Assignment Code: DA-AG-011

Logistic Regression | **Assignment**

Instructions: Carefully read each question. Use Google Docs, Microsoft Word, or a similar tool to create a document where you type out each question along with its answer. Save the document as a PDF, and then upload it to the LMS. Please do not zip or archive the files before uploading them. Each question carries 20 marks.

Total Marks: 200

Question 1: What is Logistic Regression, and how does it differ from Linear Regression?

Answer:

Question 2: Explain the role of the Sigmoid function in Logistic Regression.

Answer:

Question 3: What is Regularization in Logistic Regression and why is it needed?

Answer:

Question 4: What are some common evaluation metrics for classification models, and why are they important?

Answer:

Question 5: Write a Python program that loads a CSV file into a Pandas DataFrame, splits into train/test sets, trains a **Logistic Regression** model, and prints its **accuracy**. (Use Dataset from sklearn package)

(*Include your Python code and output in the code box below.*)

Answer:

Question 6: Write a Python program to train a Logistic Regression model using L2 regularization (Ridge) and print the model coefficients and accuracy.

(Use Dataset from sklearn package)

(*Include your Python code and output in the code box below.*)

Answer:

Question 7: Write a Python program to train a Logistic Regression model for multiclass classification using `multi_class='ovr'` and print the classification report.
(Use Dataset from sklearn package)

(*Include your Python code and output in the code box below.*)

Answer:

Question 8: Write a Python program to apply GridSearchCV to tune `C` and `penalty` hyperparameters for Logistic Regression and print the best parameters and validation accuracy.

(Use Dataset from sklearn package)

(*Include your Python code and output in the code box below.*)

Answer:

Question 9: Write a Python program to standardize the features before training Logistic Regression and compare the model's accuracy with and without scaling.

(Use Dataset from sklearn package)

(*Include your Python code and output in the code box below.*)

Answer:

Question 10: Imagine you are working at an e-commerce company that wants to predict which customers will respond to a marketing campaign. Given an imbalanced dataset (only 5% of customers respond), describe the approach you'd take to build a Logistic Regression model — including data handling, feature scaling, balancing classes, hyperparameter tuning, and evaluating the model for this real-world business use case.

Answer: