

STATISTICAL ANALYSIS AND DATA MINING

Original Article

Learning with many experts: Model selection and sparsity

Rafael Izbicki , Rafael Bassi Stern

First published: 28 October 2013

<https://doi.org/10.1002/sam.11206>

Cited by: 1

 About  Access

PDF

Tools



Share

Abstract

Experts classifying data are often imprecise. Recently, several models have been proposed to train classifiers using the noisy labels generated by these experts. How to choose between these models? In such situations, the true labels are unavailable. Thus, one cannot perform model selection using the standard versions of methods such as empirical risk minimization and cross validation. In order to allow model selection, we present a surrogate loss and provide theoretical guarantees that assure its consistency. Next, we discuss how this loss can be used to tune a penalization which introduces sparsity in the parameters of a traditional class of models. Sparsity provides more parsimonious models and can avoid overfitting. Nevertheless, it has seldom been discussed in the context of noisy labels owing to the difficulty in model selection and, therefore, in choosing tuning parameters. We apply these techniques to several sets of simulated and real data. © 2013 Wiley Periodicals, Inc. *Statistical Analysis and Data Mining*, 2013

Citing Literature

Number of times cited: 1

Luís Gustavo Esteves, Rafael Izbicki and Rafael Bassi Stern, Teaching Decision Theory Proof Strategies Using a Crowdsourcing Problem, *Journal of Statistical Software*, **71**, 4, (336), (2017).
[Crossref](#)

About Wiley Online Library

Help & Support

Opportunities

Connect with Wiley

