JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST - 2 EXAMINATIONS - 2022

M.Tech - II Semester (ECE)

COURSE CODE (CREDITS): 21M1WEC233 (3) MAX. MARKS: 25

COURSE NAME: Applied Machine Learning for IoT

COURSE INSTRUCTORS: Dr. Vikas Baghel MAX. TIME: 1.5 Hours

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

Q1. Our task is to investigate whether students will pass or fail 21M1WEC233 based on whether or not they studied, cheated, and slept well before the exam. You are given the following data for five students. There are three features, "Studied," "Slept," and "Cheated." The column "Result" shows the label we want to predict. **[CO5]**

Studied	Slept	Cheated	Result
Yes	No	No	Passed
Yes	No	Yes	Failed
No	Yes	No	Failed
Yes	Yes	Yes	Failed
Yes	Yes	No	Passed

i. What is the **entropy** $H(Result)$ at the root node? **[2]**

ii. Draw the decision tree where every split maximizes the information gain? Explain (with numbers) why you chose the splits you chose. **[3]**

Q2. **a)** The logistic function is given by $g(x) = \frac{1}{1+e^{-x}}$. Show that **[2]** **[CO1]**

$$\frac{d}{dx}g(x) = g(x)\big(1 - g(x)\big)$$

b) Consider the problem of fitting the following function to a dataset of 100 points $\{(x_i, y_i)\}$, $i = 1 \ldots 100$: **[3]**

$$y = \alpha \cos(x) + \beta \sin(x) + \gamma$$

This problem can be solved using the least squares method with a solution of the form:

$$\begin{bmatrix}\alpha \\ \beta \\ \gamma\end{bmatrix} = (X^T X)^{-1} X^T Y$$

What are X and Y ?

Q3. **a)** What are the Overfitting and Underfitting problems and how to avoid it? **[1]** **[CO1]**

b) In terms of the bias-variance tradeoff, which one (Bias or Variance) is substantially more harmful to the test error that the training error and why? **[1]**

c) Derive gradient descent algorithm's model parameters update rule for simple linear regression model: **[3]**

$$\beta_0 := \beta_0 - \frac{\alpha}{n}\sum_{i=1}^{n}\left(h_\beta\left(x^{(i)}\right) - y^{(i)}\right)$$

$$\beta_1 := \beta_1 - \frac{\alpha}{n}\sum_{i=1}^{n}\left(h_\beta\left(x^{(i)}\right) - y^{(i)}\right)x^{(i)}$$

Q4. **a)** Why KNN is known as lazy learner? **[1]** **[CO5]**

b) Suppose we have height, weight and T-shirt size of some customers and we need to predict the T-shirt size of a new customer given only height and weight information we have. Data including height, weight and T-shirt size information is shown below - **[4]**

Height (in cms)	**Weight (in kgs)**	**T Shirt Size**
158	58	*M*
158	59	*M*
158	63	*M*
160	59	*M*
160	60	*M*
163	60	*M*
163	61	*M*
160	64	*L*
163	64	*L*
165	61	*L*
165	62	*L*
165	65	*L*
168	62	*L*
168	63	*L*
168	66	*L*
170	63	*L*
170	64	*L*
170	68	*L*

The new customer 'Vikas Baghel' has **height 161cm** and weight **61kg**. Find his T Shirt Size using KNN?