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 *

Olivia Angiuli, Miyabi Ishihara, Yizhou Zhao

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 clear figures!

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

Some more quantitative methods could have been used to identify which variables have the most predictive power.

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

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

Good idea to use entire images for the CV folds, but this does mean that you can't withhold an entire image as a final untouched validation/testing set. Your final model will have used all of the data you have! Another option is to segment one or two of images into grids and use the panels of the grid as the CV folds.

It is hard to compare the individual methods to one another the way you have presented them. It would have been good to see a plot that compared LDA, QDA, RF, etc.

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

Nicely done with your Figures 10, 16, etc and good job exploring where the misclassifications occur with respect to the actual variables themselves. (fig 15)

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

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 figures were very clear (if with a bit of overplotting). Everything was explained well.

One or more things that could be improved upon

One idea is to consider adding a smoothing layer on top of your final classifier which takes the predicted labels and smooths them so that pixels next to one another are generally required to have the same label.

Other comments

Google Forms