
CSE 676 Deep Learning (Group Project)

Abstract

This project aims to help students to develop **team skills**, learn how to use **git and GitHub**, familiarize themselves with **machine learning pipelines**, reinforce the material **learned in class** and more.

1 Overview of Projects

The projects are open end, meaning you can pick any AI algorithm to solve the problem, and each group will pick **one** project from the projects ¹ provided. The projects will help reinforce the knowledge learned in course in order to have a better outcome.

Students are expected to work with each other, **equally** and **responsibly**, from the beginning of the second week of course (**Sep 5th, 2023**) till the day of presentation (**Dec 5th, 2023**). There will be points deducted if you miss the presentation and the submission deadline (**Dec 5th, 2023 11:59pm**).

The theoretical side of the course aim to have student understand different AI algorithms, hence a **report** is a must for evaluation and the application side of the course requires students to be comfortable with the deep learning framework, **PyTorch** and the programming language, **Python**.

Language, tools, resources you may need:

- **Python**: Python Tutorial
- **PyTorch**: PyTorch Official Tutorial
- **Git and Github**

2 Theoretical Side: Report and Report Grading Rubric

In order to see if the students have a good understanding of the algorithm, a report is a must. The report will be 5-7 pages, no more than 7 pages. The student **must** use this template to submit the report, failed to do so will get a **10pt** deduction. Any form of plagiarism is not allowed.

The report **must** follow the this exact format and title:

- **Title**: Group "Number": Project Name (*5pts*)
 - **Author Name** (Name of Group Members)
 - **Abstract** (Short Introduction of the Project)
- **Section 1**: Dataset (*10pts*)
 - Include a description of the dataset.
 - What data engineering used for the dataset?
 - and more ...
- **Section 2**: Model Description (Your AI Algorithm) (*25 pts*)
 - Model Architecture

¹You can **only** choose from the projects provided

- List of Models tried and why is the one the best?
- Description and Understanding of the Algorithm
- and more ...
- **Section 3: Loss Function (20 pts)**
 - What is the loss of function you choose?
 - What other loss function you tried?
 - Innovation on the loss function?
 - and more ...
- **Section 4: Optimization Algorithm (20 pts)**
 - What is the Optimization algorithm you choose?
 - What other optimization algorithm you tried?
 - Do you implement any innovations on optimization algorithm?
 - and more ...
- **Section 5: Metrics and Experimental Results (15 pts)**
 - **Show your predicted results** (This is important in the report and presentation)
 - What is the metrics you used to measure your result?
 - Compare your experiments you have done.
- **Contributions and GitHub (5 pts)**
 - Include the contribution percentage of each teammate.
 - After the submission time, make your GitHub public and include your link for your repository. (**20 Bonus Credit** ²)

In short, convince the instructor that you have done good work *by showing good result not paragraphs of words*. Submit this report as a **PDF** on the submission deadline.

3 Practical Side: Code and Implementations

The evaluation of the code is already included in the section of the report. You will need to get comfortable with PyTorch, Python and reading the official documentations.

You are **not** allowed to use pretrained models. Another thing to keep in mind: **Discuss solutions only among your teammates**. Any form of plagiarism is not allowed.

Make sure you include the GitHub link in your report and submit your work as a group.

4 Problem Statement

Choose one of the following problem statements, solve the problem statement and turn it into a project which you can showcase in your resume. The first 4 projects can have **only 6 teams** each. Sign up your project as soon as possible here.

The **milestone** is there as a grading rubric, we will not check your milestone as you are all grown ups.

4.1 PS-1: Emotion Recognition from Facial Expressions

Objective

Train a deep learning model to recognize human emotions based on facial expressions. This can be used in a range of applications from user interface design to mental health monitoring.

²All your bonus credit will be added and contribution to 15 percent of the course. In this case, $20\text{pts} \times 0.15 = 3\text{pts}$ added to the total grade

Skills Acquired Post Project

- Computer Vision
- Deep Learning
- Python Programming

Dataset

- **FER-2013:** A standard dataset for facial expression recognition including labels for emotions such as happy, sad, angry, etc.
- and a lot more ...

Milestones

- **Week 1-2:** Collect and preprocess the facial expression dataset. (5pts)
- **Week 3-4:** Implement a basic emotion recognition model using CNNs.(20pts)
- **Week 5-6:** Optimize the model for better performance.(45pts)
- **Week 7-8:** Evaluate the model's performance and integrate it with a real-time webcam feed.(30pts)

4.2 PS-2: Fake News Detection

Objective

Train a deep learning model to identify fake news articles.

Skills Acquired Post Project

- Natural Language Processing (NLP)
- Deep Learning
- Python Programming

Dataset

- **LIAR:** Contains thousands of labeled short statements from politicians.
- **FakeNewsNet:** Contains both fake and real news articles, publicly available.

Milestones

- **Week 1-2:** Data collection and preprocessing. (5pts)
- **Week 3-4:** Implement a basic fake news detection model using RNNs, LSTMs, or Transformer models.(20pts)
- **Week 5-6:** Optimize and fine-tune the models.(45pts)
- **Week 7-8:** Evaluate the model's performance, possibly integrate the model into a web application where users can test news articles for authenticity.(30pts)

4.3 PS-3: Language Translation Service

Objective

Build a neural machine translation service that can translate text from one language to another.

Skills Acquired Post Project

- Natural Language Processing
- Sequence-to-Sequence Models
- Python Programming

Dataset

- **WMT:** An annual workshop that releases datasets for machine translation tasks.
- **Europarl:** A parallel corpus extracted from the European Parliament proceedings, publicly available.

Milestones

- **Week 1-2:** Data collection and preprocessing. (5pts)
- **Week 3-4:** Implement a sequence-to-sequence model. (20pts)
- **Week 5-6:** Optimize and fine-tune the model. (45pts)
- **Week 7-8:** Evaluate and document the model's performance, possibly deploy as a web service. (30pts)

4.4 PS-4: Audio Signal Denoising

Objective

Implement a deep learning model to denoise audio signals.

Skills Acquired Post Project

- Signal Processing
- Deep Learning
- Python Programming

Dataset

- **UrbanSound8K:** Contains 8732 sound excerpts useful for training denoising algorithms.
- **Noisy Speech Database:** Created specifically for training speech enhancement algorithms, publicly available.

Milestones

- **Week 1-2:** Understand the basics of audio signal processing and set up a noisy dataset. (5pts)
- **Week 3-4:** Implement a basic denoising model. (20pts)
- **Week 5-6:** Optimize the model and compare it with traditional signal processing techniques. (45pts)
- **Week 7-8:** Evaluate and document the project, possibly include a user interface for real-time denoising. (30pts)

5 Important Deadlines

- **Start Date:** Sep 5th, 2023
- **Submission of Report:** Dec 5th, 2023
- **Presentation Day:** Week 14 and 15

6 Final Grade

The final grade of your project is evaluated as:

60 percent of your total grade of your final report + **40 percent** of your total grade of your presentation (refer milestone) = Final Grade of the Project

7 Academic Integrity

The University at Buffalo takes very seriously its commitment to principles of academic integrity as foundational to the enterprise of teaching and learning. Our academic community expects the highest ethical behavior on the part of our students in the completion and submission of academic work. Regardless of whether instruction is in-person or remote, it is every student's responsibility to complete their work in an honest fashion, upholding both the expectations of their individual instructors as well as the university's broader academic integrity policy. Only through these practices can students maintain their personal integrity and ensure UB's tradition of academic excellence.