

Model-averaged ℓ_1 -penalized regression for biological data sets

Chris Fraley, Mark Seligman
Insilicos LLC

mark.seligman@insilicos.com

Biological data sets, such as those obtained from microarray experiments, typically contain many more predictors than samples. These pose challenges for model-selection schemes, such as Bayesian Model Averaging ("BMA"), in that an appropriate data-reduction step must first be performed. We present an approach which performs BMA using ℓ_1 -penalized regression in a reduction step. This gives a model-selection method with predictive power on a par with competing methods. We have found that it is also faster and more parsimonious as well as more stable with respect to reduction parameters.