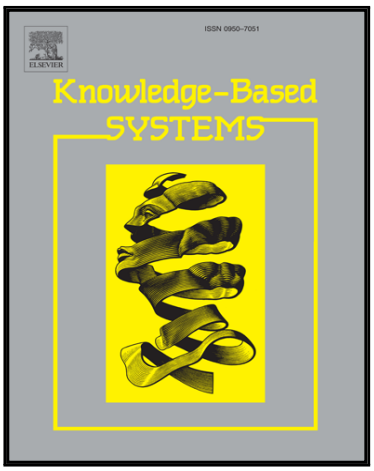

Universum Selection for Boosting the Performance of Multiclass Support Vector Machines Based on One-versus-One Strategy

Patoomsiri Songsiri, Vladimir Cherkassky, Boonserm Kijsirikul

PII: S0950-7051(18)30255-7
DOI: 10.1016/j.knosys.2018.05.025
Reference: KNOSYS 4345

To appear in: *Knowledge-Based Systems*

Received date: 29 December 2016
Revised date: 13 May 2018
Accepted date: 18 May 2018

Please cite this article as: Patoomsiri Songsiri, Vladimir Cherkassky, Boonserm Kijsirikul, Universum Selection for Boosting the Performance of Multiclass Support Vector Machines Based on One-versus-One Strategy, *Knowledge-Based Systems* (2018), doi: 10.1016/j.knosys.2018.05.025

Universum Selection for Boosting the Performance of Multiclass Support Vector Machines Based on One-versus-One Strategy

Patoomsiri Songsiri[a], Vladimir Cherkassky[b], Boonserm Kijsirikul[a,*]

[a] *Department of Computer Engineering, Chulalongkorn University, Pathumwan, Bangkok Thailand. 10330*
[b] *Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA. 55455*

Abstract

We propose a novel framework to enhance the performance of the one-versus-one support vector machine by using Universum. For solving a multiclass classification problem, one-versus-one is one of the state-of-the-art algorithms, which constructs $N(N-1)/2$ binary classifiers for an N-class problem. Each binary classifier is originally learned by two classes of data as positive and negative classes while the other $N-2$ remaining classes are ignored, even if they might also represent a hidden concept of the application domain and can help to boost the performance of the classifier. Vapnik et al. [20, 21] introduced Universum binary support vector machines to enable the use of samples that do not belong to positive and negative classes and called these samples Universum samples. However, not all Universum samples can be helpful; moreover, improper selection of Universum samples can prevent the construction of an effective binary classifier. For the construction of a Universum binary classifier in the one-versus-one strategy, there are 2^{N-2} candidate subsets of classes of Universum data; a proper selection of them can be difficult, based on the number of classes. We design an algorithm to obtain a suitable subset of classes of Universum data by applying the proposed performance measure that reflects the properties of Universum data relative to labeled training data. This measure is based on the analysis of the projection of Universum data onto the normal direction vector of the standard binary SVM hyperplane. We demonstrate experimentally that our proposed strategy outperforms existing methods.

Keywords: Support vector machine (SVM), Universum support vector machine (USVM), multi-class classification, one-versus-one

1. Introduction

The support vector machine (SVM) was proposed by Vapnik [18, 19] and originally designed to solve two-class classification problems by maximizing the margin between two-class data. Vapnik et al. [20, 21] also introduced another learning technique, called the Universum support vector machine (USVM), to enable the use of a priori knowledge from the same application domain. In this technique, samples that do not belong to positive and negative classes can be added to boost the generalization performance of the learned model. We call these samples Universum samples. In the standard SVM learning, labeled training samples are used to construct the binary classifier. While in the USVM learning, labeled training samples and Universum samples are used together to control the trade-off between the minimization of training errors (by maximizing the margin) and the maximization of the number of Universum samples lying inside the margin. However, it is not necessarily the case that all Universum samples can be helpful for Universum learning. In a previous study by Sinz et al. [16], they suggested that the performance of USVM depends on the difference between the mean of the labeled training samples and that of the Universum samples. Cherkassky et al. [4] also suggested conditions to determine the effectiveness of Universum samples. Their conditions are based on the analysis of univariate histograms of projec-

*Corresponding author. Tel.:+6622186956.
Email addresses: patoomsiri.s@chula.ac.th (Patoomsiri Songsiri), cherk001@umn.edu (Vladimir Cherkassky), boonserm.k@chula.ac.th (Boonserm Kijsirikul)