## 运用文本相似度实现主观题自动评阅【华云+中诺思】

1. 命题方向

智能计算

2. 题目类别

算法（模型）类

3. 题目名称

运用文本相似度实现主观题自动评阅

4. 背景说明

【整体背景】

AI 时代下，取景翻译，智能分词已不是新鲜的技术，初高中英语作文的自动化批改，也在一些线上学习系统中有所应用，但对于长篇大段复杂中文语义分析，仍未有较好的解决方案推出，对于大量应用主观题的公考、职业教育、资质类考试，主观题部分目前依旧依靠老师人工阅卷评分。

【公司背景】

华云数据集团连续四年被评为中国云计算、大数据独角兽。深圳市中诺思科技股份有限公司是以物联网、大数据为技术核心，提供信息化建设为主的高新技术企业，多年来双方在行业产品开发和研究有众多合作，尤其在物流、大数据相关专业的教育类产品已经比较成功，但随着教育越来越鼓励学生对知识的综合运用，而非单纯的知识水平考试，客观题的考察占比越来越低，方案设计等主观题的考察占比越来越高。传统的考评软件，仅支持客观题自动评分，主观题仍需要老师人工打分。如何将老师从批改试卷、作业这类重复性工作中完全解放出来，特编写此命题进行共同探讨。

【业务背景】

随着职业教育越来越鼓励学生对知识的综合运用，而非单纯的知识水平考试，客观题的考察占比越来越低，方案设计等主观题的考察占比越来越高。传统的考评软件，仅支持客观题自动评分，主观题仍需要老师人工打分。在教学智能化日益提升的今天，函需一款软件将老师从批改试卷、作业这类重复性工作中完全解放出来。

5. 项目说明

【问题说明】

与开放性题目不同，主观题虽然不限用户的答题内容，但仍需要求用户能回答到几个关键得分点。而由于中文的语义多样性，对于关键得分点的表述，仍可能存在多种文字描述，而即便出现关键得分点，如果语义完全错误，仍不能认为表述得分，所以需要对用户的回答语义进行正确识别，并对关键词进行匹配。学生可以针对提供的部分测试样例做训练，本实验提供的测试样例均来自教师对于学生的真实评分。

如以下示例：

问题描述：物流系统由什么组成？

答案描述：物流系统是由物流作业系统和支持物流信息流动的物流信息系统

两部分组成。

答案关键字：物流作业系统、物流信息系统

回答方式：

物流作业系统和物流信息系统 满分

物流作业和物流信息 满分

不是物流作业和物流信息 0 分

作业和信息 50%

物流作业 50%

物流管理、物流采集、物流统计 0 分

【用户期望】

（1）提供一个可推广的长文本语义相似度计算方案，提出的方案可广泛用于不同题目；

（2）分析语句的主谓宾结构，识别语句中的否定词；

（3）能够成功完成动词、形容词、副词等的双重否定或反义词的否定与原词的相似性匹配；

（4）能够针对常见的解释性语句做词语及其解释之间的匹配，如：地球（一个天体）；地球是一个天体；地球：一个天体；地球-一个天体；地球，一个天体；

（5）在题目判分中，需实现可按回答关键词命中且符合答案中的释义按点给分；

（6）以试题为例，实现以上描述的主观题自动判分（文本分析试题由企业 提供）。

6. 任务要求

【开发说明】

根据开源的汉语语料库，如 hanLP，jieba 等，对标准答案与学生答案做分析，设计一套比较标准答案与学生答案之间相似度的通用算法，使输入一系列标准答案、标准答案关键字、学生时，给出学生答案与标准答案之间的相似性，考虑每道简答题为 10 分，设计一套合理的评价机制，为学生的答案做评分，注意评分只能为整数且在 0 到 10 之间。

【技术要求与指标】

（1）能够阐明自己设计的系统的评分机制；

（2）考虑算法的效率，包括时间和空间性能；

（3）针对测试样例（部分提供给学生），统计参赛人员设计的算法得到的学生各题分数与教师提供分数之间的绝对差值，所有学生的各题分数绝对差值最小者为优。测试为开放性测试，不仅限于提供的样例数据。

【任务清单】

（1）请参赛者从物流专业职业教育角度，设计主观题自动评阅设计方案；

（2）此方案及具体功能均归参赛队伍所有，参赛者最终需要实现并演示该解决方案；

（3）方案必须在算法上有贡献，规则占比过大的视为无效。

【提交材料】

（1）项目概要介绍；

（2）项目简介 PPT；

（3）项目详细方案；

（4）项目演示视频；

（5）团队自愿提交的其他补充材料。

【开发工具与数据接口】

无要求

7. 参考信息

编程语言不限。

8. 评分要点

本赛题评分要点参考附件一：A 类企业命题统一评分标准。 除此之外，企业还会重点关注算法创新（算法创新：算法独特、新颖，创新元素多，具有技术含量，可以有效地解决此类问题）。