Your 598 ID: 5107

Title of paper: Discriminative vs Informative Learning by Rubinstein and Hastie.

What is their primary result? The authors extend discriminative and informative learning methods to more modern ones such as Naive Bayes (NB) and Generalized Additive Methods (GAM).

Why is this important? The paper provides a way of extending discriminative and informative learning to other techniques. It also proposes a method of combining the two approaches.

What are their key ideas? The authors review and compare informative and discriminative pattern classification. They show that the two approaches can be related by an application of Bayes' Rule, but that the two methods lead to different decision rules and, hence, there are tradeoffs associated to each.

In particular, they show in Theorem 1, p. 52, that NB, an informative classifier, is a spacial case of a GAM, a discriminative model by demonstrating that the induced discriminant is log additive via Bayes' Rule.

What are the limitations, either in performance or applicability? The authors argue that, contrary to what one might expect, NB and GAM may not always lead to the best classifiers. The authors perform a logspline simulation study on NB (which assumes a logspline density separately in each dimension) and GAM. The study is done on two classes: Class 1 is a complicated mixture density and class 2 is the exponential tilt (i.e., logspline discriminant) of class 1. The GAM classifier achieves a Bayes error rate of 7.2%, whereas the NB classifier does worse at 9.0%.

Furthermore, the authors note that when a small sample of training observations is available, NB classifier outperforms the GAM classifier. More specifically, in a simulation experiment with 25 training sets each containing 25 observations from each class, the NB classifier achieved an error rate of 11.1% vs 11.4% for the GAM classifier, although the model for the GAM is correct, but that of the NB is not.

Grader's 598 ID: