Business Analytics 7.25.2022

Regularization

* Adds penalties when x is too sensitive
* Basically transforms the coefficients to help reduce variability.
* THE GOAL OF THIS IS PREDICTION, NOT STATISTICAL INFERENCE.

Different Types of regularization

* L1 Regularization (LASSO or absolute value penalty)
  + Sums of Square residuals + Lambda\*(absolute value of Beta)
  + LASSO = Least Absolute Shrinkage and Selection Operation
  + Use absolute values to shrink the Betas. If an X variable is not significant nor influential in the model, it will force that Beta to zero.
* L2 Regularization (Ridge Regression, or Squared Value Penalty)
  + Sums of square residuals + lambda\*(Beta^2)
  + Asymptotically approaches zero, but will never reach it.
* Elastic Net Regularization
  + SSR + Lamba\*(abs value Beta + Beta^2)

What is Lambda here?

* Is a tuning variable, is trying to filter out noise and get to the true signal.
* If you overtune the parameter, you will filter out everything, both the noise and the station.

How to determine Lambda

* K-fold Cross validation
  + Takes your data and splits it into k # of chunks
  + Industry standard is 10 folds
* Example
  + 4 folds of data
  + Model one would take folds 2, 3, and 4 as training data and fold 1 as testing data.
  + Model two would take folds 1, 3, and 4 as training data and fold 2 will be testing data.
  + Etc. until you go through all the fold combinations
* Where does lambda come in?
  + Your models will iterate different values of lambda from 0 to 1
  + Gives mean^2 errors, then we take the average of all of the observations, which gives as a distribution of all of the models.