

Course Project, COMP / ELEC / STAT 502 Neural Machine Learning I.

A Project will be done over an approximately four-week period in the second half of the semester. Approximate dates are given below. All pertinent final dates will be posted in the Course Schedule at www.ece.rice.edu/~erzsebet/ANNcourse.html, in a timely manner.

Students will form groups of by self-organization. Group sizes will typically be 2 – 3 and will be determined depending on class size. Each group will submit a project proposal around October 27-28. It is anticipated that by that time the major ANN paradigms will have been discussed in class and students will be able to identify a project topic, either suggested by me (including data), or devise their own. I encourage addressing a problem that is related to your graduate research if applicable, and to talk with me *early*, especially about [data](#). I will review and approve each proposal (or request necessary modifications) in the order they are submitted, in scheduled project reviews with me. Projects can start as soon as I approve them. We will have one or two project presentation sessions depending on the number of groups, on the last week of classes. Project reports will be due on the last day of classes, at 11:59pm. An electronic copy of both presentation and project report will have to be submitted in Canvas. Exact scheduling and other details will be discussed in class.

Characteristics of a Successful Project

1. It will be an analysis of the behavior of an ANN algorithm (implementation as needed); or it will evaluate the application an ANN algorithm or compare a group of algorithms on a realistic data analysis problem.
2. It will be demonstrated as working, via figures, tabulated results, code, description, evaluation and discussion of the quality of the results, meeting project objectives according to declared success measures.
3. It will be written up in a *concise* (approx. 10-page) report that shows professional level of technical thoroughness, writing style, grammar and neatness.
4. It will be presented in a brief, approximately 10-minute, talk.
5. *The focus will be on ANN paradigm(s) that we studied in class, and analysis will be based on code you(r group) implemented. While for comparison of algorithms you may use canned Matlab or other algorithms, you will have to demonstrate understanding and control of those in similar terms as we require in class for the ANN-s we cover.*

Project Proposal

The Project Proposal will have the following format:

- **Statement of Problem:** a brief, one-paragraph statement indicating what the problem is that you propose to address, and why (what is the significance of it)
- **Objectives:** a brief statement of what you expect to achieve in relation to the Statement of Problem, e.g., a working algorithm, an analysis of algorithm behavior, a demonstrated analysis of data (classification, model fitting, information discovery / clustering, performance comparison, etc.), and what success measure you use.
- **Data Description:** a detailed description and demonstration of complete understanding of the data that you plan to use. Sufficient specific information will be required as under [data](#) to evaluate the feasibility of your approach to solving your stated problem using the proposed data. You will have to have the data in hand and have examined it in order to provide this information!
- **Technical Approach:** a brief outline of the *specific methods and techniques* that you plan to employ to achieve the Objectives, and how they will be used to help achieve the Objectives. This includes ANN paradigms we covered, ANN and other methods (e.g., classifiers) that you intend to use from a package, and a plan of experiments to gain the insight you are seeking.
Plan a well-focused, in-depth inquiry rather than a superficial, large set of simulations.

The Project Proposal should be short, no more than 2-3 pages including data description.

Focus on demonstrating neural learning achievements. While relevant steps of data preparation must be mentioned briefly, your project should not be dominated by data collection and transformations instead of ANN analysis.

Project Report

The Project Report should have the following components:

- Statement of Problem – as in the Project Proposal.
- Objectives – as in the Project Proposal, including a success criterion/measure.
- Data Description – as in the Project Proposal, possibly expanded to reflect any particular challenges for machine learning (based on theory or experimentation).
- Technical Approach – as in the Project Proposal but expanded as necessary to provide enough details to understand the actual processing steps, the quality and performance evaluation, and judge the merits (soundness, correctness).
- Results – Document and discuss the results achieved, in comparison to the initial objectives. Discuss how closely you met (or did not meet) your success criterion.
- Code and data – Provide the code you wrote and the data set. (In case of proprietary data please discuss with me.)
- Optional Appendices – containing, e.g., extra details of data or figures, demo, web links, etc., if desired. The report should stand on its own without Appendices.

The length of the Report should be what is *necessary* to clearly and professionally document your effort. Aim at 10-11 pages not counting appendices, and under no circumstance exceed 12 pages (single line spacing, 12pt font size, 1” margins). Appendices should not be needed for clear understanding and evaluation of your project. (I will be delighted to look at Appendices but won’t use them for grading.)

Project Presentation

Projects will be presented to the class in 10– to 12-minute time slots. PowerPoint or Beamer slides (or anything that produces nicely formatted pdf output for submission) can be used. Feedback from presentation can be used to improve the Project Report.

Deliverables

Electronic copies of the Project Presentation, Project Report, and Code/Data will be submitted to Canvas, similarly as with HW-s and Exam. Appendices are not required but may be submitted in a separate file or files (as appropriate for codes or demos, for example) but no more than one appendix for extra figures or demo if you so wish.

Project Grade

A project will have a maximum total score of 100. The front page of the project report must contain a statement, signed by each team member, indicating how much of the total percent of work was performed by each student. The total score earned by the project x (# of contributors) will be apportioned across the students according to the stated percentages. The project score will be composed as a weighted average of the scores of the Presentation and the Report (ratio TBA, Report with larger weight). We peer grade the Presentations real-time! Both will be evaluated for technical merit of the work, and format.

The more challenging and creative the project, the better the chance of receiving a high score. Thus, it may be better to partially fail at a challenging goal and be able to analyze and explain the reasons than to succeed at a trivial one.

I encourage you to ***start thinking about potential projects and data sources right at the beginning of the course***, and discuss with me any questions. See a [sample checklist](#) for describing a data set for your project (and for machine learning tasks in general).