

High-Dimensional Bayesian Classifiers

David Madigan, Columbia University

Supervised learning applications in text categorization, authorship attribution, hospital profiling, and many other areas frequently involve training data with more predictors than examples. Regularized logistic models often prove useful in such applications and I will present some experimental results. A Bayesian interpretation of regularization offers advantages. In applications with small numbers of training examples, incorporation of external knowledge via informative priors proves highly effective. Sequential learning algorithms also emerge naturally in the Bayesian approach. Finally I will discuss some recent ideas concerning structured supervised learning problems and connections with social network models.