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

Enrollment No.....

Faculty of Engineering

End Sem Examination May-2024

RA3EL08 Machine Learning

Programme: B.Tech. Branch/Specialisation: RA

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. Assume suitable data if necessary. Notations and symbols have their usual meaning.

Q.1 i. What is machine learning? **1**

(a) The selective acquisition of knowledge using computer programs
(b) The selective acquisition of knowledge using manual programs
(c) The autonomous acquisition of knowledge using computer programs
(d) The autonomous acquisition of knowledge using manual programs

ii. Which type of machine learning algorithm falls under the category of "Unsupervised learning"? **1**

(a) Linear regression (b) K-means clustering
(c) Decision trees (d) Random forest

iii. What kind of algorithm is logistic regression? **1**

(a) Cost function minimization
(b) Ranking
(c) Regression
(d) Classification

iv. What happens when the learning rate is low? **1**

(a) It always reaches the minima quickly
(b) It reaches the minima very slowly
(c) It overshoots the minima
(d) Nothing happens

v. What does CART stand for in the context of decision tree modelling? **1**

(a) Categorical and Regression Trees
(b) Classification and Regression Trees
(c) Cluster Analysis and Regression Trees
(d) Conditional and Regression Trees

vi. What is the objective of the Maximal-Margin Classifier in SVM? **1**
(a) Maximizing the number of support vectors
(b) Maximizing the margin between classes
(c) Minimizing the margin between classes
(d) Minimizing the number of support vectors

vii. Which method is used for bringing features onto the same scale by subtracting the mean and dividing by the standard deviation? **1**
(a) Min-max scaling (b) Robust scaling
(c) Standardization (d) Normalization

viii. What is the purpose of partitioning a dataset into training and test sets? **1**
(a) To handle missing data (b) To reduce dimensionality
(c) To assess feature importance (d) To evaluate model performance

ix. In time series analysis, if a model captures both trend and seasonality, it is often referred to as: **1**
(a) Seasonality without trend (b) Seasonality and trend
(c) Trend projection (d) Horizontal pattern

x. Which component of time series analysis refers to random, unpredictable fluctuations in data? **1**
(a) Trend (b) Seasonality
(c) Irregular pattern (d) Horizontal pattern

Q.2 i. List out any four applications of machine learning. **4**
ii. Differentiate between supervised, unsupervised and reinforcement learning. **6**
OR iii. Differentiate overfitting and underfitting problems encountered during machine learning. **6**

Q.3 i. Differentiate between linear regression and logistic regression. **4**
ii. Explain the concept of Linear Discriminant Analysis (LDA) in machine learning. Describe the representation of LDA models and the process of making predictions with LDA. **6**
OR iii. Compare and contrast batch gradient descent and stochastic gradient descent, highlighting their advantages and disadvantages. **6**

Q.4 i. Describe K-nearest neighbour learning algorithm with the help of suitable example. **4**
ii. What do you mean by decision tree? How does the decision tree algorithm work? Explain the attribute selection measure information gain. **6**

OR iii. Explain support vector machine. Define the terms hyperplane, support vectors, kernel, hard and soft margin. **6**

Q.5 i. Explain the following terms: **4**
(a) One hot encoder (b) Label encoder
ii. Explain the procedure for the computation of the principal components of the data. **6**
OR iii. Explain DBSCAN algorithm for density-based clustering. List out its advantages compared to K-means. **6**

Q.6 Attempt any two:
i. What are the key differences between linear trend regression and nonlinear trend regression models in time series analysis? **5**
ii. Compare and contrast moving averages and exponential smoothing techniques for time series forecasting. **5**
iii. Explain the concept of time series decomposition and its significance in forecasting. **5**

******

Marking Scheme

RA3EL08 (T) Machine Learning

Q.1	i)	C	1
	ii)	B	1
	iii)	D	1
	iv)	B	1
	v)	B	1
	vi)	B	1
	vii)	C	1
	viii)	D	1
	ix)	B	1
	x)	C	1

Q.2 i. List out any four applications of machine learning….. 1 Mark each

ii. Differentiate between
Supervised, ……2 Mark
Unsupervised and ……2 Mark
Reinforcement Learning……2 Mark

OR iii. Differentiate
overfitting ……3 Mark
and underfitting problems……3 Mark

Q.3 i. Differentiate between Linear Regression and Logistic Regression

ii. Explain the concept of Linear Discriminant Analysis (LDA) in machine learning. ……3 Mark
Describe the representation of LDA models and the process of making predictions with LDA. ……3 Mark

OR iii. Compare and contrast
batch gradient descent ……2 Mark
stochastic gradient descent, ……2 Mark
highlighting their advantages and disadvantages. ……2 Mark

Q.4 i. Describe K-nearest Neighbour learning Algorithm with the help of suitable example. ……4 Mark

ii. What do you mean by Decision tree? ……2 Mark
How does the Decision tree algorithm work? ……2 Mark
Explain the attribute selection measure information gain. ……2 Mark

OR iii. Explain Support Vector Machine. ……3 Mark
Define the terms Hyperplane……1 Mark
Support Vectors, Kernel, ……1 Mark
Hard and Soft Margin. ……1 Mark

Q.5 i. Explain the following terms:
One hot encoder ……2 Mark
b) Label Encoder……2 Mark

ii. Explain the procedure for the computation of the principal components of the data. ……6 Mark

OR iii. Explain DBSCAN algorithm for density-based clustering. ……4 Mark
List out its advantages compared to K-means. ……2 Mark

Q.6 Attempt any two:

i. differences between linear trend regression and nonlinear trend regression models in time series analysis……..5 difference 1marks each

ii. Compare and contrast
moving averages ……2.5 Mark
and exponential smoothing techniques for time series forecasting……2.5 Mark

iii. Explain the concept of time series decomposition……2.5 Mark
and its significance in forecasting. ……2.5 Mark

******