

If your forward algorithms are accurate, then they should work in reverse. I.e. if you can predict the test data targets, then you should be able to use those results to predict the training data. The error between predicting the training/validation data gives an estimate of how accurate your algorithms are. That's my theory at least, though it loses something in practice,

I did find using the variances among the various feature groups in the old data set useful for understanding problems in developing algorithms, but I haven't applied a similar analysis yet to the new data.