SuperLearner <- function(Y, X, newX = NULL, family = gaussian(), SL.library,

method = 'method.NNLS', id = NULL, verbose = FALSE, control = list(),

cvControl = list(), obsWeights = NULL, env = parent.frame()) {

CV.SuperLearner <- function(Y, X, V = NULL, family = gaussian(), SL.library, method = 'method.NNLS', id = NULL, verbose = FALSE, control = list(saveFitLibrary = FALSE),

cvControl = list(), innerCvControl = list(), obsWeights = NULL, saveAll = TRUE, parallel = "seq", env = parent.frame()) {

call <- match.call()

N <- dim(X)[1L]

Questions?

* Do we care about different error selection techniques:

Mean absolutel error

MSE vs RMSE

“Pseudo R2”

* Currently doesn’t take into effect Gaussian vs. binomial to describe the error distribution
* Cluster identification? Is this needed
* SL.Library – should this just be user specified, like a list separated by semicolons and we parse?
* Is a graphical plot of the output important?
* How exactly are people going to use this?