Least Square and Introduction to Machine Learning

Lab assignment #2 (worth 20%; deadline: Friday, November 28 at 11:59pm)

Learning objectives

After completing this lab assignment students should be able to:

- Understand and implement gradient descent
- Understand and apply logistic regression for binary classification
- Be familiar with common techniques in python
- Understand the fundamental principles of LSE
- Program practical Non-linear LSE problem using Python
- Understand and implement ML model training and hyperparameter tuning using grid search with cross validation

Preliminary Task – (5 Marks)

Similarly to Lab Assignment 1, each group must fill in the **certificate of work**, which is in the given notebook of this lab assignment. Each group is responsible for maintaining a clean Git repository similar to the Bookkeeping tasks from Assignment 1, E.g., clear project structure, consistent version control, include only necessary files etc.. Make sure the code is readable, E.g., consistent naming conventions, proper commenting, clear code structure, consistent Indentation and Spacing etc.

Task 1 – Non-linear Estimation - LSE & Gradient Decent (35 Marks)

The company Deep-AutoLab is designing a self-driving car, and one of the fundamental techniques is the ability of tracking vehicles in term of their positions. Therefore, they plan to build an algorithm of Single Point Positioning (SPP) using the measurements from Global Navigation Satellite System (GNSS) and apply the Least Square Estimation to obtain the solution. This task involves performing SPP using Least Squares Estimation (LSE) and Gradient Descent. To build this algorithm:

- Please read the background of `SinglePointPositioning` carefully at the beginning of the Non-linear LSE section and complete the questions.
- Please go through the Gradient Descent section in the notebook and complete the related questions.
- For groups of 3, you can choose to omit this **Gradient Decent TASK**. However, Gradient Descent **IS AN INTEGRAL PART OF QUIZ 2**, and it is **STRONGLY RECOMMENDED** to attempt the task, as it will greatly enhance your understanding of the concept and better prepare you for the quiz.

- Every team is required to use *Git* to collaborate on this assignment and ensure proper distribution of work.

Task 2 – Student Depression Prediction using Logistic Regression (30 marks)

This part aims at analyzing, understanding, and predicting depression levels among students.

- Please go through the Logistic Regression section in the notebook 'Assignment 2' and complete the related questions.
- Every team is required to use *Git* to collaborate on this assignment and ensure proper distribution of work.

Task 3 -MNIST Handwritten Digit Classification using MLP & CNN (30 marks)

Ever wondered how computers learn to read handwriting? This study tackles that challenge by using the famous MNIST database of handwritten digits. Leveraging its 60,000 training examples, we will construct two types of neural networks, a Multi-Layer Perceptron (MLP) and a Convolutional Neural Network (CNN), to see which one performs better at identifying the digits.

- Every team is required to use *Git* to collaborate on this assignment and ensure proper distribution of work.

Submission Materials

All the listed materials **must** be submitted in the *Lab Assignment 2 D2L* dropbox by the deadline. **Late** and **incomplete** submissions will result in a deduction of marks.

- A zip file of the `Lab_Assignment2` folder, which should include the Assignment 2 *.ipynb* file and data files
- A link to the group GitHub repository
- Readme file similar to assignment 1
- **NO REPORT** is required for this assignment

As stated from "*Additional Lab Information"*, please remember that:

- Each member is **required** to make at least one commit and push to the group repository. Failure to do so will result in a 1% **individual** course grade deduction.
- The D2L submission dropbox will close 24 hours after the deadline. All submissions in this 24-hour window are subject to a 20% lab grade deduction.
- There must not be any commits to the `Lab_Assignment2` folder in your GitHub repository after the deadline, or there will be a 20% lab grade deduction.