

数据库与信息处理

基于不完全数据的TAN学习算法

王建林¹, 王志海², 王学玲¹

1. 滨州学院 计算机科学与技术系, 山东 滨州 256600

2. 北京交通大学 计算机与信息技术学院, 北京 100044

收稿日期 修回日期 网络版发布日期 2007-12-10 接受日期

摘要 TAN算法是一种针对复杂数据且在实际中具有极强的学习能力的有效算法, 它已被广泛应用于数据挖掘、机器学习和模式识别领域。由于现实世界中的数据大多是不完全数据, 研究了怎样使TAN有效地从不完全数据中学习。首先, 用一种有效的方法直接从不完全数据中估计条件互信息, 然后应用估计条件互信息法去扩展基本的TAN算法来处理不相关数据, 最后实验比较了扩展的TAN算法和基本的TAN算法。实验结果表明, 在大多数不完全数据集上扩展的TAN算法精确性明显高于基本的TAN算法。虽然扩展的TAN算法时间复杂度高于基本的TAN算法, 但在可接受范围之内。此估计条件互信息的方法能够容易地和其它技术相结合来进一步提高TAN算法的性能。

关键词 [TAN](#) [学习](#) [不完全数据](#) [条件互信息](#)

分类号

Learning TAN from incomplete data

WANG Jian-lin¹, WANG Zhi-hai², WANG Xue-ling¹

1. Department of Computer Science, Binzhou University, Binzhou, Shandong 256600, China

2. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract

TAN is a good trade-off between the model complexity and learn ability in practice, which has been widely used in data mining, machine learning and pattern recognition etc. Since there are few complete datasets in real-world, the paper develops research on how to efficiently learn TAN from incomplete data. Firstly an efficient method that could estimate conditional Mutual Information directly from incomplete data is presented. And then the basic TAN learning algorithm is extended to incomplete data using the conditional Mutual Information estimation method. Finally, experiments are carried out to evaluate the extended TAN and compare it with basic TAN. The experimental results show that the accuracy of the extended TAN is much higher than that of basic TAN on most of the incomplete datasets. Despite more time consumption of the extended TAN compared with basic TAN, it is still acceptable. The conditional Mutual Information estimation method can be easily combined with other techniques to improve TAN further.

Key words [TAN](#) [learning](#) [incomplete data](#) [conditional mutual information](#)

DOI:

通讯作者 王建林 wxlwry@yahoo.com

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(664KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“TAN”的 相关文章](#)

▶ 本文作者相关文章

· [王建林](#)

· [王志海](#)

· [王学玲](#)