Assignment

Instructions

- Answer all questions with clarity and detail. For questions involving calculations, please show all steps clearly. Include explanations, diagrams, or visualizations wherever appropriate to support your answers.

- Submit your typed assignment as a PDF. Format your document properly and name it according to your student ID (e.g., 22-45678-3.pdf). Submit the PDF through MS Teams (inbox).

- Deadline: The assignment is due on 25th November. Submissions after this date will incur penalties.

Questions

Q1. Differentiate between the Supervised and Unsupervised Learning.

Q2. Explain how **gradient descent** is used in logistic regression for optimizing model parameters. Include the steps involved, the cost function used, and how the parameters are updated during the process.

Q3. Consider a binary classification problem where you predict whether a student passes (1) or fails (0) based on study hours. Explain how this model produces a probability value and converts it into a binary outcome using **Logistic Regression**. You can use a sample dataset use the steps as in **Q2**.

Q4. A company wants to determine whether a customer should receive a promotional discount based on *Age, Income, and Shopping Frequency*. Create a simple dataset with 10 rows and build a **decision tree** based on Information Gain for the first two levels of the tree. Interpret the results.

Q5. A model predicts whether an email is spam or not spam. After testing the model, the Confusion Matrix is as follows:

Generate the matrix from your ID. Consider the middle 5 digits of your ID as **ABCDE**

	Predicted Spam	**Predicted Not Spam**
Actual Spam	AB	10
Actual Not Spam	20	CDE

(a) Calculate the accuracy, precision, recall, and F1 score for the model.

(b) Interpret the results and discuss whether the model has a good performance.