**AIAA 2205 Assignment1**

**Overview**

In this project, you will implement a classification model from scratch. The goal is to build a complete machine learning pipeline, starting from data exploration, model derivation, training implementation, evaluation, and writing a concise report.

You are not allowed to use high-level machine learning libraries such as `sklearn`, `tensorflow`, or `pytorch`. You may use fundamental scientific computing libraries such as `numpy`, `pandas`, `matplotlib`, and `scipy` for basic matrix operations, data loading, and visualization.

We provide starter code (`main.py` and `models/logistic\_regression.py`) and two data files.

Your first task is to complete all the TODO sections so the program runs successfully. After achieving a working implementation, you are expected to experiment with improvements — for example, tuning hyperparameters, adding helper functions, or exploring additional techniques to enhance model performance.

Finally, you must submit a short report summarizing your approach, key results, and insights gained from this project.

**Project Setup**

**Folder Structure**

project\_root/

├─ main.py

├─ models/

│ └─ logistic\_regression.py

└─ data/

├─ classification\_data\_train.csv

└─ classification\_data\_test.csv

Do not rename the files or folders.

Once you have completed all required implementations, simply run the following command from the `project\_root `directory to train and evaluate your model: `python main.py`

**Submission Document**

1. **Code (15’):**

Ensure your code is well-structured and includes necessary comments for clarity. Ensure that we can successfully run your project by executing the following command from the project\_root directory: `python main.py`

1. **Project Report (Report) (15’):**

**Format:** A PDF document.

**Content:** Your report can include the following sections:

* **Problem Understanding:** Briefly describe the project you chose, its objectives, and the data.
* **Exploratory Data Analysis:** What explorations did you perform on the data? What patterns or issues did you discover through visualizations or other means?
* **Model Building and Training:** Detail how you built and trained your model.
* **Results and Analysis:** Evaluate your model's performance. You can compare the results of different models or different parameters.
* **Challenges and Solutions:** What difficulties did you encounter during the project? How did you think through and resolve these problems?
* **Conclusion and Future Work:** Summarize your project and suggest potential directions for future improvement.

Good luck!