

基于教考分离的在线考试系统设计与实现

张载平, 宋瑾钰

浙江理工大学信息学院, 浙江 杭州

收稿日期: 2022年5月26日; 录用日期: 2022年6月21日; 发布日期: 2022年6月30日

摘要

教考分离有利于教学考核公平、公正和教学质量评价科学、有效, 因此实行教考分离已经成为当前高等教育改革的趋势。针对目前教考分离执行过程中存在的主观题批阅不智能, 监考力度不够大和成绩分析不够全面等弊端, 基于技术难度及实施可行性考虑, 本文设计并实现了一个基于教考分离的在线考试系统, 包括题库管理、自动组卷、在线监考、试题辅助批阅、题目模糊搜索和自动化成绩分析等功能, 有助于缓解教师的压力, 维护考试的公平, 提高教学质量。

关键词

教考分离, 在线考试, 辅助批阅, 成绩分析

Design and Implementation of Online Examination System Based on Separation of Teaching and Examination

Zaiping Zhang, Jinyu Song

School of Information, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: May 26th, 2022; accepted: Jun. 21st, 2022; published: Jun. 30th, 2022

Abstract

The separation of teaching and examination is conducive to the fairness and justice of teaching assessment and the scientific and effective evaluation of teaching quality. Therefore, the implementation of the separation of teaching and examination has become the trend of higher education reform. In view of the disadvantages existing in the implementation of the separation of teaching and examination, such as the lack of intelligence in the marking of subjective questions, the insuffi-

cient invigilation and the insufficient comprehensive analysis of scores, based on the consideration of technical difficulty and implementation feasibility, this paper designs and implements an online examination system based on the separation of teaching and examination, including the functions of question bank management, automatic test paper generation, online invigilation, auxiliary marking of test questions, fuzzy search of questions and automatic score analysis, which helps to alleviate the pressure of teachers, maintain the fairness of examination and improve teaching quality.

Keywords

Separation of Teaching and Examination, Online Examination, Auxiliary Marking, Score Analysis

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

考试是教学过程中必不可少的重要环节,同时也是实施教学目标和教学成果检验的重要手段[1]。通过考试来检验教学成果,不仅可以对学生的知识掌握情况进行了解和评判,同时还可以得到改善教学工作、评测教学质量的宝贵信息[2]。

在习总书记对于高等教育的一系列重要论述,全国教育大会、新时代全国高等学校本科教育工作会议精神和教育部《关于深化本科教育教学改革全面提高人才培养质量的意见》[3]等思想和政策的指导下,教考分离制度已逐步成为高等教育考试制度改革的热门方向。

区别于传统考试制度的“教考合一”,教考分离制度下,任课教师仅需给学生授课,学校专门的考试机构会在授课结束后根据课程教学大纲要求进行命题,并组织实施考试、阅卷和试卷分析工作[4]。

现有的基于教考分离的在线考试系统(如超星平台等)已经实现部分题型(如选择、填空、主观题等)的在线作答和客观题自动批阅等功能,也具备一定的监考防作弊(如切后台检测和屏幕捕捉等)措施,但是仍存在诸如主观题批阅需要耗费大量时间和精力和人力资源,在线考试的监考漏洞较多和成绩分析功能单一等问题。因此,本文设计了一款基于教考分离的在线考试系统,并结合关键词算法、相似度匹配算法和瞳孔检测算法等进行实现,致力于解决在线考试系统在主观题批阅、监考和成绩分析等方面存在的问题。

2. 教考分离发展现状

传统的教考合一主要由科任教师自己独立命题和评分[5],但命题的科学性难以保证,考试的标准也不明确。多数学生会被动针对教师划定的考试范围突击学习,死记硬背来应付考试,无法为教师教学水平和教育质量的客观评价提供有效依据,从而影响了师资队伍建设和人才培养质量的提高[6][7]。

教考分离的实施,有助于更公正地评价教师的教学水平,更客观地反映学生的学习效果[5]。也能让学生认识到平时是否努力和所学课程掌握与理解程度是影响成绩的主要因素。同时,也可以提高学生的学习动力,培养学生养成良好的自主学习能力和习惯[8]。教考分离势在必行。

近年来越来越多的学校推进课程的考核方式改革,实施教考分离,但是实践结果显示,教考分离还存在许多问题:比如教考分离后存在教师教学内容和最终考试内容不一,试题库的质量不高[8]、更新缓慢,监考漏洞较大、诚信问题追溯困难,判卷耗时长、评分标准不一、成绩分析缺乏、考核功能的单一化和片面化以及总体难度难以把控等问题。

针对以上现状, 基于先进技术及其实施可行性考虑, 本系统在建立题库过程中对题目进行标签化处理; 在监考系统中设置了防作弊算法; 在批卷过程中设置了辅助阅卷功能, 可以标记出答案中的关键词, 并且给出参考分数; 在成绩分析过程设置自动生成教学建议等环节, 希望通过上述措施来更好地推动教考分离的执行, 提高考试对教学质量反馈的科学性, 促进教学质量的提高。

3. 系统需求分析与设计

3.1. 需求分析

本系统的主要服务对象为教师用户、学生用户和管理员用户。教师用户负责使用本系统开展考试、获取考查结果, 主要功能为创建考试、监考、试卷批阅和查看成绩分析。学生用户负责使用本系统参加考试、获取考试成绩, 主要功能为查看考试日历、参与考试和查看成绩分析。管理员用户负责维护并测试系统、管理题库和用户信息, 主要功能为用户管理、题库管理和考试管理。

根据用户群体的需求, Web 端向教师、学生和管理员用户提供服务。教师通过 Web 端进行账号的登录, 完成登录后即可使用题库管理、组卷、考试管理、监考、考试批阅和数据分析等操作。学生通过 Web 端进行账号的登录, 完成登录后即可使用考试、站内信、考试分析。平台管理员通过 Web 端进行账号的登录, 完成登录后即可使用包含教师和学生用户的各项功能, 还可以进行用户管理、题库管理、考试管理等操作。移动端主要为学生服务。移动端完成登录后即可使用考试、成绩分析、考试日历等操作。具体功能如图 1 所示。

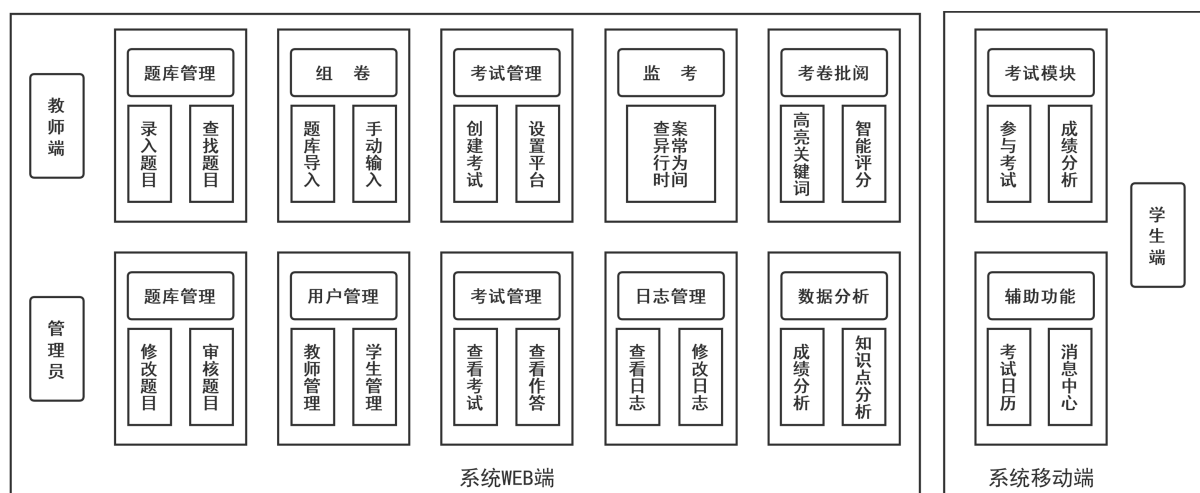


Figure 1. Function module diagram of online examination system based on separation of teaching and examination
图 1. 基于教考分离的在线考试系统功能模块图

3.2. 系统架构

系统通过 WEB 端和移动端前端系统实现用户交互。本系统架构如图 2 所示, 可分为应用层、服务存储层和计算处理层三层。

1) 应用层

应用层为用户提供了 WEB 端和移动端两大端口, 利用符合人机交互原理的交互界面为不同类型的用户提供可操作性高的应用程序。WEB 端口面向所用用户, 使用 NodeJS+Express 框架进行搭建, HTML+CSS+JS 和 Element UI 进行页面布局。移动端面向学生用户, 基于 Android 原生和 Glide 等框架进行开发。

2) 服务存储层

服务存储层实现了本系统的业务处理和信息交互, 使用 MySQL 搭建系统数据库, 使用 Apache 和 Tomcat 将数据部署到服务器上, 并通过 Spring Boot、Druid 和 ZooKeeper 等向应用层提供交互数据, 接受来自计算处理层的计算数据。

3) 计算处理层

计算机处理层实现了本系统的算法处理。该层利用 Bert 模型和 OpenCV 库等 AI 模型, 使用 Python 将捕捉到的数据进行逻辑编码, 以 JSON 的形式和服务存储层中的数据库进行数据交换。

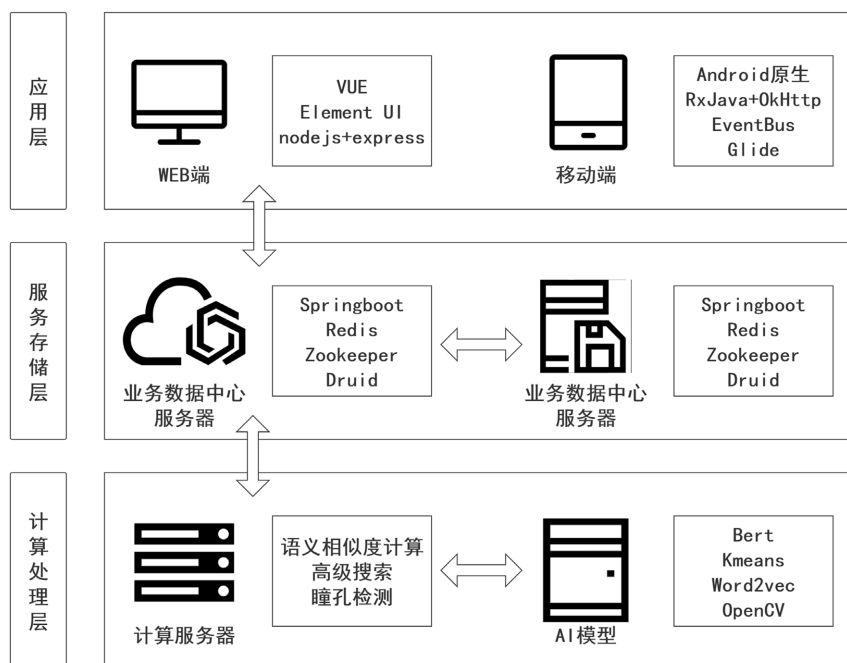


Figure 2. Diagram of the system architecture

图 2. 系统架构图

4. 系统实现

4.1. 系统 WEB 端

WEB 端作为本系统的主要端口, 为所有用户提供了对应的功能。教师登录系统后可在考试管理页面中选择添加考试, 选择组卷方式后填写相应的考试名称、平台、时间、是否开启反作弊等考试信息。随后导入考生列表, 查看考卷内容。试卷发布后, 学生可以登录系统, 选择正在进行的考试进行答题。教师端可以监测异常情况。考试结束后, 教师可以点击考试进行批改阅卷, 点击自动批改按钮则可直接对客观题进行批改, 并获取相应的分数。对于主观题, 系统将会对答案进行提取关键字, 并与学生回答关键字进行相似度分析, 对学生回答进行高亮关键字效果, 和计算相应的得分, 以供阅卷教师进行参考, 提高工作效率。系统以图表的形式, 为教师提供直观的考试综合情况, 可以查看考试的均分, 方差等基本信息, 以及考试知识点等掌握情况, 如图 3 所示。最后系统会为教师提供合理的教学建议以供教师参考。为了提升师生的使用体验, 教师在对考卷进行编辑时, 可以点击导入, 从已有的题库中选择所需的题目, 则可以自动将题目按照题目类型导入到试卷中。同时为了提高效率, 教师可以拍照上传题目, 导入系统。题库管理中还为教师提供了模糊搜索的功能, 方便查找试题。

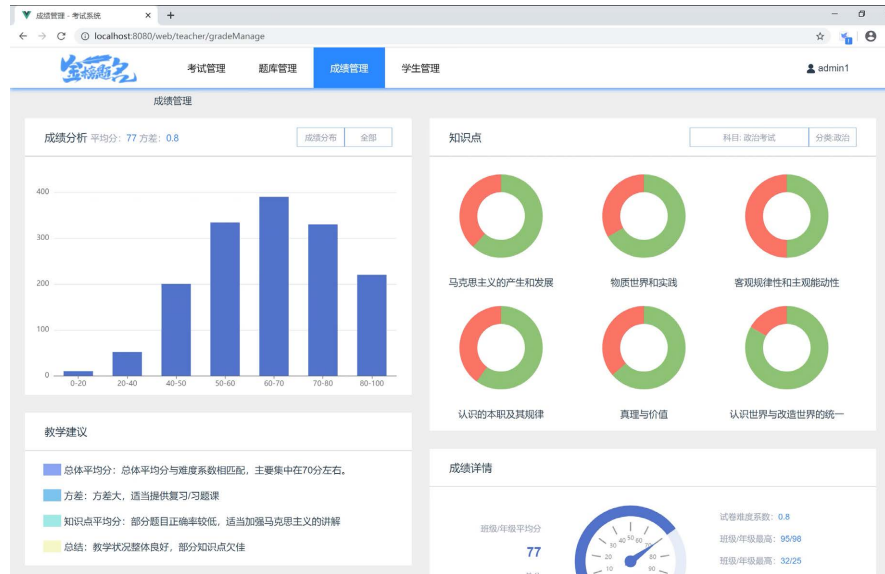


Figure 3. Performance analysis
图 3. 成绩分析

4.1.1. 试题辅助批阅

系统提供的辅助主观题批阅的功能可以有效地减轻教师批阅负担。本功能在乔亚男等[9]提出的基于特征融合的方法指导下实现, 通过融合语意和关键词的相似度, 使用模拟评分环境的 Sentence-Bert 数据集进行训练; 阅卷老师可通过系统自动高亮的考生答案的关键词, 方便提取考生的得分点; 收集一定量的人工批阅数据做训练后, 可开启自动给分的功能, 提高批阅效率。

试题辅助批阅主要通过关键词的提取和相似度匹配来实现的, 通过 Sentence-Bert 模型进行训练。具体实现过程如下:

1) Sentence-Bert 模型

本系统使用便于处理句向量的 Bert 的变体 SentenceBert 模型, 其优势在于更快的句向量计算速度。该模型使用了孪生(siamese)结构, 可以更好地使用(句子 A, 句子 B, 相似度)的元组进行训练, 如图 4 所示。

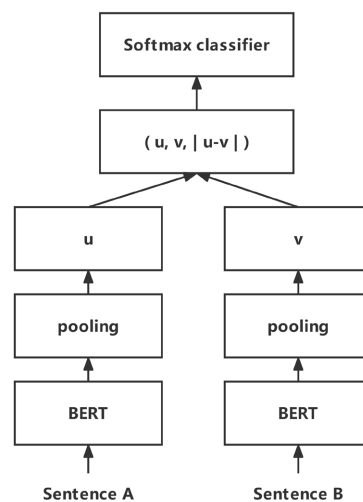


Figure 4. Diagram of the Sentence Bert structure
图 4. Sentence-Bert 结构图

2) 关键词算法

本模块在夏天[10]提出的词向量聚类加权 TextRank 的关键词抽取指导下实现。关键词算法使用 word2vec 作为 embedding, 使用 K-means 算法做聚类。在使用 TextRank 算法建立文章的词图时结合了词语间转换的权重关系进行来给出关键词, 使得可以使用较少的资源达成融合多方面权重的效果。

$$t(u) = d \sum_{v \in adj[u]} p(v \rightarrow u) t(v) + (1-d) \frac{1}{|V|}$$

其中, d 是阻尼系数, 通常取为 0.85。而 $adj[u]$ 代表 u 的相邻节点的集合, 而 $p(v \rightarrow u)$ 代表从 v 节点转移到 u 节点的概率, 称之为覆盖影响力概率。

3) 相似度匹配

在得出关键词排序后, 遍历排序后的词表, 剔除禁用词之后可得前 K 个关键词。使用关键词对应的词向量与原题目答案使用相同方法得到的关键词的词向量计算它们的最短词移距离。由此可得关键词相似度 $Similarity_{key}$ 。

$$W = \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij} \rightarrow \min$$

$$Similarity_{key} = \frac{1}{1+W}$$

其中, W 为最短词移距离, 本系统将原文的算法中的调节因子取为 1。

在获得经过训练后的 SentenceBert 模型后, 可以直接将句子转化为一个句向量。与原题目答案计算出的句向量做余弦相似度计算后可得语义相似度 $Similarity_{sentiment}$ 。

$$Similarity_{sentiment} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2 \times \sum_{i=1}^n (B_i)^2}}$$

其中, A 与 B 表示需要对比的两个句子的文本向量。

最后将上两式子赋予权重, 融合并给出评分。根据题目的特性不同, 也可以结合实际人工评分来进一步学习评分权重。在本系统中的应用如图 5 所示。

$$Similarity = \omega_1 * Similarity_{key} + \omega_2 * Similarity_{sentiment}$$

其中, ω_1 为关键词相似度的权重, ω_2 为语义相似度的权重。本系统初始的评分权重设置为各 50%, 可以根据实际情况控制权重的比例。

4.1.2. 题目模糊搜索

在题库中题目数量较大时, 由于语义表达的多样性, 仅靠关键词匹配的搜索会错失一些语义相似但表述不同的题目。模糊搜索功能可以提高教师用户的体验。具体实现过程如下:

1) 题目预处理

本功能模块利用上述的 Sentence-Bert 训练模型, 在题目被录入题库的同时将其向量化并存储下来。

2) 语义相似度计算

给出搜索的字符串后, 将其向量化, 随后与题目的句向量做余弦相似度计算。并给出排序。

3) 输出

根据排序, 剔除禁用结果后输出搜索结果。

本功能在系统中的应用如图 6 所示。



Figure 5. Auxiliary marking of test questions
图 5. 试题辅助批阅



Figure 6. Fuzzy search
图 6. 模糊搜索

4.2. 系统移动端

移动端主要为学生用户提供考试功能。若考试允许移动端进行, 学生点击进入, 会跳转至考试信息页面, 勾选阅读后方可进入答题界面。答题界面在回答当前问题后自动跳转至下一题, 也可以通过答题卡跳转至对应题目, 摄像头预览框每秒自动捕捉用户的图像上传至服务器进行作弊判断。学生在移动端内可点击已出成绩的考试跳转至成绩分析页面, 链接考试题目的知识点, 生成用户答题知识点正确率情况, 如图 7 所示。学生在移动端首页可以点击进入“考试日历”, 查看用户指定年月日之后的教务系统中的考试安排, 如图 8 所示。



Figure 7. Performance analysis (mobile terminal)
图 7. 成绩分析(移动端)

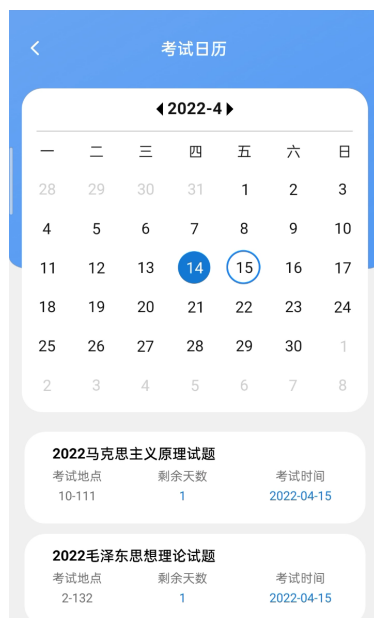


Figure 8. Examination calendar
图 8. 考试日历

当学生在移动端考试, 系统可以追踪检测瞳孔位置, 当检测到较大幅度眼球、头部和身体运动时, 可以及时把异常记录反馈给教师用户。教师用户可以通过 WEB 端监考模块对学生做提示, 也可以进一步观察学生的考试状态, 如图 9 和图 10 所示。具体实现过程如下:

- 1) 调用 OpenCV 库识别瞳孔。
- 2) 按照设定的时间间隔, 分别记录两个瞳孔位置。
- 3) 记录瞳孔位置的次数达到事先设定的阈值后, 计算两个瞳孔各自的平均位置及标准差, 重置记录。

4) 跳转至 2), 同时开始执行异常检测。

检测公式如下:

$$\text{check} = (\text{pupli} - \overline{\text{pupli}}) / \sigma(\text{pupli})$$

warning occurs when $\text{check} \geq \text{Parameter}$

其中, Parameter 是一个可调参数, pupli 为瞳孔的平均位置。

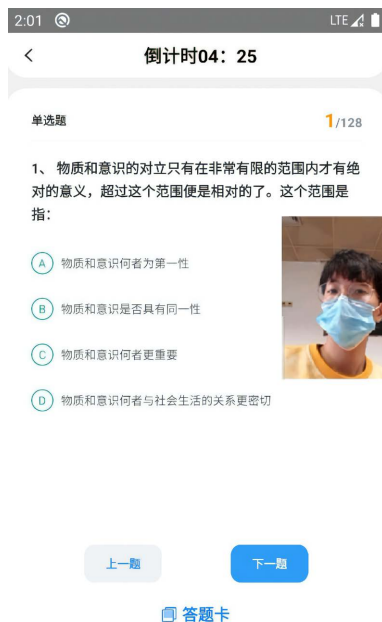


Figure 9. Pupil detection-1
图 9. 瞳孔检测-1

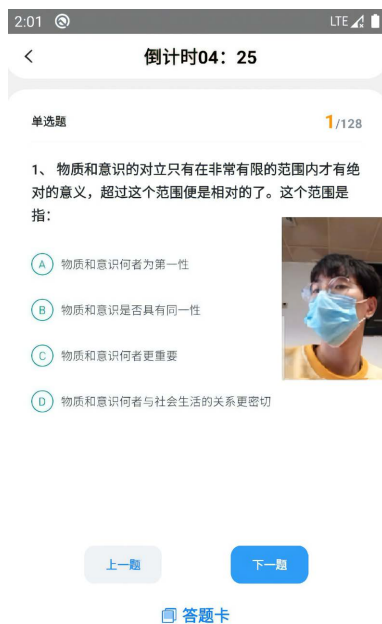


Figure 10. Pupil detection-2
图 10. 瞳孔检测-2

5. 结语

本文根据教考分离模式, 创新性地融入了人工智能 AI 算法, 提出了一种基于教考分离的在线考试系统的设计并进行实现, 为推动在线化教考分离模式的运行提供了新的思路。本系统利用关键词算法, 实现了试题模糊搜索和提取考生答案关键词等功能, 提高了教师的工作效率。在关键词算法的基础上, 利用关键词的相似度匹配算法, 实现了主观题辅助批阅功能, 提高了阅卷的效率。同时, 基于 OpenCV 库的瞳孔检测算法, 实现了考试过程中的作弊行为检测, 维护了考试的公平性。本系统提供的自动化成绩分析的功能, 为教学质量的改善提供了科学性的建议。

基金项目

浙江理工大学 2021 年高等教育科学研究课题资助项目。

参考文献

- [1] 傅军. 完善教考分离加强试题库建设与管理[J]. 考试周刊, 2007(49): 2-3.
- [2] 柴春红, 刘希军. 高等数学课程“教考分离”改革的效果分析[J]. 职业时空, 2011, 7(5): 71-72.
- [3] 曹晓菲, 王小才, 余柏林. 数学类课程实施教考分离模式的必要性[J]. 教育教学论坛, 2021(43): 136-139.
- [4] 董杨琴, 曹洪其. 以改革课程考核为切入点推进高职院校质量工程建设[J]. 南通职业大学学报, 2010, 24(2): 59-61.
- [5] 魏首柳, 柯小玲. 基于教考分离的课程试题库建设的探讨与实践[J]. 高教学刊, 2021(3): 100-103.
- [6] 王玲玲. 药学《仪器分析》课程“教考分离”考核模式初探[J]. 广东化工, 2021, 48(15): 247+254.
- [7] 龚晓燕, 杨浩. 基于随机搜索逐步优化的自动组卷算法及试题库结构的研究[J]. 考试周刊, 2007(49): 1-2.
- [8] 蔡永茂, 汤滨, 朱科铃. 对高校教考分离考试模式的思考[J]. 教育教学论坛, 2020(42): 320-321.
- [9] 乔亚男, 胡名凯, 薄钧戈. 基于特征融合的主观题智能阅卷算法研究[J]. 电气电子教学学报, 2021, 43(6): 104-111.
- [10] 夏天. 词向量聚类加权 TextRank 的关键词抽取[J]. 数据分析与知识发现, 2017, 1(2): 28-34.