

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE- SEMESTER-VII (NEW) EXAMINATION – WINTER 2024****Subject Code:3171114****Date:30-11-2024****Subject Name: Introduction of Machine learning****Time:10:30 AM TO 01:00 PM****Total Marks:70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

**MARKS**

- Q.1** (a) Write three differences between Supervised and Unsupervised Learning. **03**  
 (b) Consider the following three class confusion matrix: **04**

	<b>PREDICTED</b>		
<b>ACTUAL</b>	<b>15</b>	2	3
	7	<b>15</b>	8
	2	3	<b>45</b>

Calculate precision and recall per class. Also, calculate the weighted average precision and recall for the classifier.

- (c) Calculate the entropy of the dataset for each attribute (altitude, wind, temperature, and humidity) and determine which attribute is the best choice for the root node of the decision tree. **07**

<b>Altitude</b>	<b>Wind</b>	<b>Temperature</b>	<b>Humidity</b>	<b>Outcome</b>
High	Low	Hot	High	Crash
Low	High	Cold	Low	Safe
Low	Low	Mild	High	Safe
Medium	High	Hot	Low	Crash
High	Low	Mild	Low	Safe
Medium	High	Mild	High	Crash
High	Low	Cold	High	Crash
Low	Low	Cold	Low	Safe
Medium	Low	Mild	Low	Safe
Low	High	Hot	High	Crash

- Q.2** (a) What strategies can be employed to avoid overfitting? **03**  
 (b) Write advantages and disadvantages of k- Nearest Neighborhood algorithm. **04**  
 (c) What are the distinctions among the Filter, Wrapper, and Embedded approaches to feature selection? **07**

**OR**

- (c) List three methods for performing dimensionality reduction and provide at least two advantages and two disadvantages of this process. **07**

- Q.3** (a) Enumerate three distinct measures utilized for assessing feature relevance and redundancy. **03**

	(b)	At a certain university, 4% of men are over 6 feet tall and 1% of women are over 6 feet tall. The total student population is divided in the ratio 3:2 in favour of women. If a student is selected at random from among all those over six feet tall, what is the probability that the student is woman?	04
	(c)	Compare SVM and Neural Networks.	07
	<b>OR</b>		
Q.3	(a)	Write down advantages of Naïve Bayes Classifier.	03
	(b)	Differentiate between feature extraction and feature selection process.	04
	(c)	Write any three key properties of SVM. Enlist any two applications of SVM and write any two limitations of SVM.	07
Q.4	(a)	Enlist any three desirable properties of Clustering.	03
	(b)	Write comparison between Biological NN and Artificial NN.	04
	(c)	Explain Recurrent Neural Network with its architecture.	07
	<b>OR</b>		
Q.4	(a)	List the three most frequently employed activation functions utilized in backpropagation Multilayer Perceptrons (MLPs) along with their corresponding equations.	03
	(b)	Write categories of Clustering Algorithms.	04
	(c)	Explain Multi-Layer feed forward network with its architecture.	07
Q.5	(a)	Differentiate between Regression and Classification.	03
	(b)	Explain the process of k- fold Cross Validation technique in detail.	04
	(c)	Explain AdaBoost algorithm in detail with its advantages and disadvantages.	07
	<b>OR</b>		
Q.5	(a)	Give the reason for choosing Non Linear kernels in SVM.	03
	(b)	Enumerate two benefits and two drawbacks of the Boosting algorithm.	04
	(c)	Describe the Bagging Algorithm including its Pseudo Code, and provide advantages and disadvantages.	07

\*\*\*\*\*