

数据库、信息处理

## 基于Tri-training的半监督SVM

李昆仑, 张 伟, 代运娜

河北大学 电子信息工程学院, 河北 保定 071002

收稿日期 2008-6-24 修回日期 2008-9-16 网络版发布日期 接受日期

**摘要** 当前机器学习面临的主要问题之一是如何有效地处理海量数据, 而标记训练数据是十分有限且不易获得的。提出了一种新的半监督SVM算法, 该算法在对SVM训练中, 只要求少量的标记数据, 并能利用大量的未标记数据对分类器反复的修正。在实验中发现, Tri-training的应用确实能够提高SVM算法的分类精度, 并且通过增大分类器间的差异性能够获得更好的分类效果, 所以Tri-training对分类器的要求十分宽松, 通过SVM的不同核函数来体现分类器之间的差异性, 进一步改善了协同训练的性能。理论分析与实验表明, 该算法具有较好的学习效果。

**关键词** [半监督学习](#) [协同训练](#) [Tri-training](#) [支持向量机](#) [最小二乘支持向量机](#)

分类号

## Semi-supervised SVM based on Tri-training

LI Kun-lun, ZHANG Wei, DAI Yun-na

College of Electronic and Information Engineering, Hebei University, Baoding, Hebei 071002, China

### Abstract

One of the main difficulties in machine learning is how to solve large-scale problem effectively, and the labeled data are limited and fairly expensive to obtain. In this paper a new semi-supervised SVM is proposed. It applies Tri-training to improve SVM. The semi-supervised SVM uses a few labeled data to train few initial SVM classifiers and makes use of the large number unlabeled data to modify the classifier iteratively. Experiments on UCI dataset show that Tri-training can improve the classification accuracy of SVM and can increase the difference of classifier, the accuracy of final classifier will be higher. Although Tri-training doesn't put any constraints on the supervised learning algorithm, the proposed method uses the SVMs with three different kernel functions as the supervised learning algorithm. The different kernel can increase the difference of the three SVMs, so the performance of co-training will be better. Theoretical analysis and experiments show that the proposed algorithm has excellent accuracy and speed of classification.

**Key words** [semi-supervised learning](#); [co-training](#) [Tri-training](#) [Support Vector Machine \(SVM\)](#)  
[least square support vector machine](#)

DOI: 10.3778/j.issn.1002-8331.2009.22.034

通讯作者 李昆仑 [wzhanghbu@gmail.com](mailto:wzhanghbu@gmail.com)

### 扩展功能

#### 本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(495KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [参考文献](#)

#### 服务与反馈

- ▶ [把本文推荐给朋友](#)
- ▶ [加入我的书架](#)
- ▶ [加入引用管理器](#)
- ▶ [复制索引](#)
- ▶ [Email Alert](#)
- ▶ [文章反馈](#)
- ▶ [浏览反馈信息](#)

#### 相关信息

- ▶ 本刊中 [包含“半监督学习”的相关文章](#)
- ▶ [本文作者相关文章](#)

- [李昆仑](#)
- [张 伟](#)
- [代运娜](#)