hypothesis function & Cost Function [2 Marks]</p> <p>2. Illustrate Gradient Descent Algorithm steps. [4 Marks]</p> <p>3. Explain importance of Learning Rate with gradient descent graph [2Marks]</p>	[8]	CO3	L4
Q3	<p>a) You are given a train data set having 1000 columns and 1 million rows and your machine learning model start capturing noise of the data during training and model shows high variance which situation occurs Underfitting or Overfitting [1 Marks]. Explain why with graph [2 Marks] and give solution for the same [2Marks].</p> <p>b) i) What is the purpose of cross-validation and explain process. [3 Marks] ii) Explain K-Fold Cross Validation with block diagram [3 Marks]</p>	[5]	CO1	L2
Q4	<p>Consider database given to you to classify customers according their different purchases product. Since there are thousands of products in the dataset, which represent categories of customer. As its curse dimension dataset and your manager has asked you to reduce the dimension of this dataset to improve classification accuracy.</p> <p>a) List down advantage of dimensionality reduction of input feature in Machine Learning [Marks 3]</p> <p>b) Explain how PCA algorithm helps to find new reduced dimension dataset which help to improve classification accuracy. [5 Marks]</p>	[8]	CO2	L3
Q5	<p>Consider given dataset of customer with input features age, income and no. of credit cards. Find out response for shopping of customer Nita using K Nearest Neighbors assume K=3. (Dataset on next page given)</p>	[6]	CO4	L5

Customer Name	Age	Income in K	No.of Credit Card	Response for shopping
Amar	35	35	3	No
Dipa	22	50	2	Yes
Raj	63	200	1	No
Nita	59	170	1	No
Vijay	25	40	4	Yes
Neha	37	50	2	?

Q6 Construct Decision Tree based on Entropy and Information Gain for given dataset in which Age, Comptition, Type are attribute and Profit is output class label. **[12] CO4 L5**

1. Correctly Identify Root Node **[6 Marks]**
2. Correctly Identify Child Node **[3 Marks]**
3. Draw Final Decision Tree which classify correct Profit output label of Profit --Down or Up **[3 Marks]**

Age	Competition	Type	Profit
Old	Yes	Software	Down
Old	No	Software	Down
Old	No	Hardware	Down
Mid	Yes	Software	Down
Mid	Yes	Hardware	Down
Mid	No	Hardware	Up
Mid	No	Software	Up
New	Yes	Software	Up
New	No	Hardware	Up
New	No	Software	Up