


Name: Enrolment No:			
<p style="text-align: center;">UPES End Semester Examination, May 2025</p> <p> Course: Artificial Intelligence & Machine Learning Program: B.Tech CSE (All Splz.) Course Code: CSAI3017 </p> <p style="text-align: right;"> Semester: VI Time : 03 hrs. Max. Marks: 100 </p> <p>Instructions: Attempt all questions.</p>			
SECTION A (5Qx4M=20Marks)			
S. No.		Marks	CO
Q 1	Discuss types of machine learning models with example.	4	CO1
Q 2	Prove that $[(P \rightarrow Q) \vee (R \rightarrow S)] \rightarrow [(P \vee R) \rightarrow (Q \vee S)]$ is a contingency.	4	CO1
Q 3	Differentiate between agglomerative and divisive hierarchical clustering.	4	CO3
Q 4	Is KNN algorithm a Lazy learning algorithm? Justify your answer.	4	CO2
Q 5	Define the terms: support vector, hyperplane, and margin in the context of SVMs.	4	CO2
SECTION B (4Qx10M= 40 Marks)			
Q 6	Explain data pre-processing steps with machine learning model training for any real-world problem.	10	CO1
Q 7	A company is using a Naïve Bayes classifier to identify whether an email is Spam [1] or Not Spam [0] based on keywords. a) Discuss why is Naïve Bayes a good choice for text classification? b) If the model classifies an important email as spam, what type of error is this (False Positive or False Negative)? Justify your answer.	10	CO2
Q 8	A city police department has data about various types of crimes occurring in different areas, including time of day, crime type, and location coordinates. Discuss which clustering algorithm can be used to identify crime hotspots? Justify your answer and show its implementation.	10	CO4
Q 9	A streaming platform wants to analyze viewing patterns to recommend content and design better watchlists. a) Identify how would you represent viewing sessions for Apriori analysis? b) If the rule $\{\text{Sci-Fi, Action}\} \rightarrow \{\text{Adventure}\}$ is found, how would you use it in user recommendations?	10	CO3

SECTION-C (2Qx20M=40 Marks)			
Q 10	Discuss linear regression by taking a suitable example of your own. Support your answer with a full explanation by providing suitable Python code and taking a dataset of your own choice.	20	CO2
Q 11	<p>An e-commerce company wants to segment its users based on their purchasing behavior to personalize marketing strategies.</p> <p>a) Discuss type of clustering algorithm would you recommend for segmenting customers? Justify your choice.</p> <p>b) Explain how you measure clustering goodness?</p> <p style="text-align: center;">OR</p> <p>A bank develops a Decision Tree classifier to detect fraudulent transactions based on attributes like transaction amount, location, and time of transaction.</p> <ul style="list-style-type: none"> Justify why a deep tree might lead to a high false positive rate in fraud detection? Explain how Gini Index or Entropy can help in selecting the best split at each node? 	20	CO4