

Final Project Grading

Name of student

Yizhou Zhao

Student ID

3032130362

Readability *

	1	2	3	4	5	
Narrative unclear and/or difficult to read	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Narrative very clear and/or easy to read

Grammar *

	1	2	3	4	5	
Incorrect written grammar pervasive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Excellent written grammar

Comments about readability and grammar

Data Partitioning 

Appropriateness of data partitioning into training (choosing parameters), validation (choosing the best model), and testing (estimating performance on independent data) *

	0	1	2	3	4	
Did not partition the data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Correctly partitioned the data

Comments on data partitioning

The validation set is different from the left out fold of CV! CV should be done entirely on the training set to choose parameters then the validation set should be used to choose between all of your models and the final test set should be used to give an unbiased estimate of your final model performance.

EDA and introduction

Quality of exploratory data analysis *

	0	1	2	3	
Did not perform EDA	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Performed a thorough EDA and presented appropriate and appealing figures that highlighted the interesting parts of the data

Comments on EDA

It's tough to make 3D plots on a 2D surface but I think you did a pretty good job with Fig 2 :)

Fig 3 is nice too -- very simple.

It would have been good to see how the voxels are related to one another (e.g. there are some inherent groupings that can be seen in a correlation plot).



Appropriateness of regression methods *

	0	1	2	3	
Did not appropriately choose or implement regression methods	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Made sensible decisions in terms of choosing and implementing regression methods

Comments on regression methods (e.g. did the student try to fit the same model for each voxel or different models for each voxel - this makes more sense)

I like your choices of models but I would have liked some more context. Are you using the same model for all voxels? Will you try to obtain different regularization parameters for every voxel?

Explained each of the model selection criteria (CV, ES-CV, AIC, AICc and BIC) *

	0	1	2	3	
Did not explain the model selection criteria	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Clearly outlined what each model selection criteria does and the relative pros and cons of each criteria

Correctly implemented and compared model selection criteria (including using the correct correlation criteria rather than MSE) *

	0	1	2	3	4	
Did not compare model selection criteria	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Correctly implemented the criteria, discussed strengths and weaknesses, and provided insightful figures for the comparison

Comments on implementation and comparison of modeling and model selection criteria

figure 4 is very nice -- very clean :).

You didn't really explain what any of the criteria do or any of the pros and cons of each criteria.

For reference using MSE as the criteria for CV and ESCV doesn't quite make sense given that we are trying to maximize correlation.

The paper for ESCV actually shows that it makes more sense to maximize tau (the L1 norm of the beta coefficients) rather than lambda itself so it would have been nice to show these plots for tau instead of lambda at least for ESCV...

Evaluation of model performance and diagnostic plots *

	0	1	2	3	4	
Did not evaluate model performance	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	Thoroughly evaluated how well the models performed using correlation and provided insightful and appealing figures for diagnostic plots and model interpretation

Comments on model performance evaluation and diagnostic plots

Figures 6 and 7 are very nice (I probably would make the connecting lines a bit fainter to draw the eye to the points in figure 6).

Unfortunately you didn't really provide any diagnostic plots (e.g. residual plots or predicted vs fitted plots) or do any outlier analysis (e.g. are there any images that are strange or hard to predict?).

Interpretation of models *

	0	1	2	3	
Did not try to interpret the models	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Provided a thorough interpretation of the models

Comments on model interpretation

I like your interpretation a lot! Good idea to show the image convoluted with the wavelet!

Evaluation of model stability *

	0	1	2	3	
Did not discuss model stability	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Reasonable explored and discussed model stability

Comments on model stability evaluation

I'm not quite sure what you mean by "For each voxel, the Lasso/Ridge regression was run ten times with the 80% training data random shuffled" – are you talking about subsampling? I would like to see this done more like 100 times (10 isn't very many to get an idea of the distribution of the feature selection or correlation as the sample changes).

How well did the students model perform on val_feat (once I have completed estimated performance of all student's models you will receive a relative score out of 5) *

correlation of 0.641

Reproducibility and github stuff

Did the student provide all files and instructions in their github repo necessary for reproducing the results and report? *

	0	1	2	3	
Did not provide anything required for reproducibility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	Everything was provided and clearly named/described

Comments

I like that you have many well-named R scripts and functions

Other general comments

.....

This form was created inside of UC Berkeley.

