

CSE 5693 Machine Learning: Term Paper/Project
Due Apr 2 (Wed), 6:30pm; Canvas: Term project plan
Due Apr 30 (Wed), 6:30pm; Canvas: Term project

The term paper is based on a mini-research project. The goal is trying to improve one of the learning algorithms you have studied. You may also devise new algorithms. Empirically you will compare at least:

- your proposed improvement
- the original algorithm in the book
- (optionally/preferably) a proposed improvement in a research paper (journal or conference)

Resources for research papers: <https://cs.fit.edu/~pkc/classes/ml/resources.html>

A. Plan (10 points) [Due Apr 2, Wed]

About one page, ACM format templates for *conference* proceedings at <http://www.acm.org/publications/proceedings-template> (ignore Abstract, Categories and Subject Descriptors, General Terms, Keywords, and Copyright)

1. Introduction: motivation and problem statement
2. Related work: cite 3 research papers you plan to read OR compare/contrast existing approaches in at least 3 research papers and discuss limitations
3. Approach: your initial concrete idea, including diagrams, pseudocode, and/or equations
4. Evaluation: data (at least two sets) and criteria

References: list of at least three related research papers (including authors, title, venue, year, page numbers)

B. Paper (80 points) and presentation (10 points) [Due Apr 30, Wed]

Paper: at least 3 pages, ACM format templates for *conference* proceedings at <http://www.acm.org/publications/proceedings-template> (ignore Abstract, Categories and Subject Descriptors, General Terms, Keywords, and Copyright)

Presentation: about 15 minutes

1. Introduction: motivation and problem statement
2. Related Work (at least three research papers; compare/contrast existing algorithms and discuss limitations)
3. Approach (your idea: why it can overcome limitations of current techniques and exactly how it works, including diagrams, pseudocode, and/or equations)
4. Empirical Evaluation
 - 4.1. Evaluation criteria
 - 4.2. Experimental data and procedures (description of the data, procedures include: preprocessing, parameters used)
 - 4.3. Results and analysis (results in graphs/tables and analyzing the results) [compare with the original algorithm in the book]
5. Conclusion (summary of findings, limitations and possible improvements)

References (cited in the text; including authors, title, venue, year, page numbers)

Submission:

- (a) paper (preferably in pdf)
- (b) presentation (preferably in pdf)
- (c) source code of the algorithms
- (d) source code for different experiments (testX scripts/programs/functions like the homework assignments)
- (e) datasets [if they are too large for Canvas, put them at a website and include their links in README.txt]
- (f) README.txt on how to compile programs and run the different experiments on code01.fit.edu

Your paper will be evaluated mainly on the sophistication/innovation of your algorithm, the performance of your algorithm against the original book algorithm (and possibly an algorithm in one of the research papers), and your analysis of the results.