CSCI 416/516 Final Project

Proposal DUE: October 11, 2023, at 11:59 pm

Final Report DUE: December 8, 2023, at 11:59 pm

Overview: 20% (20 pts) of your total mark is allocated to a final project, which will require you to apply several algorithms to a challenging problem and to write a short report analyzing the results. You are allowed to collaborate with at most 2 classmates on the final project. In other words, **the maximum number of people a group can have is 3**.

Submission: If you decide to work in a group, only 1 of the group members need to submit to Blackboard. The late policy for homeworks applies to the Final Report.

Project Ideas: Below, we give a summary of some project ideas. You can choose a project idea from this list, but you can also come up with your own project idea. **You must submit a brief description (a proposal) of the proposed project idea to Blackboard**, which will be reviewed by the instructors, who will then later approve the proposal or ask you to choose a different project idea, mainly due to the simplicity/difficulty of the proposed idea. Note that your project must not be a direct application of the basic/vanilla techniques (i.e. linear regression) we discussed during class; additional task(s) must be done. In the following examples, the basic/vanilla techniques are highlighted in green, and the additional tasks are highlighted in red.

- Titanic Survival Prediction
  - Dataset: Titanic - Contains data on passengers and their survival status
  - Objective: Predict passenger survival on the Titanic.
  - Implementation: SVM using features like age, gender, etc., with feature engineering.
- Linear Regression for House Pricing
  - Dataset: Boston Housing - Information about different houses in Boston and their pricing.
  - Objective: Predict house prices based on features like number of rooms, location, size, etc.
  - Implementation: Apply linear regression on the Boston Housing dataset, and compare it against approaches such as neural network for regression.
- Image Classification

- Dataset: CIFAR-10 - Information about tiny images from 10 observed classes.
- Objective: Classify small images into 10 categories (like dogs, cars, birds).
- Implementation: A CNN pre-trained on the dataset STL-10, then continuously trained using the CIFAR-10 dataset.

Proposal: You need to submit a proposal consisting of the following information to be approved by the instructors. The proposal should be 1-2 pages long, in the IEEE Conference Template, including tables, charts, and figures, if applicable. Your proposal should consist of the following components.

- **Project Description**
  - **Problem Statement:** Begin by clearly defining the problem you plan to address in your machine learning project. Explain its significance and relevance in the context of machine learning and the course.
  - **Dataset:** Mention the dataset(s) you plan to use for your project. Provide details on the source, size, and any data preprocessing steps required.
- **Approach(es) to be Applied**
  - **Methodology:** Describe the machine learning techniques and algorithms you intend to use to solve the problem. Explain why you chose these methods and how they apply to your project.

Final Report: You are to turn in a written report summarizing the results of your project and your code. The report should be 3-4 pages long, in the IEEE Conference Template, including tables, charts, and figures. Your written final report should consist of the following components.

- **Abstract:** Provide a pithy and concise summarization of your work, briefly mentioning the motivation of your work, what was the objective you tried to solve, and the approaches you applied to solve the problem. You should also include the major evaluation results as well as their implications.
- **Introduction:** In the introduction section, provide an overview of your class project. Start by introducing the problem or topic your project addresses. Explain why it is important or relevant. Provide context by discussing any real-world applications or implications of your project. End the introduction with a clear statement of your project's objectives and goals.
- **Applied Aprroach(es):** This section is the core of your project report. Describe the approach or methodology you used to tackle the problem or achieve your project's goals. Explain the steps you took, algorithms used, data collection methods, and any other technical details. Include diagrams and flowcharts if applicable. Make it clear and concise so that someone else could replicate your project.

- **Evaluation:** In the evaluation section, present the results of your project. Describe the experiments or tests you conducted to assess the effectiveness of your approach. Include quantitative and qualitative data, graphs, tables, or any visual aids that help convey your findings. Discuss the metrics or criteria used to measure success and whether your project met its objectives.

- **Discussion:** The discussion section is where you interpret and analyze your results. Explain the significance of your findings in the context of the problem you addressed. Discuss any challenges or limitations encountered during your project and propose possible solutions or areas for future research. Offer insights, draw conclusions, and highlight the practical implications of your project.

- **Conclusion:** Provide a concise summary of your project, reiterate its significance, and mention the key takeaways. Discuss the overall success of your project and whether it achieved its goals. Mention any future work that can build upon your project.

- **References:** List all the sources, research papers, books, websites, or any other references you cited throughout your project report. Ensure proper citation formatting, such as APA, MLA, or any other style, as long as your formatting is consistent.