JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2024

M.Tech-I Semester (CSE(IS))

COURSE CODE (CREDITS): 13M1WCI331 MAX. MARKS: 35

COURSE NAME: Machine Learning

COURSE INSTRUCTORS: HRI MAX. TIME: 2 Hours

Note: *(a) All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

Q.No	Question	CO	Marks
Q1	Given X=[2,4,6,8], and Y=[3, 4, 5, 10]. If we start with b=1 and w=1, compute b and w after the first iteration if the learning rate is 0.1. What is the loss function of linear regression gradient descent? How does gradient descent update the estimate, give the general formula?	CO1	[4]
Q2	Which would you address using unsupervised learning algorithm? Explain. (a) Given emails labeled as spam/not spam, learn a spam filter. (b) Given a set of news articles found on the web, group them about the set of articles about the same story. (c) Given a database of customer data, automatically discover market segments and group customers into different market segments. (d) Given a database of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.	CO1	[4]
Q3	Describe Feature scaling, Mean normalization, Z-Score normalization in the context of the above house price prediction problem. What is the range of values generated by each approach?	CO2	[4]
Q4	Describe the differences among mean, variance and covariance in connection to studying a dataset.	CO3	[3]
Q5	Find the eigen vectors and eigen values of the following matrix. $\begin{bmatrix} 1 & 4 \\ -4 & -7 \end{bmatrix}$	CO3	[4]

Q6	Identify the best root node of the decision tree for the following table on the basis of Information Gain. Discuss all steps in detail. **Outlook** | **Temp** | **Humid** | **Wind** | **Play?** Sunny | Hot | High | Weak | No Sunny | Hot | High | Strong | No Overcast | Hot | High | Weak | Yes Rain | Mild | High | Weak | Yes Rain | Cool | Normal | Weak | Yes Rain | Cool | Normal | Strong | No Overcast | Cool | Normal | Strong | Yes Sunny | Mild | High | Weak | No Sunny | Cool | Normal | Weak | Yes Rain | Mild | Normal | Weak | Yes Sunny | Mild | Normal | Strong | Yes Overcast | Mild | High | Strong | Yes Overcast | Hot | Normal | Weak | Yes Rain | Mild | High | Strong | No	CO4	[6]
Q7	Describe the process of Gradient Boost Classification up to two rounds on the following tabular data. What is the liking for the movie by a person with th following characteristics? Likes Popcorn – Yes, Age – 50 and Favorite Color – Green Likes Popcorn | Age | Favorite Color | Loves the Movie Yes | 12 | Blue | Yes No | 87 | Green | Yes No | 44 | Blue | No	CO5	[5]
Q8	Write the algorithm for Gradient Boost Regression.	CO5	[5]