MANIPAL INSTITUTE OF TECHNOLOGY
(A constituent unit of MAHE, Manipal 576104)

I SEMESTER M.TECH. DEGREE MEDICAL INFORMATICS
END SEMESTER EXAMINATIONS JAN 2023
SUBJECT: MACHINE LEARNING (BME 5172)
(REVISED CREDIT SYSTEM)
Monday, 27 January 2023, 9AM to 12Noon

TIME: 3 HOURS   **MAX. MARKS: 50**

Instructions to Candidates:
1. **Answer FIVE full questions.**
2. **Draw labeled diagram wherever necessary**

1A. How linear regression is different from a logic regression? Discuss performance of linear regression model using RMSE. (4)

1B. Consider a feature "*x*" is normally distributed for *class-A* and *class-B*. Their prior probabilities for class-A and class-B are respectively 0.6 and 0.4. If the mean and standard deviation details of the class-A are: 50 and 2. Similarly class-B has mean and standard deviation of 70 and 2. Find P(A| x=65) value, and interpret the result. (3)

1C. State the Naive Bayesian Theorem for a *K*-class problem with *d-features*. For a given feature values as described in Table1C, calculate the posterior value *P(A|x=0, y=1)* using Naive Bayesian Theorem. (3)

Table 1C

CLASS Name	Feature x	Feature y
A	0	1
B	1	0
A	0	1
B	1	1
A	1	1

2A. What are support vectors. Explain support vector machine approach for classification. (3)

2B. Describe the design cycle approach for dementia classification with the two-class approach. Consider positive class as "**Dementia**", and Negative class as" **Normal**". (3)

2C. In case of an Arrhythmia classification problem, consider heart rate (x) and BMI (y) as the two features. The features are extracted from 5 subjects and they are: (60, 24), (71,29), (72,28), (87,29), (100,32). Use K-means algorithm to generate three clusters choosing first, third and 5th feature set as seed points . (4)

3A Write important steps of decision tree algorithm and find information gain when splitting based on **Fever** at root node. The classes are **Yes (infected)** and **No (no infection)**. (4)

Sl. No.	Fever	Cough	Breathing issue	Infected
1	No	Yes	no	no
2	No	yes	Yes	Yes
3	Yes	No	Yes	Yes
4	Yes	Yes	Yes	Yes
5	Yes	No	Yes	Yes
6	No	Yes	No	No
7	No	No	No	No

3B. Justify the following statement with an example: "*artificial neuron can be realized as processing node*" . (3)

3C. Compare the supervised learning with an unsupervised learning. (3)

4A. Design a perceptron network to classify the two-dimensional input patterns "**I**" and "**T**" . The symbol "*" indicates the data representation to be " +1 " and " • " indicate data to be " -1 ". Consider a target of "+1 " for pattern **I** and "-1" for pattern **T**. (4)

```
    *   *   *                 *   *   *
I = -   *   -           T =   -   *   -
    *   *   *                 -   *   -
```

4B. Identify the important steps that help in classification using k- nearest neighbor algorithm. (3)

4C Design a workflow for improving the pathology with digitization process and list two advantages. (3)

5A. Identify the major rules used for updating weights in the multilayer Back Propagation Neural Network (BPNN) and explain them. (4)

5B Given the following clusters compute the **inter cluster distance** using complete linkage distance: *Cluster₁*= {(30,54), (35, 57)}; *Cluster₂* = {(77,12)} (3)

5B. Construct a confusion matrix for the given two class problem having 550 samples with the following details: True positive =257, True positive=249, False positive=19, False negative= 25. Calculate the sensitivity and discuss its significance. (3)

*****