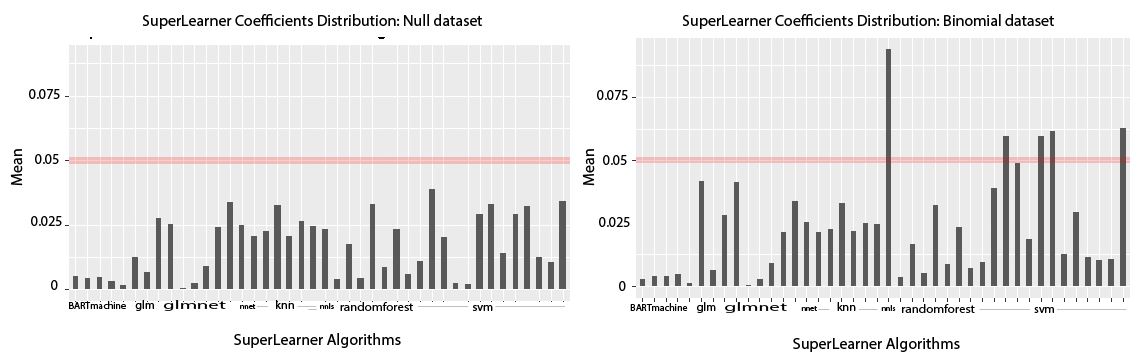
****

# Figure 4: Second Approach applied to CBDA the *Null and Binomial datasets*. Two set of M=9000 subsamples/iterations of the CBDA algorithm have been generated by sampling a *Null and a Binomial dataset* with a features/cases ratio of 1/3 (i.e., 300/100) and i) CSR=[50%-60%], ii) FSR=[10%-15%]. In both panels, the x axis shows the labels of the 42 different classes/specifications of machine learning algorithms used in the SuperLearner. The y axis displays the mean value for each of the 42 SuperLearner coefficients/weights associated to the 42 machine learning algorithms, calculated across the 9000 subsamples/iterations. The list of the 42 algorithms is the following: generalized linear model [glm] (5 istances with different parameters), neural network [nnet], non-negative least squares [nnls], Lasso and Elastic-Net Regularized Generalized Linear Models [glmnet] (5 instances, with different penalty parameters, from 0 to 1), k-nearest neighbor (4 instances), Support Vector Machine [svm] (13 istances), random Forest [randomForest] (7 istances), Bayesian Auto Regressive Tree Machine [bartMachine] (4 istances).

# *Panel A* shows the barplot generated from the CBDA analysis on the *Null dataset*.

# *Panel B* shows the barplot generated from the CBDA analysis on the *Binomial dataset*.