

Support vector machines for credit scoring and discovery of significant features

Tony Bellotti and Jonathan Crook (31 October 2006)

Credit Research Centre, School of Management, University of Edinburgh, William Robertson Building, 50 George Square, Edinburgh EH8 9JY, UK

Abstract

Logistic regression and discriminant analysis are techniques traditionally used in credit scoring for determining likelihood to default based on consumer application and credit reference agency data. We test support vector machines against these traditional methods on a large credit card database. We find that they are competitive and can be used as the basis of a feature selection method to discover those features that are most significant in determining risk of default.