Sauve, Marie; Tuleau-Malot, Christine
Variable selection through CART. (English) Zbl 1305.62154

Summary: This paper deals with variable selection in regression and binary classification frameworks. It proposes an automatic and exhaustive procedure which relies on the use of the CART algorithm and on model selection via penalization. This work, of theoretical nature, aims at determining adequate penalties, i.e. penalties which allow achievement of oracle type inequalities justifying the performance of the proposed procedure. Since the exhaustive procedure cannot be realized when the number of variables is too large, a more practical procedure is also proposed and still theoretically validated. A simulation study completes the theoretical results.

MSC:
62G05 Nonparametric estimation
62G07 Density estimation
62G20 Asymptotic properties of nonparametric inference

Keywords:
binary classification; CART; model selection; penalization; regression; variable selection

Full Text: DOI arXiv Link