ECE1513 Introduction to machine learning

Project Statement

Deadline: Sunday, December 17th, at 11:59 pm.

Coding: Use your local machine or Google Colab to create python notebook that contain all parts of the project.

Submission: You need to submit a final report following the report structure document attached with the project statement as well as a notebook, containing all the codes, demos, and results.

Late Submission: 10% of the marks will be deducted for each day late, up to a maximum of 3 days. After that, no submissions will be accepted.

Note: The project is individual work.

**Instructions**

Choose a data topic that you are interested in and analyze with the tools and techniques you have picked up. Some of the sites you can refer to for datasets are:

● www.kaggle.com/datasets

● https://datasetsearch.research.google.com/

● https://archive.ics.uci.edu/ml/datasets.php

● <https://datahub.io/collections>

You are required to analyze the dataset using Python libraries such as Pandas, and Scikit-Learn. You are free to choose the type of ML or DL problem to solve. Consider using Pytorch package in case of a DL approach.

**Guidance**

1. Choose a topic that interests you and define the ML/DL problem.

2. Source the dataset and explore the metadata (dataset information).

3. Explore and visualize the data before building the models.

4. Utilize Python’s libraries to train and validate your models.

5. Discuss your models practical use cases in your project summary.

6. Refer to the rubric in the next page for evaluation details.

**Evaluation:**

This project is graded out of 100 points using the following rubric and is worth 25% of the final grade. You may receive partial scores or a zero for unacceptable work.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Criteria | Do Not Meet Expectations  1 | Partially Meets Expectations  2 | Meets Expectations  3 | Exceeds Expectations  4 | Max Points |
| Literature survey and report structure | Unclear project abstract or does not relate to project contents presented. Includes few unrelated or no references to similar work done. | Simple project abstract or provides minimal context of the project. Includes few related references to similar work done. | Adequate project abstract or provides sufficient context of the project. Includes a few related references to similar work done and mentions relevance to presented project. | Insightful project abstract or provides deep context of the project. Includes adequate references to similar work done and highlights relevance to presented project. | 20 |
| Coding | Incomplete / wrong code submitted. Code is not structured properly, few to no comments provided. | Partially complete / correct code submitted. Code is partially structured, few comments provided. | Complete / correct code submitted. Code is partially structured, few comments provided. | Complete / correct code submitted. Code is well structured, adequate comments provided. | 40 |
| ML/DL approach | Incomplete model notebook submitted, or wrong/incomplete analysis components included. Provides no visibility to analysis. | Partially complete ML/DL pipeline submitted, or few key steps included. Provides partial visibility to analysis. Deep learning best practices not followed. | Complete ML/DL pipeline notebook submitted, or sufficient analysis components included. Provides sufficient visibility to analysis. Deep learning best practices partially followed. | Complete ML/DL pipeline notebook submitted, or insightful analysis components included. Model achieves acceptable results and follows ML/DL best practices. Notebook file is well-documented, and models are interpreted in a clear and concise manner. | 40 |