

云服务环境下的大数据多标签属性分类技术

林倩瑜

(集美大学 诚毅学院, 福建 厦门 361021)

摘 要: 云服务环境下的大数据具有多标签属性, 导致数据访问过程中的分类性不好, 为了降低大数据分类的误分率, 提出一种基于极限学习机优化神经网络的云服务大数据分类方法。采用高维相空间重构模型对随机采样的云服务大数据先验特征进行信息重排, 采用关联规则挖掘方法提取数据的标签属性特征量, 采用多维标签属性的主成分分析方法进行大数据的主特征分析, 将主特征量输入到神经网络分类器中, 结合极限学习机进行神经网络分类的加权控制, 提高数据分类的自适应性, 实现云服务大数据优化分类。仿真结果表明, 采用该算法进行云服务大数据分类的准确性较高, 误分率较低, 在云组合服务中提高了数据的召回能力。

关键词: 云服务环境; 大数据; 分类; 神经网络; 极限机学习

Big Aata Multi-label Attribute Classification Technology in Cloud Service Environment

LIN Qian-yu

(Chengyi University College, Jimei University, Xiamen 361021, China)

Abstract: Big data in the cloud service environment has multi-label attributes, which leads to poor classification in the process of data access. In order to reduce the misclassification rate of big data classification, a classifying method for cloud service big data based on the neural network optimization of extreme learning machine is proposed. The high-dimensional phase space reconstruction model is used to rearrange the information of the random sampling cloud service big data's prior feature, and the association rule mining method is used to extract the label attribute feature quantity of the data. The principal component analysis method of multidimensional label attribute is used to analyze big data's principal feature. The principal feature quantity is input into the neural network classifier, and the weighted control of neural network classification is carried out by combining with the extreme learning machine. Improve the self-adaptability of data classification, realize cloud service big data optimization classification. The simulation results show that the proposed algorithm has higher accuracy and lower misclassification rate in cloud service big data classification, and improves the recall ability of data in cloud composition service.

Key words: cloud service environment; big data; classification; neural network; limit machine learning

作者简介:

林倩瑜 女 (1985-), 硕士, 实验师. 研究方向为数据挖掘, 神经网络.

E-mail: 18005926008@163.com.