CHEE 6397 Project Proposal Instructions

Submission

- Due: Oct 19, 2023
- Upload a pdf to Canvas
- Each group only needs to submit one proposal

Contents

The proposal should include the below sections:

- Title
- Group members
- Background: Brief review of the problem you want to solve and the motivations.
- Dataset: Answer questions like: What dataset(s) you will use? How many data points? What are the features? What is the learning target?
- Plan of Work: What is/are the machine learning algorithm(s) you plan to use? What are the steps you will take to solve the problem?
- References: Optional.

Formatting

- 1 page maximum
    - excluding references (if any)
    - an optional figure can be included at the end, and it does not count towards the 1 page limit
- 1 inch margins
- 11 point font: Times, Arial, or Helvetica
- single-spaced

Use of LaTex is encouraged, but not required

- It is a good opportunity to learn LaTex if you are not familiar with it.
- You can install LaTex on your laptop, or use online LaTex editors such as Overleaf. The latter is recommended -- it is free and supports collaboration.
- This is optional; you can use MS Word and others.

Where to find datasets?

Periodic crystals

- MatMiner: https://hackingmaterials.lbl.gov/matminer/index.html

- MatSciML: https://github.com/IntelLabs/matsciml

Molecules

- Quantum machine learning: http://quantum-machine.org/datasets/

Others (not related to materials and molecules)

- Kaggle datasets: https://www.kaggle.com/datasets
- UC Irvine ML repository: https://archive.ics.uci.edu/datasets
- Scientific Data: https://www.nature.com/sdata

Google datasets search

- https://datasetsearch.research.google.com