Thanks for your comment. However, I do not believe those are valid judgements of my report or work, and my work deserves better score that this. Below is my response for the comments.

Nice job mentioning how you preprocess your data, how you decided to do or not do cross validation. In terms of leaving 4 classes out of your original 8 to avoid the effect of noises, are there other ways to make it happen without losing labels?

Response: Yes I am aware there are many resampling techniques. In fact, I’ve used several in my work. The reason for not including resampling in this assignment is that 1) resampling is a non-trivial task. It is a trial-and-error process and how to defend the new samples is always a challenge. 2) It is not the purpose of this assignment, and removing some labels does not affect the conclusion in this report.

You should be looking at testing vs cross validation data when you tune and analyze your parameters. And since you are solving for classification, you should include metrics like type 1, type 2 errors, confusion matrix, or ROC curves, into your analysis for performance.

Response: First, Precision, Recall and Confusion Matrix are all in my jupyter notebook. I took those out from final report because: 1) without knowing the meaning of features and labels, it is difficult to justify which metrics should be used, so why not pick just one and be consistent for the rest of the report? 2) I need to squeeze everything into 12 pages. With current plots, I don’t think I am able to put more into the report.

More reasons to not skip cross validation. You need to plot and analyze more than one key parameters for each of the learners, discuss and report your findings in terms of model performance, bias or variance, the effect on training/cross validation error, etc, in addition to documenting what values you find best from your experiments.

Response: Yes I agree there is a lot can be done to make this report a more in-depth analysis. However, I don’t remember in the assignment description there is a request for ‘more than one key parameters for each of the learners’. Moreover, grid search is a way to analyze more than one key parameters, which I explicitly mention in my report.

The TA might’ve missed that, but my report is written in a way to ‘report your findings in terms of model performance’, and yes in my report there are many, many paragraphs talking about ‘effect on training errors’.

Experiments on testing and training split, though it might have an effect on model performance, it's not a key parameter to an algorithm.

Response: This is what says in the assignment description: *‘the training and testing error rates you obtained running the various learning algorithms on your problems. At the very least you should include graphs that show performance on both training and test data as a function of training size (note that this implies that you need to design a classification problem that has more than a trivial amount of data)* ’. Maybe my understanding is wrong?

Nice comparison for the learners in summary, as well as quick notes about each of them. You have the general right ideas and directions for this assignment, and you've made some good observations for the algorithms, but your analysis is either light or is missing the target. It is clear you know how to conduct experiments and gather data from it. Keep up the good work, start early next time, double or triple check what's required, and I look forward to your next submission.

Response: I started this assignment 3 weeks before the deadline, and if you take a look at my jupyter notebook you will see how much effort I put in this. Yes, I also read the assignment description multiple times to understand what is really wanted. I feel really disappointed for not being credited for the efforts I put on this assignment and many questions I answer in the report, but being severely penalized by not including analyses that are either not the focus of this assignment, unclear in the assignment description, or left out because of the page constraints.