Project Proposal Handout and Rubric

Introduction

The purpose of the project proposal is to demonstrate that your team:

- knows what the goals and motivations for your project are.
- knows what dataset you will use to train your model.
- has a rough idea of the type of neural network(s) you will use.
- has a rough idea of the related work that you can build on.
- has a reasonable idea of how you will measure the success of your model, and your project?
- has a clear idea of how you will work together, and how to distribute the work fairly?

How to submit

Submit your proposal as a group on Quercus by the Project Proposal deadline. Any work that is submitted after the deadline will receive a 20% grade deduction. Proposals submitted 24 hours past the deadline will not be accepted and result in a grade of 0. Quercus submission time will be used, not your local computer time or any other screenshots that you provide. You can submit your work as many times as you want before the deadline.

This document should be written using Latex based on the course template. There is a 9-page limit (ICLR conference standard) for the main text and unlimited references is applied. If any submission exceeds to 10 pages, a 20% penalty will be applied. We do not accept any submission longer than 10 pages (100% penalty would be applied).

Proposal Rubric

The project proposal document is graded out of 40 points.

- **Introduction (4 points)**: A brief description of the motivations behind your project, the goal of your project, why it is interesting or important, and why deep learning is a reasonable approach.
  - 4/4 An introduction that clearly describes the project goal, why the project is interesting and/or useful, and convincingly describes why deep learning is an appropriate tool for the task.
  - 3/4 The introduction describes the project but does not "sell" it well or get to the point quickly, contains overly vague phrases, or minor factually incorrect information.
  - 2/4 The introduction is vague, does not sufficiently convince the reader that the project is useful, and/or that machine learning is an appropriate tool to use.
  - 1/4 The introduction does not make it clear what the specific goal of your project is.

- **Illustration / Figure (4 points)**: A figure or a diagram that illustrates the overall model or idea of your project. The idea is to make your report more accessible, especially to readers who are starting by skimming your work. For the project proposal, taking a picture of a hand-drawn diagram is fine, as long as it's legible. PowerPoint is another option. You will not be penalized for hand-drawn illustrations - you are graded on the design and illustrative power.
  - 4/4 A well thought-out figure that communicates the core idea of your project and architecture immediately.
  - 3/4 An illustration that does the job but could use some minor improvements.
  - 2/4 An illustration that does the job, but is not particularly clear, or too wordy.
  - 1/4 The illustration is significantly lacking in some respect or contains factual inconsistencies or inaccuracies.

- **Background & Related Work (4 points)**: A description of at least 5 related work in the field, to provide the reader a sense of what has already been done in this area, e.g., papers or existing products/software that do a related thing.
  - 4/4 Briefly describes 5 prior work related to your project to put your project into context. Your descriptions need not be complete but should contain important work related to the project.
  - 3/4 Background that has minor omissions or factual incorrectness, but otherwise places your project into context.
  - 2/4 Background contains too much information not related to your project or has major omissions of content on related work done in the field.
  - 1/4 Background that does not sufficiently put your project into context.

- **Data Processing (4 points)**: A description of the sources of your data. We are looking for some effort and evidence of work to collect/repurpose/clean training data vs. simply taking a known dataset. The extent of data processing may vary from project to project, and we will grade you according to your project.
  - 4/4 Clearly describes and cites sources of data, and the steps you will take to clean and format your data. The descriptions are clear enough for another classmate to follow and reproduce.
  - 3/4 Mostly clear description, but some aspects of the data processing steps are vague enough so someone attempting to reproduce your project might do something different enough to possibly affect model goal or performance.
  - 2/4 Mostly clear description, but aspects of the data processing step are incomplete.
  - 1/4 Descriptions are unclear and minimal information is provided on the dataset or proposed preprocessing steps.

- **Architecture (2 points)**: A rough description of the type(s) of neural networks that you will use. For example, RNN, CNN, etc. You won't know all the information about the number of layers and hyperparameter choices at this point, so it is okay for the description to be rough. However, you should provide enough details so another classmate could potentially take over your project.
  - 2/2 Rough description of the type(s) of neural network model(s) that you will use, and the relevant components.
  - 1/2 Some issues with the description (inconsistencies, factual inaccuracies)
  - 0/2 Unclear description of the type(s) of neural network model that you will use, or a choice that is inconsistent with your problem.

- **Baseline Model (2 points)**: Describe a simple, baseline model that you will compare your neural network against. This can be a simple machine learning model (e.g., SVM, Random Forests, etc.), a hand-coded heuristic model (that does not use machine learning), or something else. Expectations for the baseline model will vary from project to project.
  - 2/2 A reasonable choice of baseline, accompanied by a description of the baseline so that a knowledgeable classmate can either find or reproduce.
  - 1/2 A description of a reasonable baseline, but not enough detail to reproduce a version that would perform similarly.
  - 0/2 Poor choice of baseline inconsistent with the problem.

- **Ethical Considerations (2 points)**: Description of a use of the system that could give rise to ethical issues. Are there limitations to your model? Your training data?
  - 2/2 Thoughtful consideration of ethical issues in data collection, and the impact of using the model.
  - 1/2 Some consideration of ethical issues in data collection.
  - 0/2 Ethical issues in data collection are not presented.

- **Project Plan (4 points)**: Provide a table summarizing the tasks each person is responsible for and has accomplished so far, along with clear deadlines for each member. Describe how your team will work together. When will you meet? How will you communicate with each other? How will you ensure that you won't overwrite each other's code? Divide up the tasks and set deadlines for yourselves.
  - 4/4 Plan provides enough detail about the breakdown of tasks, internal deadlines, and team member responsibilities so that a new team member can replace an existing one and know roughly what their responsibilities are. Work is divided evenly amongst team members.
  - 3/4 Plan lists the breakdown of tasks, internal deadlines, and team member responsibilities, but could provide more clarity on some member's responsibilities.
  - 2/4 Plan lists the breakdown of tasks, internal deadlines, and team member responsibilities. Work is not divided evenly amongst team members or team member's responsibilities are not clearly defined.
  - 1/4 Plan misses important tasks, or has timelines that are obviously impractical (e.g., setting aside only 1 day for hyperparameter tuning).

- **Risk Register (4 points)**: Document 3-5 major/likely risks of the project, the likelihood of the risk, and what your team would do in the situation. For example, what would you do if a team member decides to drop the course? What would you do if your model training took longer than expected? Not all risks make sense for each project, so think about the risks involved with your project.
  - 4/4 A thorough analysis of the major risks, and their solutions. The risks are well aligned to the proposed project.
  - 3/4 A thorough analysis of the major risks, and their solutions. Some risks are not aligned or don't make sense for the proposed project.
  - 2/4 Good analysis, but with possible omission of serious risks.
  - 1/4 Major miscalibration of risks.

- **Link to Github or Colab Notebook (1 point)**: You are expected to store all your code in either a Colab notebook or a GitHub repository. If your team is uncomfortable with having your code public, please use a private repo and share it with your team members and the teaching team.
  - 1/1 Includes link to public Colab notebook or public GitHub repo.
  - 0/1 Omitted.

- **References (1 point)**: References should be in the templates bibTex style.
  - 1/1 At least 5 references, formatted consistently using the standard format.
  - 0/1 References are omitted or inconsistent (e.g., some use first names, other uses initials).

- **Structure, Grammar & Mechanics (8 points)** : We are looking for a document that is easy to follow, grammatically correct, and well-written.
  - 8/8 Clear, concise and well-written document. Exceeds expectations.
  - 6/8 Well-written document that has some issues with grammar, mechanics, or structure. Meets expectations.
  - 4/8 Reasonably written document with grammar, mechanics, or structural issues.
  - 2/8 Document has many issues.