Evaluation of ECE 271B Project

This document provides some guidance on how your projects will be evaluated. Beyond evaluation of your mastery of the class material, the goal is also to introduce you to the standard practices in research.

Introduction

The first thing to know is that a paper needs a hypothesis. The reason to do the work is to test the hypothesis. For example, the hypothesis can be that boosting is more suitable to classify EEG signals than the SVM. A better hypothesis is that boosting is a state of the art method for classifying EEG signals. The more sophisticated your hypothesis, the more people will care about your paper. For publication in a good conference or journal, the hypothesis must contain a fair amount of novelty. The hypothesis is usually that the method you are proposing is state of the art. We will not require this in the class, but will obviously reward novelty. If you are doing something more original, your grade will be higher. Once you have a hypothesis, you usually need a fair amount of technical work to study it. For example, you may need to derive a boosting algorithm based on a loss that is preferred for EEG signals. This is where you get to use some mathematics. Again, the more sophisticated the mathematics, the more likely people will be interested in the paper. But math is not everything. A paper with an exciting but simple hypothesis is always better than a paper full of math to support a boring hypothesis. Finally, you must test your hypothesis. Here, good papers have three types of experiments.

- Insight: a set of simple experiments that provides insight on your method. Does it do the right thing? This can be tested with synthetic data that conforms to the assumptions used to derive the method. It is the only set of experiments where you can use synthetic data. Everything else should use real data. This set of experiments is optional, if the following ones already provide all the insight on how the method works.
- Parameter tuning: in this section, you study the importance of any parameters in your method. For example, what is the best number of weak learners for the data you are using? You compare performance with multiple parameter values and report. Then, you choose the best parameter values. All methods have free parameters and you should always let the reader understand how they affect the performance of your method.
- Validating the hypothesis: once you have the best parameters for your method, you need to test the hypothesis. Is it really the best solution to the problem you are studying? To prove this, you must compare to baselines. These are solutions that already existed for the problem. For a real paper, you need to compare to the best methods in the literature and show that your method is better. For the class, we will be less ambitious. You do not need to show that you have the best method. However, you need to find baselines (i.e. other solutions) and compare your method to them.

Evaluation criteria

The first thing to know is that your paper will be evaluated "on the curve," that is against the other papers in the class. We will evaluate papers along 6 dimensions:

- writing (10 pts),
- creativity (20 pts),
- thoroughness (20 pts),
- soundness (10 pts),
- experiments (30 pts),
- references (10 pts).

Writing
Two components of writing will be evaluated. First, does the writing follow the standards of scientific publishing and does the paper provide sufficient rationale, is clear, etc.? Second, issues of grammar, spelling, formatting, etc. Some of these issues can be addressed automatically by using standard tools. For example, using the provided LaTeX style files takes care of the formatting issues. Since this is not a writing class, we are not looking for writing masterpieces. If you do a professional job, you will get full marks.

Creativity
While you are not required to go beyond the class material, many students do. This component of the grading is meant to reward those efforts. If your project reduces to the implementation of a few methods that we have seen in class, it will be average in terms of creativity and will receive half of the creativity grade. There are two ways to increase your creativity score. The first is to use what you learned in class to build some creative system. The second is to investigate techniques beyond what is taught in class. For example, you can extend the boosting algorithm by considering different weak learners or introducing a new loss function. Note that simply using a neural network to extract features and then feed these to a standard boosting algorithm does not really qualify as very creative, since this is widely used. An effective way to enhance your creativity score is to find the top conference in the research domain you are interested in and read papers from the last two years. You should find some state of the art procedure that is interesting and would be fun to experiment with. If you can improve on its performance, you may even be able to publish your work. Who knows? Notice that if you use some papers from the literature, you should clearly cite them. Do not try to pretend that the idea is yours, or fail to omit that it is not. This is an example of academic dishonesty that violates the UCSD Academic Integrity policies.

Thoroughness
Thoroughness addresses the testing of your hypothesis. If your project was on one class topic, did you compare against the other methods that we studied? Beyond that, did you compare to standard baselines in the literature? What is the most popular solution to the problem and is your method better than it? You will not be penalized if it is not, but you will be penalized if there is a sense that you did not bother testing your method thoroughly. For many problems that you would work on there are datasets available. These datasets have leaderboards, i.e. the top performers on them are known. Again, if you read recent papers on the research that these datasets support, it will be clear what these methods are. If you are building a new system, how many different solutions did you try? Did you compare their performance? In summary, did you make a convincing effort to test your hypothesis?

Soundness
This criterion evaluates the mathematical soundness of your paper. You do not need to invent new math and you will not even be rewarded for that (it will be already accounted for in your creativity score). If your presentation is mathematically sound, you will get full marks. On the other hand, if it feels like you do not really understand the issues, or you cannot make a sound presentation of your arguments, your score will be reduced.

Experiments
This criterion addresses the soundness of your experiments. Did you conduct any insight experiments? Were you thorough in your parameter tuning? Did you compare your method to various baselines, alternative methods from the class, state of the art approaches to the problem? Since there are multiple students per group, there is no excuse not to implement multiple methods, perform thorough parameter tuning, and obtain valid experimental evidence. Also, make sure that the work can be reproduce from the information given in the paper. This is what this component of the score measures.

References
In academia, it is very important that you give proper credit to the original developers of each idea. This component of the score evaluates the thoroughness of your references. Did you bother finding out who were the original proposers of the algorithms you are using? Are you referencing the original papers, not some derivative from books, etc.? Is your citation list exhaustive?