

Figure 1 Left Panel Multinomial Regression Error Rates. This figure shows error rates for all subjects. It shows cross-validation error for the subjects in the training set (solid color) and test error rates for the subjects in the test set (striped color). For each subject, this plot shows the error rate for a model that did not include that subject in its training set. We see a wide variety. Some subjects are "easy:" even if they weren't included in the training set, their predictions will be almost always correct. Other subjects are "difficult:" if the model didn't train on data from a "difficult" subject, the model will make errors recognizing that subject's activities. We were lucky with the test set; it had no "difficult" subjects. Thus the overall test error was lower than expected. The cross-validation error is a better estimate of the error rate on a bigger test set, because we would expect to see some "difficult" subjects.

Right Panel Cross-Validation Error Rates for "Difficult" Subjects, All Models. All three models had trouble with the same "difficult" subjects. The average error for these five subjects was over 15%. An ensemble model, with all three models voting on the prediction, doesn't look promising, because all the models made errors on the same subjects.