Project 2: Regression

Introduction

In this project, you will be implementing a regression method from scratch using MATLAB, based on the theory covered in Unit 3. You will need to evaluate your algorithm on some of the datasets that will be provided (you may also add your own if desired), and compare the results to the relevant built-in method in either MATLAB or Python.

Undergraduate and ETIE groups can choose one from of the following:

- Linear regression (without ε term) with gradient descent training
- K-nearest neighbors (using only Euclidean distance)
- Decision trees

Graduate groups can choose from one of the more advanced options below:

- Linear regression (including ε term) with gradient descent training
- K-nearest neighbors (with at least three types of distances)
- Random forest

A separate document will be posted later that provides hints how to approach each of these problems.

Submission Requirements

To receive full credit for this project, you must submit your complete source code (including your input and output data files) as a single ZIP archive.

In addition, you must also submit a formal report that contains the following:

- Cover page (title, names of group members, date of submission)
- Table of contents
- Problem definition (what is being done and why)
- Theoretical analysis
- Algorithm development (can include flowcharts and/or pseudocode if desired)
- Simulation results
- Discussion
- Conclusion (summarize what was done; what is some potential future work?)
- Author contributions (briefly state what each group member contributed to the project)
- Appendices containing source code, program output, and plots as appropriate

Some sample reports with similar formats (from other courses) are available on Brightspace, for your reference.

Grading Criteria

This project is worth 30 points and is based on the following:

- Completeness: *3 points*
  - Was every required part of the project addressed?
- Problem definition: *3 points*
- Theoretical analysis: *4 points*
- Correctness of results: *4 points*
- Discussion: *3 points*
- Other report sections: *2 point*
- Coding style: *2 points*
  - Use of self-contained functions
  - Documentation of code
  - Avoidance of unnecessary code duplication
- Format/organization: *3 points*
  - Is formatting consistent?
  - How easy is it to read the report?
- Spelling and grammar: *2 point*
- Steadiness of progress: *4 points*
  - Each group must update the instructor of project progress each week
  - Either in class or by email

Due Date/Time

This project is due on Brightspace by **April 11, 2021 at 11:59pm CST**.

Late submissions will receive a penalty of 10% off for each day that it is late. In some cases, it may actually be better to submit a late report that was done well rather than an on-time report that was rushed to completion.