

JAMIA HAMDARD
Department of Computer Science & Engineering
School of Engineering Sciences and Technology
New Delhi – 110062

Examination: Odd Semester Examination-2024

Paper Name: Machine Learning
Time: 3 hours

Paper Code: BTCSE DED 13 (Backlog)
Maximum Marks: 75

SECTION A

All questions are compulsory in this section.

Marking Scheme: 7×5=35

Q.No.	Questions	Marks	CO	BL
1.	Explain Interpretation and Prediction with the help of an example.	5	CO3	L2
2.	What are the different types of missing values? What are the approaches used to handle missing values?	5	CO1	L3
3.	What is Inference and Estimation in the context of machine learning?	5	CO2	L2
4.	Differentiate between Regression and Classification with an example.	5	CO3	L4
5.	What do you understand by the term 'state' and 'reward' in Reinforcement Learning?	5	CO5	L1
6.	What is feature scaling and feature transformation with reference to Feature engineering?	5	CO2	L2
7.	Explain Support vector machines and their types.	5	CO4	L2

SECTION B

Attempt any FOUR questions from this section.

Marking Scheme: 4×10=40

Q.No	Questions	Marks	CO	BL														
1.	(i) Discuss the major drawbacks and advantages of K-Nearest Neighbors Algorithm. (ii) Describe the different types of hypothesis testing with the help of an example.	10	CO4	L4														
2.	(i) What are the steps involved in performing Exploratory Data Analysis? (ii) What do you understand by variable transformation?	10	CO2	L3														
3.	A marketing company has recorded its advertisement expenditures and corresponding sales for the past five years. The data is as follows: <table><tr><th>Advertisement Spending (\$)</th><th>Sales (\$)</th></tr><tr><td>9</td><td>100</td></tr><tr><td>12</td><td>130</td></tr><tr><td>15</td><td>180</td></tr><tr><td>10</td><td>120</td></tr><tr><td>13</td><td>138</td></tr><tr><td>20</td><td>?</td></tr></table> Using regression analysis, answer the following: (i) Find the linear regression equation that models the relationship between advertisement spending and sales, where advertisement spending is the independent variable (X) and sales is the dependent variable (Y). (ii) Predict the sales for an advertisement expenditure of \$20 in the year 2019 using the regression equation.	Advertisement Spending (\$)	Sales (\$)	9	100	12	130	15	180	10	120	13	138	20	?	10	CO3	L5
Advertisement Spending (\$)	Sales (\$)																	
9	100																	
12	130																	
15	180																	
10	120																	
13	138																	
20	?																	

4.	(i) What is the entropy of this collection of training examples with respect to the target function classification? (ii) What is the information gain of a1 and a2 relative to these training examples? (iii) Draw decision tree for the given dataset. <table><tr><th>Instance</th><th>Classification</th><th>a1</th><th>a2</th></tr><tr><td>1</td><td>+</td><td>T</td><td>T</td></tr><tr><td>2</td><td>+</td><td>T</td><td>T</td></tr><tr><td>3</td><td>-</td><td>T</td><td>F</td></tr><tr><td>4</td><td>+</td><td>F</td><td>F</td></tr><tr><td>5</td><td>-</td><td>F</td><td>T</td></tr><tr><td>6</td><td>-</td><td>F</td><td>T</td></tr></table>	Instance	Classification	a1	a2	1	+	T	T	2	+	T	T	3	-	T	F	4	+	F	F	5	-	F	T	6	-	F	T	10	CO4	L6
Instance	Classification	a1	a2																													
1	+	T	T																													
2	+	T	T																													
3	-	T	F																													
4	+	F	F																													
5	-	F	T																													
6	-	F	T																													
5.	Explain the workflow of Machine learning algorithm with suitable diagram.	10	CO1	L2																												
6.	(i) Discuss in detail the different Ensemble Learning methods with examples. (ii) What are the different techniques used to handle unbalanced classes in a dataset?	10	CO5	L3																												

BL-Bloom's Taxonomy Levels: (1- Remembering, 2- Understanding, 3 - Applying, 4 – Analysing, 5 - Evaluating, 6 - Creating); CO: Course Outcome