Lab 4 GSI grading

This form contains the final scores and comments from the GSI.

The respondent's email address (**rebeccabarter@berkeley.edu**) was recorded on submission of this form.

Name of students in group *

Jordan Prosky, Hongxu Ma, Alexander Brandt

Readability

Readability of report (5 points) *

	1	2	3	4	5	
Narrative unclear and/or difficult to read	○	○	○	○	◉	Narrative very clear and/or easy to read

Grammar of report (5 points) *

	1	2	3	4	5	
Incorrect written grammar pervasive	○	○	○	○	◉	Excellent written grammar

Comments about readability

EDA & model choices

Exploratory data analysis *

	0	1	2	3	4	
Did not provide any exploratory figures or numerical summaries of the data	○	○	○	○	◉	Provided clear, relevant figures and summaries of the data

Comments about EDA

Very reasonable discussion on NDAI and relating SD to the level of confidence of labelling.

Your density plots are very clear.

Justification of variable selection *

	0	1	2	3	
Provided no figures, justification or discussion of variable selection	○	○	○	◉	Described clearly and thoughtfully which figures are best and provided insightful figures

Comments about variable selection

I would caution against using all three images in your feature selection! Perhaps one image should have been withheld entirely to test performance at the end.
Your discussion on why you chose NDAI, SD and Corr was very reasonable though.

Appropriateness of prediction methods *

	1	2	3	
Did not discuss appropriateness of methods chosen	○	○	◉	Clearly outlined the assumptions and reasons for choosing each model

Comments on classifiers

good job considering the assumptions of each method.

Model performance

Depth of exploration concerning model fit and convergence *

	0	1	2	3	4	
Did not discuss model fit	○	○	○	◉	○	Clearly described how well each model fit from a variety of different angles. Provided informative and high-quality visualizations

Thought about how to appropriately use cross-validation *

	0	1	2	3	
Did not consider CV carefully	○	◉	○	○	Clearly outlined that pixels should be grouped in some way when doing CV

Comments on model fit and convergence

I thought your quantitative assessment of model performance and parameter tuning was pretty good but it would have been good to keep one of the images for validation at the end. It doesn't seem as though you considered the implications of using standard cross-validation (e.g. by grouping nearby pixels together). Obviously you're going to do better if there are pixels in the test fold that are surrounded by pixels in the training folds. This would not be realistic if you were fitting the model to a completely separate image.

It probably would have been a good idea to make the ROC curves on a withheld test set.... An AUC of 1.0 is unrealistic...

Depth of exploration on patterns in misclassification errors *

	0	1	2	3	4	
Did not explore patterns	○	○	○	○	◉	Clearly explored and visualized patterns in misclassification errors

Comments on patterns in misclassification

Nice plots and attempted explanations of possible phenomena causing the misclassification patterns!

Justification of using model on future data *

	0	1	2	3	
Did not justify answer to whether or not the model would work well on future data	○	○	○	◉	Clearly explained why or why not the model would work on future data

Comments on using model on future data

Your discussion is clear and reasonable! "another cross-validation scheme that could have been used to perhaps create a more robust model is creating grids within each image and using a set of grids as folds for CV" -- this would have been ideal for sure!

Reproducibility

Everything was provided in order for reproducibility *

	0	1	2	3	
Did not provide all files needed	○	○	○	◉	Provided all files necessary and clearly labelled how to reproduce all analyses (i.e. which files produce what and how they all fit together)

Comments on reproducibility

Conclusion

One or more things that were well done

Your report was well done and the comparison between the performance of the methods was very clear and thorough

One or more things that could be improved upon

It is always a good idea to withhold data (e.g. an image) entirely to show performance of your final model on an independent dataset.

Other comments