JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- APRIL-2023

COURSE CODE(CREDITS): 19B1WEC636 (3) MAX. MARKS: 25

COURSE NAME: Machine Learning for Data Analysis

COURSE INSTRUCTORS: Dr. Alok Kumar MAX. TIME: 1 Hour 30 Minutes

---

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

Q.1 Briefly explain the properties of Gini Impurity. Which should be preferred among Gini impurity and Entropy? Compute the Gini index for the given datasets. [CO4] **[4 Marks]**

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay In
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

Q.2 What is pruning in decision trees, explain with example? Why it is necessary in decision tree? What are the two types of pruning used in making decision tree? [CO3] **[4 Marks]**

Q.3 Apply SVM on the given dataset. [CO3,CO4] **[4 Marks]**

$$\begin{Bmatrix} +\,Class = \binom{4}{1}, \binom{4}{-1}, \binom{8}{1}, \binom{8}{-1} \\ -\,Class = \binom{1}{0}, \binom{0}{1}, \binom{0}{-1}, \binom{-1}{0} \end{Bmatrix}$$

Q.4 What do know about Hard Margin SVM and Soft Margin SVM? Why would you use the Kernel Trick? What are the different kernel functions are used in SVM?

[CO2, CO3] **[4 Marks]**

Q.5 What is feature transformation and feature engineering? What are different feature engineering techniques? Explain with suitable examples. [CO1, CO2] **[4 Marks]**

Q.6 What is overfitting and how can it be reduced? Is overfitting high bias or variance?

[CO1, CO2] **[3 Marks]**

Q.7 What is curse of dimensionality in machine learning and how to overcome this?

[CO2] **[2 Marks]**