THE OPEN UNIVERSITY OF SRI LANKA
BSc Hons (IT) DEGREE PROGRAMME: LEVEL 06
**FINAL EXAMINATION: SEMESTER 1: 2024/2025**
**COU6308: MACHINE LEARNING**
DURATION: **TWO HOURS (2 HOURS)**

Date: **15.06.2025** Time: **1.30 pm – 3.30 pm**

**Answer FOUR Questions ONLY.**

**Q1.**

A. Define "machine learning" in your own words.

[03 Marks]

B. Why do we need to use machine learning?

[03 Marks]

C. There are three types of machine learning: Supervised Learning, Unsupervised Learning, and Reinforcement Learning.
- Define each type of machine learning.
- For **Supervised Learning** and **Unsupervised Learning**, state commonly used algorithms.
- For **Reinforcement Learning**, state its key characteristics.

[09 Marks]

D. Given the following list of real-world problems, identify which type of machine learning is most appropriate for each case. Justify your classification with a brief explanation.
**Problems:**
(i) Forecasting stock market trends using historical price and volume data.
(ii) Grouping customers based on purchasing behavior for targeted marketing.
(iii) Developing a game-playing agent that learns optimal strategies by interacting with the environment.
(iv) Classifying handwritten digits using a labeled image dataset
(v) Identifying patterns in social media usage without predefined labels.

[10 Marks]

**Q2.**

A. What are the two types of **Supervised Learning**? Briefly define and differentiate them, and give one example for each.

[06 Marks]

B. Explain the concepts of **generalization**, **overfitting**, and **underfitting** in machine learning. For each concept, provide three real-world examples that illustrate the concept clearly.

[09 Marks]

C. Given the following real-world problems, identify whether each one is a classification or a regression task. Provide a brief reason for your choice.

**Problems:**

(i) Predicting whether a telecom customer will leave the network (churn prediction).
(ii) Estimating the product-wise manufacturing labor cost.
(iii) Determining if a patient is infected with a disease.
(iv) Forecasting the monthly electricity cost for the next three years.
(v) Classifying the genre of a music track.

[10 Marks]

**Q3.**

A. Provide the general formula for a **linear regression model** and clearly label each element.

[03 Marks]

B. What is **logistic regression**? Briefly explain its purpose in machine learning.

[03 Marks]

C. What are the key parameters of the **K-Nearest Neighbors (KNN) algorithm**? Briefly explain each.

[04 Marks]

D. List three strengths and weaknesses of the K-Nearest Neighbors (KNN) algorithm.

[06 Marks]

E. Compare and contrast different types of linear models used for regression. Provide examples where applicable.

[09 Marks]

**Q4.**

A. List the parameters that affect the performance of a **Random Forest model**.

[02 Marks]

B. List the three main types of **Naive Bayes classifiers** and mention where each is used.

[03 Marks]

C. Explain the role of the alpha parameter in Naive Bayes classifiers. Why is it important

[04 Marks]

D. Define what a Random Forest is. Explain the concept of randomization in Random Forests.

[04 Marks]

E. Briefly describe the three main components of a **decision tree structure**.

[06 Marks]

F. Explain two popular methods for avoiding over fitting in decision trees. Provide examples of criteria used in pre-pruning.

[06 Marks]

**Q5.**

A. Define what a **Gradient boosted regression** tree is.

[03 Marks]

B. Explain how SVM works.

[04 Marks]

C. List and define the key parameters of an SVM model.

[06 Marks]

D. How does feature importance in Gradient Boosted Regression Trees (GBRT) compare to Random Forests, and why might GBRT ignore some features entirely?

[06 Marks]

E. List and describe three common kernel functions used in Support Vector Machines.

[06 Marks]

**Q6.**

A. What are **neural networks**?

[03 Marks]

B. Define the formula for **MLP** (Multilayer Perceptron) in the regression case and clearly label each element.

[03 Marks]

C. Explain the decision function and predicting probabilities. Differentiate them.

[04 Marks]

D. Explain what **unsupervised learning** is, and list and explain its types.

[06 Marks]

E. Draw the visual representations of the following models:

(i) Logistic Regression
(ii) Multilayer Perceptron (MLP) with a single hidden layer
(iii) Multilayer Perceptron (MLP) with two hidden layers

[09 Marks]