Roll No.

FIVTH SEMESTER – BTECH, MATHEMATICS & COMPUTING
MID-SEMESTER EXAMINATION- MAR-2022

COURSE CODE: CMCSC20
COURSE TITLE: MACHINE LEARNING **Time:** 90 minutes
Max marks:25
Instructions:
Attempt all questions.
Assume missing data if any, and mention in your answer

		Marks	CO
Q1	The following table must be used for Q1A and Q1B		

X_1	X_2	y_{ACTUAL}	$y_{PREDICTED}$
1	3	2	4
2	2	5	10
3	1	3	4
4	6	6	6
5	7	2	5

Q1A	Calculate R^2 and adjusted $\overline{R^2}$, showing the steps of calculation clearly.	2	CO2
Q1B	Calculate the weight adjustments for w_0, and w_2 with gradient descent, assuming a learning rate of 1. You can add columns to the table to show your calculations clearly.	3	CO2
Q2A	Describe a classification task that you have actually faced, where the target variable is dependent on three or more independent variables. Populate the data for 6 data points giving proper units and suitable values.	2	CO3
Q2B	Consider the following data points {(x,y)}, where y is discrete valued [1,5]. (1,1) (1.5,1) (2,1) (2.5,1) (3,5) (3.5,1) (4,1) (4.5,5) (5,1) (5.5,5) (6,1) (6.5,5) (7,1) (7.5,5) (8,5) (8.5,1) (9,5) (9.5,5) (10,5) Graph the above points and then approximate with a sigmoid. Show the True Positives, True Negatives, False Positives and False Negatives in logistic regression, assuming a threshold of 0.5. What happens when the threshold is reduced to 0.3?	3	CO1
Q3A	Using logistic regression with w_0= 2, w_1=-3, x increases from 4 to 6. By what factor does the log-odds and odds increase/decrease?	2	CO2
Q3B	From the confusion matrix given below, calculate the F1 measure, and specificity.		CO3

	Predicted No	Predicted Yes
Actual No	15	4
Actual Yes	6	18

	Compare the concepts of Entropy and Gini index for building decision	3	

85

Q4A)	tress by splitting nodes. Enumerate your points of comparison.	2	CO2
Q4B)	Derive the decision tree for the data (3 features and 2 classes) given below:	3	CO2

X1	X2	X3	Decision
A	C	E	Y
A	C	F	Y
B	D	E	N
B	D	F	N

Q5A) Given the following table, classify the new examples using 3-KNN. — 2 — CO1/CO3

X1	X2	X3	X4	Decision
1	0	1	1	Y
0	0	0	0	N
1	1	0	0	Y
0	1	1	1	N
1	0	0	0	N
1	1	1	1	?

Q5B) Write a short note on various ensemble methods. — 3 — CO2

TOTAL NO OF PAGES: 2 DATE: ROLL NO.

MID SEMESTER EXAMINATION – SEPT 2022
B.TECH COMPUTER ENGINEERING

COURSE CODE: COCSC17
COURSE TITLE: MACHINE LEARNING

DURATION 1.5 HOUR MAX MARKS: 15

Instructions:
Assume any missing data
Do not write very long answers, else you may run out of time. Justify wherever required, clearly and to the point.
Number the answers fully (such as A1a-iii)

No	Question	Marks	CO
Q1a	You have started an e-commerce site of general products. You want to deploy the following applications: (i) Planning the inventory (stocks of various products) (ii) Recommend products based on latest trends (iii) From comments about a product, segregate complaints from praise. For each of the above, decide what type of ML approach will work and why?	1.5	CO3
Q1b	Mankind wants to develop a rover for exploring another solar system. Which of the following domains may be relevant and why? NLP, Vision, ML, Robotics, Social Network Analysis.	1.5	CO3
Q2a	The ecological status Y of an area can be *balanced, unbalanced.* It depends upon (i) whether forest area x1, is less than or more than 25% of entire land (ii) energy consumption x2 is (low, high) (iii) Pollutants in biosphere x3, are in (red, green). If I want to predict the ecological status of a new region, (i) suggest the input variables and their possible values (input space) (ii) the form of the hypothesis function (ii) Calculate the hypothesis space size.	1.5	CO2
Q2b	What is the purpose of cross-validation? Given historical data of 100 instances with 20 of them positive and remaining 80 negative, suggest a suitable method for training-testing with cross-validation.	1.5	CO1
Q3a	Fisherman Somu doesn't mind if his catch of surmayi has some rohu, but he does not want to miss any surmayi. His companion Vallabh really doesn't not want rohu to mix up with surmayi, even if he misses on some of them. The following confusion matrices will be suitable for Somu or for Vallabh and why?	1.5	C23

	Actual Surmayi	Actual Rohu
Predicted Surmayi	120	75
Predicted Rohu	5	50

Q3b	Given training data of 100 GB, using stochastic gradient with 70:30 train test division and 3-fold cross validation.. It takes 1 microsecond for each of the following operations (i) Calculate loss function – 10^{-6} sec (ii) update weights using Delta learning rule – 10^{-6} sec (iii) calculating output – 10^{-6} sec. Calculate: (i) training time in hours (ii) testing time in hours (iii) field prediction time for 10 new data.	1.5	CO2

Q4a	From the Table given below showing regressor "School dropout rate" and response "employment rate" in different states of India, derive the bias and regression coefficient, from the given data.	1.5	CO2

State	School dropout rate	Employment rate
Punjab	40%	50%
Kerala	10%	93%
Bihar	60%	40%
Manipur	2%	50%

Q4b	For an application where fees of higher education is predicted based on faculty strength, faculty research profile, faculty expertise in a given field, salary paid to faculty, and faculty retentivity, which kind of regularization may be use Lasso or Ridge or a combination. Justify.	1.5	CO1
Q5a	For a fixed bias w_0=0, If the weight w_1 of the single regressor increases by 2 units, by how much will (ii) Odds increase? (iii) Log-odds increase?	1.5	CO1
Q5b	Why stochastic gradient descent / ascent may be better than considering all data points while training?	1.5	CO1

Total No. of Page: 2 4 Roll No. _____

B.Tech., V Semester, Computer Engineering, CSAI
END SEMESTER EXAMINATION December 2021

Course Code: COCSC17/CACSC17
Course Title: Machine Learning

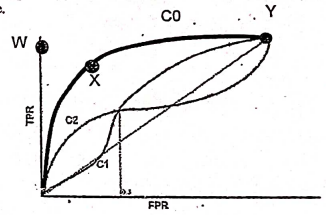
Time: 3 Hours Max. Marks : 40

Note: - Attempt all the five questions. Missing data/ information if any, maybe suitably assumed and mentioned in the answer.

Q. No.	Question	Marks	CO
Q1	**Attempt any 2 parts of the following.**		
1a	Elaborate on the following ML tasks and their type, describe suitable performance metrics, and explain how the system can gain experience for learning: (i) a handwriting recognition system (ii) forecasting the network traffic (throughput in bps) during different times of a day.	4	CO1
1b	For the following training dataset, find the linear regression model parameters for Y=a_X+b. Round the parameters to two decimal places. For the derived model, calculate R^2 and adjusted R^2. X: 1 3 5 7 9 Y: 4 5 7 10 13	4	CO1
1c	Distinguish between logistic regression as a discriminative classifier and Naïve Bayes as a generative classifier in terms of their probabilistic assumptions. Suppose you train a logistic regression classifier, and the learned hypothesis function is: $h_\theta(X) = \sigma(\theta_0 + \theta_1 X_1)$, where, $\theta_0 = 6$, $\theta_1 = -1$. Draw a graph to show the output probability. What is the log-odds when x=7?	4	CO2
Q2	**Attempt any 2 parts of the following.**		
2a	Consider the following set of training examples: (Age range in the dataset is 16 to 30 years)	4	CO3

Instance	Have Laptop	Age	Buy new laptop
1	T	16	Y
2	T	19	Y
3	T	18	N
4	F	27	N
5	F	24	Y
6	F	28	Y

Derive the Decision Tree (DT) for the above dataset showing information gain at each node split.

2b	Describe the working of the Random Forest Algorithm, explaining why aforest of stubs (collection of very small Trees) is preferred rather than a single Tree? How will the performance of the ensemble be affected if the attributes and the training examples are not distributed randomly?	4	CO3
2c	Explain how boosting can reduce both bias error and variance error. In an Adaboost ensemble, there are three DTs (1,2,3). They give the odds (probability of no-error / probability of error) as: $odds_1 = 1.6$, $odds_2 = 2.7$, $odds_3 = 0.5$. Their predictions for a given datapoint are: $\widehat{y_1} = -1, \widehat{y_2} = 1, \widehat{y_3} = -1$. Calculate the final prediction of the ensemble.	4	CO3
Q3	**Attempt any 2 parts of the following.**		
3a	(i) Given the ROC curve C0 and the operating points W, X and Y, justify which of the three is the best operating point. (ii) Given the ROC curves for two classifiers C1 and C2, compare their performance. 	4	CO2
3b	Given below are the test results of two classifiers developed to detect COVID patients. Derive their confusion matrices given that the threshold for output is at 0.6. Which of the two would a doctor use if she/he does not want to miss people with COVID, even if it means some normal people are diagnosed as COVID positive?	4	CO1

Prediction 1

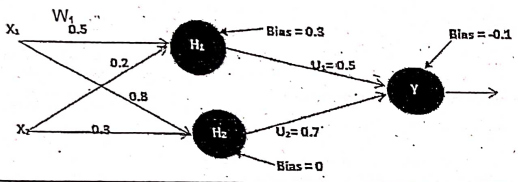
y	y pred
0	0.5
1	0.9
0	0.7
1	0.7
1	0.3
0	0.4
1	0.5

Prediction 2

y	y pred
0	0.6
1	0.9
0	0.7
1	0.7
1	0.3
0	0.8
1	0.5

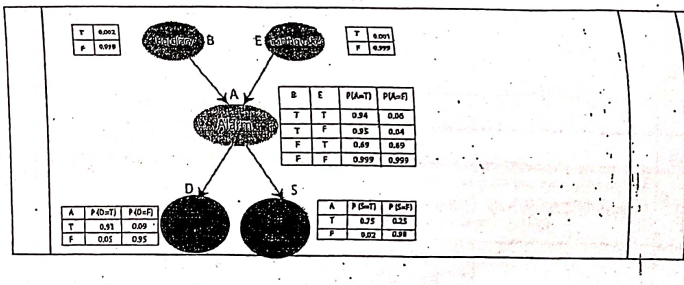
3c	What is the purpose of performing cross validation. Given the dataset {(x,y)}: { (0.1,1)(0.4,1)(0.5,0) (0.4,0) (0.9,1) (0.3,0) (0.8,1) (0.8,0) (0.2,1) (0.6,1) (0.9,1) (0.2,0) (0.3,0) (0.6,1) (0.5,1) (0.7,0) }, construct the decomposition of runs using 4-fold cross validation. Calculate the overall accuracy if the predictions for the four cycles are: (1010), (1110), (0111), (0101).	4	CO1
Q4	**Attempt any 2 parts of the following.**		

2/4.

4a	*Explain any two of the following:* *- Any one Recurrent Neural Network architecture and its application* *- Any one Long Short Term Memory network architecture, and its activation equations* *- Dropout Regularization in Convolutional Neural Networks*	4	CO1
4b	Given the MLP FNN below with sigmoid activations and training data, calculate the signal values at the input and output of each neuron, the error at the output and weight adjustment of U1 and W1 after a cycle of back-propagation. Training data = {X1= 0.2, X2=0.4, Y=0.6} 	4	CO3/5
4c	For a CNN architecture, explain the need for (i) stride >1, (ii) Maxpool layer, (iii) higher dropout rate at output classification layers, (iv) use of ReLU in inner feature learning layers.	4	CO5
Q5	**Attempt any 2 parts of the following.**		
5a	Classify a Red Domestic SUV using Naïve Bayes classification using the dataset given below:	4	CO4

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

5b	Derive the dual form of the optimized SVM, explaining its advantage over the original form. Explain how the "Kernel trick" handles non-linearly separable data.	4	CO3
5c	Given the Bayesian network shown in the Fig. below, calculate the probability that the alarm has sounded, but there is neither a burglary, nor an earthquake occurred, and Hari received a call only from Dev, but not from Sita.	4	CO4



SEMESTER M.Tech Artificial Intelligence
End -Semester Examination, December-2021

Course Code: CPAIE02

Course Title: Machine Learning

Time:03 HRS

Max Marks: 40

361
02/20

Note:- Attempt all questions. Missing data/information if any, may be suitably assumed & mentioned in the answer.

Q.No.	Questions (Attempt any two parts from each question) (4+4)=8 8X5=40 MM	Marks	CO
1a	Discuss the role of different type of learning to train the machine.	4	CO1
1b	How machine learning may be used to seek the solution of artificial intelligence related problems.	4	CO1
1c	Discuss the concept of inductive learning.	4	CO1
2a	Discuss linear classification technique with considering suitable examples.	4	CO2
2b	Discuss Artificial Neural Network with examples.	4	CO2
2c	How deep neural network play the role in machine learning?	4	CO2
3a	Discuss decision tree learning with examples.	4	CO3
3b	How the support vector machine is used to train the machine?	4	CO3
3c	Differentiate bagging and boosting techniques.	4	CO3
4a	Discuss the concept of unsupervised learning with examples.	4	CO4
4b	Compare the role of Naïve Bayes Classifier and	4	CO4

	Bayesian Belief Classifier in machine learning.	
4c	Compare K nearest neighbor and K Means clustering techniques	4
5a	Discuss the concept of reinforcement learning with suitable examples.	4
5b	What is the delta rule in reinforcement learning?	4
5c	Discuss the elements of reinforcement learning.	4

(END)

Total no. of pages - 2

Roll NO. .....

FIRST SEMESTER – M.TECH(CSE)
MID-SEMESTER EXAMINATION, SEPTEMBER, 2022

Course Code - COCSE01
Course Title - **Machine Learning**

Time-1.5hours

Max. Marks – 15

Note: - Attempt all questions. Missing data/information (if any) may be suitably assumed & mentioned in the answer.

Q. No.	Question	Marks	CO
1(a)	Explain advantages and disadvantages of Machine learning . Explain with specific example .	1.5	CO1
1(b)	Define learning rate? Explain its role in gradient descent.	1.5	CO3
2(a)	What are the difference between Liner and logistic regression	1	CO1
2(b)	Explain clarification and clustering with their use cases.	2	CO2
3(a)	What is Dimensionality Reduction and why we do so.	1.5	CO1
3(b)	Compare PCA with GA in terms of efficiency and applications. Explain algorithm steps of any one of them.	1.5	CO2
4(a)	Differntiate underfitting and over-fitting by taking a	1	CO2

159

	suitable example. Why they are avoided?		
4(b)	Why we need confusion matrix? Explain it ? Calculate accuracy, precision, recall and F1 score, with following values ▪ TP = 564 ▪ TN = 334 ▪ FP = 64 ▪ FN = 54	2	CO2
5(a)	What are the advantages of feature selection process.	1	CO3
5(b)	In case of road traffic detection which algorithm suited most SVM or Random forest and why ?	2	CO3

Code- ECCSC11

Course Title- Machine Learning and AI

Time- 1 Hr 30 Mins

Max. Marks- 15

Note: - Attempt all questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.

94/412

Q. No.	Question	Marks	CO
1a	Define Machine learning. Differentiate between the different techniques of machine learning with an example	1.5	CO1
1b	With the help of an example of prediction of a new variant of COVID-19 in a city, illustrate the ML development cycle.	1.5	CO1
2a	Given a logistic regression model, Let w_0 (or α) = -1 and w_1 (or β) = 3 . When x is doubled from 1 to 2, show by what factor does the log odds increase?	1.5	CO2
2b	A classification model has been built for Credit card fraud detection. Below is the confusion matrix: Predicted/Actual | Fair Transaction | Fraud Transaction Fair Transaction | 400 | 100 Fraud Transaction | 200 | 500 Calculate error rate, precision, recall and F1-score. Comment on the performance of the model.	1.5	CO2
3a	Given the following data (x,y) = {(1,3) (2,5), (3,7), (4,6), (5,9)} for linear regression. Calculate the mean squared error of the optimally fitting linear regression model.	1.5	CO2
3b	What is cross-validation in Machine learning? Discuss about its types.	1.5	CO1
4a	Compare lasso and ridge regularization in Linear Regression. What is the role of the regularization factor λ?	1.5	CO2
4b	Consider the following table having two attributes A1 and A2 and corresponding class label. Calculate the decision tree. S. no. | A1 | A2 | Label 1. | T | T | Class 1 2. | T | F | Class 2 3. | F | F | Class 1 4. | F | T | Class 2 5. | F | T | Class 2	1.5	CO2
5a	Consider the following data. Classify the data point X = 5.0 according to its 3-, and 5- nearest neighbors. X: 0.5, 3.0, 4.5, 4.6, 4.9, 5.2, 5.3, 5.5, 7.0, 9.5 Label: N, N, P, P, P, N, N, P, N, N	1.5	CO2
5b	What is the significance of the sigmoid function for logistic regression? Plot graphically the sigmoid function for $P(y=1) = \frac{1}{1+e^{a(bx+c)}}$, for a = 2, b=3, and c=-1.	1.5	CO2

Total No. of Pages 03 Roll No.

END SEMESTER EXAMINATION MAY 2022

B.TECH – ELECTRONICS AND COMMUNICATION ENGINEERING (ECE), IV SEMESTER

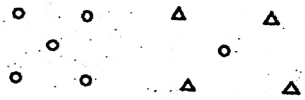
COURSE CODE: ECCSC11

COURSE TITLE: MACHINE LEARNING AND AI

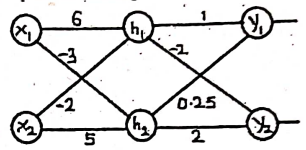
Time: 180 minutes Max marks:25

Instructions:

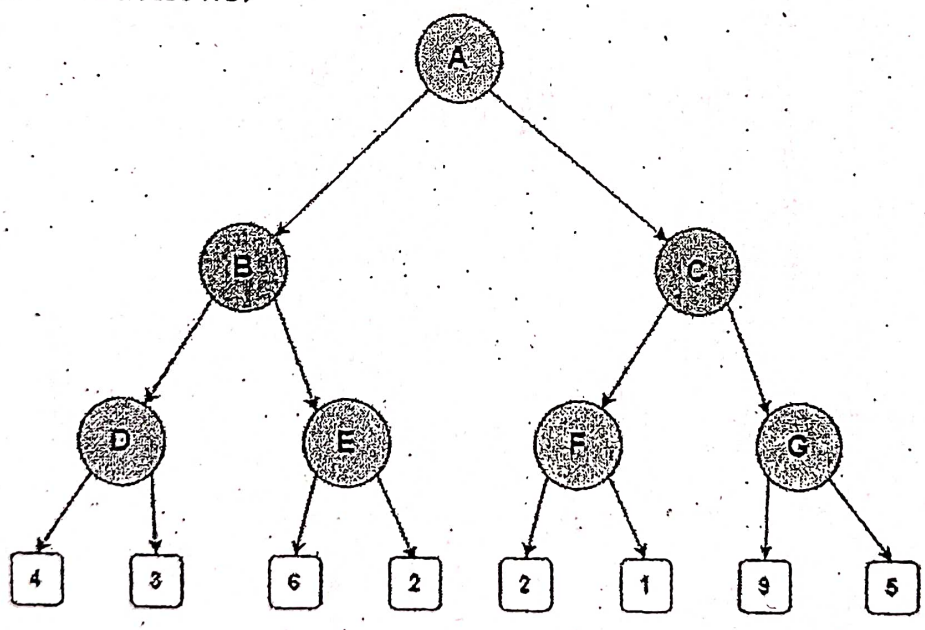
Attempt all the five questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.

Q. No.	Question	Marks	CO
Q1	**Attempt any 2 parts of the following.**		
1a	What is training, testing, and validation set? How are they used to carry out cross-validation? Explain any method in detail.	5	CO1
1b	Differentiate following: 1. Bias and variance 2. K-NN and K-means	5	CO1
1c	The given dataset shows classification data with two classes: triangle and circle.  Which of the two nearest neighbour classifiers 1-NN or 3-NN has the largest Leave-one-out cross-validation error on the dataset?	5	CO1
Q2	**Attempt any 2 parts of the following.**		
2a	Define overfitting in Decision tree. How can we avoid it?	5	CO2
2b	Cluster the following eight points (with (x, y) representing locations) into three clusters: A1 (2, 10), A2 (2, 5), A3 (8, 4), A4 (5, 8), A5 (7, 5), A6 (6, 4), A7 (1, 2), A8 (4, 9). Assume A1, A4 and A7 as the initial cluster centres. Draw the plot and execute at least 2 passes.	5	CO2
2c	Summertime is festival time: Ooty Flower Festival (OFF), Baisakhi (BAISKH), Shimla Summer Festival (SSF) and many more. To help you choose which of the festival you choose to visit, develop a model based on the data below. How would you predict *Like* given the *Festival = OFF, Visited before = N, Duration = 30 days?*	5	CO5

S. No.	Festival	Visited before	Duration (in days)	Like
1	OFF	Y	14	N
2	BAISKH	Y	1	N
3	OFF	N	14	N
4	SSF	N	30	N
5	SSF	N	30	Y
6	OFF	N	14	Y
7	BAISKH	N	1	Y
8	BAISKH	Y	1	Y

Q3	**Attempt any 2 parts of the following.**		
3a	Over the neural network given below, execute a forward and backpropagation pass. The training data is $x_1, x_2 = [3,1]$ and $y_1, y_2 = [1,0]$  	5	CO1
3b	Define perceptron. Can you represent the following Boolean function with a single perceptron with logistic activation? If yes, show the weights. If not, explain why not.	5	CO3
3c	What do you mean by convolution operation? If an input image is converted to a matrix of size 256x256 and a kernel/filter of size 3x3 with a stride of 1 and no padding is used in a CNN. Then, what will be the size of the convoluted matrix?	5	CO3
Q4	**Attempt any 2 parts of the following.**		
4a	i) Differentiate between deductive, abduction and inductive reasoning with appropriate examples. ii) Check whether following are equivalent. $p \rightarrow (q \cup r) \equiv (p \rightarrow q) \cup (p \rightarrow r)$	5	CO4
4b	Consider the following statements (S_i) and check if the arguments (A) are valid or not. In case of invalid argument, give the most valid conclusion which can be derived. i) S_1: If you work hard, you will pass the exam. S_2: You are not working hard. A: You will not pass the exam. ii) S_1: If you work hard and you have talent, you will become musician. S_2: If you become musician, you will be happy. S_3: You are not happy. A: You did not work hard.	5	CO4
4c	Convert the following statements into predicate logic. i) Every car on expressway is fast and dangerous. ii) Every red fast car on expressway is dangerous.	5	CO4

A	B	F(A,B)
1	1	0
0	0	0
1	0	1
0	1	0

	iii) Every student of college has computer or he is friendly with someone who has computer. iv) Some boys in the class are taller than all the girls.		
Q5	**Attempt any 2 parts of the following.**		
5a	i) When would best first search perform worse than a simple breadth first search? ii) Is minimax procedure depth-first or breadth first? Justify your answer.	5	CO5
5b	A 3-foot-tall monkey is in a room where some bananas are suspended from the 8-foot ceiling. He would like to get the bananas. The room contains two stackable, movable, climbable 3-foothigh crates. Give the initial state, goal state, successor function and cost function for getting the bananas.	5	CO5
5c	A game tree is as follows:  Which nodes would not be examined using alpha-beta pruning procedure? Explain the steps.	5	CO5

COURSE CODE: CMCSC20 MID-SEMESTER EXAMINATION- MAR-2022

COURSE TITLE: MACHINE LEARNING **Time:** 90 minutes

Max marks:25

Instructions:

Attempt all questions.

Assume missing data if any, and mention in your answer

		Marks	CO
Q1	The following table must be used for Q1A and Q1B		

X_1	X_2	y_{ACTUAL}	$y_{PREDICTED}$
1	3	2	4
2	2	5	10
3	1	3	4
4	6	6	6
5	7	2	5

Q1A	Calculate R^2 and adjusted $\overline{R^2}$, showing the steps of calculation clearly.	2	CO2
Q1B	Calculate the weight adjustments for w_0, and w_2 with gradient descent, assuming a learning rate of 1. You can add columns to the table to show your calculations clearly.	3	CO2
Q2A	Describe a classification task that you have actually faced, where the target variable is dependent on three or more independent variables. Populate the data for 6 data points giving proper units and suitable values.	2	CO3
Q2B	Consider the following data points {(x,y)}, where y is discrete valued [1,5]. (1,1) (1.5,1) (2,1) (2.5,1) (3,5) (3.5,1) (4,1) (4.5,5) (5,1) (5.5,5) (6,1) (6.5,5) (7,1) (7.5,5) (8,5) (8.5,1) (9,5) (9.5,5) (10,5) Graph the above points and then approximate with a sigmoid. Show the True Positives, True Negatives, False Positives and False Negatives in logistic regression, assuming a threshold of 0.5. What happens when the threshold is reduced to 0.3?	3	CO1
Q3A	Using logistic regression with w_0= 2, w_1=-3, x increases from 4 to 6. By what factor does the log-odds and odds Increase/decrease?	2	CO2
Q3B	From the confusion matrix given below, calculate the F1 measure, and specificity.		CO3

	Predicted No	Predicted Yes
Actual No	15	4
Actual Yes	6	18

	Compare the concepts of Entropy and Gini index for building decision	3	

(85)

Q4A)	tress by splitting nodes. Enumerate your points of comparison.	2
Q4B)	Derive the decision tree for the data (3 features and 2 classes) given below:	3

X1	X2	X3	Decision
A	C	E	Y
A	C	F	Y
B	D	E	N
B	D	F	N

Q5A) Given the following table, classify the new examples using 3-KNN. 2

X1	X2	X3	X4	Decision
1	0	1	1	Y
0	0	0	0	N
1	1	0	0	Y
0	1	1	1	N
1	0	0	0	N
1	1	1	1	?

Q5B) Write a short note on various ensemble methods. 3

Total no. of pages - 2

Roll NO. .....

SECOND SEMESTER – M.TECH.(Bioinformatics)
MID-SEMESTER EXAMINATION, Feb- March, 2022

Course Code - **BTBID08**
Course Title - **Machine Learning**

Time-1.5hours

Max. Marks – 15

Note: - Attempt all questions. Missing data/information (if any) may be suitably assumed & mentioned in the answer.

<table>
<tr><th>Q. No.</th><th>Question</th><th>Marks</th><th>CO</th></tr>
<tr><td>1(a)</td><td>Define Machine learning . Explain with specific example .</td><td>1</td><td>CO1</td></tr>
<tr><td>1(b)</td><td>List and explain perspectives and issues in machine learning.</td><td>2</td><td>CO3</td></tr>
<tr><td>2(a)</td><td>What are the difference between Liner and logistic regression</td><td>1</td><td>CO1</td></tr>
<tr><td>2(b)</td><td>Use K-Means Algorithm to create two clusters for-
<table>
<tr><td>Data point</td><td>x</td><td>y</td></tr>
<tr><td>A</td><td>2</td><td>2</td></tr>
<tr><td>B</td><td>3</td><td>2</td></tr>
<tr><td>C</td><td>1</td><td>1</td></tr>
<tr><td>D</td><td>3</td><td>1</td></tr>
<tr><td>E</td><td>1.5</td><td>0.5</td></tr>
</table></td><td>2</td><td>CO2</td></tr>
<tr><td>3(a)</td><td>Explain Find-S algorithm with an example .</td><td>1.5</td><td>CO1</td></tr>
</table>

(14.5)

3(b)	List down the problem domains in which Decision Trees are most suitable.	1.5	C
4(a)	Describe Hypothesis space search in decision tree learning.	2	C
4(b)	Why we use Gradient descent algorithm.	1	C
5(a)	What are the advantages of random forest over decision tree .	1	C
5(b)	Calculate entropy when total sample size is 20 in which positive samples are nine and rest are negative samples.	2	CO

END SEMESTER EXAMINATION

B. TECH – MATHEMATICS AND COMPUTING, VI SEMESTER

Course Code: CMCSC20
Course Title: Machine Learning

Time: 3 Hours

Max. Marks : 40

Note: -

1. **Attempt all the five questions, any two parts from each question.**
2. **All descriptive answers should be written point-wise clearly.**
3. **Missing data/ Information if any, maybe suitably assumed & mentioned in the answer.**

Que No	Question	Marks	CO
Q1a)	***Answer any two (2) parts from 1a, 1b,1c*** Explain in points, the **significance** of the covariance-variance matrix and **why** standardization is required for PCA? (please note the question - you need not write what these concepts are). Derive the covariance – variance matrix for the following data: X1 X2 4 -200 14 150 8 600	4	CO2
Q1b)	Enumerate any three properties of Principal components. There are 4 features in a data table. If the Eigen values of the principal components are 0.668, 0.141, 0.333, 0.05, justify which components would you like to discard or keep.	4	CO2
Q1c)	Explain clearly how cross validation can be used to ensure that training is not biased towards (i) specific portions of the data (ii) one particular decision class	4	CO2
Q2a)	***Answer any two (2) parts from 2a, 2b, 2c*** Given the following table of immediate rewards R and next state S' for different initial states S_i and actions A_j, derive the Q-table after following trials: (1) S1, then A3, (2) S2 then A4 (3) S3 then A2 (4) S3 then A3 (5) S2 then A4 (6) S4 then A3 Assume Discount factor = 0.8 and Learning Rate = 1. Use a single Q-Table and show updates by marking the trial number	4	CO1
Q2b)	Choose a machine learning paradigm (or a combination) for developing each of the following applications, and justify your	4	CO3

	Actions A (Reward R, Next State S')			
From Initial State S	A1 Stop	A2 Turn Left,	A3 Turn Right	A4 Take U-turn
S1	10, S1	20, S2	40, S3	-10, S1
S2	0, S2	50, S3	-20, S1	30, S1
S3	5, S3	20, S2	20, S1	70, S4
S4	200, S4	35, S2	25, S3	0, S1

	choice point-wise: 1. Identifying the rotten apples from a basket of apples 2. Arranging the logistics for organizing an event 3. Segregating e-story-books on the basis of their genres		
Q2c)	Explain the impact of (i) increasing K, and (ii) decreasing K on the performance of the KNN algorithm. For the Data: {X1,X2,Y} = (2,3,1), (3,7,1), (4,3, 0), (5,0,1), (6,4,0), (7,1,0) Let the second input variable X2, be twice as important as X1. Find: (2.5, 1, ?) with K=3	4	CO1
Q3a)	***Answer any two (2) parts from 3a, 3b, 3c*** Given the following data, use Bayesian classification to predict whether the new candidates will be selected or rejected: Samples of marks for class "Selected" : 80, 60, 73, 90, 95, 60 Samples of marks for class "Rejected" : 70, 55, 60, 50, 70, 82 No of selected students who were PG: 4 out of 6 No of rejected students who were PG: 2 out of 6 No of selected students who were UG: 5 out of 6 No of rejected students who were UG: 1 out of 6 New Candidate: Marks: 85, UG	4	CO1
Q3b)	Give four important points of differences between the approaches in learning with Decision Trees and with Bayesian learning	4	CO3
Q3c)	Given the following table, show in **clear steps**, the evolution of centroids in 2-means clustering through 2 cycles. {X1, X2} = (2,3), (10,7), (6,7), (1,3), (4,5), (9,12)	4	CO1
Q4a)	***Answer any two (2) parts from 4a, 4b, 4c*** Show the output of convolution of the following image and kernel, using same padding with stride =1 horizontally, and valid padding with stride = 2 vertically: Image 1 0 1 1 0 1 Kernel 1 0 1 0 0 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 0 1 0 1	4	CO2
Q4b)	A SVM uses the Polynomial Kernel with a high value of C and a low value of Gamma. Explain what impact these have on the working of the algorithm and its performance.	4	CO1
Q4c)	Differentiate between the primal and dual representations on the optimization in SVM.	4	CO2
Q5a)	***Answer any two (2) parts from 5a, 5b, 5c*** Describe the drop-out method of regularization in CNN (explain pointwise)	4	CO2
Q5b)	Give four possible configurations for RNN, and explain briefly one example application for each configuration.	4	CO3
Q5c)	Explain the vanishing gradient problem in RNN. With figures, explain how any LSTM architecture addresses this problem.	4	CO3

NETAJI SUBHAS UNIVERSITY OF TECHNOLOGY

Mid Semester Examination – February-March 2022
MTech II Semester (EE)

Subject : Machine Learning
Time : 1 Hr 30 Mins

Paper : EEELE14
MM : 15

NOTE : Attempt all questions. Missing data/ information, if any, may be suitably assumed & mentioned in the answers.

Q. No.	Questions	Marks	CO
1	A. Define Machine Learning. B. Explain with the help of a figure, categories and sub-categories of machine learning.	1.5+1.5	1
2	A. What is linear Regression? B. Explain Regression with the help of an example.	2+1	1
3	A. What are the basic data types in machine learning? B. Give an example of each one of them.	2+1	1
4	A. What is Box plot? Explain by an example. B. Define Confusion Matrix, Precision and Recall.	2+1	1
5	A. Explain Gradient Descent Algorithm. B. Explain K-fold cross validation technique.	2+1	2

Total no. of Pages-1

Roll no. .....

SIXTH SEMESTER B. Tech.

Branch: Information Technology

END-SEMESTER EXAMINATION, May- 2022

Course Code- ITITE24

Course Title- Machine Learning and Data Analytics

Time- 3.00 Hours

Max. Marks- 40

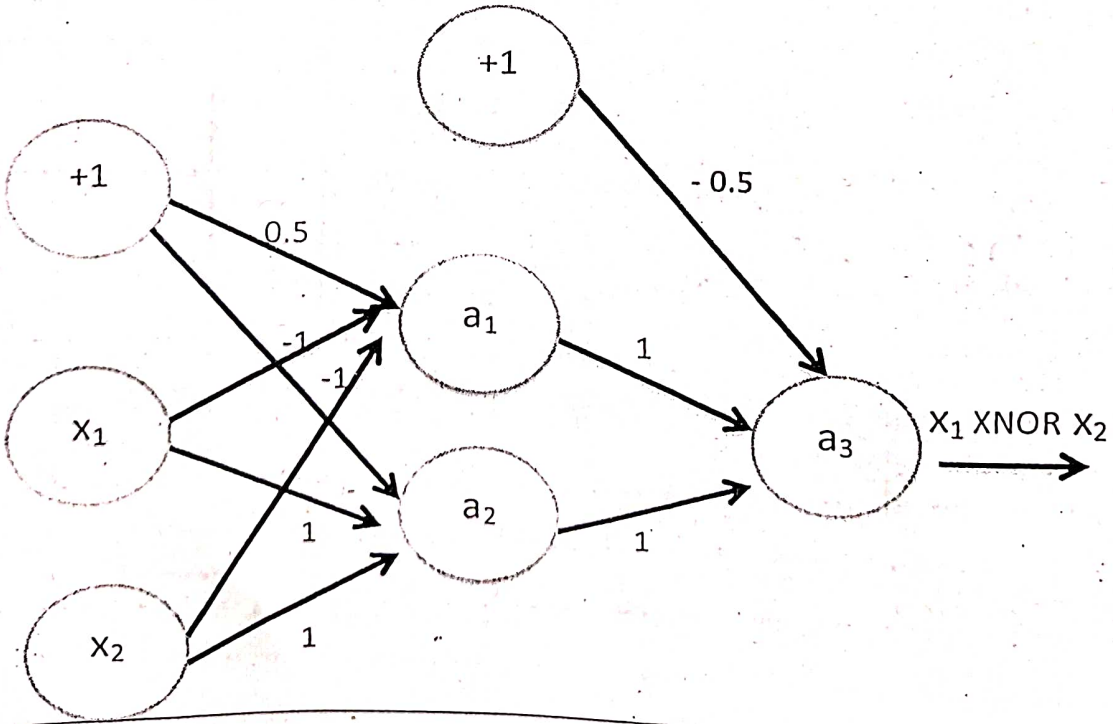
Note: - Attempt All questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.

Q.No.	Question	Marks	COs
Q1	**Attempt any 2 parts of the following**		
1a	Define Machine Learning. Explain with specific example.	4	CO1, CO5
1b	Explain the role of Dataset in Machine Learning. If you want to generate the Dataset what care, you will take in Features selection.	4	CO1, CO2
1c	Explain the difference between lazy learner and eager learner.	4	CO1, CO2
Q2	**Attempt any 2 parts of the following**		
2a	Explain various ways by which the performance of the Model can be measured.	4	CO2, CO3
2b	Generate Confusion Matrix Using Bayes Classifier for Multiple features, Multiple values and Multiple classes for any application.	4	CO3, CO5
2c	How sentiment Analysis is performed. Explain with example	4	CO1, CO5
Q3	**Attempt any 2 parts of the following**		
3a	Explain the concept of Principal Component Analysis (PCA). Write the steps to compute. How training and testing is performed.	4	CO3, CO5
3b	Explain the concept of Linear Discriminant Analysis. How it helps in Machine Learning. Write the steps to compute.	4	CO4, CO5
3c	How Dimensionality Reduction and classification is performed in Machine Learning. Explain with example	4	CO3, CO5
Q4	**Attempt any 2 parts of the following**		
4a	With example illustrate the working of Gradient Descent for Linear regression	4	CO2, CO5
4b	Explain the concept of Multivariate Linear regression	4	CO2, CO4
4c	Explain the difference between Gradient Descent and Stochastic Gradient Descent.	4	CO4, CO5
Q5	**Attempt any 2 parts of the following**		
5a	Explain how normalization is performed in Machine Learning. Explain with example	4	CO2, CO5
5b	Explain how unsupervised learning is performed. Explain the k-Means Clustering algorithm.	4	CO4, CO5
5c	Write short Notes on a) Support Vector Machine b) Logistic Regression	2+2	CO3, CO4, CO5

4a		4	CO3

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Create Frequency table and Likelihood table for above example

4b	What are the benefits of laplace smoothing. Explain with formula.	4	CO2
4c	Compare generative and discriminative models by taking research examples.	4	CO5
Q 5	*Attempt ANY 2 of the following parts.*		
5a	What is problem with gradient descent . How to resolve that problem.	4	CO3
5b	Explain multi-layer neural network by taking any research application. Explain why back propagation algorithm is required.	4	CO5
5c	For a XNOR function as given in the figure below, activation function of each node is given by: $$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & otherwise \end{cases}$$ Consider X_1 and $X_2 = 0$, what will be the output for the above neural network? 	4	CO4

Total no. of Pages 4 Roll no ....

(140)

Vth Sem.-B.Tech. END-SEMESTER EXAMINATION, NOV.-DEC.,2022

Course Code- COCSC17 Course Title- Machine Learning

Time- 03 Hours Max. Marks- 40

Note: - Attempt all the five questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.

Q.No.	Attempt any 2 parts of the questions as BELOW	Marks	CO
1(a)	The following dataset contains loan information. It can be used to predict whether a borrower will default (Yes) or not (No). Use the Naïve Bayes method to determine whether a loan: X=(Home Owner = No, Marital Status=Married, Income=High) should be classified as a Defaulter or not. Show the steps of deriving the probability P(Yes|X), and P(No|X) clearly.	4	CO5

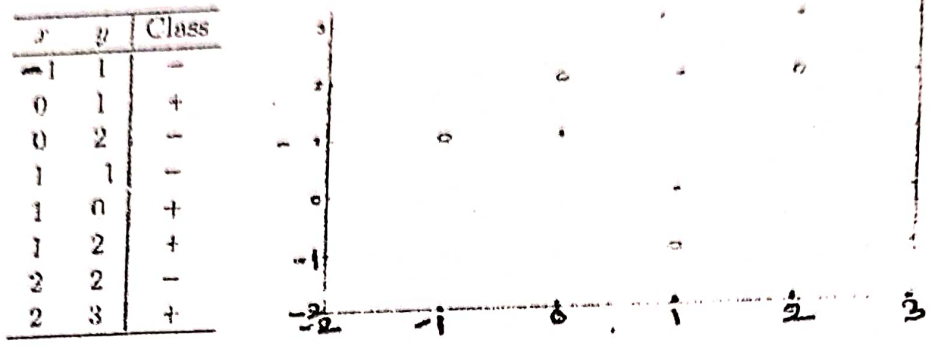
Tid	Home Owner	Marital Status	Annual Income	Defaulted Borrower
1	Yes	Single	High	No
2	No	Married	High	No
3	No	Single	Low	No
4	Yes	Married	High	No
5	No	Divorced	Low	Yes
6	No	Married	Low	No
7	Yes	Divorced	High	No
8	No	Single	Low	Yes
9	No	Married	Low	No
10	No	Single	Low	Yes

1(b)	Consider the training examples shown in the following table for binary classification problem on potential buyers of a computer system.				4	CO5

age	income	student	credit_rating	buys_computer
youth	high	no	fair	no
youth	high	no	excellent	no
middle_aged	high	no	fair	yes
senior	medium	no	fair	yes
senior	low	yes	fair	yes
senior	low	yes	excellent	no
middle_aged	low	yes	excellent	yes
youth	medium	no	fair	no
youth	low	yes	fair	yes
senior	medium	yes	fair	yes
youth	medium	yes	excellent	yes
middle_aged	medium	no	excellent	yes
middle_aged	high	yes	fair	yes
senior	medium	no	excellent	no

Derive the decision tree levels using the concept of either Entropy or Gini Index.

1(c)	Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 are actually, sick. For the sick people, a test was positive for 620 and negative for 380. For the healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the precision, recall, F1-measure for the data.	4	CO5
2(a)	Which ensemble method will you use for the following use cases? Justify and explain how the final output is derived in each case. 1. Obtaining fast classification using a combination of Decision trees and Naïve Bayes 2. Reducing both variance and bias error	4	CO4
2(b)	Explain the need for: (i) Leaky ReLU activation function in intermediate stages of CNN, instead of sigmoid. (ii) Softmax activation function in the output stage of CNN	4	CO4
2(c)	What effect can the following anomalies in training data have on the final ML model? (i) Training data does not have enough variety of feature combinations (ii) Training data is heavily biased towards a particular decision class (iii) Training data is too less in quantity (iv) Training data is noisy having untreated, erroneous data	4	CO4
3(a)	In each of the following use cases, choose which method you will use among Decision Tree and Naïve bayes. Justify. (i) The training data is less and has many continuous valued features (ii) The training data has many features which are partially interdependent (iii) You are not very sure which features are more important (iv) You want to understand which features are relevant for predicting a given output class.	4	CO3
3(b)	Given two clusters with following values of a single feature for all data points, calculate the sum of intra-cluster distances and the inter-cluster distance: Cluster 1: 2,4, 9,18 Cluster 2: 20,23,40 Would you prefer that point 18 went into the Cluster 2? Justify.	4	CO3
3(c)	The predicted and actual outputs of a training set input to Adaboost are as follows: Predicted: 1, 1, 0, 1 — Actual: 1, 1, 1, 1 Calculate the initial weights and the altered weights after the first iteration of the four data points.	4	CO3
4(a)	Consider the following data, given in the Table and scatter plot, where x and y are the 2 input variables and Class is the dependent variable.	4	CO2

	 Suppose, you want to predict the class of new data point x=1 and y=1 using Euclidean distance using 3-NN. To which class this data point belong?			
4(b)	Explain how the vanishing gradient problem can occur in Feed Forward Multi-Layer Neural Network (FFMLNN) with Backpropagation.	4	CO2	
4(c)	Given a training set of 1000 datapoints with 9 features. Explain clearly how the input population would vary for 4 weak experts using: (i) Random forest (ii) Adaboost	4	CO2	
5(a)	Explain the function of each of the given stages of a CNN and calculate the number of neurons generated in each stage: (i) Convolution stage 1: Image size is 127 X 127 X 3, There are 12 filters each of size, 7 X 7, Valid padding is used, stride is 2. (ii) Maxpool stage: The above stage's output is input to a maxpool layer with 12 filters, filter size 5 X 5, stride 2.	4	CO1	
5(b)	Enumerate any three properties of Principal components in PCA. There are 4 features in a data table. If the Eigen values of the 4 principal components are 0.70, 0.65, 0.33, 0.05, justify which components would you like to discard or keep.	4	CO1	
5(c)	What is the Inductive Learning Hypothesis. Explain the following Inductive biases, giving an example for each: (i) Restriction bias (ii) Preference bias	4	CO1	

x	y	Class
−1	1	−
0	1	+
0	2	−
1	1	−
1	0	+
1	2	+
2	2	−
2	3	+

Roll No:.....

V[th] Sem. B. Tech –Course work

END-SEMESTER EXAMINATION, Nov, 2022

urse Code: CACSC17

urse Title: Machine Learning

ne: 3 Hours

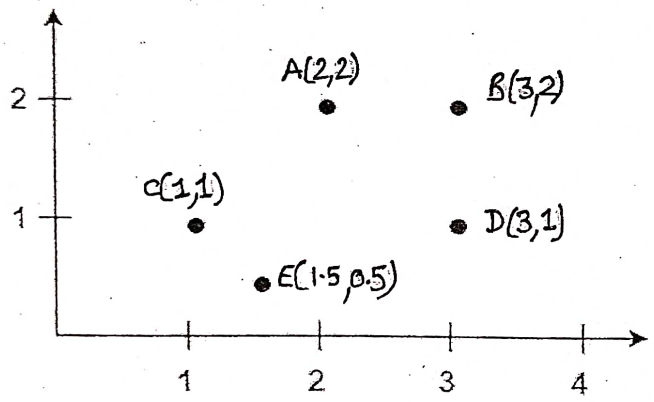
Max. Marks: 40

te: **Attempt all the five questions. Missing data / information (if any), may be suitably assumed and ntioned in the answer.**

Question No.	Attempt any 2 parts of the following	Marks	CO
Q 1			
1a	What is learning system? Explain various categories of learning in machine learning.	4	CO1
1b	Differentiate general and specific hypothesis with suitable example. Write the candidate elimination algorithm.	4	CO1
1c	Write the steps in find S algorithm. Explain unanswered question by find S algorithm.	4	CO1
Q 2	**Attempt any 2 parts of the following**		
2a	Explain ensemble learning. How does boosting classifier works on dataset.	4	CO2
2b	What is cross validation? Explain various types of cross validations.	4	CO2
2c	What are various learning curves used for comparing learning algorithms. Explain at least two of them in detail.	4	CO2
Q 3	**Attempt any 2 parts of the following**		
3a	What is probably approximately correct (PAC) learning. What will be minimum number of samples for a concept to be learned which can be described by conjunctions of up to 10 Boolean literals with 95% probability that a hypothesis be learned with error less than 10%.	4	CO3
3b	Explain VC dimension with example? Why is VC dimension used?	4	CO3
3c	Explain sample complexity for finite hypothesis spaces with suitable example.	4	CO4
Q 4	**Attempt any 2 parts of the following**		
4a	Explain First Order Inductive Learner. Explain its working with performance evaluation measure.	4	CO3
4b	What is perceptron model? Explain the working of perceptron.	4	CO4
4c	Explain recurrent neural network. What are various advantages and disadvantages of recurrent neural network?	4	CO5
Q 5	**Attempt any 2 parts of the following**		
5a	Explain support vector machine. Why is SVM kernel important? Write advantages of SVM.	4	CO5
5b	Explain generative and discriminative models of training. Differentiate generative and discriminative models.	4	CO4
5c	Explain logistic regression. What are various types of logistic regression?	4	CO5

Course Code- COCSE01 Course Title- Machine Learning
Time- 03 Hours Max. Marks- 40

Note: - Attempt all the five questions. Missing data/ information if any, maybe suitably assumed & mentioned in the answer.

Q.N.	Questions	Marks	CO
Q1	*Attempt ANY 2 of the following parts.*		
1(a)	Define hyper parameters? Explain their significance by taking a suitable application.	4	CO1 CO5
1(b)	Write the Implementation code of Principal component analysis for dimensionality reduction using python.	4	CO3
1(c)	Explain curse of dimensionality by taking a suitable example?	4	CO1
Q2	*Attempt ANY 2 of the following parts.*		
2(a)	Why we use clustering? Justify by using K-Means Algorithm on the following data 	4	CO4
2(b)	Define the concept learning . Explain the task of concept learning .	4	CO1
2(c)	Explain Find - S algorithm with an example .	4	CO2
Q3	*Attempt ANY 2 of the following parts.*		
3(a)	Describe Hypothesis space search in decision tree learning.	4	CO2
3(b)	Why we use gradient decent? Explain the problem with the gradient descent? How it can be tackled.	4	CO3

4(b)	Explain the process and
4(c)	Define Entropy in machi twenty in which positive
Q5	*Attempt ANY 2 of the fol*
5(a)	What do you mean by e applications?
5(b)	Compare the following 1. Recurrent neur 2. SVM and Decisi
5(c)	Develop a Supervised Write its python code a

3(c)	Describe the characteristics of back propagation algorithm .	4	CO2
Q4	*Attempt ANY 2 of the following parts.*		
4(a)	Illustrate the working of multi layer perceptron by taking an application of computer vision	4	CO1 CO2
4(b)	Explain the process and advantage of DBSCN with a suitable diagram.	4	CO4
4(c)	Define Entropy in machine learning? Find entropy when total sample size is twenty in which positive samples are nine and rest are negative samples .	4	CO1
Q5	*Attempt ANY 2 of the following parts.*		
5(a)	What do you mean by explainable AI? Explain its potential challenges and applications?	4	CO1, CO5
5(b)	Compare the following by taking suitable applications 1. Recurrent neural networks and convolutional neural networks 2. SVM and Decision Tree	4	CO2 CO3
5(c)	Develop a Supervised Machine Learning Model for the Prediction of COVID-19. Write its python code also.	4	CO2

Total no. of Page:2 Roll no:

B.E (COE)

MID SEM Examination: September,2019

Course Code: CED17

Course Title: Machine Learning

Time:1.30 hours Max Marks:25

Note: All questions are compulsory

Q1. Given below is a Decision Table showing whether majority of people surveyed, would like to settle in a given city, given its parameters.

(i) Suggest a method to fill the missing value and fill it accordingly

(ii) Discretize the "Index of industrial development" column so that there are two (binary) values with maximum information gain

(iii) Make the decision tree with ID3.

(2+2+4=8 marks)

City	Index of industrial development (0..1)	Cost of living	Political Stability	Pollution level	Settle?
Tokyo	0.94	High	High	Low	Y
Delhi	0.8	Low	High	High	Y
Vancouver	0.6	Moderate	High	Low	N
New York	0.9	High	High	High	Y
Hongkong	0.7	Moderate	Low	High	N
Poland	0.55	Low	Low	Low	Y
Vienna	0.42	Moderate	Low	Low	Y
Prague	0.32	Low	High	Low	Y
Hyderabad	0.78	Moderate	High	High	Y

Q2. In the following, it is desired to describe whether a person is ill. We use a representation based on conjunctive constraints (three per subject) to describe individual person. These constraints are "running nose", "coughing", and "reddened skin", each of which can take the value true ('+') or false ('–'). We say that somebody is ill, if he is coughing and has a running nose. Each single symptom individually does not mean that the person is ill.

(i) Specify the space of hypotheses that is being managed by the version space approach. To do so, arrange all hypotheses in a graph structure using the more-specific-than relation.

(ii) Apply the candidate elimination (CE) algorithm to the sequence of training examples specified in the following Table and name the contents of the sets S and G after each step.

Training	running nose	coughing	reddened skin	Classification
d_1	+	+	+	positive (ill)
d_2	+	+	–	positive (ill)
d_3	+	–	+	negative (healthy)
d_4	–	+	+	negative (healthy)
d_5	–	–	+	negative (healthy)
d_6	–	–	–	negative (healthy)

Table1: List of training instances for the medical diagnosis task

(3+4= 7 marks)

Q3. Attempt any five parts

a. On a particular data set, we use the ensemble method approach for building a predictor and achieve state of the art performance. Is it possible for some of the individual models in this ensemble to have poor performance as measured on the training data? Explain why?

i. Yes

ii. No

b. With regards to bagging and boosting, which of the option(s) among

the following are true? And why?

i. The different learners in bagging can be trained in parallel
ii. The different learners in boosting can be trained in parallel
iii. Individual classifiers in bagging are trained with all data points in the training set
iv. Individual classifiers in boosting are trained with all data points in the training set

Which of the following tasks can be best solved using Clustering? And why?

i. Predicting the amount of rainfall based on various cues
ii. Detecting fraudulent credit card transactions
iii. Training a robot to solve a maze

I am the marketing consultant of a leading e-commerce website. I have been given a task of making a system that recommends products to users based on their activity on Facebook. I realize that user-interests could be highly variable. Hence I decide to

- First, cluster the users into communities of like-minded people and
- Second, train separate models for each community to predict which product category (e.g. electronic gadgets, cosmetics, etc) would be the most relevant to that community.

The first task is a/an _____ learning problem while the second is a/an _____ problem.

Choose from the options and justify:

i. Supervised and unsupervised
ii. Unsupervised and supervised
iii. Supervised and supervised
iv. Unsupervised and unsupervised

e. Differentiate between Reduced error pruning and Rule post pruning.
f. Differentiate between Inductive learning bias in Concept learning and Decision trees.
g. Explain the pros and cons of Occam's razor.

(2+2+2+2+2=10 Marks)

_____GOOD LUCK_____

www.vtucodespace.in