# REVA UNIVERSITY
### Bengaluru, India

## Seventh Semester B. Tech Semester End Examination December 2018
## Course Title (Subject): Machine Learning and Applications
## Course Code: BTCS15F7100 (Qp code – 32721)

Time   3 Hours

Max. Marks: 100
Marks

**Note:** *Answer ONE FULL question from each Unit*

### UNIT – I

1.  a) Identify the machine learning algorithms which can be implemented for the below applications and explain briefly  **5**

    i. Identification of other books purchased along with Tom Mitchell's Machine Learning.

    ii. Help the management to segregate the students according to the geographical location.

    b) Consider designing a program to learn to play checkers. Identify the target function and choose the representation of the target function and explain how to adjust the weights.  **12**

    c) Machine learning involves learning from specific examples and generalize them. FIND-S algorithm is one such algorithm used. Explain the FIND-S algorithm and apply for the given sample of dataset.  **8**

| Origin | Manufacturer | Color | Decade | Type | Example Type |
|--------|-------------|-------|--------|------|-------------|
| China | Lenovo | Gold | 2008 | Smart Phone | Positive |
| China | Oppo | Gold | 2008 | Smart Phone | Positive |
| Finland | Nokia | Red | 2008 | Basic Phone | Negative |
| China | One plus | Red | 2008 | Smart Phone | Positive |
| China | Vivo | White | 2018 | Smart Phone | Positive |

### OR

2.  a) Machine learning is being looked as the driver of the next industrial revolution. According to Tom Mitchell, define machine learning and explain the terminologies of definition with any two suitable examples.  **5**

    b) Mr. Ram wants to implement Concept learning to predict sport can happen or not. So help him by defining the Concept learning and also describe the notations and General-to-specific ordering of Hypothesis involved in concept learning task by considering EnjoySport  **10**

| Sky | Temp | Humid | Wind | Water | Forecast | Enjoy Sport |
|-----|------|-------|------|-------|----------|-------------|
| Sunny | Warm | Normal | Strong | Warm | Same | Yes |
| Sunny | Warm | High | Strong | Warm | Same | Yes |
| Rainy | Cold | High | Strong | Warm | Change | No |
| Sunny | Warm | High | Strong | Cool | Change | Yes |

c) Machine learning is spreading its tentacles in almost all the fields. List out the fields in which machine learning has been applied. — 5

d) For a developer to develop a machine learning system, illustrate the basic design issues and approaches. — 5

## UNIT – II

3. a) Clinical data collected to diagnose a disease has a lot of redundant and irrelevant attributes and might also contain noise. So, there is a need to reduce the number of features or attributes. Discuss the two types of dimensionality reduction with mathematical steps and compare them. — 10

b) Enumerate the classification algorithm and list out four applications with proper explanation. — 7

c) A teacher wants to find out that how exam score of students are correlated with their number of hours studied for the exam. So help the teacher how exam score is associated with number of hours studied with any of the correlation coefficient method. — 8

| Hours Studied | Exam Score |
|---|---|
| 2 | 58 |
| 4 | 32 |
| 5 | 63 |
| 7 | 87 |
| 3 | 67 |
| 1 | 45 |
| 6 | 68 |

### OR

4. a) The data collected from Mangalyaan has a lot of attributes, and all of these are not necessary as some data will be redundant. Also the attributes are not necessary as it becomes computationally expensive. Hence the number of attributes are to be reduced or extracted. Outline the Linear Discriminant Analysis which is used for feature extraction and derive the optimization function of LDA method. — 10

b) Explain two different methods of Feature selection with formulas and compare them. — 7

c) Both Factor Analysis and Principal Component Analysis are used for feature extraction. Outline how objective of Factor Analysis is different from PCA and mention the mathematical steps involved to find out the new dimension of the features. — 8

## UNIT – III

5. a) Given the daily expenditures of food (X1) and clothing (x2) of five persons. Classify the persons into two clusters based on the attributes. And outline the algorithm which is used. — 7

| Person | Food (X1) | Clothing(X2) |
|---|---|---|
| A | 2 | 4 |
| B | 8 | 2 |
| C | 9 | 3 |
| D | 1 | 5 |
| E | 8.5 | 1 |

b) Astronomers are using decision tree to classify the distant objects in the sky based on the size, color, magnitude and temperature. The objects need to be labelled as star, galaxy and nebula etc. Explain the different terminologies, mathematical equations and stopping criteria involved in constructing a decision tree using ID3 algorithm.　　10

c) Expectation-Maximization (EM) algorithm is used in maximum likelihood estimation where the problem involves two sets of random variables of which one, X, is observable and the other, Z, is hidden. Illustrate with an example Expectation maximization with hidden values.　　8

**OR**

6. a) A company produces tissue papers used in Biological sciences companies. A data collected through survey from clients. Expert in that area said quality mainly depends on Acid durability, Strength. Identify and explain the algorithm which can help us to classify whether the new tissue paper is good or bad from the given sample dataset and also predict the class for the tissue paper which is having acid durability as 3 and strength as　　7

| Type | Acid durability | Strength | Class |
|------|-----------------|----------|-------|
| Type 1 | 7 | 7 | Bad |
| Type 2 | 7 | 4 | Bad |
| Type 3 | 3 | 4 | Good |
| Type 4 | 1 | 4 | Good |

b) Clustering is one of the commonly used Preprocessing step in machine learning applications. Outline about Hierarchical clustering and different methods in it with equations and differentiate between them with a neat diagram.　　8

c) Construct a decision tree considering S1 to S6 instances, to predict whether viewer will read or skip the content of the link sent to him for instances S7 & S8.　　10

| Sample | Author | Thread | Length | Location | Action |
|--------|--------|--------|--------|----------|--------|
| S1 | Known | New | Long | Home | Skips |
| S2 | Unknown | New | Short | Work | Reads |
| S3 | Unknown | Old | Long | Work | Skips |
| S4 | Known | Old | Long | Home | Skips |
| S5 | Known | New | short | Home | Reads |
| S6 | Known | old | Long | work | Skips |
| **New Sample** | | | | | |
| S7 | Known | New | Short | Work | ??? |
| S8 | Unknown | New | Long | Work | ??? |

**UNIT – IV**

7. a) Neural networks are inspired by the working of neurons. In artificial neural network, instead of neurons perceptron's are used. Describe perceptron with a neat diagram. Identify the Boolean function which cannot be implemented by single perceptron and Enumerate the reason.　　7

b) Outline the importance of Back Propagation Algorithm. Write the steps involved in Back Propagation Algorithm.　　10

c) Learning can be categorized into supervised, unsupervised and reinforced learning. With suitable example define the significance of reinforcement learning. Distinguish reinforcement learning problem from other function approximation tasks.  8

**OR**

8. a) Illustrate the operation of Q learning algorithm with an example.  10
   b) Write about the characteristics for which Neural Network is mostly used.  5
   c) NAND and NOR are the two universal gates using which any digital circuit can be realized. Implement the working of NAND gate using neural networks.  3
   d) Outline how an agent interacts with its environment in reinforcement learning.  7

***