Please check that this question paper contains 09 questions and 02 printed pages within first ten minutes.

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

Uni. Roll No. 1705398

Program: B.Tech. (Batch 2018 onward)
Semester: 6th
Name of Subject: Introduction to Machine Learning
Subject Code: PCIT-114
Paper ID: 17206

Time Allowed: 03 Hours **Max. Marks: 60**

NOTE:

1) Parts A and B are compulsory
2) Part-C has Two Questions Q8 and Q9. Both are compulsory, but with internal choice
3) Any missing data may be assumed appropriately

Part – A

[Marks: 02 each]

Q1.

a) How machine learning is different from traditional programming?

b) What is Overfitting, and how can you avoid it?

c) How do you handle missing or corrupted data in a dataset?

d) What are unsupervised Machine learning techniques?

e) Compare classification and regression.

f) What is Principal Component Analysis?

Part – B

[Marks: 04 each]

Q2. What is 'training Set' and 'test Set' in a Machine Learning Model? How Much Data Will You Allocate for Your Training, Validation, and Test Sets? Justify.

Q3. Explain the Confusion Matrix with Respect to Machine Learning Algorithms.

Q4. Considering a long list of Machine Learning algorithms, given a data Set, How do you decide which one to use? Explain.

Q5. With proper example elaborate Hierarchical Aglomerative clustering in detail.

Q6. Explain the intuition behind Logistic Regression in detail.

Q7. A doctor knows that Cold causes fever 50% of the time – Prior probability of any patient having cold is 1/50,000 – Prior probability of any patient having fever is 1/20. If a patient has fever, what's the probability he/she has cold?

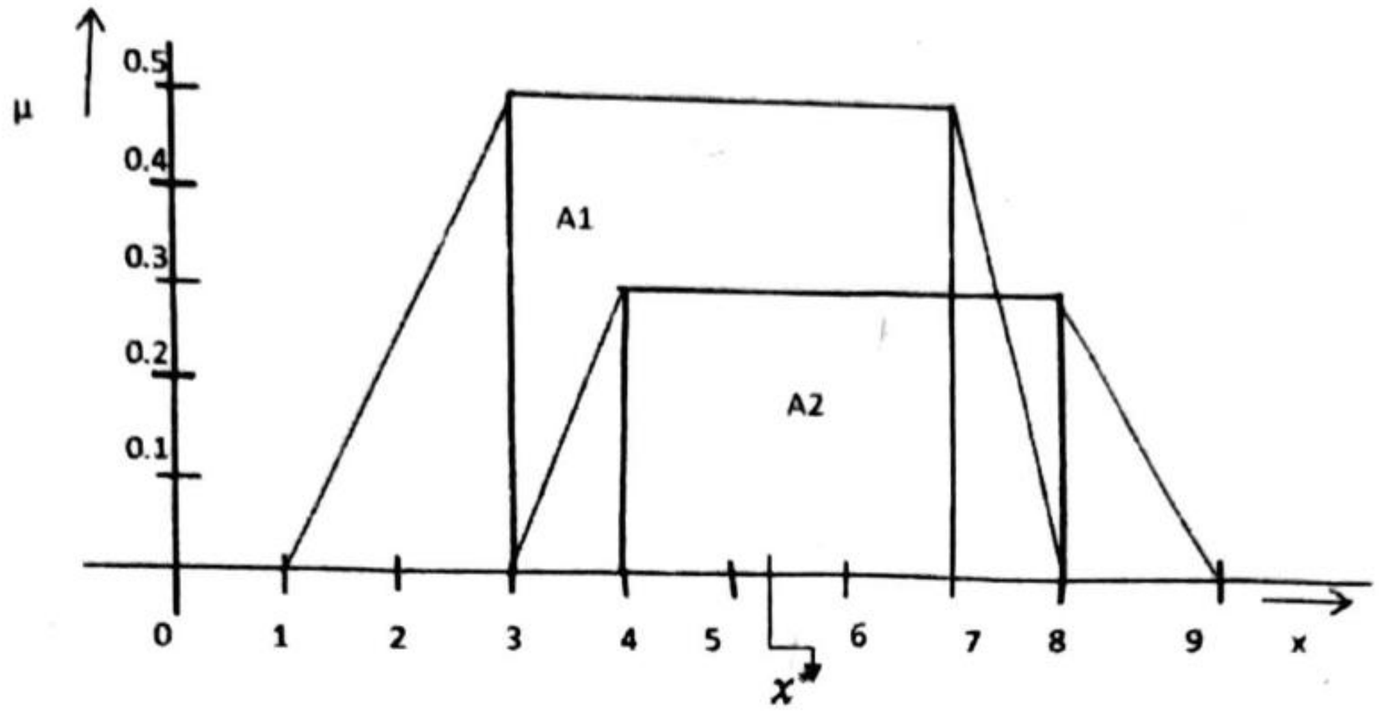
Part – C

[Marks: 12 each]

Q8. Compare Artificial Intelligence, Machine Learning and Deep Learning in detail with proper example.

OR

Calculate the deffuzzified value using Center of Sums (COS) Method.



Q9. Elaborate the steps involved in K-means clustering and K- Mode Clustering. Compare both the clustering techniques with examples in detail.

OR

Draw a decision tree for the following set of training examples. Do we require feature scaling for decision trees? Is it possible to have more than one decision tree for same training sample? Explain.

Day	Weather	Temperature	Humidity	Wind	Play?
1	Sunny	Hot	High	Weak	No
2	Cloudy	Hot	High	Weak	Yes
3	Sunny	Mild	Normal	Strong	Yes
4	Cloudy	Mild	High	Strong	Yes
5	Rainy	Mild	High	Strong	No
6	Rainy	Cool	Normal	Strong	No
7	Rainy	Mild	High	Weak	Yes
8	Sunny	Hot	High	Strong	No
9	Cloudy	Hot	Normal	Weak	Yes
10	Rainy	Mild	High	Strong	No

****