Total No. of Questions: 6 *Total No. of Printed Pages:3*

Enrollment No.....

Faculty of Engineering

End Sem (Odd) Examination Dec-2022

IT3EA07 Machine Learning

Programme: B.Tech. Branch/Specialisation: IT

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d.

Q.1 i. Application of Machine learning is _____. **1**

(a) Email filtering (b) Sentimental analysis

(c) Face recognition (d) All of these

ii. _____ is the machine learning algorithms that can be used with labeled data. **1**

(a) Regression algorithms (b) Clustering algorithms

(c) Association algorithms (d) All of these

iii. The supervised learning problems can be grouped as_____. **1**

(a) Regression problems (b) Classification problems

(c) Both (a) and (b) (d) Clustering

iv. The unsupervised learning problems can be grouped as _____. **1**

(a) Clustering (b) Association

(c) Both (a) and (b) (d) None of these

v. Which of the following is not a supervised learning? **1**

(a) PCA (b) Naive Bayesian

(c) Linear Regression (d) KNN

vi. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? **1**

(a) KNN (b) Random Forest

(c) Regression (d) Classification

vii. High entropy means that the partitions in classification are- **1**

(a) Pure (b) Not pure

(c) Useful (d) Useless

P.T.O.

viii. How do you handle missing or corrupted data in a dataset? **1**
(a) Drop missing rows or columns
(b) Replace missing values with mean/median/mode
(c) Assign a unique category to missing values
(d) All of these

ix. Which of the following is a disadvantage of decision trees? **1**
(a) Factor analysis
(b) Decision trees are robust to outliers
(c) Decision trees are prone to be overfit
(d) None of these

x. What is true about Machine Learning? **1**
(a) Machine Learning (ML) is the field of computer science
(b) ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method
(c) The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention
(d) All of these

Q.2 i. Write different type of machine learning applications. **2**
ii. Differentiate between classification & regression algorithm. **3**
iii. What are different types of machine learning algorithms? Explain each type in details. **5**
OR iv. What is 'training set' and 'test set' in a machine learning model? How much data will be allocated for training, validation, and test sets? Explain with examples. **5**

Q.3 i. What is overfitting and underfitting? **3**
ii. Explain decision tree algorithm. **7**
OR iii. Explain Naïve Bayes algorithm. **7**

Q.4 i. What is unsupervised learning? **3**
ii. What is clustering? Explain K-Means algorithm. **7**
OR iii. What is dimensionality reduction? Explain Principal Component Analysis. **7**

Q.5 i. What is Neural Network? **3**
ii. Explain Tensorflow and Keras python machine learning libraries in detail. **7**
OR iii. Explain feed forward, back propagation and RNN. **7**

Q.6 Attempt any two:
i. Evaluating machine learning algorithms and model selection with suitable example. **5**
ii. Explain ensemble technique bagging and boosting in detail. **5**
iii. Explain deep learning concept. **5**

******

Marking Scheme
IT3EA07 Machine Learning

Q.1	i.	Application of Machine learning is _____. (d) All of these		**1**
	ii.	_____ is the machine learning algorithms that can be used with labeled data. (a) Regression algorithms		**1**
	iii.	The supervised learning problems can be grouped as_____. (c) Both (a) and (b)		**1**
	iv.	The unsupervised learning problems can be grouped as _____. (c) Both (a) and (b)		**1**
	v.	Which of the following is not a supervised learning? (a) PCA		**1**
	vi.	Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? (b) Random Forest		**1**
	vii.	High entropy means that the partitions in classification are- (a) Pure		**1**
	viii.	How do you handle missing or corrupted data in a dataset? (d) All of these		**1**
	ix.	Which of the following is a disadvantage of decision trees? (c) Decision trees are prone to be overfit		**1**
	x.	What is true about Machine Learning? (d) All of these		**1**
Q.2	i.	Four type of machine learning applications. 0.5 mark for each		**2**
	ii.	Differentiate between classification & regression algorithm. 1 mark for each difference		**3**
	iii.	Types of machine learning algorithms Each type in details	2 marks 3 marks	**5**
OR	iv.	'Training set' and 'test set' Data will be allocated for training, validation, and test sets Examples.	2 marks 1 mark 2 marks	**5**
Q.3	i.	Overfitting Underfitting	1.5 marks 1.5 marks	**3**
	ii.	Decision tree algorithm Explanation	4 marks 3 marks	**7**
OR	iii.	Naïve Bayes algorithm. Explanation	4 marks 3 marks	**7**
Q.4	i.	What is unsupervised learning?		**3**
	ii.	Clustering K-Means algorithm.	2 marks 5 marks	**7**
OR	iii.	Dimensionality reduction Principal Component Analysis	2 marks 5 marks	**7**
Q.5	i.	What is Neural Network?		**3**
	ii.	Tensorflow Keras python machine learning libraries	3.5 marks 3.5 marks	**7**
OR	iii.	Feed forward Back propagation RNN	2 marks 3 marks 2 marks	**7**
Q.6		Attempt any two:		
	i.	Machine learning algorithms Model selection	2.5 marks 2.5 marks	**5**
	ii.	Ensemble technique bagging Boosting	2.5 marks 2.5 marks	**5**
	iii.	Deep learning concept. Example	4 marks 1 mark	**5**

******