

**Weston, Jason; Elisseeff, André; Schölkopf, Bernhard; Tipping, Mike**

**Use of the zero-norm with linear models and kernel methods.** (English) Zbl 1102.68605

*J. Mach. Learn. Res.* 3, No. 7-8, 1439-1461 (2003).

**Summary:** We explore the use of the so-called zero-norm of the parameters of linear models in learning. Minimization of such a quantity has many uses in a machine learning context: for variable or feature selection, minimizing training error and ensuring sparsity in solutions. We derive a simple but practical method for achieving these goals and discuss its relationship to existing techniques of minimizing the zero-norm. The method boils down to implementing a simple modification of vanilla SVM, namely via an iterative multiplicative rescaling of the training data. Applications we investigate which aid our discussion include variable and feature selection on biological microarray data, and multiclass classification.

**MSC:**

**68T05** Learning and adaptive systems in artificial intelligence

Cited in **1** Review  
Cited in **76** Documents

**Software:**

[Delve Datasets](#)

**Full Text:** [DOI](#)