**Super Learner Algorithm (SL)**

The super learner function takes a training set pair (X,Y) and returns the predicted values based on a validation set. With the V-fold cross validation theorem, this function could extent the candidate learner selector to include weighted averages of the candidate learners.

**Algorithm:**

Step 0: train each candidate learner on entire dataset.

Step 1: Split data into v blocks.

Step 2: train each candidate learner.

Step 3: predict the outcomes in the validation block based on the corresponding training block candidate learner.

Step 4: model selection and fitting for the regression of the observed outcome onto the predicted outcomes from the candidate learners.

Step 5: evaluate super learner by combining predictions from each candidate learner with

Usually, the default method on estimating the coefficients for the super learner and the model to combine the individual algorithms in the library is NNLS.

**NNLS (The Lawson-Hanson algorithm for non-negative least squares):**

by solving the problem with the constraint x

In fact, if we have m different predictions and n different subjects, we could seem A as matrix of predictions of these m subjects by n methods and b is the real value of this m subject. And then, find x (the coefficients of different learners), which lead to min error.

In R, there always has a prescreen algorithms as well, which first rank the variables in X based on either a univeriate regression p-value of the ransomForest variable importance. A subset of the variables in X is selected based on a pre-defined cut-off. With this subset of the X variables, the algorithm in SL.Library are then fit in.