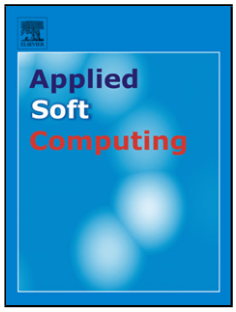
Accepted Manuscript



Title: A Robust Semi-Supervised SVM via Ensemble Learning

Author: Dan Zhang Licheng Jiao Xue Bai Shuang Wang Biao Hou

PII: S1568-4946(18)30044-9
DOI: https://doi.org/doi:10.1016/j.asoc.2018.01.038
Reference: ASOC 4685

To appear in: *Applied Soft Computing*

Received date: 9-1-2017
Revised date: 23-1-2018
Accepted date: 26-1-2018

Please cite this article as: Dan Zhang, Licheng Jiao, Xue Bai, Shuang Wang, Biao Hou, A Robust Semi-Supervised SVM via Ensemble Learning, *<![CDATA[Applied Soft Computing Journal]]>* (2018), https://doi.org/10.1016/j.asoc.2018.01.038

A Robust Semi-Supervised SVM via Ensemble Learning

Dan Zhang*, Licheng Jiao, Xue Bai, Shuang Wang, Biao Hou

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education,

International Research Center for Intelligent Perception and Computation,

Joint International Research Laboratory of Intelligent Perception and Computation,

Xidian University, Xian, Shaanxi Province 710071, China

Abstract

Semi-supervised learning is one of the most promising learning paradigms in many practical applications where few labeled samples are available. Among such learning models, semi-supervised support vector machine (SVM) is a typical and prominent one. However, the typical semi-supervised SVM cannot avoid estimating the distribution of positive and negative samples. In this paper, we put forward a method with the name of EnsembleS3VM which deals with the unknown distribution by ensemble learning. It builds a semi-supervised SVM model composed of base learners based on different disturbance factors, and raises an ensemble method based on clustering evaluation means. Meanwhile, it presents a combination of two multi-classification strategies in order to reduce the running time and enhance the classification accuracy simultaneously. The proposed method can deal with semi-supervised classification problems even with unknown distribution or unbalanced data. Experiments on UCI datasets prove the effectiveness of ensemble strategy and the robustness under different sample distributions. We also apply the proposed algorithm to a practical application, i.e., ground cover classification for polarimetric synthetic aperture radar images which is a typical but difficult semi-supervised classification problem.

Keywords: semi-supervised SVM, disturbance factor, base learner, ensemble learning

*Corresponding author. E-mail address: see2011_dan@126.com (D. Zhang).