

ASSIGNMENT 9
FEATURE ENGINEERING STUDIO
PROFESSOR RYAN BAKER
CONSTRUCT VALIDITY
“ONE WHO VISIONS MUST BE STEEPED IN DATA”
DUE IN CLASS, MONDAY NOVEMBER 11

The goal of this assignment is to leverage some outside perspectives on your problem.

Step 1: Create at least three new features (but no upper limit) based on what you learned during the poster session. What are these features' cross-validated goodness? (Correlation, A' , or Kappa)

Step 2: Create the best possible model of your construct that you can, involving some subset of all your features. What is this model's cross-validated goodness? (Correlation, A' , or Kappa)

Step 3: Give a friend a list of the features you are using. Explain your problem to them. Have them select the features which they think are likely to genuinely be useful. (Don't tell them your results in advance). It can't be all of the features, but it can be any subset of them. Note that it's best to have a domain expert do this, but in case you don't have one handy, any intelligent person is fine (for this assignment). Now create the best possible model of your construct that you can, involving some subset of the features your friend chose. What is this model's cross-validated goodness? (Correlation, A' , or Kappa)

Step 4: Which model is better? How much better is it?

Step 5: Email me your report and be ready to discuss this in class.