

# Heuristics for Feature Selection in Mathematical Programming Discriminant Analysis Models

Konstantinos Falangis and John Glen

Management School, University of Edinburgh

## Abstract

In developing a classification model for assigning observations of unknown class to one of a number of specified classes using the values of a set of features associated with each observation, it is often desirable to the classifier on a limited number of features. Mathematical programming (MP) discriminant analysis methods for developing classification models can be extended for feature selection. By using a mixed integer programming (MIP) model in which a binary variable is associated with each training sample observation, features can be selected using classification accuracy as the selection criterion, but the binary variable requirements limit the size of problems to which this approach can be applied. Heuristic feature selection methods for problems with large numbers of observations are developed in this paper. These heuristic procedures, which are based on the feature selection MIP model, are then applied to a number of credit scoring datasets.