On High-Dimensional Classification for Sparse Signals

DAWIT G. TADESSE

Department of Mathematical Sciences, 2815 Commons Way,
University of Cincinnati, Cincinnati, OH, 45221-0025, USA
Email: tadessdt@ucmail.uc.edu

MARK CARPENTER

Auburn University, USA

SUMMARY

In this paper we conduct a rigorous performance analysis of the two linear methods for high-dimensional classification, Independence Rule (or Naive Bayes) and Fisher discriminant both in theory and simulation. We know that, for the normal population model, when all the parameters are known Fisher is optimal and Naive Bayes is suboptimal. But in this paper we give the conditions under which Naive Bayes is optimal. Through theory and simulation, we further, show that Naive Bayes performs better than Fisher under broader conditions. We also study the associated feature selection methods.

Keywords and phrases: Feature Selection; Fisher discriminant; High-Dimensional classification; Misclassification error rate; Naive Bayes; Sparse Signals.

2010 Mathematics Subject Classification: 62H30