Automatic Interaction Detection for Longitudinal Data

Jacky Galpin*
School of Statistics and Actuarial Science
University of the Witwatersrand
South Africa
jacky@galpin.co.za

Tree based methods for detecting variables predictive of a continuous or categorical dependent variable are well developed, and available in several statistical packages, for the case of a single dependent variable. The case of multiple dependent variables has also received some attention in the literature, but little has appeared for the case of longitudinal data. Several of these methods use techniques such as boosting and bagging to ensure stability of the tree. Not using these methods allows the technique to identify outliers in many cases. This paper examines the use of the CHAID (Chi-Squared Automatic Interaction Detection) and XAID (eXtended Automatic Interaction Detection) methodologies (Hawkins and Kass, 1982) to investigate misspecification of model assumptions in the analysis of longitudinal data.

CHAID XAID outlier misspecification

Reference:
Hawkins, D. M. and Kass, G.V. (1982). 'Automatic Interaction Detection' in *Topics in Applied Multivariate Analysis,* Ed D. M. Hawkins, Cambridge University Press.