

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- APRIL-2023

COURSE CODE (CREDITS): 18B1WCI634 (2)

MAX. MARKS: 25

COURSE NAME: Machine Learning

MAX. TIME: 1 Hour 30 Minutes

COURSE INSTRUCTORS: Dr. Yugal Kumar, Dr. Abhilasha Sharma & Dr. Vipul Sharma

Note: All questions are compulsory. Marks are indicated against each question in square brackets.

- Q1. CO-3
- a. Which should be preferred among gini impurity and entropy? List down the problem domains in which decision trees are most suitable. [2,3]
- b. Use ID3 algorithm to draw the decision tree for the given dataset.

Instance	A1	A2	Class Label
1	T	T	+
2	T	T	+
3	T	F	-
4	F	F	+
5	F	T	-
6	F	T	-

- Q2. CO-3
- a. Why can't we use linear regression in place of logistic regression for binary classification? [2,2,1]
- b. Can we solve the multiclass classification problem using logistic regression? If yes then how?
- c. What is the difference between logistic regression and SVM?

- Q3. CO-4
- a. What do you mean by bias variance tradeoff? [2,3]
- b. What is the difference between correlation and regression? Is it possible to apply linear regression for time series analysis?

- Q4. CO-4
- a. What are support vectors in case of SVM's? [1,3,1]
- b. Derive the objective function for soft SVM? Also explain the significance of ξ .
- c. What do you mean by kernel trick? Give its significance.

[P.T.O]

Q5. CO-5

[2,3]

- a. What is K in case of KNN? What will happen to decision surface if we keep on increasing the value of K? Also explain how we can compute the optimal value for K.
- b. Given a training dataset as shown below. Find out what kind of sport Angelina will play if k factor for KNN is 3 and her age is 5. (Use Euclidean Distance Metric)

Name	Age	Gender	Class of Sports
Ajay	32	0	Football
Mark	40	0	Hockey
Sara	16	1	Cricket
Zaira	34	1	Cricket
Sachin	55	0	Hockey
Rahul	40	0	Cricket
Pooja	20	1	Hockey
Smith	15	0	Cricket
Laxmi	55	1	Football
Michael	15	0	Football
