

数据库、信息处理

基于k-means聚类的神经网络分类器集成方法研究

李凯, 常圣领

河北大学 数学与计算机学院, 河北 保定 071002

河北省机器学习与计算智能重点实验室, 河北 保定 071002

收稿日期 2008-4-30 修回日期 2008-8-11 网络版发布日期 接受日期

摘要 针对差异性是集成的必要条件, 研究了基于k-means聚类技术提高神经网络分类器集成差异性的方法。通过训练集并使用神经网络分类器学习算法训练许多分类器模型, 在验证集中利用每个分类器的分类结果作为聚类的数据对象; 然后应用k-means聚类方法对这些数据聚类, 在聚类结果的每个簇中选择一个分类器代表模型, 以此构成集成学习的成员; 最后应用投票方法实验研究了这种提高集成学习差异性方法的性能, 并与常用的集成学习方法bagging、adaboost进行了比较。

关键词 [差异性](#) [集成学习](#) [分类器](#) [聚类](#)

分类号

Study of ensemble method of classifiers for neural networks based on k-means clustering

LI Kai, CHANG Sheng-ling

School of Mathematics and Computer, Hebei University, Baoding, Hebei 071002, China

Key Lab of Machine Learning and Computational Intelligence of Hebei Province, Baoding, Hebei 071002, China

Abstract

Aiming at diversity being a necessary condition of the ensemble learning, this paper studies the method for improving diversity of the neural networks ensemble based on k-means clustering technique. This paper proposes a selecting approach that is first to train many classifiers through training set with neural network algorithm, and uses the result by the classifiers from validation set for clustering. And then this paper uses the k-means algorithm to cluster the data set from the result and selects a classifier model from every cluster to make up of the membership of the ensemble learning. At last, this paper studies the performance of ensemble method by using vote method and compare performance with bagging and adaboost methods.

Key words [diversity](#) [ensemble learning](#) [classifier](#) [clustering](#)

DOI: 10.3778/j.issn.1002-8331.2009.22.039

通讯作者 李凯

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(374KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“差异性”的 相关文章](#)

▶ [本文作者相关文章](#)

· [李凯](#)

· [常圣领](#)