Guru Nanak Dev Engineering College, Ludhiana
(Department of Information Technology)
MSE Question Paper

Guru Nanak Dev Engineering College, Ludhiana			
Department of Information Technology			
Program	B.Tech.(IT)	**Semester**	6
Subject Code	PCIT-114	**Subject Title**	Introduction to Machine Learning
Mid Semester Test (MST) No.	2	**Course Coordinator(s)**	Er. Jaskiran Kaur
Max. Marks	24	**Time Duration**	1 hour 30 minutes
Date of MST	24th May 2023	**Roll Number**	2004428

Note: Attempt all questions

Q. No.	Questions	COs, RBT level	M
Q1	What do you understand by noise in data? How it affects the results?	CO1, L1	2
Q2	Differentiate between Multi-class and multi-label classification.	CO2, L4	2
Q3	Explain Artificial Neural Networks in detail. Also explain the three activation functions used in Neural Networks.	CO1, L2	4
Q4	Identify the first splitting attribute for decision tree by using ID3 algorithm for the followingdataset.	CO5, L3	

Major	Experience	Tie	Hired?
CS	programming	pretty	NO
CS	programming	pretty	NO
CS	management	pretty	YES
CS	management	ugly	YES
business	programming	pretty	YES
business	programming	ugly	YES
business	management	pretty	NO
business	management	pretty	NO

Q5	Compare and contrast KNN and K-means Clustering.	CO2, L4	4
Q6	a) Describe Support Vector Machine Classification in Machine Learning. b) Consider a fictional dataset that describes the weather conditions for playing a game of golf. Given the weather conditions, each tuple classifies the conditions as fit("Yes") or unfit("No") for playing golf.	CO3, L2 CO6, L5	2 6

	Outlook	Temperature	Humidity	Windy	Play Golf
0	Rainy	Hot	High	FALSE	No
1	Rainy	Hot	High	TRUE	No
2	Overcast	Hot	High	FALSE	Yes
3	Sunny	Mild	High	FALSE	Yes
4	Sunny	Cool	Normal	FALSE	Yes
5	Sunny	Cool	Normal	TRUE	No
6	Overcast	Cool	Normal	TRUE	Yes
7	Rainy	Mild	High	FALSE	No
8	Rainy	Cool	Normal	FALSE	Yes
9	Sunny	Mild	Normal	FALSE	Yes
10	Rainy	Mild	Normal	TRUE	Yes
11	Overcast	Mild	High	TRUE	Yes
12	Overcast	Hot	Normal	FALSE	Yes
13	Sunny	Mild	High	TRUE	No

Calculate the decision for today = (Sunny, Hot, Normal, False) using Naïve Bayes algorithm.

Course Outcomes (CO)Students will be able to:

1. Apply Supervised Learning, Unsupervised learning, Deep Learning, Visualization techniques.
2. Recognize and formalize a task as a machine learning problem.
3. Interpret and present the predicted model.
4. Identify suitable algorithms to tackle different machine learning problems.
5. Apply machine learning algorithms to real datasets.
6. Make powerful and accurate predictions.

Lower Order Thinking Levels (LOTS)			Higher Order Thinking Levels (HOTS)		
L1	L2	L3	L4	L5	L6
Remembering	Understanding	Applying	Analyzing	Evaluating	Creating

Guru Nanak Dev Engineering College, Ludhiana

Department of Information Technology

Program	B.Tech.(IT)	**Semester**	6
Subject Code	PCIT-114	**Subject Title**	Introduction to Machine Learning
Mid Semester Test (MST) No.	1	**Course Coordinator(s)**	Er. Jaskiran Kaur
Max. Marks	24	**Time Duration**	1 hour 30 minutes
Date of MST	15-02-2024	**Roll Number**	

Note:Attempt all questions

Q. No.	Questions	COs, RBT level	Marks
Q1	Define the usage of train_test_split() function.	CO1, L1	2
Q2	Examine the formula for distance of a points (x_1, y_1) from a line ax+by+c=0 and find the distance of the points (-5, -7) from the line 3x-4y-56=0.	CO4, L5	2
Q3	Discuss atleast four evaluation metrics for regression.	CO1, L3	4
Q4	Explain Linear Regression and obtain the regression equation of Y on X from the following: x: 2, 4, 6, 8 y: 3, 7, 5, 10	CO2, L2	4
Q5	Differentiate amongst Supervised, Unsupervised and Reinforcement Learning with suitable examples.	CO1, L4	4
Q6	Analyze the application of linear regression and determine the regression coefficients using the provided data: Y: 9, 10, 13, 14, 16 X1: 1, 3, 4, 6, 7 X2: 10, 14, 15, 18, 20	CO3, L4	8

Course Outcomes (CO) *Students will be able to:*

1. Apply Supervised Learning, Unsupervised learning, Deep Learning, Visualization techniques.
2. Recognize and formalize a task as a machine learning problem.
3. Interpret and present the predicted model.
4. Identify suitable algorithms to tackle different machine learning problems.
5. Apply machine learning algorithms to real datasets.
6. Make powerful and accurate predictions.

Please check that this question paper contains 09 questions and 02 printed pages within first ten minutes.

[Total No. of Questions: 09] [Total No. of Pages: 02]

Uni. Roll No. .....

Program: B.Tech. (Batch 2018 onward)
Semester:.6th
Name of Subject: Introduction to Machine Learning
Subject Code: PCIT-114
Paper ID: 17206.

Time Allowed: 03 Hours **Max. Marks: 60**

NOTE:

1) Parts A and B are compulsory.
2) Part-C has Two Questions Q8 and Q9. Both are compulsory, but with internal choice.
3) Any missing data may be assumed appropriately.

Part – A **[Marks: 02 each]**

Q1.

a) Define Machine Learning (ML) and its significance in modern technology.
b) What distinguishes regression from classification in the context of ML?
c) Define overfitting and underfitting and its impact on model performance.
d) What is a perceptron?
e) Define clustering in ML and give an example of its application.
f) Define Fuzzy Logic and explain its significance in ML.

Part – B **[Marks: 04 each]**

Q2. A real estate company needs to predict house prices based on features like bedrooms, size, location, and age. They have historical data available. Elaborate on which regression model would you choose and why?

Q3. A company aims to recognize handwritten digits (0-9) from scanned documents using a labelled dataset. Which neural network would you choose and why? Also, elaborate on the network architecture and training process.

Q4. Compare and contrast Supervised Learning vs. Unsupervised Learning by outlining their goals, methodologies, and use cases.

Q5. Analyze the impact of choosing different types of membership functions on the performance of a Fuzzy Logic system.

Q6. Elaborate in detail the steps involved in the Fuzzification process in Fuzzy Logic.

Q7. Compare and contrast the K-means and hierarchical clustering algorithms, discussing their advantages and disadvantages.

Part – C **[Marks: 12 each]**

Q8. Explain the concept of ML and how it differs from traditional programming. Discuss the various applications of ML across different domains and elaborate on why ML is considered the future of many industries.

OR

Elaborate on the steps involved in the hierarchical clustering process, using AGNES and DIANA as examples. Provide a detailed comparison between the two methods.

Q9. An email service provider aims to develop a robust spam detection system using a labeled dataset of spam and non-spam emails. Discuss in detail the selection of a classification technique for this task. Explain your data preprocessing steps, model training, and evaluation metrics. Describe how you would handle high-dimensional data and any potential challenges in implementing the spam detection system.

OR

Explore the concept of fuzzy set theory in depth, explaining how it differs from traditional theory and its significance in handling uncertainty in data. Provide examples illustrating the representation of fuzzy sets and discuss the process of fuzzification, including linguistic variables and fuzzy membership functions.

****