CSCI 5090/7090 Final Project
Mehdi Allahyari

1 General Guidelines

The final project is an opportunity for you to try machine learning techniques on interesting real-world problems. What we expect to see is an idea (task) that you clearly describe, implement and test on a dataset. Projects can be carried out individually, or in teams of two students. You can discuss your ideas and approach with the instructor, but obviously the final responsibility to define and execute an interesting piece of work is yours. The final project is worth 20% of your grade, which will be divided in two deliverables: *proposal* and *final report*.

Final projects will be evaluated by the following three criteria:

1. Technical Depth: How technically challenging was what you did?

2. Scope: How broad was your project? How many aspects, angles, variations did you explore?

3. Presentation: How well did you explain what you did, your results, and interpret the outcomes? Did you use the good graphs and visualizations? How clear was the writing?

2 Project Proposal

You must turn in a project proposal (hardcopy and electronic copy) on **February 14th**. The hardcopy must be turned in at the beginning of the class and the pdf file on Folio.

Project proposal format: Proposals should be one page maximum. Include the following information:

- Proposal title

- Dataset

- Project idea. This should be approximately two paragraphs.

- Software you will need to write.

- Papers to read. Include 13 relevant papers. You will probably want to read at least one of them before submitting your proposal.

- Teammate: will you have a teammate? If so, whom? Maximum team size is two students. One proposal per team. We expect projects done in a group to be more substantial than projects done individually. **Note: graduate students must work on individual projects.**

3 Final Report

You will submit your final report on **April 25th**. You should submit both the hardcopy at the biginning of the class and electronically via Folio. Name your electronic submission *Project-your-group-number.pdf*.

Your write-up should 8 pages maximum including the figures, tables and references. Your submission must include at least two figures which graphically illustrate quantitative aspects of your results, such as training/testing error curves, learned parameters, algorithm outputs, etc. Your submission should follow the generally accepted style of paper writing:

- Introduction (motivation)

- Problem definition

- Proposed method:

    i. Intuition− How it compares with previous work.
    ii. Desription of your approach

- Experiments:

    i. Details of your experiments; observations and resutls

- Conclusion: summary of what you discovered

- References: at least 3 references to previous published papers or book sections.

4 Project ideas

You are encouraged come up with your own ideas for projects. For instance, you can look for data for a problem that you are interested in and try out different machine learning algorithms to predict something interesting about the data.

You may also look at the various current and past competitions on Kaggle (`https://www.kaggle.com/competitions`) for inspiration.

You are welcome to explore and expand upon project ideas listed below or use these as a starting point for brainstorming ideas.

- courtesy: Andrew Ng, link

- courtesy: Carlos Guestrin, link

- courtesy: Vibhav Gogat, link

- courtesy: Ilyas Cicekli, link

- courtesy: Tom Heskes, link

- courtesy: Rob Schapire, link

- courtesy: Jason Brownlee, link