Please check that this question paper contains__09____ questions and ___02___ printed pages within first ten

[Total No. of Questions: 09] [Total No. of Pages: 02.]

Uni. Roll No. …….

Program: **B.Tech**
Semester: **6**
Name of Subject: **Introduction to Machine Learning**
Subject Code: **PCIT-114**
Paper ID: **17206**

Time Allowed: 02 Hours **Max. Marks: 60**

NOTE:

1) Each question is of 10 marks.
2) Attempt any six questions out of nine
3) Any missing data may be assumed appropriately

07-07-21(E)

Q1. Explain the work flow of a machine learning project with the help of a diagram. Why machine learning is the future?

Q2. Write down the truth table and draw the decision tree to represent the following Boolean functions:
(i) A XOR B
(ii) Parity function of four Boolean attributes, A, B, C, and D
**Hint: (Parity =(T) if number of true inputs and false inputs are even, Parity = (F) if number of true or false inputs are odd)*
Is it possible to simplify these trees**?**

Q3. Consider the training examples shown in the following Table for a binary classification problem:

Instance	A1	A2	A3	Target Class
1	T	T	1	+
2	T	T	6	+
3	T	F	5	-
4	F	F	4	+
5	F	T	7	-
6	F	T	3	-
7	F	F	8	-
8	T	F	7	+
9	F	T	5	-

(a) What is the entropy of this collection of training examples with respect to the Positive (+) class?
(b) What are the information gains of A1 and A2 relative to training examples?
(c) What is the best split (among A1 and A2) according to the information gain?
(d) What is the best split (between A1 and A2) according to classification error rate?

(e) What is the best split (between A1 and A2) according to the Gini index?

Q4. Differentiate between simple linear regression and multiple linear regression. Find out the predicted values of 'y' in the following data using simple linear regression. Draw the scatter plot for the given data and write down the regression equation and RMSE value for the obtained model.

x	1	2	3	4	5	6	7	8	9	10	11
y	5	7	4	8	9	6	5	8	10	12	14

Q5. What is the general principle of an ensemble machine learning method? How bagging is different from the boosting? Write down the pseudo code of Random Forest machine learning method clearly mentioning the difference of random forest as classifier and random forest as regressor.

Q6. The following table gives data set about stolen vehicles. Using Naive Bayes classifier, Classify the new data {Color: Red, Type: SUV, Origin:Domestic**}**

Color	**Type**	**Origin**	**Stolen**
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

Q7. Why clustering algorithms belongs to the category of unsupervised learning. Describe K-Means, K-Modes and Gaussian Mixture Model in terms of the following criteria:

- Shape of clusters that can be determined
- Input parameters that must be specified
- Limitations

Q8. Draw a block diagram to represent a Fuzzy System. Explain the process of fuzzification and defuzzification through proper examples.

Q9. Draw schematic diagrams and write examples for the following fuzzy membership functions:

(a) trimf (b) trapmf (c) gaussmf (d) gbellmf

****

For office use only:

(To be typed by paper setter, and submitted along with the question paper)

Separate Page to be provided by Paper Setter without revealing paper setter id

Program:.B.Tech..... Semester 6.....

Name of Subject: Introduction to Machine Learning

Subject Code: PCIT-114

Paper ID: 17206.

Detail of CO Mapping and Revised Bloom's Taxonomy (RBT) Level of the questions									
Question No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
CO number	2	2	1	3	1	5	4	6	6
RBT Level*	L2	L3	L4	L5	L3	L5	L4	L4	L2

Description of RBT Levels

RBT Level Number	L1	L2	L3	L4	L5	L6
RBT Level Name	Remembering	Understanding	Applying	Analyzing	Evaluating	Creating

* Since examination is conducted in open book pattern, so there should be no question with RBT level L1 in the question paper.