*ENGG*6600: Machine Learning with Applications*
Course Final Project

Prof. Shawki Areibi
School of Engineering, University of Guelph
Fall 2025

Start Date: Monday, Week #9, Fall 2025
Due Date:Monday, Week #13, Winter 2025

1 Introduction

The ENGG*6600 course project is an excellent opportunity for students to apply state-of-the-art machine learning methods to different application areas including VLSI, CAD, Health, Finance, Environment, and many others.

Aside from intrinsic benefit, course projects are an opportunity to demonstrate your machine learning knowledge and potentially add to a portfolio you can use for job or graduate school applications. Previous students have even submitted outstanding course projects for publication!

1.1 Aims of the Project

The aim of this project is to allow students to experience research in machine learning. Students will implement an existing or novel machine learning model on a dataset suggested by the instructor, or data of their own choosing. Students will communicate their research results in a technical report style or even a publication paper style, typical of machine learning conferences, the main mode of dissemination in this field.

- What we expect to see is a simple but original idea that you describe clearly, relate to existing methods, implement, and test on a small scale problem.
- To do this you will need to **write code**, run it on some data, make some figures, read a few background papers, collect some references, and write a few pages describing your algorithm and results.
- As a rough rule of thumb, spend about as much time doing the work as you would have studying for an exam and a few hours writing it up after that (instead of actually writing the exam).

2 Logistics

2.1 Project Group

- Project groups will have two students.
- Students are expected to contribute equally to the project.

2.2 Marking Scheme

Projects will be judged on the basis of visual appearance, the grammatical correctness and quality of writing, and the visual appearance and readability of any graphics, as well as their contents. You are encouraged to have a classmate or peer review your report prior to submission (researchers often do this).

The project grade (30% out of final mark) is further broken down to:

1. Proposal – 3%
2. Python Code (Notebook) – 2%
3. Final Report write-up – 25%

The following criteria will be taken into account roughly equally when marking reports:

1. Clarity of problem statement and description of approach.
2. Discussion of relationship to previous work and references.
3. Design and execution of experiments.
4. Figures/Tables: easily readable, properly labeled, informative.
5. Analysis of the results.
6. Documentation of your code.

3 Advice

- Be Honest! You are not being marked on how good the results are. It doesn't matter one bit if your method is better or worse than the ones you compare to. What matters is that you clearly describe the problem, your method, what you did, and what the results were.
- Be modest! Don't pick a project that is way too hard. Usually, if you select the simplest thing you can think of to try, and do it carefully, it will take much longer than you think.
- Be careful! Don't do foolish things like test on your training data, set parameters by cheating, compare unfairly against other methods, include plots with unlabeled axes, use undefined symbols in equations, etc. Do sensible crosschecks like running your algorithms several times, leaving out small parts of your data, adding a few noisy points, etc. to make sure everything still works reasonably well. Make lots of pictures along the way.
- Have fun! If you pick something you think is cool, that will make getting it to work less painful and writing up your results less boring.

4 Final Report

The final report will count for 85% of the student's grade. One grade will be assigned to the entire project group. The goal of the final report is to have students practice technical writing in order to convey the contributions of their project to a audience with an adequate machine learning background. Reports will be judged both on their technical quality (mastery of machine learning concepts) and writing proficiency. Here is the rubric for the final report:

- The final report should be between 15-20 pages. It will not be graded if it exceeds 20 pages.
- The report must be submitted as a pdf to Courselink. It could be written with LaTex using the techreport style files or word (5% adherence to style requirements).
- The report must consist of the following sections (5% adherence to section requirements):
    1. **Abstract**. At a high level, summarize what your problem is, what methods you used, and your results. An abstract that is shorter and more concise is better.
    2. **Introduction**. What is your problem? Why is it important? What are the contributions if any?
    3. **Background**. Introduce necessary background for reader to understand the report.
    4. **Related work**. What have people previously done in regards to your problem? What work is related? (This is a great place to summarize some of the papers you wrote about in-depth for the Literature Review assignment).
    5. **Methodology**. What is the overall methodology used in the project? What machine learning models are utilized and why?
    6. **Dataset**. Describe what dataset(s) you are using, where these came from, and some basic properties of the dataset.
    7. **Results and Evaluation**. What are your results? How did you evaluate these results?
    8. **Conclusion & Future Work**. What can you conclude from your project? What did you learn? What are future directions? Are there any real-world implications from your work?
    9. **References (bibliography)**. List citations here. Use any reasonable style file to cite certain works.
- Projects must adequately cite at least 5 research articles (5%).
- Additionally, the report must also contain the following:
    1. At least two figures. These figures could show the overall methodology, or results, interesting analyses, exploration of the features, an overview of the data and modeling pipeline, etc. (20%)
    2. At least one table. The table could consist of data statistics, results, etc. (10%)
- Writing mechanics: grammar and typos (10%).
- Writing clarity: high-level writing style and arguments conveyed effectively (10%).
- Mastery of machine learning concepts (30%). Evaluation on this category could include (but is not limited to):
    1. Proper train-test split (or train-dev-test split or cross-validation)
    2. Proper data-driven selection of hyper-parameters

3. Comparing models against a baseline (e.g. predicting the majority class)
4. Comparing more than one machine learning model
5. Proper use of a machine learning package (e.g. sklearn or other library) or development of a new machine learning model.

- Novelty of ideas (5%). A truly masterful project will include novel ideas that have not been tried before by others.

4.1 Report Format

Name your report ENG6600_F24_Project_Group#_LastNames.pdf

Below is the general format of the report required:

1. The **front page** (i.e. title page) should contain only the following:
   - Course #, Course Name and Date
   - Your Group # .. Your Names and IDs
   - Project title.
2. **Introduce the problem** to be solved:
   - **Problem Statement**,
     (a) Briefly describe the problem solved in the project.
     (b) **Assumptions and Constraints**.
3. **Literature Review**:
   - Briefly summarize any work (papers) related to your work.
   - Explain how your work is different (if any).
4. **Methodology**:
   (a) Explain how you implemented your solution.
   (b) **Block diagram** of the methodology.
   (c) **System Overview & Justification of Design**
      i. Give an overview of the system to be designed.
      ii. Briefly explain how the system works and reasons behind the design.
5. **Results**:
   (a) Results Obtained (Tables, figures, ..).
   (b) Include Plots and Graphs.
   (c) Discuss the results obtained and explain why you got these results.
6. **Conclusions**:

(a) Give a final conclusion on the work you pursued in your project.

(b) Highlight any future work.

7. **Python Code**

   (a) **Do not include Python Code within your report.**

   (b) Submit a separate file for your code (i.e. a Jupyter Notebook).