DS 8013: Deep Learning
Course Project Guidelines
Due April 25 (Thursday), 2023 **11:59PM**

Project topics must be approved by me. I may be able to provide some suggestions on project topics, or you can choose a topic that is interesting to you. Students must submit a one-page project proposal by March 14, 11.59pm (EST). In your proposal, please clearly specify what type of project you intend to do: Research, Replication or Empirical (see below for more information). In this one-page proposal, you should summarize the project scope with proper citations/references to the relevant studies (e.g., in a similar way to the ones that you see in this document). For instance, if you are planning to replicate/extend a previously published paper, you can first summarize the main ideas in the paper, and then provide details on your overall ideas about how to do a project based on this paper. All students should submit a final project report (as a pdf file) by April 25, 11.59pm (EST).

Here are some potential paths you can pursue for the project:

- *Research-focused project*: Incorporate deep learning and/or reinforcement learning into your current line of research, or any other research area of your interest. Another way of doing research-focused project would be to read papers in a particular area of application, experiment with different deep learning and/or reinforcement learning models and algorithms and reporting your findings.

- *Replication-focused project*: Pick a (relatively recent) paper published in a "good journal", which includes some computational work on deep learning and/or reinforcement learning. You may implement existing ideas in the paper, propose extensions, and provide computational tests and observations.

- *Empirical analysis-focused project*: Finding a practically relevant (non-tabular) data from online repos –e.g., Kaggle, UCI, timeseriesclassification.com– and performing a detailed empirical analysis with DL methods). Some examples on dataset relevancy:

  - "*Alzheimer disease classification from MRI scan images*": relevant; "*dogs vs cats image classification*": not relevant.
  - "*Product review sentiment classification*": not relevant; "*information extraction from product reviews (e.g., to learn useful product information)*": relevant.

Many nice, recent papers on deep learning and reinforcement learning can be found at arXiv.

Some research papers explored by previous students are provided below.

- Mohammadjafari et al., Using ProtoPNet for interpretable Alzheimer's disease classification (2021) [link].

- Lai et al., Modeling long- and short-term temporal patterns with deep neural networks (2018) [link].

- J. Liu et al., Dynamic pricing on e-commerce platform with deep reinforcement learning (2019) [link].

- S. Shah et al., Neural approximate dynamic programming for on-demand ride-pooling (2019) [link].

Final Deliverables

Project has two final deliverables (not following the submission guidelines will result in point deductions):

- PDF report prepared via LaTeX or Word/Google doc, named DS8013_project_report_lastName1_lastName2.pdf — lastName1 is for the first group member, lastName2 is for the second group member (if applicable).

  - You are required to convert your report to pdf if it is prepared in Word.

- Zip file of all your codes and documentation, named DS8013_project_codes_lastName1_lastName2.zip.

  - All the codes should be in python files (jupyter notebooks are not accepted).
  - As a part of documenting the implementation, you are required to include a file in your zip folder named README.md. This file should include information on the implementations, including how to run the codes (e.g., what is the main file, which arguments to pass etc), specific libraries that are needed to be implement as well as library requirements (e.g., tensorflow, keras versions etc). You can find information on how to prepare markdown files online, e.g., see
    * https://medium.com/@saumya.ranjan/how-to-write-a-readme-md-file-markdown-file-20cb7cbcd6f
    * https://github.com/awslabs/gluon-ts – a sample well prepared markdown included in a github repo!

- The clear and detailed commenting on the implementation/codes likely means higher grade!

- All the submissions (proposal, intermediate report, final report and the zip file) should be done via D2L. Project submission folders are under Assignments in the Assessment tab.

Report Format

Your written report should contain everything about the study (e.g., paper(s) selected for replication), your implementation and the experimentation process. There is no page minimum or maximum; just write enough to thoroughly describe your process, and *no more*. Remember the report is also graded on clarity and conciseness. At the very least, your report should contain the following information (the section outline is provided as an example, but highly recommended to be followed):

Section 1. **Introduction**: Explain the research problem in a detailed manner. Aim to answer questions such as

  - What is the motivation and background for the research problem?
  - Why is the proposed solution method/algorithm appropriate for the problem?

Section 2. **Literature review**: A summary of the relevant studies. A literature review is expected regardless of the project type, e.g., if you are doing an *empirical* project.

Section 3. **Methods**: A detailed description of the underlying methodology in your own words. If you proposed any modifications to the original study, what are those (be as specific as possible). Also, disclose data sources: if it is publicly available, provide references, if not, how did you generate synthetic data.

Section 4. **Experimental setup**: A detailed information on the experiments, hyperparameters used and if you performed hyperparameter tuning, which hyperparameters you experimented with.

Section 5. **Results**: A detailed comparison of your results with those of the original study (if applicable). Provide detailed explanation and interpretation of the results you obtain from your analyses. Use plots and tables whenever appropriate. Make sure the plots and tables you use are clear and easy to read (with proper captioning and referencing).

Section 6. **Conclusions**: Provide particular insights from your project. Aim to answer questions such as

  - What is the take away message from the study and the project?
  - What are the weaknesses of the study and your project (i.e., is there any strong modeling assumptions that might not be valid in practice)? What are the threats to validity?
  - What are the future research directions?

For the projects that aim to replicate a paper or paper(s), a sample report is provided in D2L (see DS8013_project_sample_replication_paper.pdf).

Grading Scheme

In grading literature review-focused projects, the grading scheme provided in Table 1 would not apply. Instead, the grading will be based on items such as "Report", "Depth of analysis", and "Insights from the studies" (extra points from "Novelty" is not applicable in this option, so choosing literature review option as the course project is less advantageous for collecting bonus points).

Table 1: Tentative point distribution over 100 pts (†: A one page proposal, describing the project scope and deliverables.)

Item	Points
Report (proposal†: 5 pts, final report: 45 pts)	50 pts
Implementations/documentation	20 pts
Soundness of the results and comparisons	30 pts
Novelty	+10 pts
Individual project	+5 pts
Nicely written report in LaTeX	+5 pts