

Subsumption resolution: an efficient and effective technique for semi-naive Bayesian learning

Fei Zheng · Geoffrey I. Webb · Pramuditha Suraweera ·
Liguang Zhu

Received: 31 October 2007 / Accepted: 30 November 2011 / Published online: 22 December 2011
© The Author(s) 2011

Abstract Semi-naive Bayesian techniques seek to improve the accuracy of naive Bayes (NB) by relaxing the attribute independence assumption. We present a new type of semi-naive Bayesian operation, Subsumption Resolution (SR), which efficiently identifies occurrences of the specialization-generalization relationship and eliminates generalizations at classification time. We extend SR to Near-Subsumption Resolution (NSR) to delete near-generalizations in addition to generalizations. We develop two versions of SR: one that performs SR during training, called eager SR (ESR), and another that performs SR during testing, called lazy SR (LSR). We investigate the effect of ESR, LSR, NSR and conventional attribute elimination (BSE) on NB and Averaged One-Dependence Estimators (AODE), a powerful alternative to NB. BSE imposes very high training time overheads on NB and AODE accompanied by varying decreases in classification time overheads. ESR, LSR and NSR impose high training time and test time overheads on NB. However, LSR imposes no extra training time overheads and only modest test time overheads on AODE, while ESR and NSR impose modest training and test time overheads on AODE. Our extensive experimental comparison on sixty UCI data sets shows that applying BSE, LSR or NSR to NB significantly improves both zero-one loss and RMSE, while applying BSE, ESR or NSR to AODE significantly improves zero-one loss and RMSE and applying LSR to AODE significantly improves zero-one loss. The Friedman test and Nemenyi test show that AODE with ESR or NSR have a significant zero-one loss and RMSE advantage over Logistic Regression and a zero-one loss advantage over Weka's LibSVM implementation with a grid parameter search on categorical data. AODE with LSR has a zero-one loss advantage over Logistic Regression and comparable zero-one loss with LibSVM. Finally, we examine the circumstances under which the elimination of near-generalizations proves beneficial.

Keywords Classification · Naive Bayes · Semi-naive Bayes · Feature selection · AODE

Editors: Mark Craven and Johannes Fürnkranz.

F. Zheng · G.I. Webb (✉) · P. Suraweera · L. Zhu

Faculty of Information Technology, Monash University, Clayton, Vic. 3800, Australia
e-mail: Geoff.Webb@monash.edu