

数据库与信息处理

LSA在中文短文自动判分系统中的应用研究

李莉, 张太红

新疆农业大学 计算机与信息工程学院, 乌鲁木齐 830052

收稿日期 修回日期 网络版发布日期 2007-6-29 接受日期

摘要 对潜在语义分析 (Latent Semantic Analysis, LSA) 的理论基础进行了介绍, 研究了潜在语义分析在中文短文写作自动评分领域的应用方法。从136名大学生的短文写作试卷着手, 对比了不同的语义空间构造方法和不同数据标准化方法对机器自动评分结果的影响, 探讨了SVD的作用和奇异值个数K的取值规律, 比较了LSA对不同类型学生的短文写作自动评分结果的差异。通过与两名教师对学生短文写作评分的比较表明, 使用机器对主观题进行自动评分是可行的, 该方法为自动化考试系统试题多样性提供了有效的解决方案。

关键词 [潜在语义分析](#) [奇异值分解](#) [主观题自动判分](#)

分类号

Application researches of latent semantic analysis in Chinese essay auto scoring system construction

LI Li, ZHANG Tai-hong

College of Computer and Information Engineering, Xingjian Agricultural University, Urumqi 830052, China

Abstract

The basic idea of Latent Semantic Analysis (LSA) is described in the introduction of the paper. The application of LSA to the Chinese writing essays' auto scoring is studied in this paper. The influence of machine auto scoring results of different data standard processing and different semantic space construction methods are compared basing on the analysis of 136 college students' writing test paper. The effect of Singular Value Decomposition (SVD) and the rules for definition of the K numbers of singular values are discussed. In the last part of the paper, two teachers validate the results. It shows that it is feasible to score Chinese subjective test by machine auto scoring system. The method proposed in this paper offers a new option for variety test questions of auto test system.

Key words [Latent Semantic Analysis \(LSA\)](#) [Singular Value Decomposition \(SVD\)](#) [machine auto scoring in subjective test](#)

DOI:

通讯作者 李莉 [E-mail: 25580920@sina.com.cn](mailto:25580920@sina.com.cn)

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(1095KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“潜在语义分析”的相关文章](#)

▶ [本文作者相关文章](#)

· [李莉](#)

· [张太红](#)