

## Course Project

### Introduction

The purpose of the course project is to provide students an opportunity to study and work on a topic in-depth so as to obtain experience with either (1) practical experience in implementing **Neural Networks** in a real system, or (2) application of Neural Networks to develop an interesting novel application system. **This is a group project with a maximum team size of Three students. Also, COMP 5630 students can only partner with another COMP 5630 student, while COMP 6630 students can only partner with COMP 6630 students.**

### Grading Criteria

Your project will be graded primarily based on the following weighting scheme:

- Project Proposal Presentation - 10%
- Project Proposal Report - 20%
- Project Final Presentation - 30%
- Project Final Report - 40%

### Requirements and Topics

You will work on a multi-class classification problem, e.g., image classification, text classification, etc. which can be solved by **Multi-Layer Perceptrons (MLP)**. **For backpropagation, you cannot use any library functions, i.e., you have to implement the backpropagation algorithm by yourself.** You will design and implement MLP in the context of a real-world problem and collect/build the necessary training/testing dataset as part of this project.

Some topics of interest where MLP-based classification may be applied are given below:

- Recommendation Systems
- Image Classification
- Search & Information Retrieval
- Social Network Analysis
- Sentiment/Emotion Classification
- Health Document Classification
- Legal Document Classification
- Spam filtering
- Customer behavior prediction
- Ad click-through rate prediction
- Song Genre Classification
- Product categorization
- Malware classification
- Anomaly detection problems such as fraud detection

## **Project Proposal Report**

You must write a two-page proposal before you actually begin your project in-depth. These will be submitted via Canvas. In the proposal, you should (1) address the following questions, (2) include the names and email addresses of all the team members, and (3) identify the coordinator of the project in case the team has more than one member, who would take the primary responsibility of coordinating the work of all team members; the coordinator is also our primary contact for providing feedback about the project.

As long as these questions are addressed, the proposal does not have to be very long. A couple of sentences for each question would be sufficient. You should focus on the following in your proposal:

- Define clearly what problem will you work on. (What is the primary application domain? What are the inputs? What are the outputs?)
- How MLP can be applied to solve this problem?
- Briefly identify who the users of your classifier will be, and why it is important to implement this classifier.
- What are the potential challenges you may face in this project?
- What is your dataset? How will you create/build your dataset?
- How do you plan to build it? Identify what technologies you plan on leveraging to implement your software. This may be programming languages, supporting libraries, etc.
- What hyper-parameters will be involved in your classifier? How would you fine-tune these hyper-parameters?
- How will you demonstrate the usefulness of your classifier?
- Provide a rough timeline to show when you expect to finish what. List a couple of milestones if possible (they can be tentative).

## **Project Proposal Presentation**

This will be a 10-minute presentation based on your project proposal report.

## **Project Final Report**

At the end of the semester during the final exam week, every project team must submit a project report (10 pages minimum with 11 pt. font). The main artifact to deliver here is the software itself, which will make up half of the grade for your report.

You, however, must also submit a written report. This report should serve two purposes: (1) it should provide a description of the MLP algorithm implemented to solve the task in sufficient mathematical details (with backpropagation calculation formulas), and (2) it should provide long-form documentation in the form of a tutorial that describes how to use your software. Your

software description should be approximately 2-3 pages and should describe implementation challenges.

Please also provide benchmarks of your classifier against other methods (either other implementations of the same method, or implementations of a similar method or baseline; you can use any libraries for the baseline methods). Clearly describe your evaluation metric and the formulas used to compute that metric. Based on the results, summarize your work, draw conclusions if possible, and discuss how you think the work can be further improved/extended.

In your documentation, you should provide enough details that a new user could utilize your software. There is no page limit here, but you should attempt to be as detailed as is necessary, but not overly verbose. Your documentation should be practical. (No, your code is not documentation.) We should be able to follow your tutorial and actually be able to use your software.

Because of the nature of the final deliverable for such a project, a strong emphasis will be placed on your actual software itself. You will need to provide a link to your project, which should be made available as open-source software under an appropriate license of your choosing (e.g., GPL, Apache, MIT, etc.).

**Grading.** We will focus on four main areas: (1) [20%] Problem description (clarity, completeness); (2) [20%] MLP theory and implementation (clarity, correctness); (3) [20%] Software documentation (clarity, completeness); (4) [40%] Results of multiple benchmark methods and discussion of findings/observations.

Each project group needs to submit a single report. **Please make sure to include, for each member of your team, a couple of sentences to describe what he/she did exactly for the project. If you are a 6630 student, you need to write the report using ACM Sigconf Latex format.**

## **Project Final Presentation**

This will be a 20-minute recorded presentation (Panopto Recording) based on your project's final report.