



北京大学

硕士研究生学位论文

题目： 软件工程课程的学生个性化
评价系统设计与实现

姓 名：	柳家奎
学 号：	2001210362
院 系：	软件与微电子学院
专 业：	软件工程
研究方向：	智能科技
导师姓名：	朱郑州 副教授

二〇二三年四月

摘要

学生学习评价指利用一定的技术手段,以学生为对象,通过具体的指标解读学生的学习质量或状态。新时代教育评价改革指出要完善学生科学、德育、教育和学业等方面评价;同时,随着信息技术和课程教学深度融合,混合式教学模式成为高校当下较为热门和有效的教学方式。针对当前混合式教学模式下的学生评价个性化程度不足这一问题,本文以软件工程课程为研究对象设计一种学生个性化评价系统,结合最近发展区和成果导向教育理念,从多个维度对学生进行针对性评价,在一定程度上解决学生学习评价个性化程度不足的问题,提升了教学质量。本文的主要贡献有:

(1) 生成数据驱动的学生个性化评价报告。本文所设计的数据驱动的个性化评价报告生成算法融合了 OBE 和最近发展区等前沿教育理论,从学生知识掌握度、课程素养和学习风格三个维度对学生进行个性化评价,基于学习数据实现多维特征构造并使用基于 KNN、朴素贝叶斯、决策树、SVM、随机森林、Xgboost 等多种机器学习算法的融合模型对学生知识掌握度进行分析设计;使用基于学生数据的评分规则算法对课程素养和学习风格评价进行设计,进而生成高质量的学生个性化评价报告,提高了教学质量和课堂满意度。

(2) 生成学生成绩预测报告。在成绩预测算法设计上,使用多种机器学习算法模型基于多维度学生特征进行模型预测对比实验,并以此为基础,设计实现基于决策树、随机森林和 MLP 前馈神经网络的成绩预测集成融合模型,且学生成绩预测准确度高。

(3) 完成系统开发和教学应用实验,使用最新技术框架编码实现学生个性化评价系统并应用于《软件工程》课程教学项目。应用结果显示,软件工程课程的学生个性化评价系统界面简单直观,操作方便,系统运行正常且对学生的评价和预测结果符合设计预期;应用后的学生章节平均成绩和课堂满意度有所提升,学生对评价报告的准确性和满意度较高,系统应用效果好,具有推广价值。

通过以上研究,实现面向软件工程课程的学生个性化评价系统,该系统评价有着较高的个性化程度,能够一定程度解决学生评价个性化程度不足这一问题,其相关应用能帮助教师因材施教,提升教学质量,对学生能强化正反馈效果,提升对软件工程课程教学满意度,激发学生自主学习动机,为成果导向的个性化教学设计持续改进提供解决方案。

关键词: 个性化评价, 成绩预测, 机器学习, 最近发展区, 软件工程

Design and Implementation of Student Personalized Evaluation System for Software Engineering Courses

Liu JiaKui (Software Engineering)

Directed by Zhu ZhengZhou

ABSTRACT

Student learning evaluation is the use of certain technical means to interpret the quality or status of students' learning through specific indicators. The new era of education evaluation reform points out the need to improve student evaluation in scientific, moral, educational, and academic aspects. At the same time, with the deep integration of information technology and curriculum teaching, blended learning has become a popular and effective teaching method in universities. In response to the problem of insufficient personalization in student evaluation under the current blended learning mode, this paper designs a personalized student evaluation system for the software engineering course, combining the Zone of proximal development and Outcome based education concepts to provide targeted evaluations of students from multiple dimensions. This helps to some extent to solve the problem of insufficient personalization in student learning evaluation and improve teaching quality. The main contributions of this paper include:

(1) Generating data-driven personalized student evaluation reports. The data-driven personalized evaluation report generation algorithm designed in this paper integrates cutting-edge educational theories such as Outcome based education and Zone of proximal development, Personalized evaluation of students is carried out from three dimensions: knowledge mastery, course literacy, and learning style. Multi-dimensional feature construction is achieved based on learning data, and a fusion model using multiple machine learning algorithms such as KNN, Naive Bayes, Decision tree, SVM, Random forest, Xgboost, etc.is used to analyze and design student knowledge mastery. A rating rule algorithm based on student data is used to design course literacy and learning style evaluation. Generate high-quality personalized student evaluation reports, which improved teaching quality and student classroom satisfaction.

(2) Generating student performance prediction reports. In the design of score prediction algorithm, a variety of machine learning algorithm models are used to carry out model

prediction contrast experiments based on multi-dimensional student characteristics. Based on this, a performance prediction integration fusion model based on Decision tree, Random forest and MLP feedforward neural network is designed and implemented, and the accuracy of student score prediction is high.

(3) Completing system development and teaching application experiments. The latest technology frameworks are used to code and implement the personalized student evaluation system, which is applied to the "Software Engineering" course teaching project. The application results show that the personalized evaluation system interface for students in software engineering courses is simple and intuitive, easy to operate, the system runs normally, and the evaluation and prediction results for students are in line with the design expectations. After application, the average score of student chapters and classroom satisfaction have been improved, and students have high accuracy and satisfaction with the evaluation report. The system application effect is good, and it has promotion value.

Through the above research, a personalized student evaluation system for the software engineering course is realized, which has a high degree of personalization in evaluation and can solve the problem of insufficient personalization in student evaluation to some extent. Its related applications can help teachers teach according to students' aptitudes, improve teaching quality, and strengthen positive feedback effects on students, thereby enhancing students' satisfaction with software engineering course teaching and stimulating their motivation for independent learning. It provides a solution for continuously improving outcome based personalized teaching design.

KEY WORDS: Personalized evaluation, Performance prediction, Machine learning, Zone of proximal development, Software engineering

目 录

第一章 绪论.....	1
1.1 研究背景及意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究问题.....	2
1.1.3 研究意义.....	2
1.2 国内外研究现状.....	3
1.3 研究内容.....	7
1.4 论文组织结构.....	8
第二章 相关理论和技术	9
2.1 经典教育理论.....	9
2.1.1 最近发展区理论.....	9
2.1.2 成果导向教育.....	10
2.2 学生评价算法模型.....	10
2.2.1 基于分类模型的学生分类.....	10
2.2.2 基于回归模型的学生成绩预测.....	11
2.3 本章小结.....	13
第三章 需求分析和系统概要设计	14
3.1 系统需求分析.....	14
3.1.1 学生个性化评价功能需求分析.....	15
3.1.2 学生成绩预测功能需求分析.....	16
3.1.3 其他功能需求分析.....	17
3.1.4 非功能性需求分析.....	17
3.2 系统概要设计.....	18
3.2.1 体系结构设计.....	18
3.2.2 系统功能模块设计.....	18
3.2.3 系统总体架构.....	19
3.2.4 业务流程设计.....	20
3.3 本章小结.....	21
第四章 学生个性化评价系统算法设计	22
4.1 总体方案设计.....	22
4.2 实验数据.....	23
4.2.1 数据预处理.....	23
4.2.2 多维特征构造.....	25
4.3 学生知识掌握度评价算法设计	27

4.3.1 成果导向的课程知识教学目标确定.....	27
4.3.2 学生知识水平最近发展区诊断.....	28
4.3.3 评价算法设计.....	29
4.4 学生课程素养评价算法设计	36
4.4.1 评价指标设计.....	37
4.4.2 评价标准设计.....	38
4.4.3 评价算法设计.....	39
4.5 学生学习风格评价算法设计与实现	40
4.5.1 评价指标与评价标准设计.....	41
4.5.2 评价算法设计.....	42
4.6 学生成绩预测算法设计.....	43
4.6.1 特征选择.....	44
4.6.2 成绩预测算法设计.....	45
4.7 本章小结.....	51
第五章 学生个性化评价系统的实现与应用	52
5.1 系统实现.....	52
5.1.1 系统编码设计.....	52
5.1.2 系统整体架构实现.....	52
5.1.3 数据上传与查询功能实现.....	52
5.1.4 个性化评价报告功能实现.....	53
5.1.5 学生成绩预测报告功能实现.....	54
5.2 系统测试.....	54
5.2.1 系统功能性测试.....	54
5.2.2 系统非功能性测试.....	55
5.3 系统应用效果评估.....	56
5.4 本章小结.....	62
第六章 总结与展望	63
6.1 总结.....	63
6.2 展望.....	63
6.2.1 系统应用前景展望.....	63
6.2.2 学生评价研究展望.....	64
参考文献.....	65

第一章 绪论

基于成果导向教育（Outcome Based Education, OBE）理念的混合式软件工程课程教学在教学培养目标上设计了知识教学目标和职业素养培养等目标，在帮助学生理解和应用软件开发基本理论方法的同时注重学生综合素养的培育。以此为背景的学生评价在帮助教师完成课程培养目标的过程中起着至关重要的作用。传统的学生评价方式中，对教师依赖较高，更注重学生的课程成绩达成度，个性化程度不足，在全面反映学生的实际学习状态和课程素养上具有一定的困难，对教学效率的提升不足。新的学生评价模式要全面采集和分析学生在混合式教学课程中的数据，有针对性的评估学生的知识水平、课程素养和学习风格等方面，提升评价的个性化。因此，需要设计和开发一种基于学生数据的软件工程课程学生个性化评价系统，以此对成果导向教育的软件工程课程教学建设提供有益帮助。

1.1 研究背景及意义

1.1.1 研究背景

本文研究依托 2022 年北京大学教学新思路 2.0 项目，项目名称为：“基于 OBE 持续改进的教学评价——以《软件工程》为例”，项目研究旨在通过应用多模态学生数据，将人工智能算法与教学相融合，支持动态适应性的学习过程以及多维度学生综合能力评价，目前该项目已经结题并获得优秀结题项目。

新时代教育评价改革指明要以全面发展的方式对学生评价进行改革，强调科学、德育、教育、学业等全方位评价，新时代教育教学改革背景下，探索针对学生教育的个性化评价的重要性不言而喻。关注学生个性化发展是当今教育改革的重要方向，个性化教育尊重个体差异，提供多样化的教育方式，而学生个性化评价是个性化教育中的重要环节之一。同时，线上线下结合的混合式教学模式成为高校创新教学模式的主流，数字化技术在此类教学模式中的深度应用使得大量学生学习数据被保存或记录，亟需利用这些学生数据来提升学生评价的个性化程度。基于上述背景，当前某大学软件工程课程基于 OBE 教学的学生评价面临部分问题：

（1）混合式教学模式大量学生数据未被挖掘利用，该教学模式使用多个教学平台，教师难以同时监控多个平台下学生的学习状态，难以做到因材施教。

（2）课程设计基于成果导向理念，但该理念还未融合到日常学生评价过程；教学设计理念尊重学生个体差异，但缺乏针对个体差异的个性化诊断方法，个性化程度不

足。

(3) 课程测试和作业等任务多, 但学生难以及时得到针对性和个性化的反馈, 课程规模较大, 教师在教学过程和期末都难以对学生做出准确和全面的评价。

1.1.2 研究问题

本文针对当前混合式教学模式下, 对学生的评价个性化程度不足这一问题进行研究, 当前学生学习评价个性化程度不足体现在:

(1) 评价在内容方面主要评价学生的学业成绩达成情况, 个性化和针对性不足, 缺乏对课程实际知识教学目标和综合素养教学目标的融合, 大多成果导向教育的课程设计仅停留在教学过程, 没有融入学生评价, 评价缺少对不同个体不同发展水平的个性化诊断; 在内容反馈上多以直观数据反映学生指标状态, 缺乏完整的个性化评价报告。

(2) 在评价维度方面, 多以单一的评价维度反映学生某一方面的状态, 缺乏基于多个维度并应用不同评价方法进行的全面个性化评价,

(3) 在评价方法上, 大多采用过程性的评价, 很少使用过程性和终结性评价相结合的联合评价方式, 难以使评价完整贯穿整个教学过程。

本文结合最近发展区理论和成果导向教育理论, 从评价效果、个性化程度和评价模式上一定程度解决当前学生学习评价研究中所存在的不足。

1.1.3 研究意义

在传统学生学习评价模式上, 教师依据学生课程成绩或日常课堂表现给与的主观性评价, 这种学习评价模式存在客观性不足、个性化不足、评价效率低、难以形成全面性的评价等问题。在混合式教学模式推广以及信息技术与教学深度融合后, 学生学习评价在客观性和评价效率上有较大提升, 进一步减轻了教师工作量, 但对学生的评价仍然存在反馈延迟, 线上线下分立, 评价维度少等问题。基于上述背景, 本文所研究的软件工程课程的学生个性化评价系统具有以下研究意义:

(1) 推进个性化教育, 尊重学生个体差异, 帮助教师因材施教。通过学生数据分析, 给予不同水平、不同发展区间的学生针对性的评价, 评价由教师生成、编辑和发布, 使教师动态掌握每一个学生在知识掌握度, 课程素养和学习风格上的状态, 便于因材施教。

(2) 提升教师教学效率, 提高整体教学质量。帮助教师更及时准确的了解学生学习状态, 教师能够针对学生的批量过程性评价对教学方案进行调整, 例如评价显示大批学生出现某章节的掌握度过高或过低时, 教师对教学难度或课程内容进行调整。此外, 评价报告的自动生成能够大幅减少教师工作量, 提高教学过程中的互动性和学生

满意度。相关研究可以重点应用到混合式教学领域,提升线上智能化评价系统的效果,本研究实例将率先应用到某高校《软件工程》日常课程教学,丰富课堂个性化教学内容。

(3) 激发学生自主学习动力和意愿。结合课程教学目标和学生最近发展区诊断,给予学生贯穿整个教学过程的过程性和终结性相结合的个性化学习评价,使学生及时获得正反馈,明确发展方向和区间,激发自主学习动力,该评价过程是成果导向教育教学设计的重要环节。

本文研究充分考虑了教学评价对学生学习目标达成和个体发展方向的影响,扩充了评价维度,结合多维的学生学习数据使整个评价过程得到完整的数据支持。能够一定程度上为当前学生个性化评价领域存在的问题提供参考或解决方案。

1.2 国内外研究现状

学生个性化评价是关注学生独特发展特性的评价,是摒弃唯分数论评价的一种多样化评价的方式,个性化旨在尊重个体间差异,帮助学生个性化的发展。当前对学生评价的大部分研究在评价模型的建立,评价的应用以及评价维度上都具有一定的个性化程度,值得深入研究和借鉴。

(1) 学生评价的运用

学生评价从对评价结果的运用上可以划分为用于改进的“形成性评价”(Formative Assessment)和用于考评的“终结性评价”(Summative Assessment)两种^[1],如表 1.1 形成性与终结性评价对比。

表 1.1 形成性与终结性评价对比

序号	评价分类	说明	特点
1	形成性评价	用来诊断	重视教育和发展、不间断的动态过程、帮助教学改进、尊重不同个体的不断发展
2	终结性评价	用来考核	形成某种判断、衡量学习水平或教学水平、评价方式具有多元化

形成性评价是为了学习改进而进行的评价,其主要关注和收集学生学习过程并进行分析,教师可以依据形成性评价结果对教学过程进行改进以提高效果^[2],其在教学过程中一般发生在终结性评价前。部分学者对形成性评价的指标和策略进行研究,探究不同形成性评价机制在教学中的应用。例如学者马志强^[3]对在线同侪互评这一形成性评价策略进行研究,并对其现有研究方法、工具和效果等进行了综述,证明后续研究侧重于网络教学和教师评价的结合。翟苗等^[4]学者对高校混合式教学形成性评价指标进行

了研究,设计了高校混合式教学形成性评价指标模型,为混合式教学的学生评价方案提供参考。张晓刚^[5]基于程序开发类课程实践,探索混合式教学下的形成性评价体系,从三个方面对形成性评价体系进行了构建。部分学者通过应用实验证明了形成性评价效果的优越性,例如曹国刚等^[6]对形成性评价在软件工程相关专业课程中的应用进行研究和实践,实验结果显示形成性评价的教学模式在效果上优于传统的终结性评价。

在终结性评价中,评价结果一般作为判断标准来对学生的学习状况或教师的教学质量进行考评,其主要评价形式包括考试,考察等。传统的终结性评价单独使用难以达到更好的效果,大多学者对终结性与形成性的联合评价进行研究,例如刘雅^[7]通过研究计算机基础课程中的形成性评价与终结性评价联合评价,证明使用联合评价来完善考核反馈机制能够达到更好的效果。学者 Zhang Yi^[8]基于教师可持续发展理念和互动分析系统,从形成性评价和总结性评价两个方面对课堂教学进行评价,为终结性评价提供新思路。部分学者在传统终结性评价基础上创新,探究新的终结性评价方式,例如学者 Suaco P 等^[9]对在线学习期间使用协作评估的终结性评价方式进行研究,证明此类终结性评价有助取得好的学习成绩和高质量的产出,且效果不受年纪和性别差异影响。

总之,而终结性评价与过程性评价并不是完全对立的关系,一些评价方式可以既是形成性的又是总结性的,既有联系又有区别,在部分教学模式下可以相互转化。

(2) 学生评价方法

如表 1.2 数据挖掘与适应性测试对比,学生评价在评价方法上可以划分为学习数据挖掘和学生适应性测试。

表 1.2 数据挖掘与适应性测试对比

评价方法	优点	缺点
适应性测试	1. 通过环境调整适应学生个性特征 2. 能够实现过程性和诊断性的导学 3. 能够随时进行测试和诊断	1. 学习诊断人力时间成本高 2. 测试问题难以考虑个体差异 3. 受个体主观因素影响较大
数据挖掘	1. 实现自动化、智能化的评价 2. 较好的利用信息化教学数据 3. 真实性、客观性较强	1. 对学生数据集有较高要求 2. 准确性易受较多因素影响 3. 部分评价维度难以量化

学生适应性测试是根据学生个体水平差异进行不同难度的测试,能够提升学生评估准确性,两种评价方法在应用上各有优缺点,将数据挖掘和传统适应性测试相结合的智能化适应性测试在学生评估的个性化程度上有较好的效果。牟智佳等^[10]使用布鲁姆教学目标分类和 Q 矩阵理论,基于学生测评数据构建评价模型并进行用户验收测试,能够个性化评价学生测试结果。学者贾积有等^[11]基于最近发展区等理论和最新人工智

能技术,设计实现了学生智能评测和辅导系统,智能性的为每个学生提供个性化、适应性的测试。

学习数据挖掘手段进行评价的方法在混合式教学或线上教学模式下有广泛应用,其核心是使用针对课程的改进型数据挖掘方法提升评价的准确度,例如学者刘进^[12]提出了从标准化到数据化的评价方法转向,论证了传统评价研究方法的不足,提出基于大量数据的评价研究方法并验证可行性。学者马辛已^[13]基于 K-means 聚类技术对大量教学数据进行挖掘并建立学生画像,实现对学生综合能力的精准评价。周楠等^[14]使用深度学习技术对学生行为视频信息和表情检测数据进行挖掘,为学习者学习行为识别和教学效果评估提供了解决方案。

(3) 不同教学模式下的学生评价

学生评价针对不同的教学模式,又主要分为课堂学习评价(Classroom Learning Evaluation),混合式学习评价(Blending Learning Evaluation)和在线学习评价(Online Learning Evaluation)。课堂学习评价是基于学生课堂学习行为和数据进行评价的过程。学生导向、诊断及激励功能是课堂教学评价的主要价值体现^[15]。学者张志红针对当前课堂参与式评价窄化、缺乏有效评价工具和个性化需求关注不足的问题,就参与式课堂评价的理念和策略进行研究,提出共享学习目标、获取目标达成证据、调动积极情感等解决措施^[16]。王楠对学生课堂关注度评价进行研究,针对测评方法缺乏数据集和准确率等问题,构建了有效的数据集和学生课堂专注度联合评价模型^[17]。线上学习评价方面,魏勃颀等^[18]基于线上学习过程性评价形式单一这一问题,提出基于隐马尔可夫模型的多维评价方法实现对学习状态持续跟踪,实验验证表明该模型下的线上学习多维评价结果能有效反映学习状态过程性变化。学者郑勤华^[19]等通过学习分析技术进行学生在线学习建模与评价,为在线评价模型提供科学可行的模型解决方案。

在信息化教学时代,混合式教学更能够为学生提供个性化学习的计划,成为不断成为促进学生学习绩效提升的最佳方式。混合式学习评价是对课堂学习和在线学习评价的融合。学者 Wu M^[20], 牟智佳^[21]等通过研究学生混合式教学下的多模态数据,构建了学习成绩预测模型,模型实现基于机器学习算法,以此进行个性化学习结果预测和混合式学习表现评估,取得了较好效果。张玉欣等^[22]构建了涵盖线上、线下指标的评价体系,设计基于 PNN 神经网络的学习评价模型并进行仿真实验,为混合式教学下的过程化、多样化的学生评价提供了一种可靠方案。Yanfei Shen 等^[23]基于 CSE 评价模式(Center for Study of Ealuation, CSE)探讨了混合式教学模式下以学生为中心、过程、多元评价为导向,融合能力、态度、过程和结果的全面性评价,并尝试在大学英语教学中构建混合式学习评价模型,实验结果表明该评价模式能够提升积极性和学习能力,引导学生自主高效学习。

(4) 学生评价维度

如表 1.3 常见学生评价维度所示, 学生评价在评价维度研究上可主要分为学生学习效果评价、学生满意度评价、学生学习风格评价、学生综合素质评价、学习结果预测和其他混合维度评价等。

表 1.3 常见学生评价维度

序号	评价维度	常用理论或方法	特点
1	学习效果评价	神经网络, 成果评估体系	需要建立学习效果评估体系
2	学生满意度评价	美国高校学生满意度测评	适合使用问卷量表形势进行评估, 同时能够反映教学质量
3	学习风格评价	Felder- Silverman 学习风格模型	适合使用问卷量表形式进行评估或在量表特征上改进
4	综合素养评价	学生画像、五育并举、中国学生核心素养	需要依据不同教学要求建立综合素养评价模型
5	学习结果预测	机器学习、知识图谱	对学生数据集有较高要求, 改进算法模型不断提升预测准确率, 反映影响学习成绩的关键因素

学生学习效果评价研究方面, Li Deming 等^[24]通过研究基于神经网络算法的学生课堂学习效果评价来建立学生学习成果的评估体系, 采用 BP 弹性梯度算法完成评价模型的生成, 达到了很好的训练效果。Na Liu 等^[25]针对线上学习效果评价中数据模糊的问题, 提出一种基于 MOOC 的学习效果评价模型实现线上学习效果评价。实验结果显示模型评分与专家评分基本一致, 精度达到百分之九十。

在学习风格评价, 相关研究主要在经典学习风格模型上改进, 例如学者 Christos Troussas 等^[26]提出了一种结合视觉、听觉、阅读/写作和 vark 学习风格模型的学生模型, 引入混合工具和利用刻板印象进一步增强学生的个性化需求和偏好。Lwande Charles 等^[27]建立了一种从学生学习管理系统中评价学生学习风格和认知特征的方法, 把 Felder- Silverman 模型和认知特质模型结合来识别学生行为, 结果表明这种对学习风格和认知特质的评价能够为行为相似的学习者相互求助提供有利环境和参考。

学生满意度评价研究主要用于验证教学效果, 学者 Baherimoghadam Tahereh 等^[28]通过评价学习风格和一般自我效能感对学生线上学习满意度的影响表明, 发现加工信息的积极维度与满意度水平有显著关系, 自我效能感与满意度水平存在中度显著关系。岳俊芳等^[29]对中国人民大学网络教育进行设计实验, 构建了远程学习者二维满意度评价指标体系和评价方法。

学生综合素质评价研究方面, 柴唤友等^[30]针对现有理论概念问题, 对学生综合评价提出综合素养的新概念并构建综合素养理论模型, 内容包含自主发展、文化基础和

社会参与,为学生综合素养评价的系统开发提供理论依据。张治等^[31]提出使用德智体美劳全要素数据来刻画学生综合素质的成长画像,阐明了基于数字画像的综合素质评价对于促进学生发展的价值。王怀波等^[32]构建智能技术下包含多元数据、理论模型、教育理念和育人导向的综合素养评价的实践框架,一定程度上解决我国学生综合素质评价过程存在的条目化、分立性、忽略人的发展、真实性和客观性不足等问题。

学生学习结果预测研究较为依赖学生的特征数据集,成绩预测能够找出影响学生课程成绩的关系因素。例如 Alsariera Yazan 等^[33]通过考察现有的方法和预测学生成绩的关键特征,证明了 ANN 模型的优点,同时证明学业情况、数量、内部评估和家庭情况是重要的预测特征。学者陈曦等^[34]针对高等教育成绩预测问题,基于课程知识图谱和传统协同过滤算法,证明知识图谱可以作为信息补足,使预测误差显著下降。罗杨洋等^[35]针对混合教学在线数据对学生成绩进行预测难度大这一问题,提出使用增量学习算法搭建成绩预测模型,验证了增量学习随机森林算法有较好的预测准确率和稳定性。Mengash H A 等^[36]使用数据挖掘技术基于高中成绩、录取成绩和一般能力倾向测试成绩等特征预测申请者大学成绩从而支持大学招生决策,结果显示,入学考试是最重要的参考标准且人工神经网络模型有着更高的预测准确率。

在学生评价上还存在学生注意力评价等其他维度的相关研究,如 Qingshan Deng 等^[37]提出基于机器学习方法的学生注意力测量框架,利用支持向量机对学生状态的自动分类,对实际的学生注意力评价具有重要价值。学者 Zafari Mostafa 研究开发一个基于机器学习的系统^[38]来评价影响学生表现的最主要因素,通过识别相关因素实现教育改进。

面向软件工程课程的学生个性化评价研究在评价方法、应用领域和评价维度上基于上述研究基础,针对某大学软件工程这一课程教学,设计和实现了个性化程度更高的评价系统,为混合式教学模式下的学生智能化、个性化评价提供一种解决方案。

1.3 研究内容

本文通过上述学生学习评价的研究综述,拟使用过程性与终结性相结合的联合评价,并使用多种机器学习对学生数据进行挖掘,结合前沿教育理念评价并针对当前软件工程混合式教学评价的特点,设计一种学生个性化评价系统,本文主要研究内容和工作有:

(1) 针对学生学习数据和行为数据生成学生个性化评价报告,引入最近发展区和成果导向教育理论模型,确认评价维度和计算规则,基于软件工程学生数据构造多维特征并使用基于线性模型、树模型和集成模型等多种分类算法的融合模型以及基于数据评分的评价算法生成学生个性化评价的核心内容和关键词,使用基于个性化评价话

术模板的文本生成模式生成完整学生评价报告。

(2) 基于多维特征数据生成学生成绩预测报告,使用实验筛选后的机器学习回归算法并进行集成学习融合达到更好的预测效果。以此对学生课程成绩和最终章节知识掌握程度做出预测,并综合生成学生成绩预测报告。

(3) 完成个性化评价教学系统的实际开发测试,使用最新技术框架编码实现学生教学个性化评价系统,实现个性化评价报告和成绩预测报告的生成,审核,发布和查询,实现学生数据的上传和查询等其他功能,完成系统功能测试以及维护,并应用于《软件工程》教学项目,研究应用后对课程教学质量的影响效果。

1.4 论文组织结构

本文在论文组织结构上分为:

第一章 绪论。

第二章 相关理论和技术介绍。本章主要分析最近发展区理论和成果导向教育理论在本文背景下的应用。在技术介绍方面主要介绍用于实现学生评价分类和学生成绩预测功能的常见机器学习算法模型。

第三章 需求分析和系统概要设计。本章对学生个性化评价系统进行需求分析,主要包含学生个性化评价的整体业务功能需求、个性化评价功能需求、成绩预测功能需求和其他功能需求。之后对学生个性化评价系统进行概要设计,主要对个性化评价系统的体系结构、主要功能模块设计、总体架构和业务流程设计等进行说明。

第四章 学生个性化评价系统功能设计与实现。本章在系统功能设计上确立了以生成个性化评价报告和成绩预测报告为主的核心功能。本章详细说明学生知识掌握度、课程素养和学习风格三个维度评价的分析设计,并完成基于多种机器学习算法的融合模型对学生知识掌握度的评价实验和基于决策树、随机森林、前馈神经网络的集成学习融合模型对学生成绩的预测实验,在上述基础上,通过基于个性化评价设计模板的评价生成算法生成学生评价。

第五章 学生个性化评价系统的实现与应用。本章对总体系统架构、个性化评价报告生成功能、成绩预测报告生成功能、学生数据上传查询等其他功能进行了编码实现,对个性化评价系统的运行进行功能测试和其他性能测试后完成部署,并应用于《软件工程》课程教学,并对比测试实际应用对教学质量提升的效果。

第六章 总结与展望。

第二章 相关理论和技术

2.1 经典教育理论

本文所研究的个性化评价系统是成果导向教育的软件工程课程教学设计中的重要一环,评价始终围绕教学成果目标设计。为尊重学生差异,增加个性化程度而引入最近发展区理论,使过程性评价更好为学生指明发展方向。本章将对个性化评价系统设计所参考的教育理论和算法模型进行介绍。

2.1.1 最近发展区理论

心理学家列维·维果茨基基于教育心理学提出了最近发展区理论 (Zone-of-Proximal-Development, ZPD),该教育心理学理论认为学生的最近发展区是他们独立问题解决的实际发展水平和经过教师指导后的潜在发展水平之间的区域,是实际发展水平与潜在发展水平的差距^[39]。该教育理论目前已经被广泛应用于教育实践中,尤其是个性化教学方面。该理论为教师提供了一种新的教学方法,使他们能够更好地适应学生的能力水平,帮助学生更好地实现个性化学习。此外,最近发展区理论还为教育技术的开发和应用提供了理论基础,可以更好地帮助学生掌握新的知识和技能。

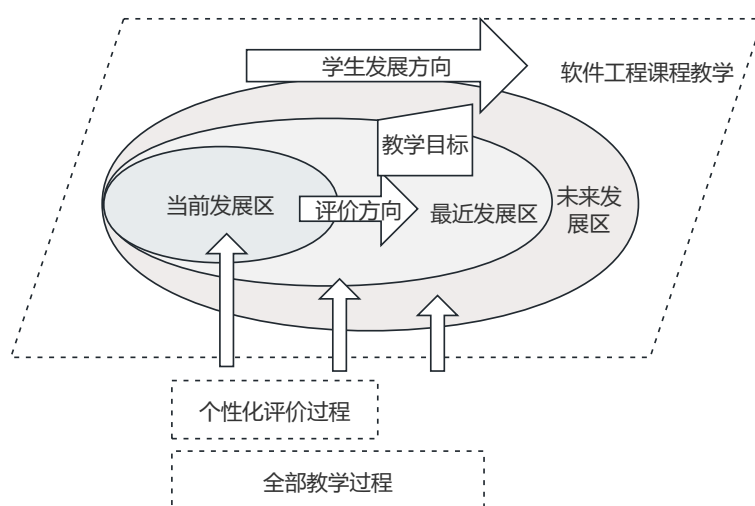


图 2.1 面向最近发展区的个性化评价过程

将最近发展区融入学生个性化评价能够进一步促进学生自主学习,其指出要为学生设置合理的发展目标,目标过高将损害学生的积极性,不易于完成;目标过低则不能激发学生学习能力的提升。如图 2.1 面向最近发展区的个性化评价过程,首先要确认软件工程的课程教学目标,以教学目标为导向确认学生在知识与能力水平的发展方向,

将个性化评价贯穿学生的发展过程，诊断不同学生个体当前发展水平与潜在发展水平。

对学生最近发展区进行诊断，学者刘宁^[40]等提出利用数据技术对学生的进行学习数据进行采集分析，以此诊断学生的最近发展区，其研究基于传统的动态评估方法，提出了数据支持下嵌入学习过程的动态评估，数据技术的介入能够解决传统模式下区域策略和方法论困难以及时间、人力成本高等问题，将评估来源从单一样本转为多模态数据，将评估内容从知识技能转向多元量化表征，在诊断方法上提出依据核心教学内容编制测试题目以及分析学生过程数据，通过逻辑回归和机器学习算法计算学生在知识、应用和创新三个维度的分值，以可视化的系数诊断学生已有的能力表现和潜在发展的能力表现。

2.1.2 成果导向教育

成果导向教育（Outcome based education, OBE）由 Spady 于 1981 年提出，该教育理念在引进国内后逐步影响国内高校的教学模式，其指出课程教学设计以学习目标和结果为导向，以持续改进的方式提升工程教育质量不断提升，教学核心完成从内容到学生的转变。

学者刘强等^[41]研究了软件工程课程教学设计并基于 OBE 对教学方案进行了相关改进，结合软件工程专业培养目标进行课程学习目标的确立，并以此改进教学内容和实践，在课程学习目标评价上结合了学生在期末考核、课堂表现、单元作业、课程项目、团队贡献上的表现，达到了知识与能力双重考核的效果，在评价指标上包含各考核项目的目标、实际达成度、均值和期望达成率。上述内容为学生个性化评价提供了参考，本文中评价面向的软件工程课程采用 OBE 教学模式，主要设立知识和素养两个维度的教学目标，评价数据来源为学生期末考核、课堂表现、单元作业、单元测试、线上学习行为和课程项目，在评价设计上将培养目标与实际达成情况进行比对，由此产生学生个性化评价。

2.2 学生评价算法模型

机器学习或深度学习等算法模型在学生评价中主要有两方面应用，一是基于分类算法模型对学生从某一维度进行分类；二是基于回归算法模型对学生成绩进行预测，本章对两类算法模型进行介绍。

2.2.1 基于分类模型的学生分类

机器学习模型针对学生某一维度进行分类评价有较广泛的应用，例如 E. H. Yossy 等^[42]使用分类方法评价学生素养，基于学生表现数据使用 K 近邻、分类和回归树、

朴素贝叶斯、Adaboost、极限树、伯努利贝叶斯和随机森林七种机器学习方法，结果显示随机森林分类准确率为 89.78%，高于其他模型。树模型在学生分类问题上通常具有较好的效果，部分学者对树模型进行了深入探究实验，例如 A. A. Supianto 等^[43]单独研究了决策树的分类效果，包括随机树、REP 树和 C4.5 决策树，使用计算机学院的数据作为案例，结果显示 C4.5 分类算法的平均准确率更高，达 77.01%。如表 2.1 常见分类模型说明所示，本节对常见机器学习分类模型进行简要介绍，相关模型为学生评价算法模型的设计实现提供了参考。

表 2.1 常见分类模型说明

分类模型	说明
Logistic Regression	通过输入特征与概率函数相结合来预测类别
Decision Tree	多决策分类，每个决策都是基于某个特征的阈值
Support Vector Machine	基于线性分类器来解决非线性分类问题，将数据映射到高维空间
K-Nearest Neighbor	计算与目标最近的 K 个训练样本的标签来预测目标类别
Naive Bayes	基于贝叶斯定理，利用先验概率和条件概率来预测目标类别
Random Forest	多个决策树的集成学习，通过投票或均值确定最终的预测结果
Gradient Boosting	多个弱分类器的集成学习，迭代训练每个分类器逐步提高准确率
Neural Network	多层神经元模型，通过前向传递、反向传播算法训练并进行分类

2.2.2 基于回归模型的学生成绩预测

使用算法模型对学生成绩进行预测的相关研究主要包括通过机器学习和深度学习模型针对不同学生数据进行学习成绩预测。例如娄增辉等^[44]使用 K 近邻、决策树、全连接神经网络和循环神经网络对 16 个特征维度的学生数据进行成绩预测，结果显示两种神经网络模型准确率更高。S. D. A. Bujang 等^[45]使用决策树、支持向量机、朴素贝叶斯、K 近邻、逻辑回归和随机森林六种机器学习模型实现基于 SMOTE 采样技术和两种特征选择方法的多类预测模型改进，结果显示随机森林能够达到更好的改进效果。Qin Jingliang 等^[46]基于 3555 名学生的公共数据集，使用支持向量机、多层感知器和增强的长短期记忆网络（Long Short-Term Memory, LSTM）模型进行学生成绩预测，实验结果表明改进的 LSTM 模型准确性更高。

在个性化评价教学系统中，基于学生数据对学生最终学业成绩做出准确预测是一个典型的回归问题。基于相关研究，对常见的机器学习和深度学习回归模型进行总结说明。如表 2.2 常见回归模型对比所示，本节针对常见回归模型的优缺点和适用范围进行分析，为学生成绩预测在模型选择和设计上奠定基础。

表 2.2 常见回归模型对比

算法名称	优点	缺点	适用范围	应用领域
Linear Regression	原理简单 计算速度快	异常值敏感 拟合能力受限	特征数量较少 具有线性关系	基于线性的结果预测
Ridge Regression	防止过拟合 抗干扰能力强	引入偏差 计算量大	特征数量较多 存在多重共线性	预测、评估 自然语言处理等
Lasso Regression	泛化能力强 减少过拟合	结果稀疏 训练实践长	高维特征数据集 自变量有相关性	预测、评估 图像分割等
Support Vector Regression	泛化性能好	核函数和参数 选择依赖经验	非线性问题	统计, 预测 语音识别等
Decision Tree	可解释性强 多型数据处理	容易过拟合 局部优化	非线性数据 大型数据集	预测、评级等
Random_forest	鲁棒性高	训练时间长	高维度数据 大型数据集	预测、评估 自然语言处理等
Multilayer Perceptron	灵活性高 特征自动提取	训练时间长 容易过拟合	非线性问题	诊断、预测 图像处理、语音识别
CNN	扩展性好	难以解释	空间时间相关性	图像处理、音频处理 自然语言处理
RNN	自适应序列长 实时预测	梯度消失 长序列限制	序列化数据 非线性关系	自然语言处理 时间序列预测

(1) 线性模型及改进

线性回归 (Linear Regression) 用于建立两种或多种变量之间的线性关系。其目标是通过最小化误差平方和来找到最优的系数, 可以用于预测连续的数值型数据。岭回归模型 (Ridge Regression) 是通过对系数进行惩罚来解决线性回归模型中的过拟合问题。岭回归模型可以通过引入 L2 正则化项来实现, 使得模型在训练过程中倾向于选择较小的系数。套索回归模型 (Lasso Regression) 与岭回归模型类似, 也是为了解决线性回归模型中的过拟合问题, Lasso 回归模型引入了 L1 正则化项, 与岭回归模型不同的是, Lasso 回归模型可以将某些系数压缩为 0, 从而实现特征选择的功能。

(3) 非线性模型

支持向量回归模型 (Support Vector Regression, SVR) 是一种非线性回归模型, 通过将自变量映射到高维空间中, 可以通过核函数来处理非线性关系。支持向量回归模型通过最小化间隔误差来找到最优的超平面, 从而进行预测。树模型主要包括决策树

和基于决策树的集成学习模型。决策树回归 (Decision Tree) 是一种基于树形结构进行预测的回归算法, 它通过将数据集递归地分成不同的子集, 每个子集都对应于一个节点, 从而构建出一棵树。树的每个叶子节点表示一个预测值, 输入不同自变量, 沿着树的分支进行遍历, 最后到达一个叶子节点来得到预测结果。随机森林回归 (Random Forest) 是一种基于决策树的集成学习算法, 通过同时构建多个决策树来提高模型的预测精度和稳定性。在随机森林回归中, 每个决策树都是基于随机选择的特征子集进行构建的, 这样可以减少特征之间的相关性, 并增加模型的泛化能力。

(4) 网络模型

多层感知机 (Multilayer Perceptron, MLP) 回归是一种基于多层感知器的回归模型, 它属于神经网络模型的一种。MLP 是一种前向反馈神经网络, 由多个神经元组成的输入层、隐藏层和输出层依次排列, 可以进行非线性的回归任务。CNN 回归 (Convolutional Neural Network Regression) 是一种基于卷积神经网络的回归模型, 在网络结构上, 其通常包含多个池化层、卷积层和全连接层, 前两层可以提取出高维数据的特征, 全连接层则用于回归输出。RNN 回归是一种基于循环神经网络 (Recurrent Neural Network, RNN) 的回归模型, 它可以用于序列数据的建模和预测。RNN 是一种具有记忆能力的神经网络, 能够对序列数据进行建模, 从而对序列数据的未来进行预测。

2.3 本章小结

本章对个性化评价系统研究所引入的经典教育理论、教育心理学理论和机器学习算法模型进行了介绍。首先介绍了最近发展区理论及其在学生评价中的应用; 其次介绍了成果导向教育理论在软件工程课程教学方案设计上的应用, 为基于课程教学目标的个性化学生评价设计提供了参考。在机器学习算法模型上, 介绍了常见的分类模型和回归模型, 为后续系统融合模型的研究设计奠定了基础。

第三章 需求分析和系统概要设计

本章对本文阐述的学生个性化评价系统进行需求分析，主要包含学生个性化评价的实际业务需求、系统功能需求、非功能需求等。第二部分是对学生个性化评价系统的总体设计，主要包括体系架构设计，功能模块设计和业务流程设计等。

3.1 系统需求分析

本文所设计与实现的个性化学生学习评价系统，是基于《软件工程》课程数字化改革建设需求，主要系统功能需求包含生成学生个性化评价报告和生成学生成绩预测报告。

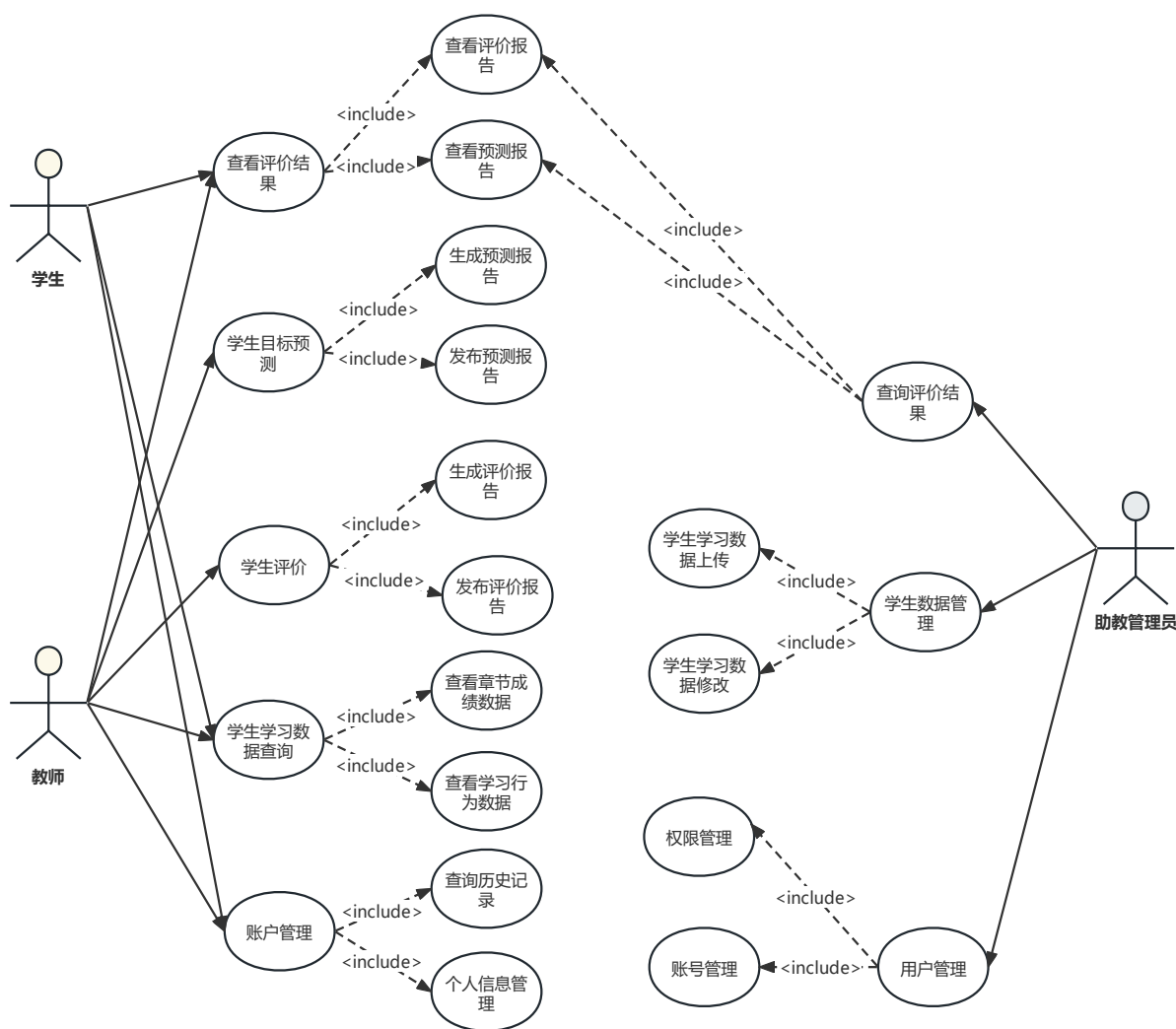


图 3.1 个性化评价系统用例图

在学生个性化评价报告生成上,对近三年软件工程课程教学网,课堂派,Moodle 学习平台,课程人工记录数据存档等学生数据进行挖掘,从学生知识结构、课程素养和学习风格多个角度建立学生评价模型,确立评价指标和标准,生成学生的个性化评价报告;同时,利用学生日常成绩和行为表现数据为特征进行期末课程成绩预测和学生章节掌握度预测,生成学生成绩预测评价报告。

基于上述需求分析可得该系统用例图,如图 3.1 个性化评价系统用例图所示,该系统主要有三类用户,软件工程授课教师、选课学生、课程助教每类用户具有单独的权限和操作。教师用户主要负责对学生进行个性化评价(包含生成和发布评价报告两类操作),对学生进行学习成绩预测(包含生成和发布预测报告),查询学生学习数据(包括学生章节学习成绩和学生学习行为数据),查看学生评价结果(包括查看评价报告和查看预测目标)以及个人账户管理(管理个人信息和查询历史记录)。学生用户主要查看评价结果(包括查看评价报告和查询预测目标),查看个人学习数据(包括章节成绩数据和学习行为数据)以及进行个人账户管理。助教管理员用户主要包括学生数据管理(学生数据上传和修改),查询评价结果以及课程学生管理。

3.1.1 学生个性化评价功能需求分析

个性化评价模块用于满足对学生的知识掌握程度、课程素养、学习风格等方面做出个性化评价,是该系统的核心功能模块。该模块主要用户需求包括生成个性化评价报告,编辑评价报告、发布评价报告和查看个性化评价报告。该功能由教师选中需要评价的学生信息,点击生成个性化评价报告,教师可对评价报告中的文本和关键词做出审核或修改编辑,随后可点击发布评价,发布后学生可登录查看评价报告,对不同用户群体,成绩预测模块设置不同的操作权限。如表 3.1 个性化评价报告功能需求所示。

表 3.1 个性化评价报告功能需求

组成部分	评价类型	教师权限	学生权限
知识掌握度评价	过程性评价	章节教学完成后生成、编辑、发布	查看
课程素养评价	过程性评价	生成、编辑、发布	查看
学习风格评价	终结性评价	生成、编辑、发布	查看

上述功能在关键理论、技术和实现效果上有以下需求,如图 3.2 学生个性化评价模块功能需求所示。学生知识掌握度评价为过程性评价,依据章节教学进度贯穿整个教学过程,在评价目标确立上依托成果导向教育理论,在评价内容上基于最近发展区理论对学生的当前知识掌握水平和潜在知识掌握水平进行评价,为教师对学生的指导方向以及学生下一步发展提供参考,在学生当前知识水平分类上使用特征数据集对多种机器学习模型进行训练并设计融合模型实现分类结果的准确输出。学生课程素养和

学习风格评价分别基于《中国学生发展核心素养》和所罗门学习风格模型改进，在学生素养评估和风格判别上基于学生数据进行评分计算。

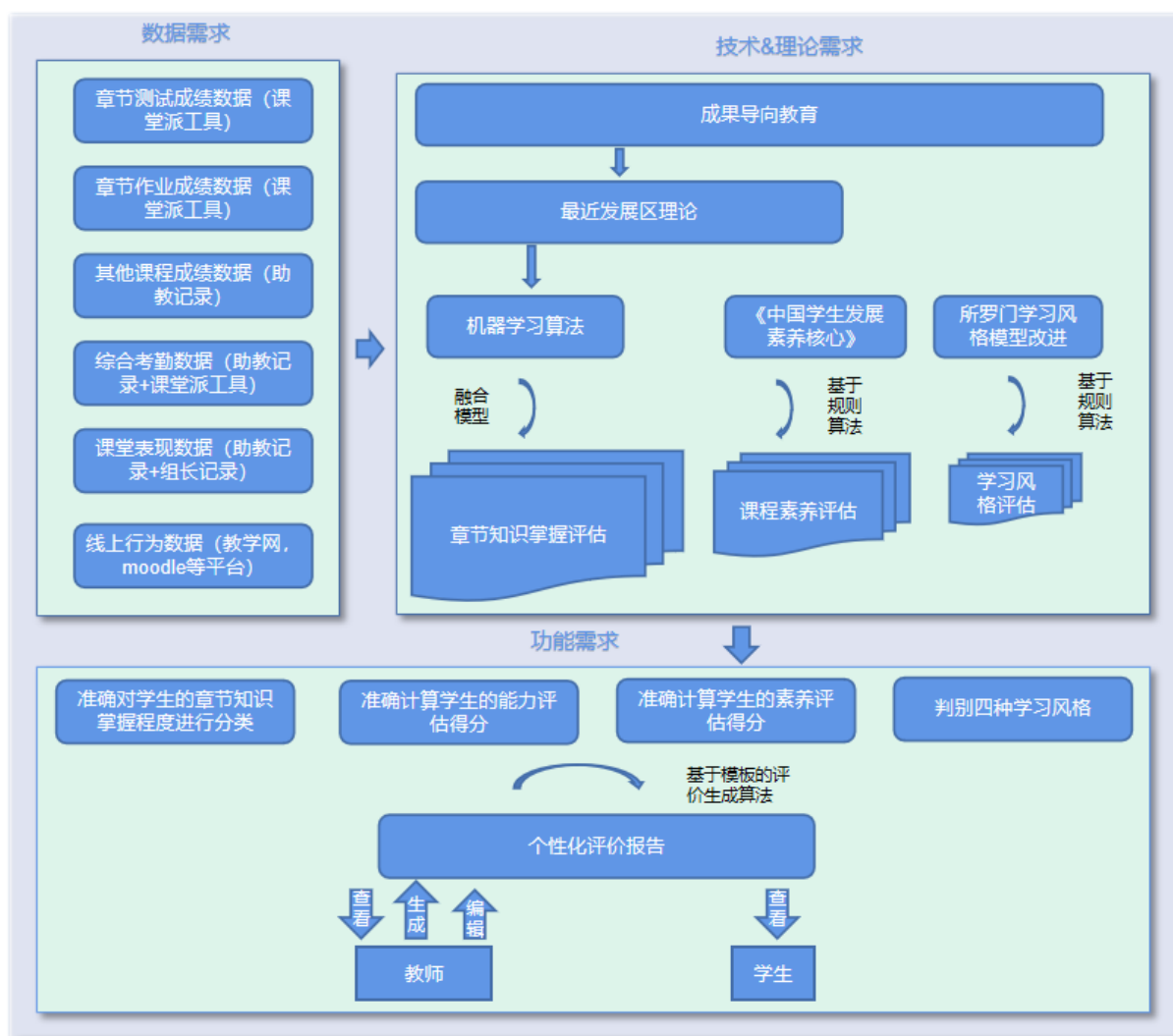


图 3.2 学生个性化评价模块功能需求

3.1.2 学生成绩预测功能需求分析

学生成绩预测模块用于满足对学生最终课程成绩及《软件工程》各章节最终知识掌握度进行评估预测，如图 3.3 学习成绩预测模块功能需求所示，该模块的功能主要包括生成成绩预测分析报告，发布成绩预测报告和查看成绩预测报告。该功能由教师点击生成学生成绩预测评价报告，该模块调取学生数据通过机器学习融合模型处理后得学生最终学业成绩预测和各章节最终知识掌握度分类，在教师完成审核后发送给学生，学生登录后即可查看，评价显示内容包括最终预测课程成绩、当前预测排名、各章节掌握度预测。对于教师、学生和助教三类用户群体，与个性化评价模块类似，对成绩预测模块设置不同的操作权限。

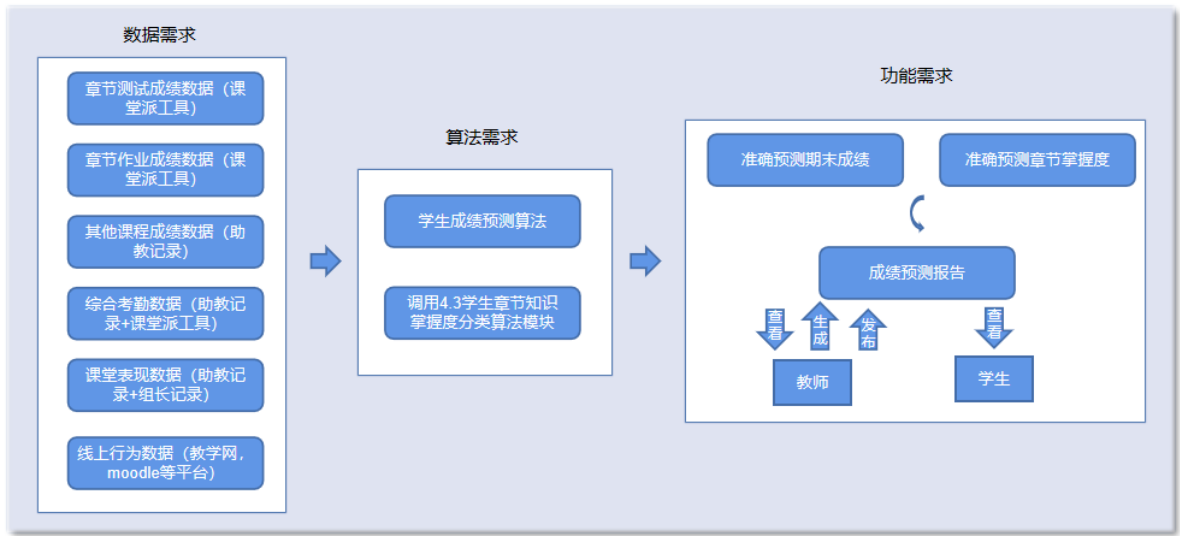


图 3.3 学习成绩预测模块功能需求

3.1.3 其他功能需求分析

其他功能需求主要包含数据上传、查询、展示、个人信息管理和用户管理等。数据上传查询功能主要包括学生学习数据上传、学生学习数据分析、学生学习数据查看等功能。该课程教学过程中产生的其他学生其他信息，如所获课程相关奖励、课堂互动频次等其他信息由课程助教通过上传实现。学生数据由系统分类整理后以可视化形式在教师和学生界面展示，如表 3.2 其他功能需求，上述功能在设计上主要有以下需求。

表 3.2 其他功能需求

需求	权限归属	说明
数据上传	管理员	为管理员提供数据上传界面
数据查询	教师	为教师提供关键词模糊查询
数据分析展示	教师、学生	为教师和学生提供学生数据的可视化展示
个人中心	教师学生	为教师和学生提供个人信息管理
用户管理	管理员	为管理员提供用户管理功能

3.1.4 非功能性需求分析

个性化评价系统非功能性需求是除功能需求以外的特性，是为了满足《软件工程》课程实际课程使用需要而应该具有的需求，主要包括系统的性能需求等。

（1）系统兼容性和吞吐量

为满足实际课程需求，系统在兼容性上应当适用于当前常见的浏览器、计算机系统和硬件。系统吞吐量用指系统承受教师和选课学生访问的程度，具体为单位时间成

功传输的数据重量，主要与用户并发量相关，学生评价系统需要满足教师和学生需求的快速响应，保证界面操作流畅。根据《软件工程》课程业务量计算，平均并发用户数和并发用户峰值数量级均在十的三次方以内，在系统上线后使用压力测试工具对系统进行吞吐测试，保证在软件工程课程并发量下系统的流畅性和稳定性。

（2）系统响应时间

该系统响应时间在于用户发布指令到系统把结果以可视化的形式展现出来为止，反映着用户对系统的直观印象。系统响应时间如果较慢，会影响师生的使用积极性，该系统主要在于提升页面的响应速度，主要手段在使用外部的 JavaScript 和 CSS，二次访问时直接使用浏览器缓存等方式，使系统页面的最长响应时间小于 5 秒。

3.2 系统概要设计

3.2.1 体系结构设计

在系统结构设计上如图 3.4 B/S (Browser/Server) 结构，使用基于浏览器访问和服务器的 B/S 结构。本文所设计的学生个性化评价系统在考虑用户本地操作权限，系统安全性，访问量和成本后考虑使用基于浏览器访问和服务器的 B/S 结构进行设计。

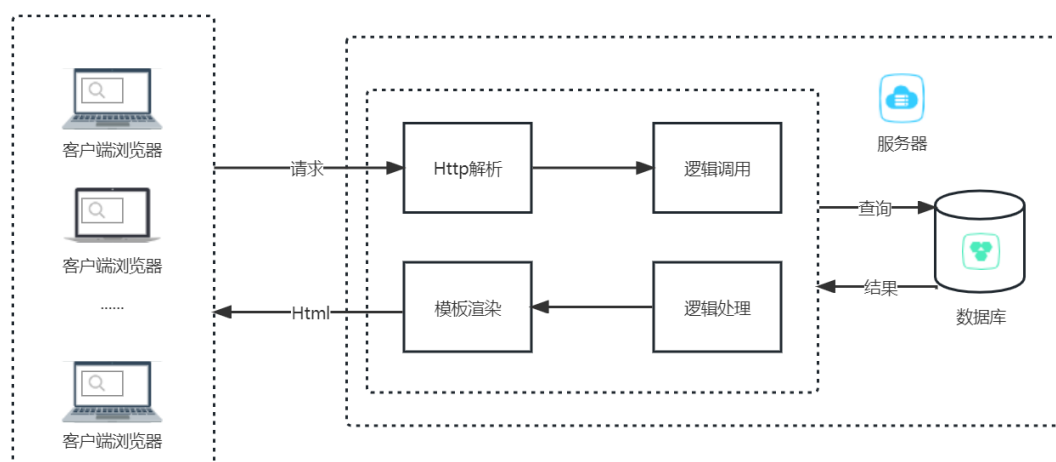


图 3.4 B/S (Browser/Server) 结构

在该系统中 B/S 结构有着较大优点，系统使用 HTTP 协议进行数据交互，软件工程课程的用户只需要使用浏览器访问服务端即可，后端的控制权限和管理权限更高，此外，客户端不再需要后续的更新维护和升级。个性化评价系统的大量数据不需要在客户端处理和计算，只反馈数据信息和文本信息。

3.2.2 系统功能模块设计

个性化评价模块和学习成绩预测模块为核心功能模块，如图 3.5 核心功能模块设计所示。其他功能模块还包括数据上传与查询模块、用户管理模块等。首先是个性化评价模块，该模块是系统核心模块，用于生成和管理学生个性化评价报告，主要包括评价报告的生成；评价报告的审核、编辑、发布、评价报告的查询等功能。然后是学习成绩预测模块，该模块用于生成学生学习成绩预测报告，主要包括学生成绩预测结果生成、审核、发布和查询等。其次是数据上传与查询模块，最后是管理员模块，这个模块用于管理系统的访问权限和管理用户信息。

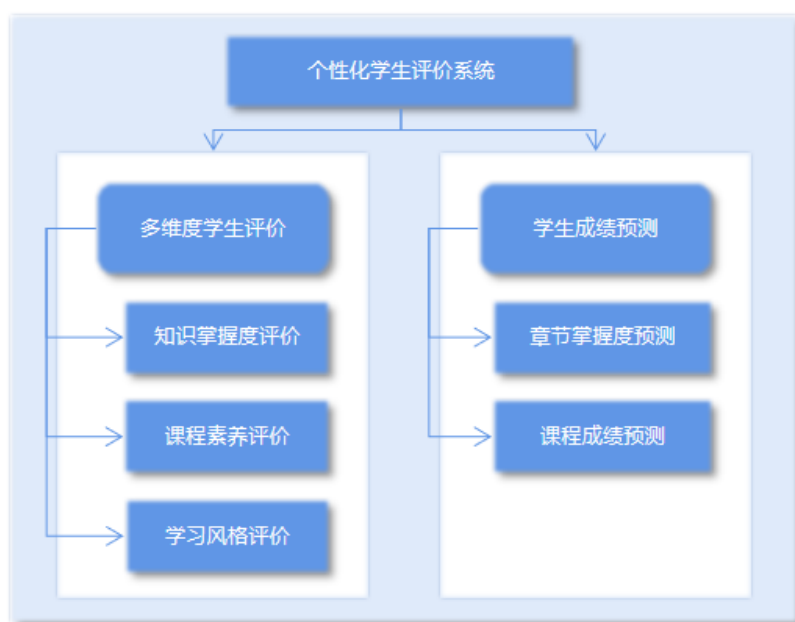


图 3.5 核心功能模块设计

3.2.3 系统总体架构

如图 3.6 系统总体架构所示，学生个性化评价系统从底至上分别是数据访问层（Data-Access-Layer）、业务逻辑层（Business-Logic-Layer）和表现层（Presentation-Layer），参照传统结构模式进行三层架构设计。本系统的表现层采用 Web 界面实现用户交互；学生，教师和助教可以使用浏览器访问系统界面。在本层教师可以通过调用个性化评价模块为指定学生生成个性化评价报告，同时可以对报告做出修改、审核和发布操作；还可以调用成绩预测模块生成学生学习成绩预测报告；所有用户在本层则可以使用数据查询模块查询学习数据。业务逻辑层调取本层传递的用户请求后进行处理，并将处理结果返回给用户。本系统的业务逻辑层包括个性化评价模块、学生成绩预测模块、数据上传与查询模块和用户管理模块四大模块，核心功能主要由前两个模块实现，包括学生数据的处理、成绩预测结果的生成和评价报告的生成等。数据获取依赖数据访问层，业务逻辑层需要与其进行交互，进行学生数据的获取和修改。本系统的底层为数据访问层，主要包括数据存储、数据管理和数据访问的实现，为业务层提供数据的读写、

存储、查询等操作。本系统中的数据层主要对学生学习数据，用户数据进行存储，在数据库设计上本系统使用的是 MySQL 进行存储。

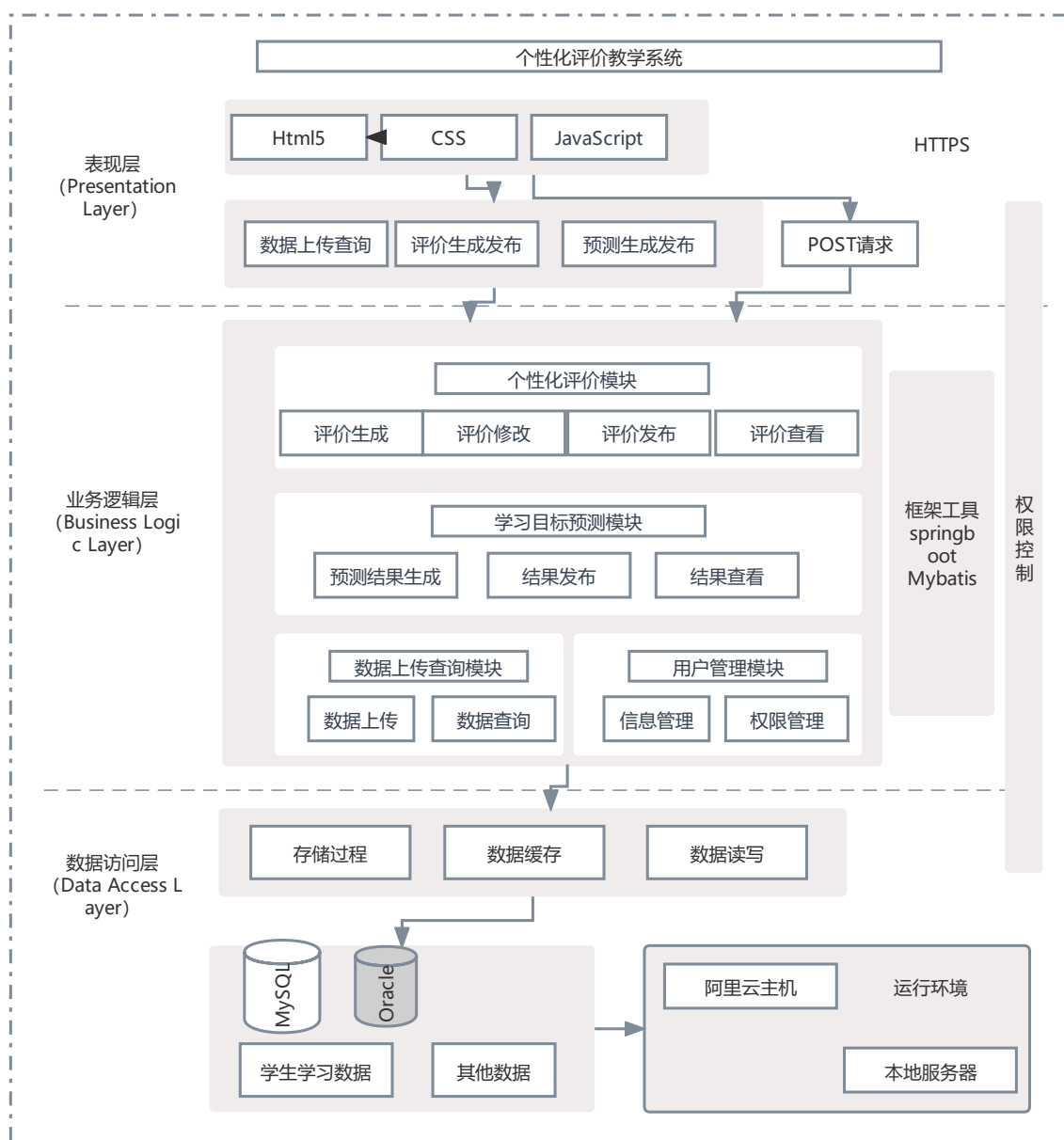


图 3.6 系统总体架构

3.2.4 业务流程设计

系统整体流程可参考图 3.7 系统业务流程示意图。本节在业务流程设计介绍上主要按照参与对象、动作、次序、输入、输出的模式进行说明，教师在登录系统后可查看我的课程，选择当前软件工程课程后查询当前选课学生的详细学习数据，选择指定学生生成评价报告或预测报告，系统调用数据库和算法模块并进行处理后将结果返回给

教师；教师可对评价内容进行编辑修改，教师发布后将结果写入数据库。学生登录查看自身评价报告时，时调用数据库返回评价结果。

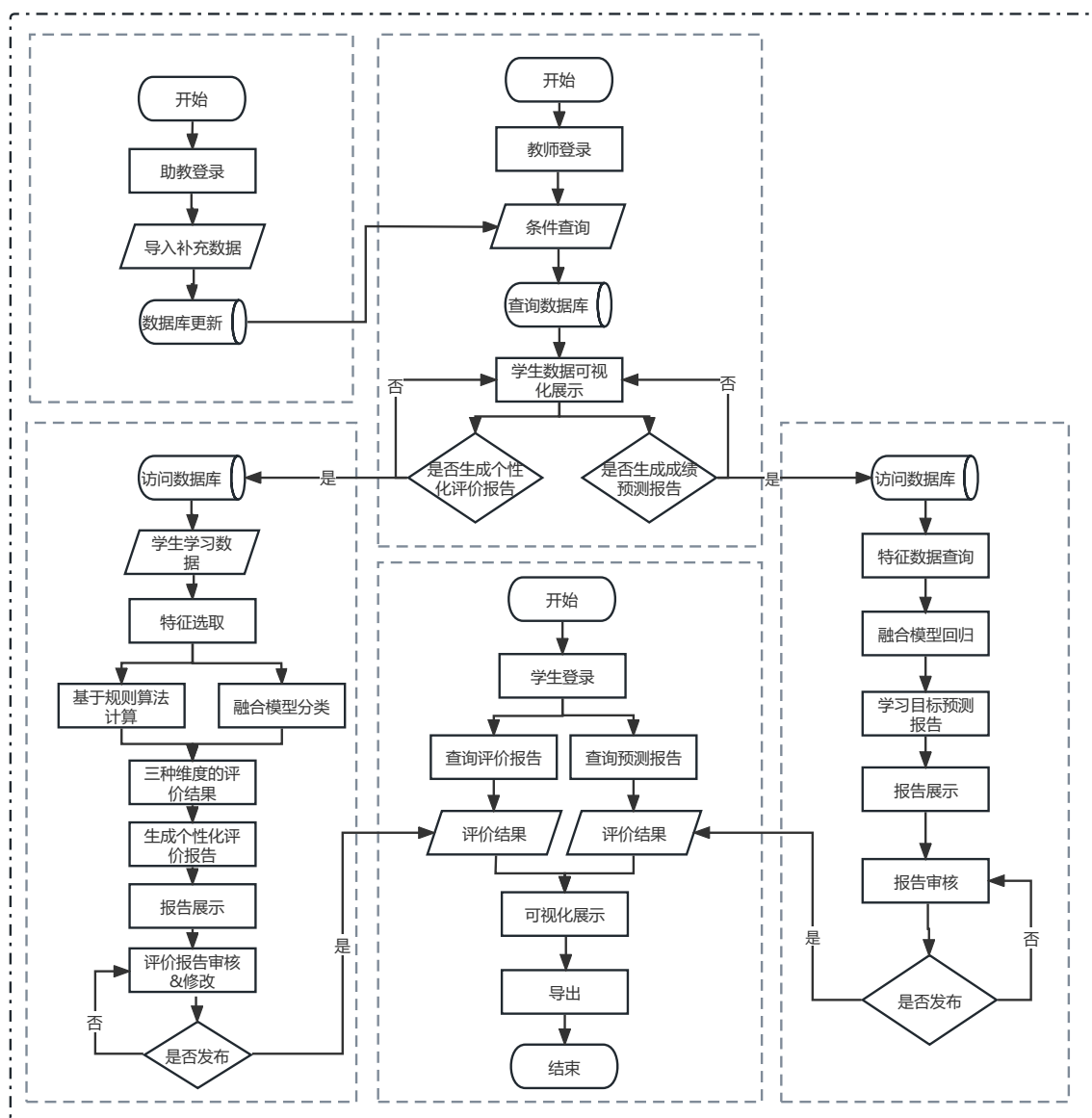


图 3.7 系统业务流程示意图

3.3 本章小结

本章首先对个性化评价系统的总体需求和两大核心功能模块需求进行了描述，对个性化评价模块和成绩预测模块在用户功能、关键理论技术和实现效果等方面进行了说明，对系统非功能需求进行了阐述。其次对系统体系结构、主要模块和三层架构设计进行介绍，为系统编码实现提供支撑。最后是对系统业务流程的介绍，明确了系统的开发和使用需求。

第四章 学生个性化评价系统算法设计

依据对本文所研究的学生个性化评价系统的需求分析以及概要设计，在系统功能设计上确立了以生成个性化评价报告和学生成绩预测报告为主的核心功能。本章将阐述如何整合学生数据确立评价维度和构建多维特征，说明各维度学生评价算法的设计与实现以及如何依据相关理论和机器学习模型设计实现学生评价，生成针对不同个体学生的个性化评价报告和成绩预测报告。

4.1 总体方案设计

(1) 评价方案设计

如图 4.1 系统总体研究方案设计所示，学生数据获取来自《软件工程》课程教学所使用的多个线上平台以及助教人工补充数据，主要包含课堂教学数据、多个平台的线上行为数据和成绩数据。系统在经过整合导入学生数据库用于数据查询和生成评价报告，使用机器学习融合模型对学生章节知识掌握度进行评价；使用整体数据生成基于评价标准和评分规则的课程素养和学习风格两个维度的评价，生成包含知识、素养和风格评价的个性化评价报告，同时选取部分数据作为特征经过集成学习融合模型生成成绩预测报告。

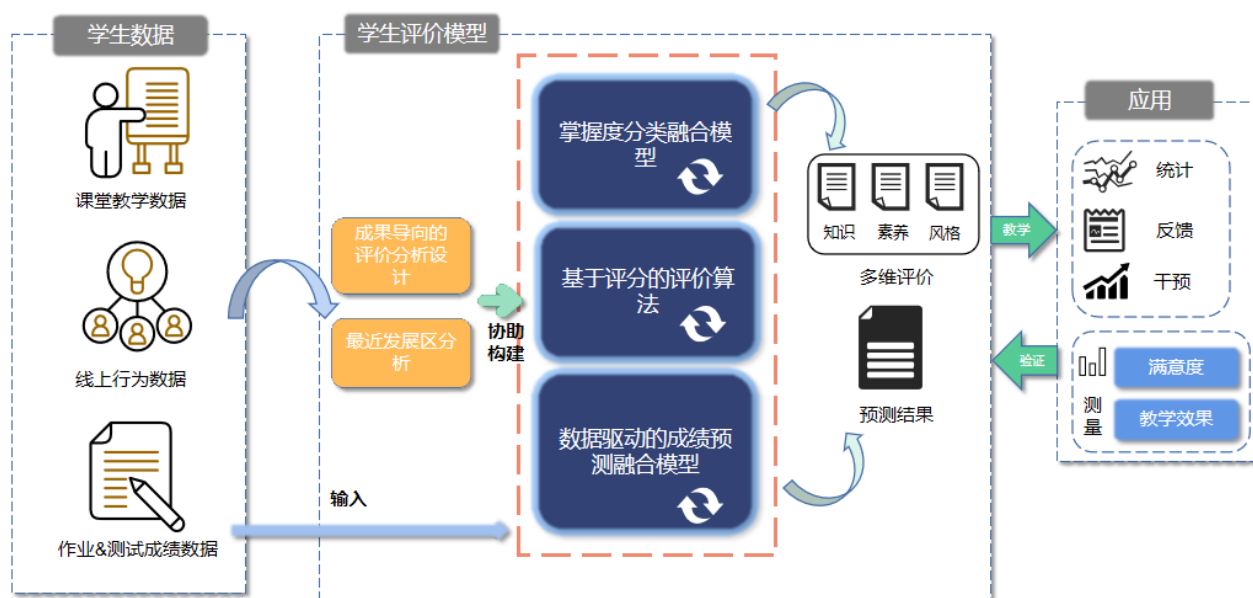


图 4.1 系统总体研究方案设计

(2) 评价内容设计

在学生个性化评价组成上包含多维度学生评价和学生成绩预测两块内容，多维度

学生评价又由学生知识结构掌握度、课程素养和学习风格评价构成；学生知识掌握的评价为过程性评价，通常在章节教学完成后对学生章节课程知识点的掌握程度进行评价，可视化展示使用知识结构雷达图；课程素养评价为过程性评价，对教学过程中学生学习行为所蕴含的课程素养进行评价；学习风格评价为终结性评价，根据学生整个学期的学习行为进行判别，帮助学生对自身学习风格的认知；学生成绩预测评价为过程性评价，在课程教学的重要时间点对学生成绩做出预测，便于教师及时进行学习干预。

4.2 实验数据

该系统实验数据来自某大学《软件工程》课程实际教学数据，数据获取自 2019 至 2020 第一学期、2020 至 2021 第一学期、2021 至 2022 第一学期以及 2022 至 2023 第一学期《软件工程》课程，线上教学数据从课程所使用的某大学教学网、课堂派教学工具、Moodle 在线学习平台、头歌平台等多个线上教学平台获取，课堂教学数据来自学生作业信息存档，课程助教人工记录信息等。整理上述数据作为系统实验的数据集。

4.2.1 数据预处理

将采集后的数据进行汇总和分组，分组后的详细数据范围包括学生基本信息（包括姓名、学号、邮箱等）；学生考勤数据 P_i （课程出勤率、章节完成率、作业完成率、测试完成率、话题参与率等）；学生成绩信息 S_i （各章节课后测试成绩、各章节课程作业成绩、章节互评成绩、一句话软件工程思政报告得分、大作业成绩，期末考试成绩、最终课程得分等）；课堂互动数据 N_i （弹幕数、受奖励次数、互动回答次数）；学生线上访问记录 T_i （在线时长、平台登录次数、资料下载次数、推进内容点击次数等）；学生角色信息 R_i （组内担任角色、项目担任角色）；moodle 系统学生线上行为数据 B_i （观看历史、开始、暂停、重播、快进、快退、观看时长、隐藏页面、恢复页面、全屏、ESC 退出、静音、关闭静音、调节音量、视频结束、用户在线测试等详细操作记录）。

如表 4.1 学生特征数据编号，对于上述学生数据进行初步清洗，去除与学生评价无关的属性，同时为方便后续的特征选取和模型训练实验进行，对数据按照分组编号。进一步对这些基础数据进行预处理，进而解决基础实验数据的重复、不准确、不完整等问题。首先是数据清洗，检查学生数据是否存在缺失值、在《软件工程》课程数据中存在大量显示为空值或文本未提交、未批改；将相关的数据用及格线分数常量填充、删除部分缺失值比例较大的样本数据，保留数据较完整的个体样本；对于异常值和重复值，进行填充处理或删除。将布尔型的数据属性转化为 0,1 值，之后将来自不同平台的数据整合到一个数据集中，确保数据格式和属性一致，在本系统中主要是将人工统计数据

和平台数据进行整合。

表 4.1 学生特征数据编号

编号	数据	编号	数据
p1	课堂出勤率	s22	面向对象分析文档得分
p2	作业完成率	s23	结构化分析互评成绩
p3	课前测试完成率	s24	面向对象分析互评成绩
p5	课后测试完成率	s25	面向对象设计互评成绩
p6	话题参与率	s26	系统实现互评成绩
p7	头歌平台闯关完成率	s27	项目管理互评成绩
n1	发布弹幕数	s28	软件工程新技术互评成绩
n2	互动回答次数	r1	是否担任组长
n3	受奖励次数	r2	项目担任角色
s1	软件工程概述章节测试成绩	t1	moodle 在线时长
s2	可行性研究章节测试成绩	t2	教学网在线时长
s3	结构化分析章节测试成绩	t3	powerhighlight 登录次数
s4	面向对象分析章节测试成绩	t4	资料下载次数
s5	面向对象设计章节测试成绩	t5	推荐视频点击次数
s6	系统实现与维护章节测试成绩	b1	在线检测
s7	项目管理章节测试成绩	b2	观看历史
s8	软件工程新技术章节测试成绩	b3	开始操作
s9	软件工程概述报告得分	b4	暂停操作
s10	可行性研究报告得分	b5	重新开始操作
s11	需求规格说明书得分	b6	快进操作
s12	详细设计说明书得分	b7	快退操作
s13	概要设计说明书得分	b8	隐藏页面
s14	编码练习得分	b9	恢复页面
s15	进度计划编制得分	b10	全屏
s16	软工新技术报告得分	b11	关闭全屏
s17	思政报告得分	b12	静音
s18	课程大项目得分	b13	关闭静音
s19	期末考试成绩	b14	调节音量
s20	软件工程概述互评成绩	b15	重新开始
s21	可行性研究互评成绩	b16	视频结束事件

结合数据完整性以及该项目所需要的数据内容,经过筛选和处理后保留了 519 位学生详细数据,共 992046 个学生数据片段。部分学生行为数据如图 4.2 清洗后的学生数据片段所示。

面向对象分析文档	详细设计说明书	概要设计说明书	需求规格说明书	可行性研究报告	软件过程模型讲解	软件维护学后测	软件维护学前测	编码练习(正式版)	系统实现(学后)	系统实现(学前)	编码练习	系统实现(编码+测试基础)	系统实现(编码+测试)	面向对象设计	面向对象分析	互评——详细设计说明书	互评——概要设计	互评——需求规格	互评——可行性研
90	90	85	87	90	90	4	50.6	3	60	50	56.5	57	60	76	90	39	60	60	60
90	90	80	90	85	60	7	50.6	4	7	50	50.0	50	56	50	75	50	45	90	48
90	90	80	90	85	60	10	100.8	4	10	70	77.9	84	98	73	90	50	85	60	52
92	91	75	89	92	96	7	55.4	2	60	50	72.6	75	66	53	80	70	65	100	94
92	90	95	85	90	95	3	40	4	8	20	84.8	91	67	52	100	50	60	70	58
91	90	100	82	90	90	10	0	4	10	20	88.5	81	66	92	100	20	55	85	60
90	90	85	87	90	90	3	40.7	1	60	60	60.9	75	77	97	55	20	20	100	60
95	94	90	90	90	96	10	40.6	3	10	90	90.6	54	68	91	90	60	50	75	98
90	90	90	90	90	90	8	5	5	10	80	60.3	67	94	63	100	90	65	75	60
90	90	80	90	85	60	6	400.2	4	8	60	61.1	95	81	94	85	80	80	90	60
92	94	93	90	90	92	3	40	5	60	60	52.1	81	74	72	60	60	65	55	92
96	95	90	90	90	95	10	90.6	5	8	50	87.0	55	76	63	100	60	55	60	98
87	90	78	70	70	60	7	70.6	0	0	0	60.0	95	81	50	95	50	75	90	60
90	90	80	90	85	60	3	40	4	5	20	74.0	60	71	93	65	0	80	95	88
87	90	78	70	70	60	5	0	5	10	0	77.8	80	70	70	80	20	75	60	60
90	90	80	90	85	60	7	25.4	4	8	80	83.5	96	62	75	95	0	80	60	56
90	90	90	92	90	95	6	405.4	2	7	20	99.1	80	83	59	75	80	80	70	58

图 4.2 清洗后的学生数据片段

4.2.2 多维特征构造

传统的学生分类或预测模型大多数从单一的学生成绩、学生课程任务的完成度或单一的学生适应性测试结果进行特征选取,单一维度的数据造成的结果是受测试评分的量规,授课内容的难度等因素影响较大,难以形成全面准确的预测或评价。学生章节的实际掌握程度和期末课程成绩通常受多种因素的影响,因此本节拓宽数据特征的维度并进行细化。如图 4.3 多维度特征构造所示,结合 4.2.1 中学生原始数据,将原始数据集划分为课堂参与度、章节报告、章节测试、线上参与度和线上学习行为五个维度,深入挖掘学生学习数据。

课堂参与度维度主要由学生课堂出勤率、互动答题次数、受奖励次数等数据构成。出勤率数据来自课堂派平台考勤、其余数据来自每一小组分配的课堂助教进行记录。这些数据能够很大程度反映学生的学习态度,对于最终学业成绩有重要影响,是评价模型设计的重要特征。

章节测试维度主要由单个章节的课前测试成绩和课后测试成绩构成,能够显著反映学生课程知识学习前后的差别,也是章节掌握度评价和其他维度评价的主要特征。

章节报告维度主要由每个章节的课程展示作业成绩和其他报告成绩等数据构成,是对章节知识应用效果的检验,能够反映学生知识掌握程度、实践创新素养,并影响最终成绩。

线上参与度维度主要包括头歌教学平台,教学网,课堂派等线上平台记录的学生

资料下载率、登录次数、访问时长、推荐内容访问次数等数据，能够一定程度反映学生的自主学习能力和课程认同感，是学生知识和信息获取的重要反映，可作为学生课程素养和学习风格评价的主要特征。

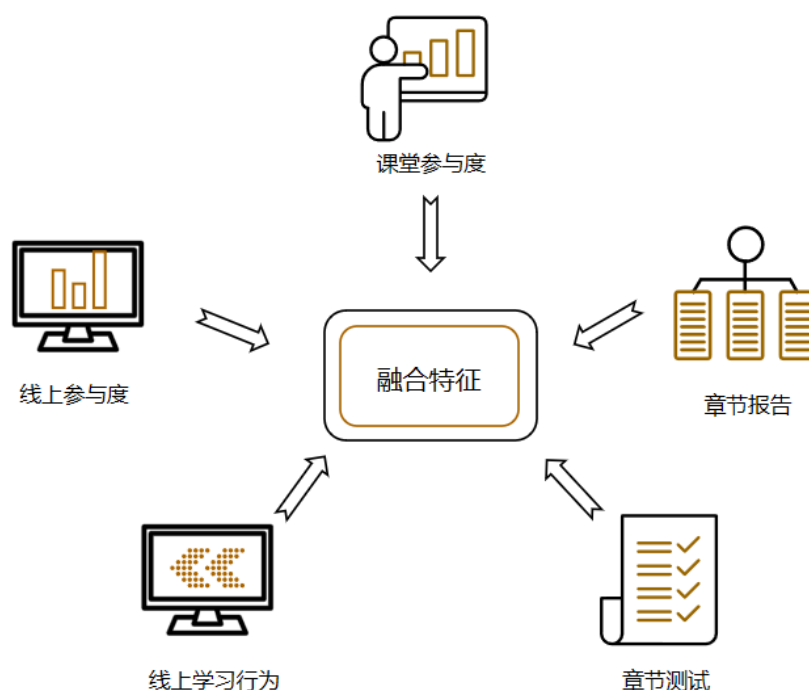


图 4.3 多维度特征构造

线上行为数据维度主要包括学生在线上学习平台 moodle 系统的行为记录，具体行为可参考 4.21 数据预处理，在平台导出的原始数据中，学生行为由快进、快退、暂停等单个元行为记录，每个学生的行为数据片段可能包含上百个记录序列。对于这些字符型文本串，无法直接作为特征反映学生的学习状况，也无法作为模型训练的输入，需要一定的设计方法将其转化为可表示的数值或向量且表示方法要有可解释的意义。学生线上行为数据记录可视为文本数据，因此本节使用将文本进行向量转化的词袋模型对行为数据进行特征提取。

使用词袋模型进行向量转化，在特征的意义主要考虑学生某一行为出现频率的影响，由于序列较多，暂不考虑前后顺序关系的影响。由于文本字符较长，首先对文本型字符串进行处理，将原始数据每个样本进行编号，对 17 种关键学生行为赋予 0 到 17 的数字编号，比如使用 1 表示快进，使用 2 表示快退，使用 3 表示暂停，使用 4 表示重播等，则处理后的学生行为向量可表示为 $[x_1, x_2, x_3, x_4, \dots]$ ， x_i 表示单个学生某种行为的出现频次。对于处理后的学生行为数据，还需要对单个行为在整体的权重进行评分计算。采用 NLP 中的常用统计方法 TF-IDF 进行处理，其计算方法如式（4.1）。

$$IDF(\omega) = \log \frac{\text{学生数}}{\text{出现行为}\omega\text{的学生数} + 1} = \log \frac{n}{t + 1} \quad \text{式(4.1)}$$

该算法可以用来评估某一行为在整体学生行为数据集中的重要性，可以用来提取最为重要的关键行为特征。 TF 表示某一学习行为 h 在当前学生 S_i 的行为字符串中的出现频次， IDF 叫做逆文本频率，其中 n 为数据集中学生总数， t 为出现该行为的学生个数。因此，逆文本频率的计算方式如式（4.2）。

$$TF - IDF = TF * IDF \quad \text{式(4.2)}$$

通过上述算法处理后包含权重的学生行为向量可以作为输入用于模型训练，可调用库方法对数据进行处理，如算法 1：输出权重向量的 TF-IDF 算法所示。

算法 1：输出权重向量的 TF-IDF 算法

输入：线上学习行为数据 $behavior = ['xxx']$

输出：学生行为向量及权重 $X.toarray()$

1. `TfidfVectorizer` # sklearn 库方法
 2. `behavior = ['xxx']` # 学习行为数据集初始化
 3. `vectorizer = CountVectorizer()`
 4. `X = vectorizer.fit_transform(behavior)`
 5. `print(X.toarray())` # 输出行为特征向量
-

4.3 学生知识掌握度评价算法设计

本节对基于 OBE 教学设计的软件工程课教学目标进行说明，并基于教学目标对学生的章节知识掌握水平进行自动分类，将分类结果作为学生最近发展区诊断的依据，为当前处于不同发展水平的学生提供个性化的学习评价，学生知识掌握度评价为过程性评价。

4.3.1 成果导向的课程知识教学目标确定

依据教育部标准，即《普通高等学校本科专业类教学质量国家标准》（后续简称《国标》）在软件工程专业人才培养目标中指出，要培养学生掌握基本原理与技术，应用于系统分析设计、系统实现维护、系统开发管理等工作^[47]。依据相关标准和软件工程师学科教学内容，本节对成果导向的软件工程课程知识教学内容进行详细划分，如表 4.2 《软件工程》课程知识结构划分，在知识教学内容上可参照教学章节划分，共包含可行性研究、面向对象分析、结构化分析、面向对象设计、系统实现与维护、项目

管理、软件工程新技术、软件工程概述八个章节。

表 4.2 《软件工程》课程知识结构划分

软件工程概述	面向对象分析与设计
软件危机 软件过程 SDLC	面向对象分析 面向对象设计
可行性研究	系统实现与维护
经济可行性 技术可行性	编码规范 系统测试 系统维护
软件需求分析	软件项目管理
业务需求 用户需求 功能需求	成本管理 质量管理 进度管理
结构化分析与设计	软件工程新技术
总体设计 详细设计	智能编程 缺陷修复

结合《国标》软件人才培养目标和课程教学培养方案,本节对课程知识教学成果目标进行说明,如下表 4.3 软件工程课程知识教学成果目标所示,知识结构为课程教学的主体内容,课程教学与学生评价均以知识教学成果目标为导向。

表 4.3 软件工程课程知识教学成果目标

编号	知识结构	教学成果目标
1	软件工程概述	掌握软件过程,软件危机,软件安全等知识
2	可行性研究	掌握技术可行性、经济可行性等研究方法和理论
3	需求分析	掌握分析方法和需求规格说明书撰写
4	系统分析与设计	掌握结构化分析,总体设计,详细设计知识和方法
5	面向对象分析与设计	掌握面向对象分析与设计方法
6	系统实现与维护	掌握软件开发,软件测试,软件维护等工程能力
7	软件项目管理	掌握进度计划,成本管理知识,熟练使用项目管理工具
8	软件工程新技术	在工程实践中使用新技术,了解软件工程前沿技术或方法

4.3.2 学生知识水平最近发展区诊断

基于教学目标对学生的章节知识掌握度进行分类,如表 4.4 学生章节知识掌握度分类说明,参考历史数据,依据教师最终对学生章节成绩打分情况,将学生的知识掌握度分为 A 优秀、B 良好、C 一般和 D 待改进,阐述对学生课程知识教学目标的完成情况。学生掌握度在数据上的体现主要在章节测试成绩、作业成绩、考勤和线上学习

行为数据。基于章节知识掌握度对学生知识水平最近发展区进行诊断,如表 4.5 最近发展区诊断说明所示,对掌握度为优秀的学生进行评价激励,教师参与引导持续学习;对掌握度为良好的学生进行评价激励,教师参与引导自主学习;对掌握度为一般的学生进行评价提醒,教师和助教介入,及时干预学习路径;对掌握度为待改进的学生进行评价警示,教师和助教干预,及时修正学习路径。

表 4.4 学生章节知识掌握度分类说明

章节掌握度	说明	主要影响因素
A 优秀	达到教学目标水平,教师评分在区间[90-100]	章节作业成绩、测试成绩、课堂出勤率、线上学习完成度
B 良好	基本达到教学目标水平,评分在区间[75-90)	章节作业成绩、测试成绩、课堂出勤率、线上学习完成度
C 一般	未达到教学目标水平,评分在区间[60-75)	章节作业成绩、测试成绩、课堂出勤率、线上学习完成度
D 待改进	未达教学目标水平,亟待教师干预,评分低于 60	章节作业成绩、测试成绩、课堂出勤率、线上学习完成度

表 4.5 最近发展区诊断说明

当前知识掌握度	最近发展区	潜在发展水平	措施
A	→	持续学习	评价激励;引导持续学习
B	→	A	评价激励;引导自主学习
C	→	B	评价提醒;教师、助教干预
D	→	C	评价警示;教师、助教干预

4.3.3 评价算法设计

本小节对学生章节知识掌握度评价的算法设计进行说明。用于训练的学生原始数据主要包括个人基本信息、学生考勤数据、各章节课程测试成绩、各章节课程作业成绩、大作业成绩、课堂互动数据(弹幕数、受奖励次数、互动回答次数)学生平台学习行为数据记录等。具体指标可参见 4.2 实验数据处理部分。为准确判断学生的章节掌握度,对每个维度的特征选取如下:学生课堂参与度维度包括了出勤率、互动回答次数、受奖励次数等;学生线上参与度维度包括学生发布弹幕数、章节线上课程观看时长、章节线上资料完成率、头歌平台闯关完成率、线上平台在线时长、推荐内容下载次数等;学生章节报告维度包括了章节报告成绩,互评成绩等;学生章节测试维度包含课前测试成绩和课后测试成绩等;线上学习行为维度使用处理后的线上学习行为向量。

如图 4.4 知识掌握度评价模型所示, 在学生《软件工程》知识掌握度评价算法上, 本章通过基于多个机器学习分类模型对章节知识掌握度评价进行实验后进行模型融合, 并依据分类结果进行个性化评价。

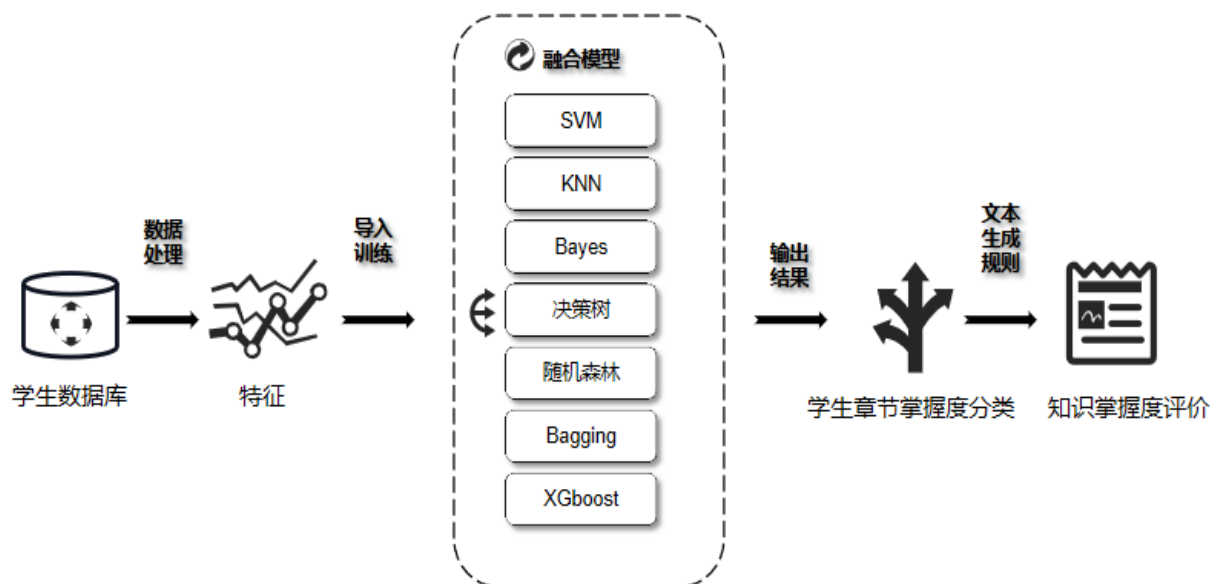


图 4.4 知识掌握度评价模型

本节通过构建多个不同类型的机器学习分类模型并进行对比实验, 将学生数据集拆分为分成包含 415 个学生数据的训练集和包含 104 个学生数据的测试集, 用训练集来训练模型, 并用测试集来评估各模型的准确性, 使用 sklearn 中的 metrics 库来计算准确率, 最终评估模型的性能。下面是对所使用机器学习模型的介绍。

(1) SVM 分类模型

由于在二维空间上难以将学生章节知识掌握情况进行非常准确的切分, 因此考虑使用支持向量机来找到分类平面, 对于超平面 $ax + b = 0$, 样本点 (x_i, y_i) 到超平面的几何间隔距离计算如式 (4.3)。

$$\gamma_i = y_i \left(\frac{\omega}{\|\omega\|} \cdot x_i + \frac{b}{\|\omega\|} \right) \quad \text{式(4.3)}$$

给定整个训练集 $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$, SVM 的训练目标为最大化超平面关于所有样本点的几何间隔的最小值, 如式 (4.4)。

$$\gamma = \min_{\omega, b} \gamma_i, \max_{\omega, b} \gamma_i \left(\frac{\omega}{\|\omega\|} \cdot x_i + \frac{b}{\|\omega\|} \right) \geq \gamma, i = 1, 2, \dots, n \quad \text{式(4.4)}$$

学生章节知识掌握度属性被划分为四类, 使用一对剩余 (One vs Rest, OVR) 的多分类方式, 即分别将四个类别: 优秀、良好、一般和待改进作为正样本, 其余三类为负样本, 对于每一个数据将分别丢到四个分类器中进行得分检验, 将结果划分到得分高

的类别。SVM 模型的实验结果将用作其余模型结果的参照。本节将使用 `sklearn` 库实现 SVM 模型，在核函数上使用了高斯核函数。

(2) KNN

KNN 模型通过计算对象间距离作为各个对象之间的非相似性指标，本节在 K 值选择上采用交叉验证法，计算使用欧氏距离。

(3) 朴素贝叶斯分类器

在上述分类器的基础上，考虑使用概率分类的方式，使用有监督学习算法朴素贝叶斯进行章节掌握度分类，每一个学生的样本数据满足独立同分布特点，通过统计每个条件组合下结果的概率进行分类，判别规则如式 (4.5)。

$$O_{nb}(x) = \underset{c \in Y}{argmax} P(c) \prod_d^{i=1} p(x_i | c) \quad \text{式(4.5)}$$

(4) 决策树分类模型

由于 SVM 使用高维核函数进行分类，具有一定的黑箱性。为了能够进行规则提取，在上文的基础上选用了决策树模型。决策树在构成时会优先选择能使分类效果最好的特征，对于本节中的多维特征具有较好的适应结果。本节使用基尼系数、信息熵等指标来衡量各个特征的分类效果。为了防止一个叶节点中只有一个数据样例的情况出现，设置分裂纯度阈值、最大深度、节点最少样本数量和叶子节点个数上限等来进行剪枝。

(5) 随机森林分类模型

为了在决策树模型的基础上进一步提升，本节又选用了随机森林模型。其包含多个决策树的分类器，学生章节掌握度分类类别的最终输出的结果取所有决策树输出的类别的众数。

本节在构造单棵决策树的时候，从整体的训练集中进行抽样，仅选取部分数据以及部分的数据属性进行构造，防止决策树之间的相关性过高，导致随机森林退化为决策树，设置森林包含 10 棵决策树，特征选择时使用了基尼系数，设置结点的最小样本数量为 30，最大深度为 13，使用了 `sklearn` 框架实现的随机森林模型。

(6) Bagging 集成学习分类

如图 4.5 Bagging 有放回重采样训练方法，采用多个分类器进行随机采样的训练组合，本节中使用 `scikit-learn` 中的 `BaggingClassifier` 元估计器作为子分类器，`max_samples` 和 `max_features` 均设置为 50%，采用有放回的抽取方法，最终的分类结果由多数分类器表决决定。

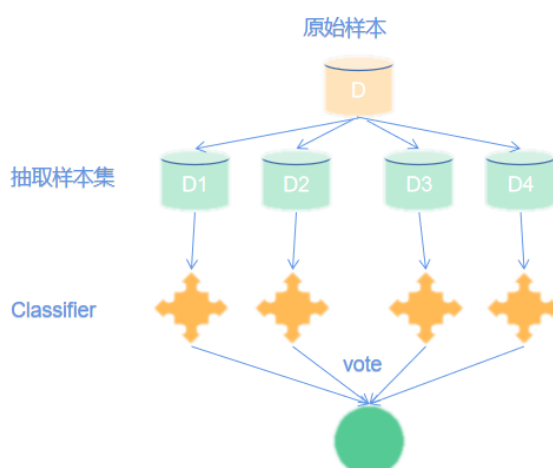


图 4.5 Bagging 有放回重采样训练方法

(7) XGBoost 分类模型

Xgboost 模型实现对学生成绩的预测回归通过最优分裂点进行优化，底层由决策树实现，它使用集成学习方法及并行模式加快计算速度，通过集成多个弱学习器来使得整个模型具有更高的预测能力。每个弱学习器即决策树都是在前一个决策树的基础上进行构建的，如图 4.6 Boosting 方法示例，在算法训练过程中，XGBoost 由多个分类树组成，利用字母 k 表示树的数目， f 函数则代表树的结构，则分类模型可表示为式 (4.6)。

$$\hat{y}_i = \sum_{k=1}^k f_k(x_i) \quad \text{式(4.6)}$$

对于上述 K 个弱分类器，则第 i 个样本的最终输出可表示为目标函数可表示为 \hat{y}_i ；则添加损失函数和正则化项后的目标函数可以表示为式 (4.7)。

$$O(\theta) = \sum_i^n G(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad \text{式(4.7)}$$

为了避免单个决策树的回归精度有限，XGBoost 核心思想是用每一弱分类器的预测结果去拟合上一个分类器预测结果的梯度或残差，可表示为式 (4.8)。

$$\hat{y}_i^{(j)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(j-1)} + f_t(x_i) \quad \text{式(4.8)}$$

我们可以得到 j 棵树模型的预测结果，在数值上等于前面 $j-1$ 棵树的预测结果，加上第 j 棵树的表现。那么对于 j 棵树我们的目标函数如式 (4.9)。

$$O(j) = \sum_{i=1}^n G\left(y_i, \hat{y}_i^{(j-1)} + f_j(x_i)\right) + \Omega(f_j) + N \quad \text{式(4.9)}$$

之后则根据泰勒级数对函数进行近似处理，使损失函数值为常数，此时目标函数只和如何划分树相关。后续将对一棵树进行定义并构建树。为提升模型效果，增加一步特种工程，使用模块 `sklearn.feature_selection` 剔除低方差特征，设置最大树 `max_depth` 和决策树个数 `n_estimators` 均使用默认值避免过拟合或者欠拟合，之后进行训练。

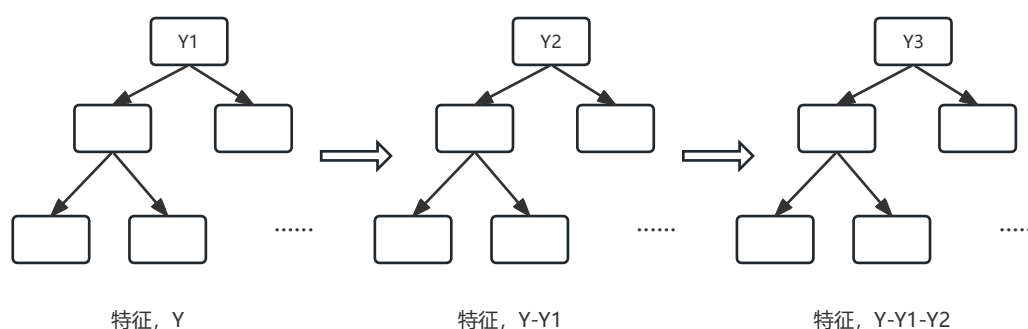


图 4.6 Boosting 方法示例

为进一步提高模型的正确率，避免个别机器学习模型的过拟合或者欠拟合，采用上述分类模型进行集成学习融合模型设计，提高模型整体性能。

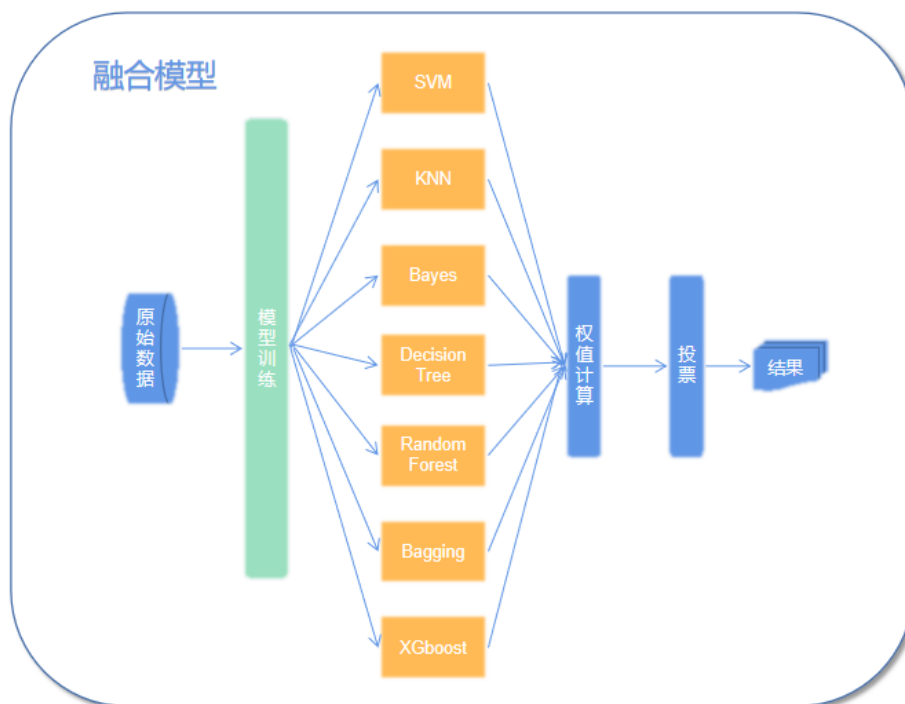


图 4.7 融合模型计算方法

由于将学生章节掌握程度进行四分类，单纯使用 vote 法进行模型融合可能存在重复值，因此在模型融合上选择先对分类器进行加权平均再投票的方法，如图 4.7 融合模型计算方法。

对于每个分类模型，计算在测试集上的准确率，使用 softmax 函数对准确率进行 0-1 映射作为每个分类器权重，将四种分类结果分别乘以分类器权重，得到的值作为票数，依据少数服从多数的原则，将票数最大的结果作为输出并计算在测试集上的准确率。如算法 2：知识掌握度融合模型算法所示。

算法 2：知识掌握度融合模型算法

输入：训练好的七种分类模型 models = []；测试集；训练集

输出：学生章节掌握度分类结果 results

1. models=[DecisionTreeClassifier(),RandomForestClassifier(),KNNClassifier(),BayesClassifier(),BaggingClassifier(),SVMClassifier(),XGboostClassifier()# 定义融合模型
 2. accuracies = []# 训练每个分类器并在测试集上计算准确率
 3. for model in models:
 4. model.fit(X_train, y_train)
 5. accuracies.append(model.score(X_test, y_test))
 6. weights = np.exp(accuracies) / np.sum(np.exp(accuracies))# 使用 softmax 函数对准确率进行 0-1 映射得到每个分类器的权重
 7. predictions = np.zeros((X_test.shape[0], len(models)))# 对测试集进行预测并根据分类器权重进行加权平均和投票
 8. for i, model in enumerate(models):
 9. predictions[:, i] = model.predict(X_test)
 10. predictions_weighted = np.sum(predictions * weights, axis=1)
 11. results = np.zeros_like(predictions_weighted)
 12. for i in range(len(predictions_weighted)):
 13. results[i] = np.argmax(np.bincount(predictions[i], weights=weights))
 14. print(results)# 输出融合模型的结果
-

如表 4.6 各模型评价指标结果，对于融合模型性能，本节使用了分类准确率指标进行评价。准确率计算可以表示为 $(TP+TN)/N$ ，分子分母分别代表正确分类的样本数和总体样本数。对于各模型的实验结果，可以看出使用融合模型对学生章节掌握度进行分类预测拥有较高的正确率。

在使用融合模型完成对学生知识掌握度分类后确认每类学生的详细评价内容，如表 4.7 知识掌握度评价话术所示，以课程系统实现教学章节为例，话术中包含对学生当前章节知识水平和潜在发展水平评价，帮助学生找准最近发展区，明确发展方向。

表 4.6 各模型评价指标结果

模型	accuracy
SVM	74.36%
KNN	76.03%
Bayes	66.67%
Decision Tree	87.18%
Random Forest	89.74%
Bagging	85.90%
XGboost	87.03%
融合模型	89.95%

表 4.7 知识掌握度评价话术

掌握度	当前水平评价话术	潜在发展水平评价话术
A	根据你的学习状态和各项测试，我们判断你能够熟练掌握软件开发、测试和维护的理论基础和工程实践，完整掌握章节所学知识	你已达成该章节课程教学目标要求，你的下一步目标是全面发展，提升持续学习能力，关注自身其他教学目标的达成度，保持领先状态
B	你的各项数据表现良好，我们判断你较好的掌握了系统开发的课程知识，稳定在了较高水平，在小组项目的开发、测试、或运维上有着关键贡献	下一阶段，你需要完善全部线上学习过程，独立完成工程实践，在课堂测试中达到满分状态，相信你在章节各方面表现均能达到优秀
C	经过学情分析可以猜测你目前具有一定开发经验，部分掌握了开发测试理论和方法，但参与度不足，还有较大进步空间	教师和助教会及时联系你，下一步希望你、能够独立重复完成编码学后测试，融入课程项目，成功负责项目开发，测试或运维的任一环节，积极参与线上学习
D	请注意！你的实践经验不足，还未掌握软件开发，测试和维护方法；缺乏独立编码能力，缺乏学习动力	教师和助教会及时与你联系，纠正你的学习路径，下一步希望你能够独立重复完成编码测试，融入课程项目，成功负责项目开发，测试、运维的任一环节

在学生知识掌握度评价生成算法设计上，将全部教学章节的知识掌握度话术导入评价模板，评价报告的实现将采用基于模板的评价生成算法，依据章节掌握度分类结果为不同的学生选择评价话术，对学生当前发展水平和潜在发展水平进行说明，实现对不同学生的个性化评价。如算法 3：学生知识掌握度评价生成算法所示。

算法 3：学生知识掌握度评价生成算法

输入：基于学生数据的章节知识掌握度分类结果 `result_list[][]`；评价话术模板 `comments[]`

输出：学生章节知识掌握度评价语句

```

15. →String[][] comments = {} // 定义评价模板
16. →int[][] studentGrades = {}; // 学生评级结果
17. commentsMap.put()...//输入或修改评价话术
18. level_evaluation ← "在章节知识掌握度上"
19. for (int i = 0; i < comments.length; i++)
20. for (int j = 0; j < comments[i].length; j++)
21. String[] commentArr = commentsMap.get((i+1) + "章节" + getcomment(j));
22. if (commentArr == null) {
23. continue;
24. evaluation += commentArr[result_list[i][j]] + "\n"; // 根据分类结果添加话术
25. System.out.println(level_evaluation); // 输出评价

```

在生成学生知识结构评价的同时，系统将根据分类结果自动生成学生的软件工程知识掌握结构图，帮助学生了解实际课程章节的掌握情况，查漏补缺。以某学生数据为例，生成结果如图 4.8 某学生软件工程知识掌握结构图。

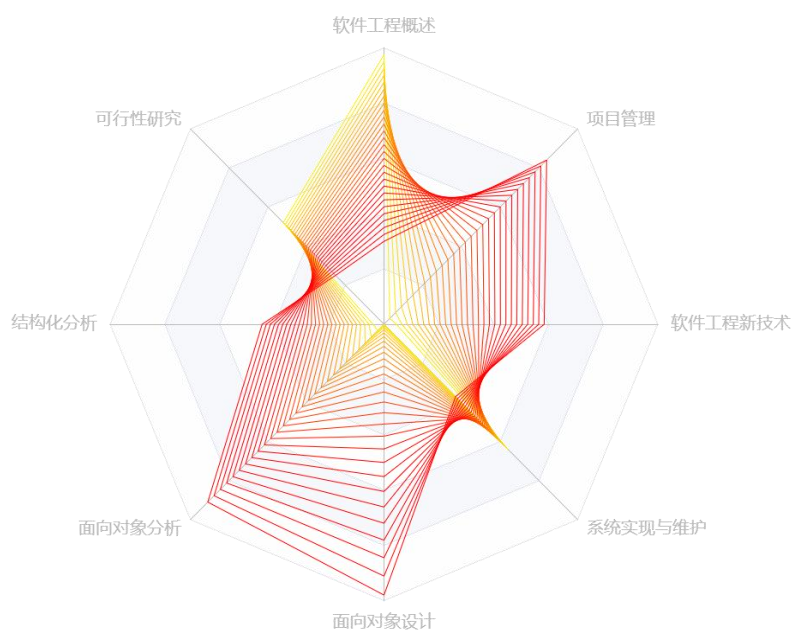


图 4.8 某学生软件工程知识掌握结构图

4.4 学生课程素养评价算法设计

教育部在 2020 年印发了《高等学校课程思政建设指导纲要》，指出要将思政教

育贯穿高校教育培养体系，推进课程思政建设，充分发挥课程的育人作用^[48]。考虑《软件工程》课程教学目标设计，学生课程素养评价设计在评价指标的选择上涵盖学生核心素养，专业认同感评价，在软件工程思政教学体系中具有全面性和代表性。在评价标准制定上参照评价指标的要求，设定评分标准和评价描述，既能够体现个性化的差异，又能够反映学生的优点和不足，为学生提供改进的方向。在评价方法上将在相关评价量表的基础上改进，设计基于学生数据的个性化评价方法，使评价能够更加客观地反映学生的课程素养水平。在评价结果的应用上，将学生课程素养评价融入个性化的评价报告，激发学生的主动参与和反思。

4.4.1 评价指标设计

以《软件工程》课程中对软件人才的职业和素养培养目标为参考，在评价指标设计上选择学生核心素养，专业认同感评价作为评价指标。

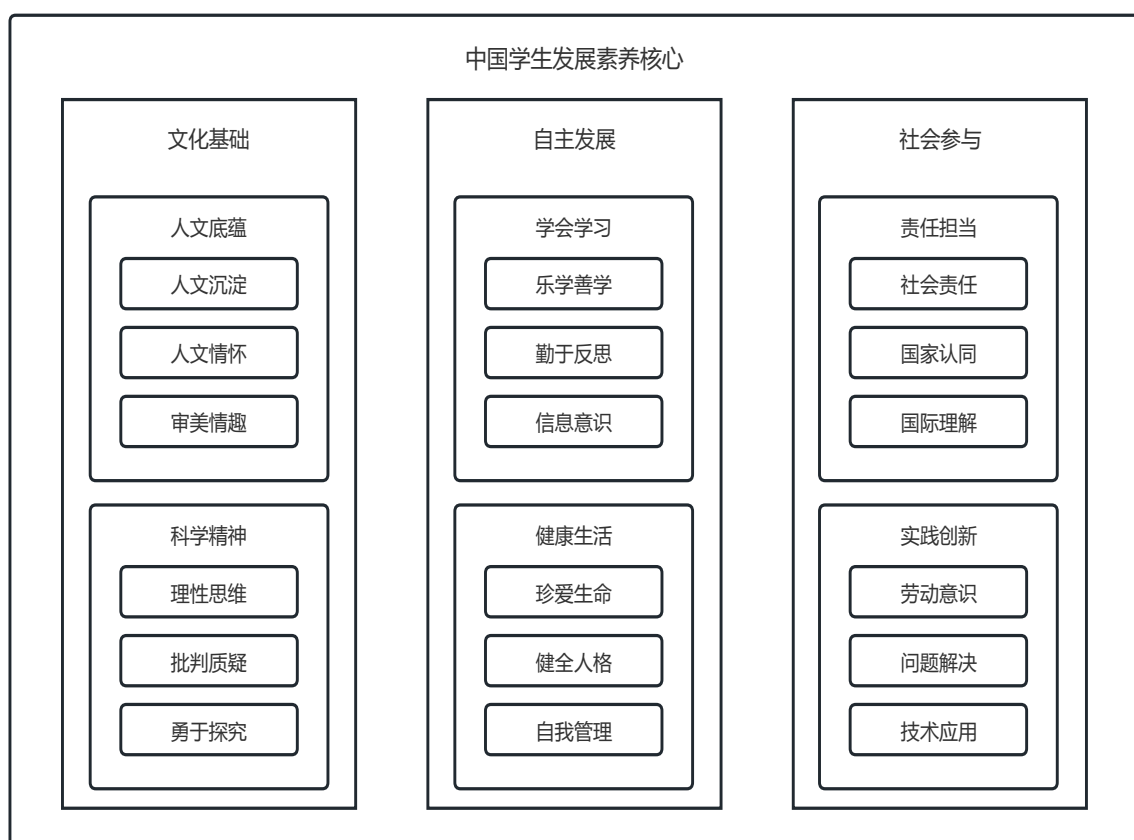


图 4.9 中国学生发展核心素养

如图 4.9 中国学生发展核心素养^[49]所示，教育部于 2016 年发布相关研究成果，《中国学生发展核心素养》的理论和指标设计为评价学生客场核心素养提供了内容和

关键词。

如表 4.8 课程素养评价指标与维度所示, 在评价指标与维度设计上依据社会心理学所探讨的态度构成, 将学生专业认同划分为三个方面, 分别为专业情感、专业行为和专业认知三个方面^[50]。包含核心发展素养的人文、科学、学习、生活、责任、创新等六个主要维度。

表 4.8 课程素养评价指标与维度

评价指标	评价维度	说明
中国学生发展核心素养	人文底蕴	在学习知识过程中的价值取向
	科学精神	学习科学知识形成的行为表现, 如批判质疑、理性思维等
	学会学习	学习方法和进程调控方面的表现, 反思能力和信息意识等
	健康生活	学生认识自我, 发展身心的表现, 主要是自我管理
	责任担当	处理与团体关系的价值取向, 在课程中体现在集团责任
	实践创新	解决问题挑战等方面的表现, 如问题解决、技术应用等
	专业认知	对课程的了解程度, 包括培养目标、课程安排、应用价值
专业认同感	专业情感	对课程的情感态度, 包括获得感、热情等等
	专业行为	对专业学习的积极行为, 例如参与度、投入程度

4.4.2 评价标准设计

评价标准的设计用来衡量学生在软件工程课程学习中综合素养水平和专业认同感。传统的学生综合素养和专业认同感测评通常以问卷量表的形式进行, 本文是基于学生学习数据的个性化评价研究, 如表 4.9 学生课程素养评分计算方法, 在学生专业认同感问卷和学生核心素养李克特五分评价量表的基础上改造, 将常见量表问题转化为基于数据的评分, 对相关评价标准进行量化。

表 4.9 学生课程素养评分计算方法

维度	常见量表问题	基于数据的评分构成(百分制)
人文底蕴	我已掌握专业知识	作业均分*0.3+报告均分*0.4+附加测试*0.3
	我已掌握专业规范	
科学精神	我有理性思维	提问次数/10*0.4+一句话软件工程评分*0.4+展示报告得分*0.2
	我敢于质疑批判	
	我勇于探究	

续表 4.9 学生课程素养评分计算方法

维度	常见量表问题	基于数据的评分构成（百分制）
学会学习	我认真踏实 我信息处理能力强 我善于团队协作	章节互评平均分*0.3+资料完成率*40+项目管理实践得分*0.3
健康生活	我能够自我管理	出勤次数/课程次数*100
责任担当	我愿意担任项目经理 我愿意担任组长	担任项目经理(T:+50, F:+0) 担任组长 (T:+50, F:+0)
实践创新	我能够独立解决问题 我能够应用新技术 我能理论用于实践	项目编码成绩*0.4+附加技术测试成绩*0.4+项目管理实战得分*0.2
专业认知	我了解课程领域成果 我了解领域发展方向 我了解课程章节结构	软件工程概述报告得分*0.6+软件工程概述章节测试得分*0.4
专业情感	我对课程很有热情 我相信能学好课程 我尽可能完成课程要求	出勤次数/课程次数*30+总体完成率*50+推荐内容点击次数/15*20
专业行为	我经常阅读推荐资料 我经常参考课程互动 我能利用课程知识思考	视频观看时长/总时长*0.3+资料完成率*20+课程互动次数/10*30+参与讨论次数/5*20

4.4.3 评价算法设计

本节对学生思政情感评价主要包含学生综合素养评价与学生专业认同感评价，共计九个维度，依据学生数据综合计算出每个维度的评分后进行评价。在评价方法与评价话术上采用正向勉励的评价手段。作为师生间的对话行为，正向勉励具有语言行，互依性，导向型等属性，反应了怎样勉励学生和勉励学生的方向^[51]。如表 4.10 勉励型评价话术设计所示，基于正向勉励的评价方式，将学生课程素养方面的表现分为两类：

（1）对于评分高于 50 的维度，认为学生在该维度上有较好的表现，采用赞许型的评价话术满足学生自我实现的内在需求。学生都有自我实现的内在需求，被赞赏是实现学生潜能的有利条件，帮助学生有信心解决问题和突破自我，教师评价中要包含对学生的积极赞赏，便于“积极的自我意向”的形成。

（2）对于评分低于 50 分的维度，认为学生在该维度上的表现还有较大提升空间，可以给予期待型的评价。教师给予学生期待型的评价话语，给予建设性的意见，鼓励学

生向响应目标努力，这种期待包括情感，意志等对学生个性化发展有意义的方面^[52]。

表 4.10 勉励型评价话术设计

勉励型评价	评价话术设计示例
赞许型（评分>50）	<ol style="list-style-type: none"> 1. 你在 xx 方面有着非常突出的表现，证明你是一个综合素质优秀的学生，拥有这一品质能够帮助你快速学习软件工程的知识与技能 2. 老师关注到你 xx 方面表现格外突出，这一优点是软件工程课程学习的关键素养，请你再接再厉，保持领先
期待型（评分<50）	<ol style="list-style-type: none"> 1. 在 xx 的养成上，你还有着很好的发展前景和进步空间，在后续的学习中，老师和助教会帮助你在 xx 上加深认知，期待你的表现 2. xx 方面是你之后的努力方向，及时与老师沟通，相信你的成绩会更优秀

本节情感价值评价的实现将采用基于模板的评价生成算法，评价模板中将包含评价维度相关的关键词，依据数据为不同的学生选择评价话术，满足不同学生的个性化评价诉求。如算法 4：学生课程素养评价生成算法所示。

算法 4：学生课程素养评价生成算法

输入：基于学生数据计算得到的九个维度评分 scores[]；每个维度赞许型和期待型的评价话术模板 judge[][]

输出：学生课程素养评价语句 Quality_evaluation

1. $\rightarrow \text{judge}[][] = \{\dots\{"在\ xx\ 方面\ 有着\ 非常\ 突出\ 的\ 表现\dots", "关注\ 到\ 你\ xx\ 方面\ 表现\ 突出\dots",\dots\},\{"在\ xx\ 的\ 养成\ 上,\ 还有\ 着\ 很\ 好\ 的\ 发展\ 前\ 景\ 和\ 进\ 步\ 空\ 间\dots", "xx\ 方面\ 是\ 你\ 之\ 后\ 的\ 努力\ 方\ 向\dots",\dots\}\dots\};$
2. $\rightarrow \text{Quality_evaluation} = "你在课程素养方面, ";$
3. for (int i = 0; i < 9; i++)
4. int length = judge[i].length;
5. int j = (int) (Math.random() * length); #随机选择语句
6. if (scores[i] > 50)
7. Quality_evaluation += judgements[i*2][j] + "。 \n"; #加入赞许型评价
8. else
9. Quality_evaluation += judgements[i*2+1][j] + "。 \n"; #加入期待型评价
10. $\rightarrow \text{System.out.println}(\text{Quality_evaluation});$

4.5 学生学习风格评价算法设计与实现

学习风格评价为终结性评价，是对学生本学期课程的诊断。本章在学生在学习风格

评价上参照心理学家 David A. Kolb 和 Roger Fry 所研究形成的所罗门学习风格量表, 学生的学习风格由学习风格问卷(Index of Learning Styles Questionnaire, ILS)推断而出^[53], 从信息加工、感知、输入、理解四个方面对学习风格进行划分, 共划分为四组, 每组包含基本对立的两种类型, 在信息加工上划分活跃型和沉思型; 在信息感知上划分感悟型和直觉型; 在信息输入方面划分视觉型和言语型; 在信息理解上划分序列型和综合型。本节在所罗门学习风格量表的基础上对学生进行学习风格评价。如表 4.11 学生学习风格偏好所示, 学习风格划分四个组别, 8 种类型偏好。

表 4.11 学生学习风格偏好

组别	类型	偏好
1	活跃型	先做后想型, 积极做事, 讨论
	沉思型	安静思考, 独立工作
2	感悟型	结构化、有序化、理性化知识转换
	直觉型	灵活创新, 不喜重复, 擅长新概念和抽象
3	视觉型	转化信息多来自课堂图像、影片和演示
	言语型	擅长通过文本和言语转换信息
4	序列型	偏好分点学习, 线性完成学习过程
	综合型	偏好先获取宏观视角, 大幅获取知识

4.5.1 评价指标与评价标准设计

本节在参考学习风格量表的基础上通过对学生数据进行分析, 调整和修改学习风格模型, 避免了学习风格问卷带来的主观影响, 使其适用于某大学软件工程教学模式下的学生学习风格评定。

本节将通过分析学生学习行为对学生学习风格的影响来对模型进行修改, 将由问卷结果反应学生学习风格的评价方式调整为由学生学习行为数据评价学生学习风格, 其工作包括总结学习行为模式与学习风格间的关系并进行验证, 最终以调整后的模型作为学生学习风格评价的基础。

在 Felder-Silverman 量表中, 同一组的两种学习风格类型在信息加工、感知、输入、理解等特征是上是不同或者相反的, 因此某一项数据在对同一组两种学习风格的评定上应当互斥或属于不同的区间范围。例如以第一组活跃型和沉思型风格为例, 学生课程访问次数 T_course 越多, 学生课程发布弹幕数 $N_barrage$ 越多, 则学习风格越偏向活跃型; 反之则偏向沉思型。

根据 ElHaddioui Ismail^[54]等对学习风格与行为模式研究, 以及相关文献对行为序

列与学习风格关系之间的研究验证, 本文使用了信度较高的学习风格问卷量表对学生进行学习风格调查, 将通过学习风格问卷所计算出的具有相同学习风格的学生进行分组, 再采集学生学习数据并依据评价规则计算, 最终将计算结果与设定的区间阈值比对后进行修正, 如表 4.12 学习风格与学习行为阈值所示, 本节将结合《软件工程》课程数据特征对学生学习风格评价与学习行为和成绩表现设置阈值区间。

表 4.12 学习风格与学习行为阈值

学习风格	学习行为或表现	阈值区间	
		(+)	(-)
活跃型 (+) 沉思型 (-)	学生线上课程登录总次数 T_course	>20	<7
	学生发布弹幕总数 N_barrage	>10	<3
	课堂互动记录次数 N_inter	>6	<2
	是否担任项目负责人 Y_charge	=1	=0
感悟型 (+) 直觉型 (-)	抽象知识章节访问时长/总访问时长 P_abstract	<10%	>30%
	实验章节访问时长/总访问时长 P_test	<10%	>50%
	记忆型章节测试均分 G_memory	>85	<80
	抽象型章节测试均分 G_abstract	<70	>90
视觉型 (+)	线上视频访问次数 N_video	>15	<5
言语型 (-)	文本教案下载次数 N_text	>6	<2
序列型 (+)	点击快进次数/总点击次数 P_fast	>40%	<10%
综合型 (-)	总览章节访问次数/总访问次数 P_overview	<10%	>30%

4.5.2 评价算法设计

对于上述学习行为表现及对应的学习风格所设定的阈值, 需要设定相应的评价规则来计算学生的学习风格; 以组一活跃型和沉思型学习风格为例, 在本文获取的学生数据中, 其主要受 4 个学生行为或表现 h1、h2、h3、h4 的影响, 若学生行为 h 属于活跃型 (+) 的阈值区间, 则将 h 记为 1, 若学生行为 h 属于沉思型 (-) 的阈值区间, 则将 h 记为 0, 否则记为中间值 0.5; 以此为规则计算学生学习风格 O_j 的值, 可表示为式 (4.10)。

$$O_j = \sum_{i=1}^n h_i / n \quad \text{式(4.10)}$$

上述公式中 n 表示每一组学习风格收到几种学生行为或表现的影响, 以此计算得出的每一组的 O_j 值在区间 $[0,1]$ 之间, 当值在区间 $[0,1/3)$ 时, 则表示计算出的学生学习风格为该组中 (-) 代表的学习风格; 当值在区间 $(1/3,2/3)$ 时, 表示该学生在本组的学习风格特征中表现不明显, 可记为平衡型; 同理, 当值在区间 $(2/3,1]$ 时, 计算出的学生学习风格为该组中 (+) 所表示的学习风格, 如算法 5: 学生学习风格评价生成算法所示。

算法 5: 学生学习风格评价生成算法

输入: 基于学生数据计算得到的四个维度评分 `learningStyleArray[]`; 学习风格评价话术模板 `evaluationTemplate`

输出: 学生学习风格评价语句 `Style_evaluation`

```

1.  →learningStyleArray[]
2.  →evaluationTemplate = "经过对你学习行为数据的评估, 我们猜测你在信息加工维度的学习风格为 %s ;  "+"在信息感知维度, 你的风格是 %s ;  "+"在信息输入维度, 你的风格是 %s ;  "+"在信息理解维度, 你的风格是 %s 。";
3.  String[] processingTypes = {"活跃型", "平衡型", "沉思型"};
4.  String[] perceptionTypes = {"感悟型", "平衡型", "直觉型"};
5.  String[] inputTypes = {"视觉型", "平衡型", "言语型 I"};
6.  String[] understandingTypes = {"序列型", "平衡型", "综合型"};
7.  String processingType = getType(learningStyleArray[0], processingTypes);
8.  String perceptionType = getType(learningStyleArray[1], perceptionTypes);
9.  String inputType = getType(learningStyleArray[2], inputTypes);
10. String understandingType = getType(learningStyleArray[3], understandingTypes);
11. String evaluation = String.format(evaluationTemplate, processingType, perceptionType, inputType, understandingType);
12. System.out.println(Style_evaluation); {#输出评价语句}
13. private static String getType(int value, String[] types)
14. if (value < 1/ 3)
15. return types[0];
16. else if (value > 2/ 3)
17. return types[2];
18. else
19. return types[1];

```

4.6 学生成绩预测算法设计

学生成绩预测为过程型评价，能够帮助学生进行学习预警。如图 4.10 学习成绩预测模型，学生学习成绩预测模型将在学生评价模型的基础上，重点分析学生知识掌握程度以及学生学习行为与章节成绩间的关系，预测学生的最终学业成绩，该功能为学生个性化评价提供重要依据，是教学评价中的重要一环。

本节将重点讲解通过多种机器学习算法模型分别训练和实现预测任务，对比实际效果，并基于上述模型实现学生期末成绩预测的融合模型算法设计。

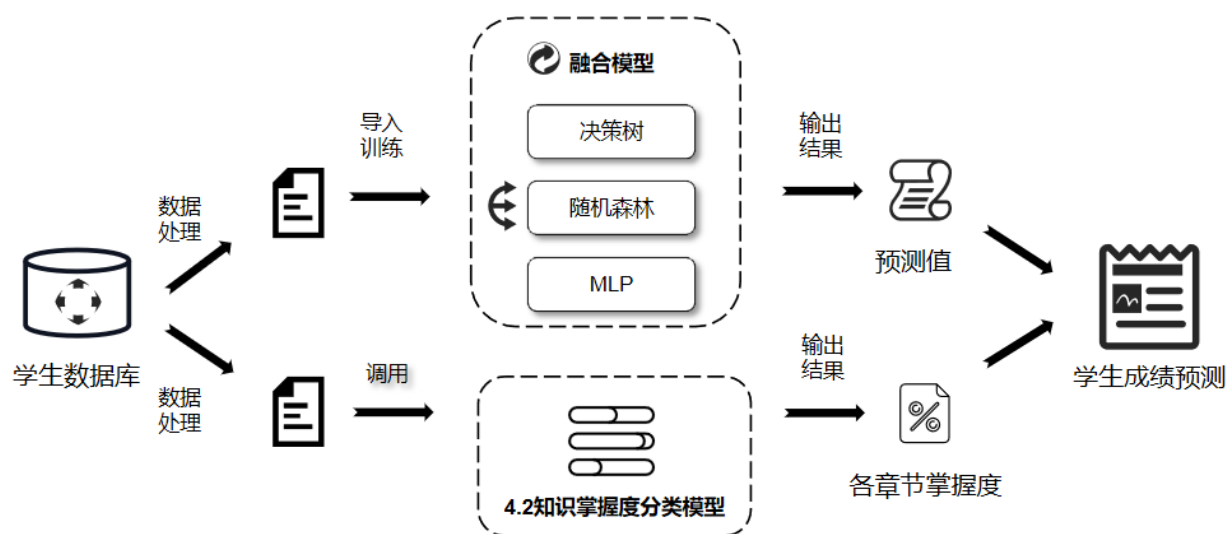


图 4.10 学习成绩预测模型

4.6.1 特征选择

对 4.2 中清洗后的数据进行特征选择。由于学生数据片段过多，在部分特征设计上采用多个数据片段综合计算得出单个特征的模式，例如学生线上操作行为可处理为单个向量，学生互评信息可处理为互评意见数量等。

为保证预测模型的准确性，在特征选择避免过度拟合和提高模型效果，使用数据挖掘工具对可处理为连续变量的特征数据进行筛选，计算变量与结果之间的相关性，如图 4.11 部分特征相关性分析结果，在移除不重要变量后根据实际情况和课程要求，本节选取了如下特征变量对于学生章节表现属性 S_{ij} ，选用特征①各章节报告成绩 S_{1j} ；②各章节测试成绩 S_{2j} ；③各章节互评意见数 S_{3j} ；对于学生课堂行为属性 p_i ，选用特征④互动答题次数 p_1 ；⑤缺勤率 p_2 ；⑥受奖励次数 p_3 。对于学生线上行为属性 b_i ，选用特征⑦教学网在线学习时长 b_1 ；⑧powerhighlight 平台在线时长 b_2 ；⑨moodle 在线时长 b_3 ；⑩线上行为向量 b_4 。各特征变量的取值均为实数类型。

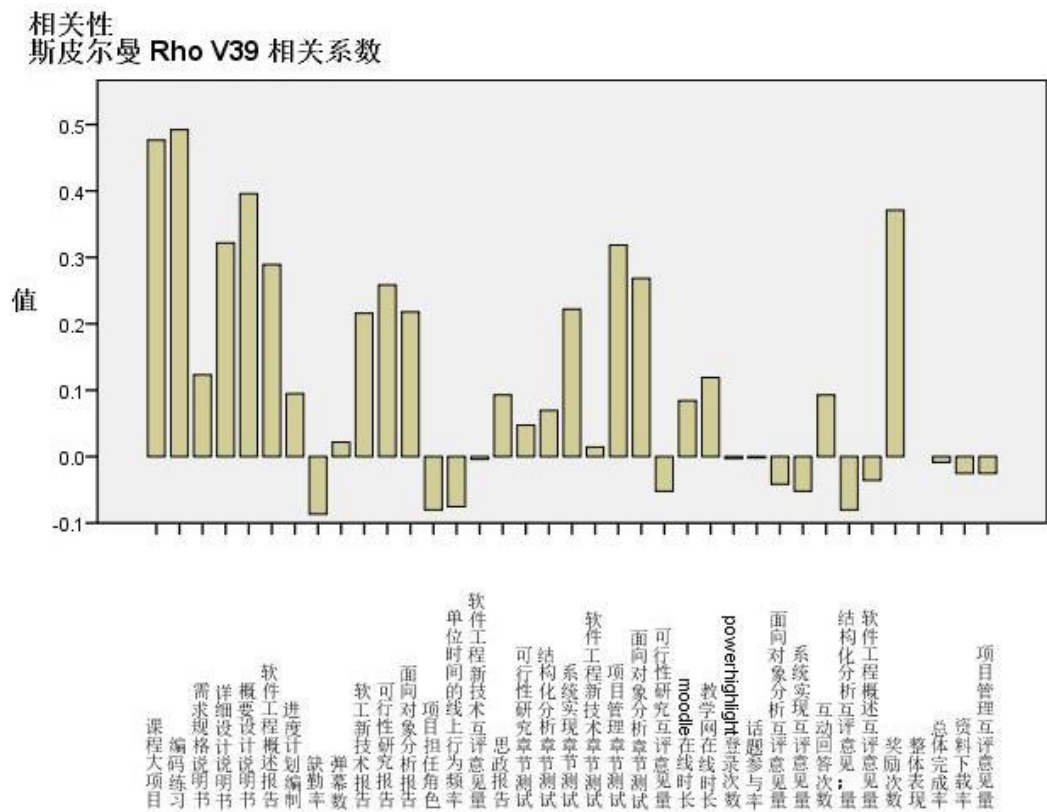


图 4.11 部分特征相关性分析结果

4.6.2 成绩预测算法设计

本节将选用多个机器学习回归模型如表 4.13 常机器学习回归算法模型特性对比，并进行预测效果实验，选用拟合效果较好的模型进行模型融合进一步实验比对。

表 4.13 常机器学习回归算法模型特性对比

算法名称	模型类型	模型特点	训练算法
linear_regression	判别模型	拟合线性关系	最小化分割面距离和
Ridge 岭回归	判别模型	正则化、防止过拟	L2 正则化方法
Lasso 套索	判别模型	正则项、有偏估计	正则化、交叉验证
Decision_tree	判别模型	分类树	正则化的极大似然估计、剪枝
SVR 支持向量机	判别模型	分类超平面，核技	极小化正则合页损失
Random_forest	判别模型	随机性、集成方法	基于决策树的集成学习方法
XGBoost	判别模型	基于 CART 的集成	极小化正则集成损失、剪枝
MLP 多层感知机	判别模型	分离超平面	误分点到超平面距离极小化

将特征数据整理导入后数据装载，分别使用上述算法模型进行训练，特征评价标

准均设置为 MSE，计算输出训练后的模型在测试集上的均方根误差、均方误差、平均绝对误差和决定系数，并进行模型间比较。对于全部模型的训练，均将目前已经清洗完成的 519 位学生特征数据划分为包含 80% 学生数据的训练集和包含 20% 学生数据的训练集。

（1）使用 linear_regression 线性回归模型

学生期末成绩与特征之间具有一定线性相关性，对学生成绩预测使用线性回归模型，使用参数 w 和 b 作为求解损失函数最小时的函数， w 代表特征 x_i 前系数， b 为常数。通过构建如下损失函数，对满足均方误差最小的参数进行求解，可以表示为式(4.11)。

$$(w^*, b^*) = \arg \min_{(w, b)} \sum_{i=1}^n (wx_i + b - y_i)^2 \quad \text{式(4.11)}$$

模型优化方法则采用梯度下降法，通过对 w 和 b 求偏导不断更新自变量。之后对参数 `normalize`、`fit_intercept`、`copy_X` 和 `n_jobs` 进行调参，将数据集导入后进行训练。

（2）使用 Ridge 岭回归模型

Ridge 岭回归可以更好的解决学生成绩预测中多重共线性的问题，在模型方法上，使用 `auto` 的求解器使用并使用交叉验证方法来选择最优岭回归系数。

（3）使用 Lasso 套索模型

Lasso 与 Ridge 的差别主要体现在损失函数正则项中，如图 4.12 Lasso 等高线与约束域所示，Lasso 方法由于约束域为方形可能存在切点在坐标轴上的结果，使维度权重存在为 0 的可能，可以剔除部分特征，由于学生数据特征数远少于样本数，参数求解选用坐标轴下降法。

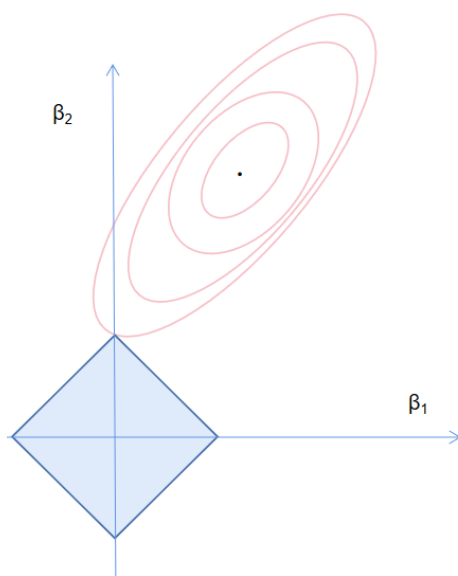


图 4.12 Lasso 等高线与约束域

(4) 使用 SVR 支持向量机模型

SVR 在解决学生成绩这一回归问题上是通过找到平面来满足集合数据到平面的距离最近。在模型参数上,使用网格搜索法对超参数 `svr__C`、`svr__kernel`、`svr__gamma` 等进行参数最优取值,核函数选择高斯核函数进行模型训练。

(5) 使用 Decision_tree 决策树模型

在使用决策树模型解决学生成绩预测问题的过程中,可以将决策回归树视为具有较多分类的分类问题,原理如图 4.13 训练后的决策树模型。

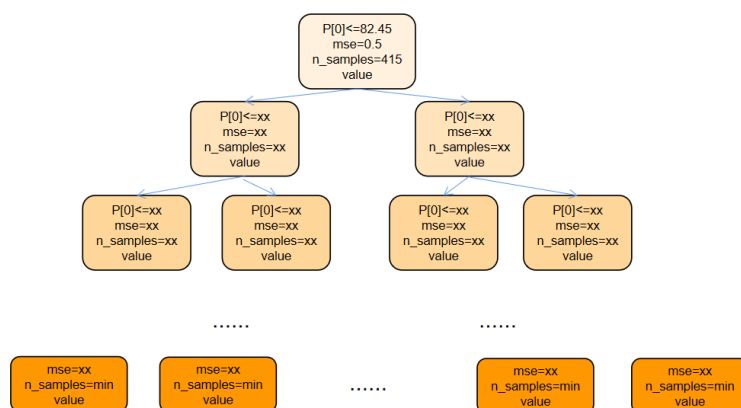


图 4.13 训练后的决策树模型

(6) 使用 Random_forest 随机森林模型

Random_forest 是由多个决策树构建的集成学习模型,对学生成绩的预测将使用多个决策树的均值,较适合对于学生期末成绩的回归问题,参数优化上选择导入 sklearn 的网格搜索模块优化超参数组合,逐步缩小 `n_estimators` 等参数范围。

(7) 使用 MLP 前馈神经网络进行预测

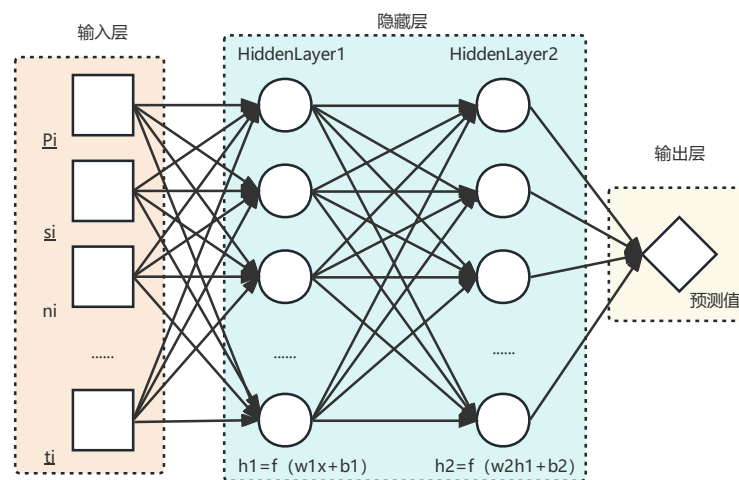


图 4.14 四层人工神经网络结构

在学生学习成绩预测过程中，将面临影响学生学习成绩特征之间权重难以估算的问题，上述机器学习模型中大部分是使用线性组合来回归，这里引入人工神经网络 MLP 搭建解决非线性预测的神经网络，层与层直接为全连接，输入层为包含 4.2 中选取的数据特征所组成的多维向量，隐藏层为两层，两层的输出分别使用 $f(w_1x+b_1)$ 和 $f(w_2(w_1x+b_1)+b_2)$ 来表示，在激活函数 f 选择上使用 sigmoid 函数，使神经网络能够逼近非线性函数，该模型为 HiddenLayer+HiddenLayer+线性层映射到单个标量的四层 MLP 模型，如图 4.14 四层人工神经网络结构所示。

针对上述线性模型、树模型和神经网络模型的实验，进行预测效果对比，参考参数为测试集上得出的 MSE 、 $RMSE$ 、 MAE 和 R^2_score 。在下列模型性能指标公式中， n 代表测试集样本个数， $y_{test}^{(i)}$ 代表模型计算得出的学生成绩预测值， $\hat{y}_{test}^{(i)}$ 代表样本中的真实学生成绩， $\bar{y}_{test}^{(i)}$ 为成绩均值。

MSE 为均方误差，代表学生成绩预测值与实际成绩之间差距的平方和的均值， MSE 越小则代表预测效果越接近真实值， MSE 的计算逻辑如公式 (4.12)。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{test}^{(i)} - \hat{y}_{test}^{(i)})^2 \quad \text{式(4.12)}$$

$RMSE$ 为均方根误差，为解决不同问题量纲不同对均方误差的影响，使用开方后的预测值与实际值之间差距的平方和均值， MSE 与 $RMSE$ 呈现正相关，都能表示预测的准确度，其计算如公式 (4.13)。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{test}^{(i)} - \hat{y}_{test}^{(i)})^2} \quad \text{式(4.13)}$$

MAE 为平均绝对误差，通过计算差值的绝对值来避免正负抵消的影响，在学生预测准确性的判别中与前两种指标起到相同的作用，其计算如公式 (4.14)。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{test}^{(i)} - \hat{y}_{test}^{(i)}| \quad \text{式(4.14)}$$

R^2 为确定系数，上述三种指标能够反映预测模型的性能，但当面临不同计量单位的预测问题时难以直观比较或判断，比如对于等级制和百分制的学生成绩进行预测效果对比时，前三种模型在结果上难以评估。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{test}^{(i)} - \hat{y}_{test}^{(i)})^2}{\sum_{i=1}^n (y_{test}^{(i)} - \bar{y}_{test})^2} \quad \text{式(4.15)}$$

引入 R^2 是找到样本中学生成绩均值，其计算可表示为公式(4.15)。 R^2 趋近1，则表明预测结果靠近真实值，模型效果越好；趋近0则表示预测值都接近均值，没有预测意义。

表 4.14 各模型预测结果对比

模型	MSE	RMSE	MAE	R^2
Linear_regression	20.95	4.58	2.67	0.78
Ridge	20.87	4.57	2.67	0.78
Lasso	21.51	4.64	3.16	0.77
SVR	29.82	5.46	3.73	0.69
Decision_tree	15.24	3.90	1.23	0.84
Random_forest	14.13	3.76	1.86	0.84
MLP 神经网络	15.61	3.95	1.25	0.83
融合模型	13.70	3.70	1.23	0.86

实验结果参考表 4.14 各模型预测结果对比，从上述模型在均方误差、均方根误差、平均绝对误差和决定系数上的结果可以看出，决策树模型，随机森林模型和多层感知机模型的 MSE 数值更低， R^2 均高于 0.8，模型效果较好且具有更高的准确性，更适用于本研究的数据环境下预测学生学业成绩。

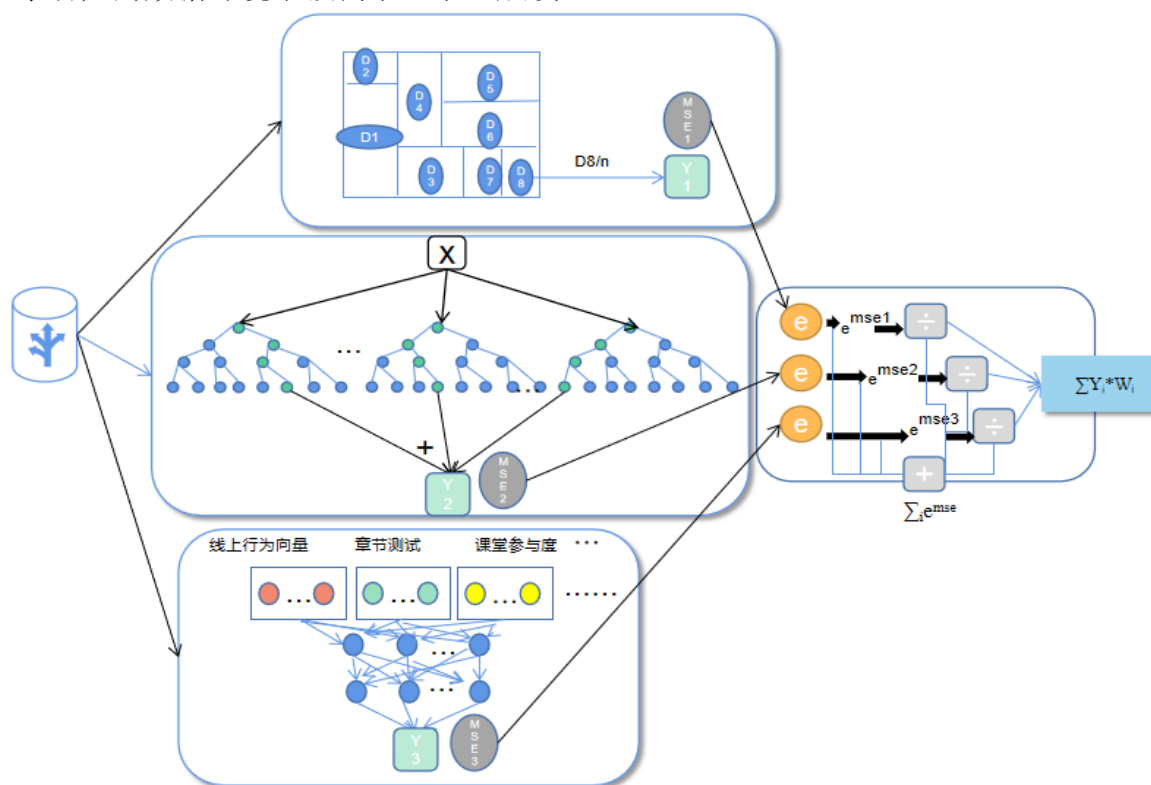


图 4.15 决策树、随机森林、MLP 融合模型

在上述实验结果的基础上, 为进一步提高预测的准确率, 基于预测效果较好的 Decision_tree、Random_forest 和 MLP 神经网络模型创建融合模型。如图 4.15 决策树、随机森林、MLP 融合模型所示。在模型融合上选用加权平均法的融合方法, 依据不同模型的效果进行权重分配, 具体实现为计算三个模型 MSE 的倒数, 并将这些倒数存储在一个权重列表中, 使用 Softmax 函数处理这些倒数, 得到三个和为一的权重。最后将这些权重将三个模型的预测值融合在一起, 并输出融合模型的预测值, 如算法 6: 学生学习风格评价生成算法所示。

算法 6: 学生学习风格评价生成算法

输入: 训练好的三种模型 `models = [DecisionTreeRegressor(), RandomForestRegressor(), MLPRegressor()]`; 数据库中的学生数据

输出: 学生预测成绩 `ensemble_preds`

```

1. def load_data():#数据库导入训练集和测试集数据
2.     conn = mysql.connector.connect()
3.     cursor.execute("SELECT * FROM train_data")
4.     return train_data, test_data
5.     models = [DecisionTree (), RandomForest (), MLP ()]# 定义模型列表
6.     train_data, test_data = load_data()#将数据分成特征和标签
7.     mse_list = []# 计算每个模型在测试集上的均方误差
8.     for model in models:
9.         mse = mean_squared_error(test_labels, preds)
10.        mse_list.append(mse)
11.    weight_list = []
12.    for mse in mse_list:# 计算三个 MSE 的倒数
13.        weight = 1 / mse
14.        total_weight = sum(weight)# 使用 softmax 函数处理三个 MSE 的倒数, 得到权重
15.    for i, model in enumerate(models):# 使用权重融合三个模型的预测值
16.        preds = model.predict(test_features)
17.        ensemble_preds += weight_list[i] * preds
18.    print(ensemble_preds)# 输出融合模型的预测值

```

基于以上实验和设计的融合模型后续可以依据其他机器学习模型的实验效果融入更多模型进一步提高成绩预测的准确率, 还可以通过更改数据集和模型参数来适用于不同的回归预测问题。对融合模型在测试集上的计算结果进行损失函数计算, 并将结果与上述多种单独机器学习回归模型进行对比, 实验结果显示, 融合模型的均方误差和决定系数为 13.7 和 0.86, 均优于决策树, 随机森林和 MLP 神经网络模型, 模型效果

和准确性较好，能够适用于本文所研究的学生个性化评价系统的成绩预测模块。

4.7 本章小结

本章对学生个性化评价算法设计与实现进行了详细说明。首先对实验数据和学生多维度特征构造进行了详细阐述；针对软件工程的数据特征，使用多种机器学习融合算法模型设计实现基于教学目标对学生的章节知识掌握水平进行自动分类，生成知识掌握度评价；使用基于学生数据的评分规则算法设计生成课程素养评价和学习风格评价并综合生成个性化评价报告。使用基于决策树、随机森林和 MLP 神经网络的融合算法模型设计预测学生成绩并生成成绩预测报告。

第五章 学生个性化评价系统的实现与应用

本章将对学生个性化评价系统的总体架构、个性化评价报告生成功能、学生成绩预测报告生成功能以及学生数据上传查询等其他功能进行了编码实现，将本文前文所设计的系统组件和评价算法模型集成到系统中。最后，对个性化评价系统的运行进行全面测试，并将系统部署到云服务器，用于《软件工程》课程实际教学使用并测试系统应用效果。

5.1 系统实现

5.1.1 系统编码设计

本节主体框架采用了主要采用了 SpringBoot、Mybatis、VUE 前端框架等工具对前端和服务端相关功能进行实现，使用 Java、Python、HTML5+JS、SQL 等语言对系统进行开发。在 Java 应用程序框架上选用了最新基于 Spring Framework 的 Spring Boot 框架，该框架和传统的 SSM 框架（Spring 和 Spring MVC 和 MyBatis）相比能够更大限度地提高开发效率，SSM 框架在开发过程中需要手动配置数据源、事务、缓存等较多细节，而 Spring Boot 框架能够实现自动配置，基于 Spring 快速创建应用程序。系统实现后的部署使用“云服务器 ECS”，实例规格选择 ecs.c5.xlarge（4 核 8G 内存）、操作系统选择 CentOS 7.4，随后通过配置安全组来限制访问虚拟机的 IP 地址和端口等信息。使用 FTP 工具将代码上传至虚拟机，最后购买域名并进行域名解析。

5.1.2 系统整体架构实现

该系统整体采用 B/S 架构，前端页面由 HTML5 和 Javascript 语言编写，使用基于 node.js 的包管理工具和基于 Vue 3 的组件；系统服务端程序主要使用 Java，Python 和 SQL 语言完成，开发工具主要使用 Eclipse 集成开发工具，数据库使用开源轻量级数据库 MySQL；同时使用 Python 开发完成算法模型实现，工具使用 PyCharm 集成开发环境，为便于学生数据展示以及学生评价效果更加直观，也使用了 Highcharts、DataTables 等可视化插件。

5.1.3 数据上传与查询功能实现

本章所设计的学生数据上传和查询功能在实现上基于所创建的 Spring Boot 控制器类，该类将处理来自前端的请求，并将数据传递给服务层类，在服务层中编写上传和查询方法实现上传和查询功能，并使用 Spring Data JPA 操作数据库。如图 5.1 学生数据

查询界面所示，在教师用户的数据查询界面上包含对全部学生数据的展示以及模糊查询功能。教师可浏览学生学习数据，在选中指定学生后可为该学生生成个性化评价报告或成绩预测报告。

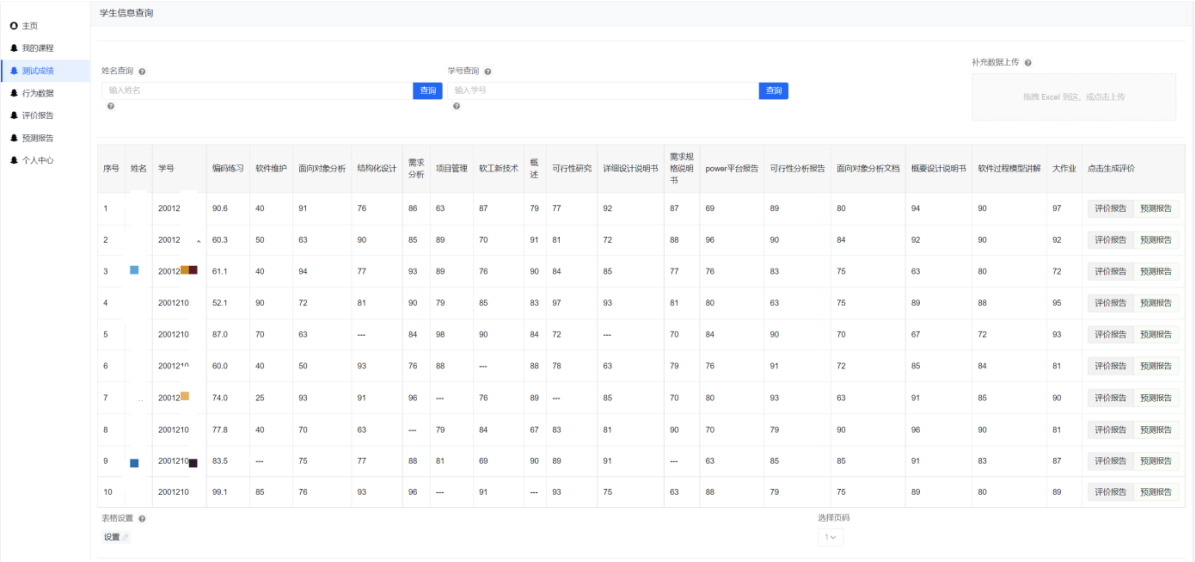


图 5.1 学生数据查询界面

5.1.4 个性化评价报告功能实现

如图 5.2 评价报告生成界面所示，在个性化评价功能界面将生成指定学生的个性化评价报告内容，同时将通过可视化图表的形式展示统计后的学生数据详情，教师可直接编辑修改报告内容，并在审核通过后发布，发布后学生看登录查看自己的个性化评价报告。



图 5.2 评价报告生成界面

5.1.5 学生成绩预测报告功能实现

如图 5.3 成绩预测报告界面所示，在学生模板预测界面将生成指定学生的成绩预测报告，同时将通过知识掌握度雷达图的形式展示学生的章节知识掌握情况，教师确认后可发布报告，发布后学生可登录查看。



图 5.3 成绩预测报告界面

5.2 系统测试

本节对个性化评价教学系统进行了测试。在功能性测试上主要检验数据图表，评价报告等内容的生成和显示是否符合设计目标；检验评价结果是否符合模型预测结果。在非功能性测试上主要检测对常见操作系统和浏览器的兼容性以及系统使用性能能否满足《软件工程》课程教学使用需要。如下表 5.1 系统测试环境，在完成测试环境搭建后针对本地和云服务器进行测试。

表 5.1 系统测试环境

配置	环境参数
内存	4G/8G/16G
操作系统	Window7/Window10/Window11/CentOS8.2
处理器	Intel XeonE5-2682 v4/Intel Corei5/Intel Corei7

5.2.1 系统功能性测试

如表 5.2 系统功能性测试信息, 在系统功能测试上主要对个性化评价功能, 成绩预测功能和系统其他子功能进行测试。学生个性化评价功能测试主要测试系统用户的操作权限, 业务操作流程上是否符合系统设计规范, 测试系统页面是否正常响应, 组件是否正常加载; 在个性化评价报告生成、成绩预测报告生成、学生信息上传查询、图表数据展示等功能是否完备, 生成报告文本结构和组成是否符合设计预期, 数据显示是否准确等。

表 5.2 系统功能性测试信息

测试内容	设计目标	实际测试结果
知识结构评价文本	包含计算得出的关键词, 基于评价模板和基于融合算法模型生成的评价语句通畅	章节知识结构评价准确, 与算法模型预测结果一致, 文本结构和流畅度符合预期
情感价值评价文本	评价语句通畅, 包含章节情感价值内容和学生学习态度或品德评价	与模型结果一致, 文本结构和流畅度符合预期
学习风格评价文本	评价语句通畅, 对四组学习风格均给予评价或判断	与模型预期结果一致, 文本结构和流畅度符合预期
成绩预测报告文本	评价语句通畅, 对学生最终成绩和各知识教学目标掌握度均做出预测	预测结果与模型一致, 内容完整, 文本结构和流畅度符合预期
评价报告编辑	教师能够对报告编辑	报告文本可直接在 web 端修改, 保存正常
知识结构雷达图	对章节知识掌握情况得分绘制雷达图	显示正常, 与模型计算一致
总体完成率饼状图	各章节线上课程完成度的分布情况	显示正常, 与数据库结果一致
章节报告柱形图	显示指定学生章节报告成绩分布情况	显示正常, 与数据库一致
测试成绩折线图	显示指定学生测试成绩的分布情况	显示正常, 与数据库一致
数据详情展示	显示全部或指定学生的全部数据	显示正常, 与数据库一致
学生数据上传	以 csv 文件导入数据库	上传成功, web 端展示正常
数据查询	能够依据学号或姓名查询学生数据	支持模糊查询, 数据显示正常

5.2.2 系统非功能性测试

针对个性化评价教学系统的实际应用场景需要, 其非功能性测试如表 5.3 系统非功能测试信息所示。本次测试旨在测试个性化评价教学系统的兼容性、响应时间和吞吐量等。测试环境的服务器配置为 ECS 服务器 4 核 8G, CentOS7.4, 数据库服务器为

MySQL8.0, 测试数据使用 2022 年下半年软件工程课程数据子集; 测试场景为多个并发用户同时访问系统。对运行测试结果预估要求平均响应时间小于 5 秒。测试的实际结果显示线上平均响应时间小于 4s, 最大 6.2 秒; 错误率为 0.3%。因此在本测试场景下, 系统兼容性良好, 响应时间和错误率都达到预期目标。

表 5.3 系统非功能测试信息

测试内容	场景	测试结果
兼容性测试	macOS Ventura13.0.1, Microsoft Edge	页面元素加载正常, 响应正常
	macOS Ventura13.0.1, Google Chrome	页面元素加载正常, 响应正常
	Microsoft Windows 10, Microsoft Edge	页面元素加载正常, 响应正常
	Microsoft Windows 10, Google Chrome	页面元素加载正常, 响应正常
	Microsoft Windows 11, Microsoft Edge	页面元素加载正常, 响应正常
	Microsoft Windows 11, Google Chrome	页面元素加载正常, 响应正常
	iOS 16, FireFox Focus	页面元素加载正常, 响应正常
性能测试	Android 12, Google Chrome	页面元素加载正常, 响应正常
		首页响应: 1500ms
	工具: WebPageTest; Jmeter4.0	数据查询页面响应: 1900ms
	服务器: 4 核 8G, CentOS 7.4 64 位	评价报告生成页面响应: 4300ms
	数据库服务器: MySQL 8.0	预测报告生成页面响应: 3600ms
压力测试		评价查阅页面响应: 2600ms
		CPU 使用率: 小于 64%
		内存使用率: 小于 78%
	工具: Apache JMeter	平均响应时长: 3900ms
		QPS: 167 ; 并发数: 650
		错误率: 0.3%

5.3 系统应用效果评估

面向软件工程教学的学生个性化评价系统能够为课程评价提供高效便捷的手段, 让教师能及时掌握学生的学习状态, 同时给予学生反馈, 从知识、素养、风格等多个角度为学生提供个性化的评价内容, 较好的提高了教学效率, 促进软件工程授课教师“因材施教”。老师在生成学生评价报告和预测报告后能够及时掌握学生的章节知识掌握程度、学生课程素养和学习风格等个性化特征, 了解学生的当前发展情况和潜在发展情况, 在保证基于数据的学生评价具有客观真实性的同时, 教师能够对报告进行修改

编辑,融入指导建议,通过有针对性的沟通指导并调整教学策略帮助学生步入学习“最近发展区”。学生通过查看教师发布的评价报告,能够受到勉励性评价的激励,提升自我实现的内在需求,同时了解自身在课程知识,专业能力、学习态度情感和学习风格等方面的特点或不足,通过加强自主学习完成学业提升。



图 5.4 系统应用效果评估方法

面向软件工程教学的学生个性化评价系统于 2022 年应用于某大学软件工程课程混合式教学,为课程授课教师和课程助教提供学生评价辅助功能。如图 5.4 系统应用效果评估方法,为验证该系统的教学评价效果,本文采用问卷调查和两学期章节成绩对比分析的形式对 2021 年(未使用评价系统)和 2022 年(使用评价系统)某学校软件工程课程的教学效果进行对比统计。

信度效度检测结果 2021

可靠性统计		
克隆巴赫 Alpha	基于标准化项的克隆巴赫 Alpha	项数
.748	.751	6

KMO 和巴特利特检验		
KMO 取样适切性量数。		.725
巴特利特球形度检验	近似卡方	230.272
	自由度	15
	显著性	.000

信度效度检测结果 2022

可靠性统计		
克隆巴赫 Alpha	基于标准化项的克隆巴赫 Alpha	项数
.802	.824	6

KMO 和巴特利特检验		
KMO 取样适切性量数。		.696
巴特利特球形度检验	近似卡方	94.655
	自由度	15
	显著性	.000

图 5.5 问卷信度效度检测结果

问卷在相关问题和量表设计上主要采用经典的李克特五分量表形式,在统计系统对教学效果的影响上,主要设计指标为学生的满意度统计。学生满意度调查能够翻译学生对课程教学的满意程度,而把学生自身成长收获感知与其客观情况反应进行联系更能反应教育教学活动的效果^[55];而问卷中学生满意度统计部分设计参考经典学生满意度评价量表,诺埃尔-列维茨 (Noel-Levitz) 公司发布的“美国高校学生满意度测量量表^[56],采用 6 个维度统计学生对课程的满意程度,分别是学生对课程的整体满意度、评价报告准确性、课程内容量、课程难易程度、教师助教个性化反馈状况和期末考试难度。在完成全部教学过程后对 2021 和 2022 年某大学软件工程结课学生发放问卷,如图 5.5 问卷信度效度检测结果所示,为保证问卷的准确有效,使用 IBM SPSS 工具对问卷进行信度和效度检验。

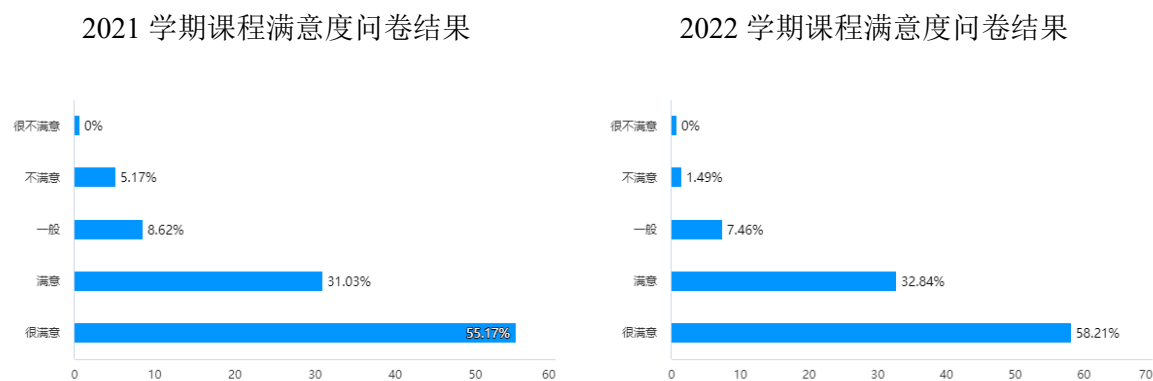


图 5.6 课程整体满意度对比

使用学生个性化评价系统后学生对课程整体满意度如图 5.6 课程整体满意度对比所示,两学期的满意度打分均值分别为 4.36 和 4.46,结果显示第二学期的课程整体满意度较高。

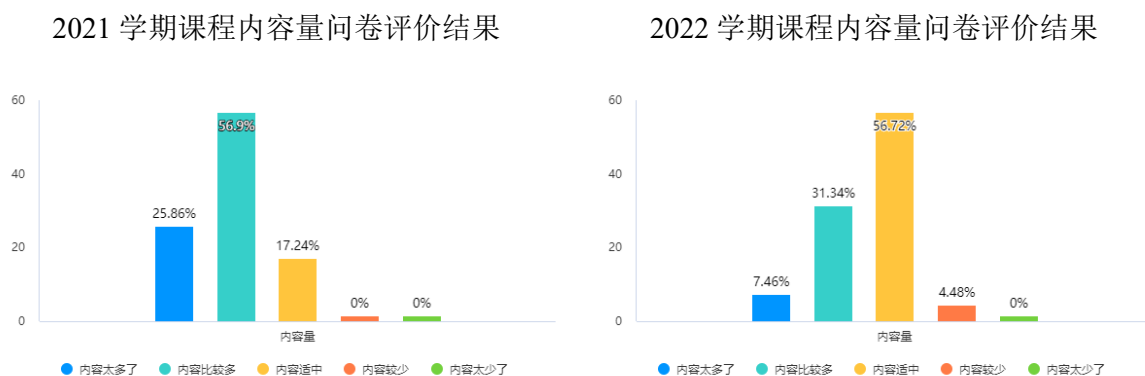


图 5.7 学生对课程内容量评价对比

学生个性化评价系统中包含学生对课程资料的下载和推荐内容的点击情况,包含

学生的课程知识整体完成率统计,结合评价系统的信息统计和评价报告情况,教师能对课程内容量不断进行调整,依据图 5.7 学生对课程内容量评价对比显示,2021 和 2022 年评分均值分别为 1.91 和 2.62,显然 2022 年的评分更接适中值,在课程内容量设置上更为合理。

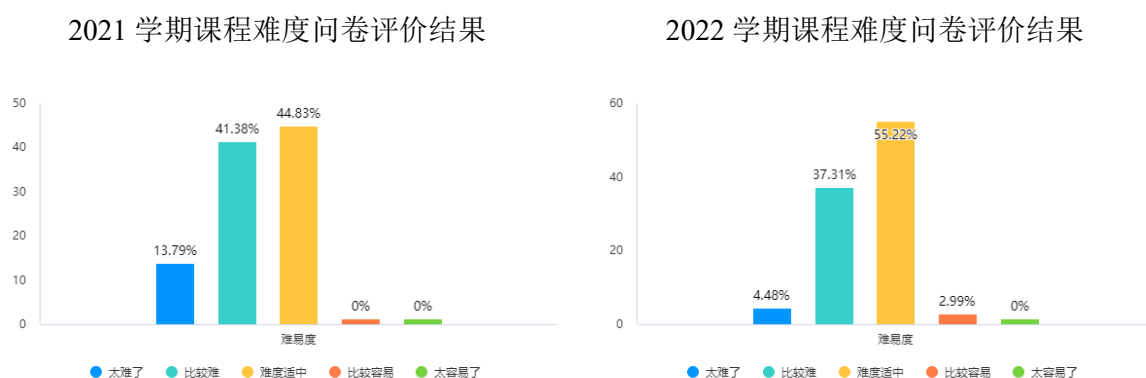


图 5.8 学生对课程难度评价对比

学生个性化评价系统中的学生知识结构评价为过程性评价,教师完成章节主要内容教学后可及时了解学生的章节知识掌握程度,对测试和教学任务难度进行调整,如图 5.8 学生对课程难度评价对比显示,第二学期学生对课程难度的评价在均值上更接近适中。

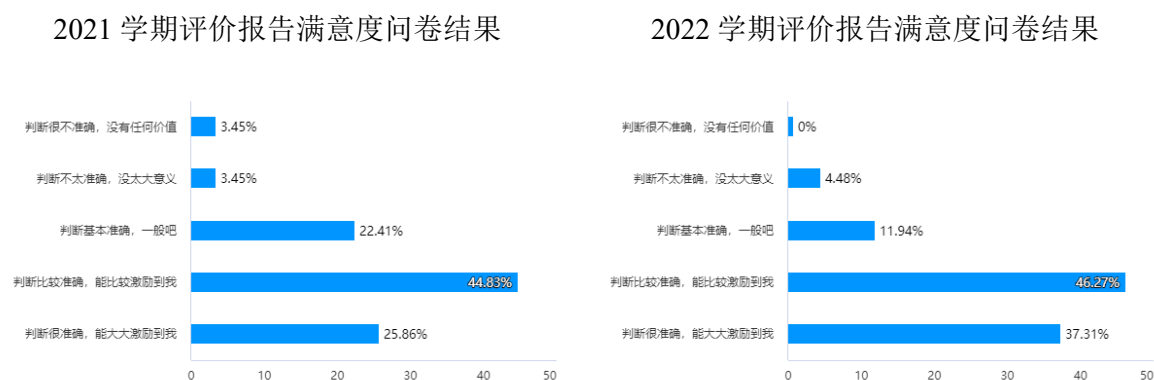
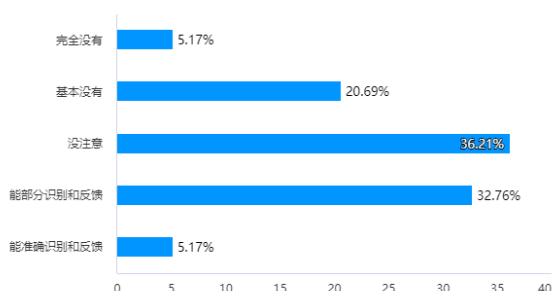


图 5.9 学生对个性化评价报告满意度对比

课程第一学期的评价报告生成由助教统计学生数据信息,教师为每位学生撰写评价后通过邮件发送,数据信息仅包含章节测试成绩,考勤和大作业成绩,评价内容主要为学生考试成绩和日常表现。第二学期的评价由学生个性化评价系统生成和发布。如图 5.9 学生对个性化评价报告满意度对比,两个学期的满意度均值分别为 3.86 和 4.17,结果显示第二学期的评价准确程度和激励效果提升,学生对个性化评价的满意度更高。

2021 学期学生对反馈满意度问卷结果



2022 学期学生对反馈满意度问卷结果

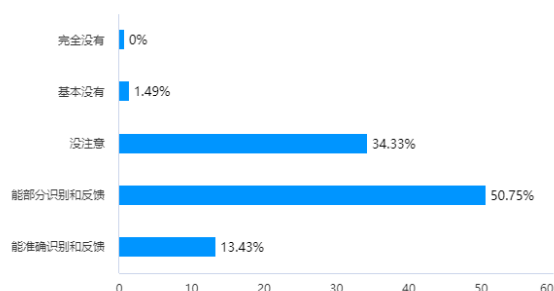
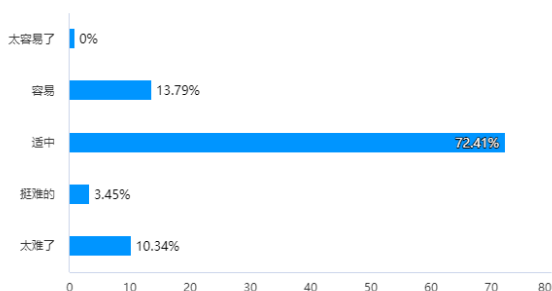


图 5.10 学生对教师和助教反馈满意度对比

个性化评价系统应用价值在于能够辅助教师及时给予学生个性化的学习评价反馈，教师可根据评价反馈调整教学策略，提升教学效率。如图 5.10 学生对教师和助教反馈满意度对比，两个学期的满意度均值分别为 3.12 和 3.78，结果显示第二学期中大部分学生都认为教师给予过评价反馈，且识别和反馈的准确性提高。

2021 学期期末试卷难度评价问卷结果



2022 学期期末试卷难度评价问卷结果

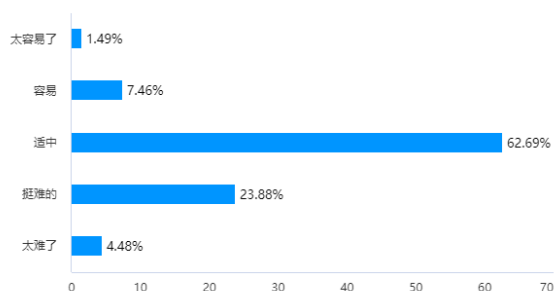
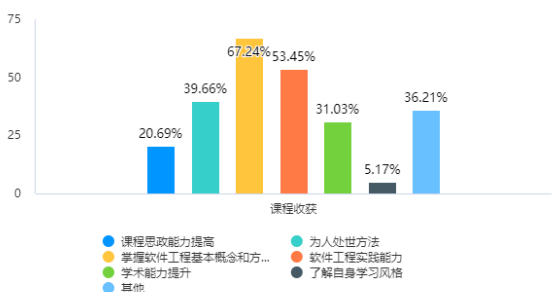


图 5.11 学生对期末试卷难度评价对比

2021 学期学生课程收获问卷结果



2022 学期学生课程收获问卷结果

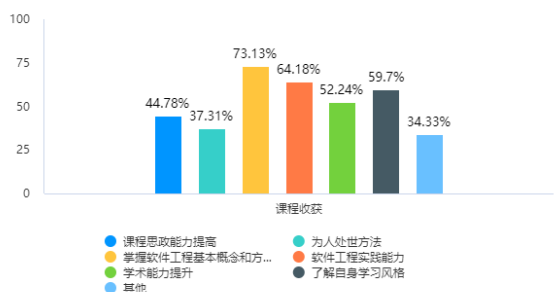


图 5.12 学生课程收获评价对比

教师通过学生个性化评价系统了解学生每个章节的掌握程度和分布情况，对期末

考试试卷章节内容比例和难度进行调整,如图 5.11 学生对期末试卷难度评价对比,两学期的评分方差分别为 0.765 和 0.724,第二学期学生对期末试卷难度评分方差小于第一学期,更符合正态分布结果,证明调整后的试卷难度更加合理。

学生个性化评价报告中加入了对学生学习风格和情感素养的评价元素,能够帮助学生评估自身学习风格,协助教师对学生的综合素养进行评估,进行针对性的思政教学或综合素质培养。从图 5.12 学生课程收获评价对比可以看出,第二学期学生在思政能力提升和自身学习风格收获上有大幅度提升。

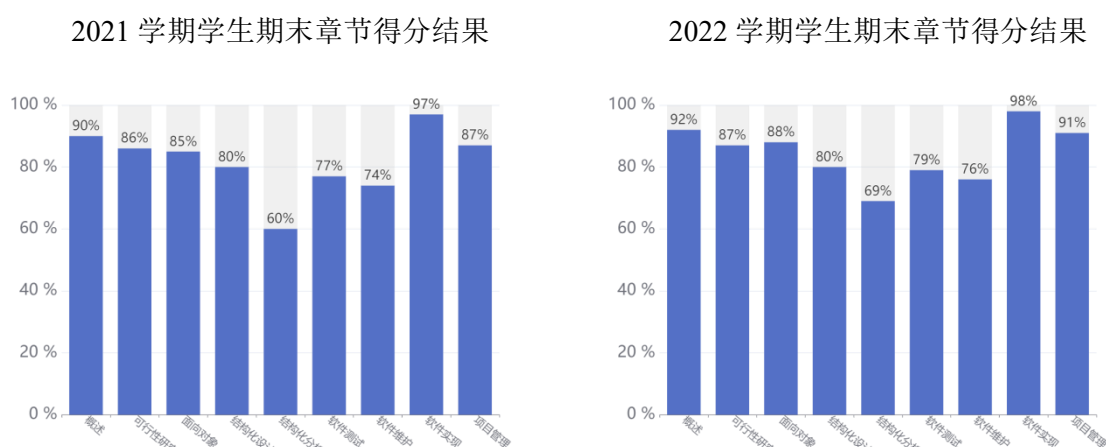


图 5.13 学生期末章节得分对比

如图 5.13 学生期末章节得分对比,学生个性化评价系统应用后学生在期末各章节内容平均得分均有不同程度上升,学生平均成绩水平高于上学期。

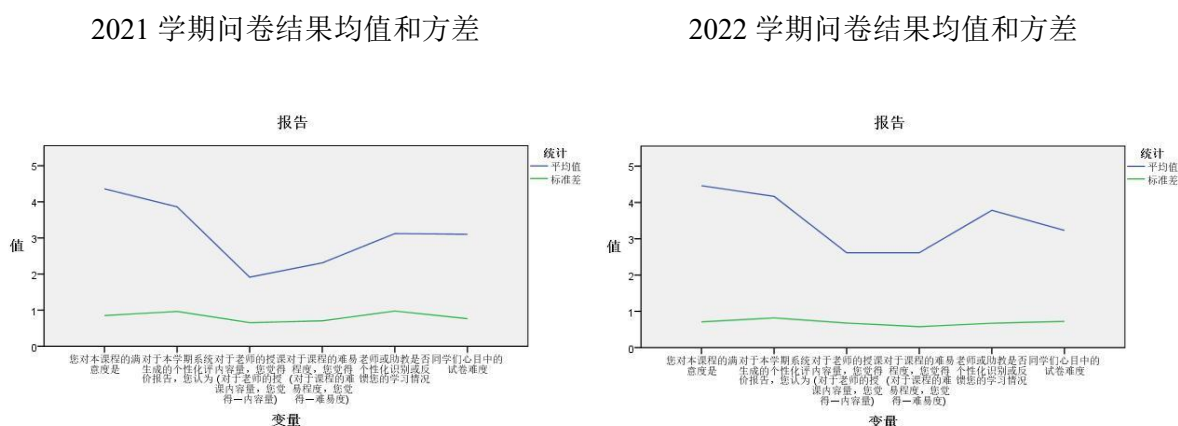


图 5.14 问卷结果均值和方差对比

如图 5.14 问卷结果均值和方差对比显示,在引入学生个性化评价系统后,对于六个维度的满意度问卷调查结果均有不同程度的提升。基于上述全部对比结果,证明该评价系统应用后提升了学生对软件工程课程的整体满意度,能够帮助教师依据学生评价及时调整教学内容和难度,能够给予学生及时准确的个性化反馈,促进成果导向的课程教学目标完成,在提升软件工程课程教学质量上具有一定的效果。

5.4 本章小结

本章编码实现了软件工程课程的学生个性化评价系统，搭建了完整的测试环境对系统各模块和子功能进行了测试，之后对系统性能等非功能性指标进行了测试。结果显示评价报告和预测报告生成、发布等系统功能均正常，符合设计预期；系统性能符合要求。最后系统应用前后效果进行了对比评估，证明评价具有较高的个性化程度和准确性，系统在提升教学质量上具有一定效果。

第六章 总结与展望

6.1 总结

本文通过对某大学软件工程课程进行研究实验,针对如何挖掘学生各个学习阶段和场景下的数据并给予学生个性化评价进行研究。本文在第一章总结了学生评价领域的研究状况,提出结合最近发展区理论和成果导向教育理论提出设计和实现一款学生个性化评价系统,并使用该系统为混合式教学模式下的《软件工程》课程提供个性化评价报告生成和学生成绩预测等服务。

本文主要贡献如下:

(1) 本文基于 KNN、决策树、随机森林、SVM、Xgboost、朴素贝叶斯、Bagging 多种机器学习算法的融合模型对学生知识掌握度进行实验分析,实现对学生章节知识掌握度的较准确分类和判断。在评价设计上引入 OBE、成果导向教育等前沿理论对不同发展水平的学生从多个维度进行基于学生数据的针对性评价并综合生成个性化评价报告,增加了学生评价的个性化程度。

(2) 在学生成绩预测上,本文在对比实验后使用基于决策树、随机森林和前馈神经网络的集成融合预测模型实现对学生课程成绩的准确预测,并综合生成学生成绩预测评价报告。

(3) 在上述研究的基础上,本文使用 Springboot 框架编码实现学生个性化评价系统并应用于 2022《软件工程》教学课程项目,并设计问卷量表调查应用后的课程教学效果,问卷结果和成绩统计结果显示系统应用后,整体学生章节平均成绩和课堂满意度均有所提升,且教师能够根据批量评价结果调整教学内容多少以及教学难度,提升教学质量;学生在课程教学过程中得到了较为准确和及时的反馈,对软件工程课程的教学质量提升起到了良好的效果。

6.2 展望

6.2.1 系统应用前景展望

从应用前景上,本文所探究的是面向某高校软件工程课程的学生个性化评价,在评价维度和标准设计上参照软件工程课程教学模式的需要,在未来的学生评价应用上,可以拓宽到其他学科或其他平台的教学过程,形成通用型的学生评价系统。例如该学生个性化评价系统将在下学期应用实验于某高校项目管理和软件工程技术专题课程以继续探究和改进评价效果;对于提高学生个性化评价系统在其他领域的通用性,有以下改进和展望:

（1）更加智能化的评价生成方法

学生个性化评价系统在评价文本的生成上可以借助文本生成模型实现，例如在系统中使用循环神经网络（RNN）或变换器（Transformer）等模型，前提是能够积累大量人工标注的评价数据集，使每个样本包含一个评价文本和对应的学生特征，基于文本生成模型的评价则更依赖人工审核来确保评价的适用性。

（2）提升系统的可扩展性和可定制性

本文所研究的学生个性化评价系统依据软件工程课程知识教学目标和人才培养要求设计，在未来需要可扩展性和可定制性，以适应不同高校、老师和学科的需求。系统可以支持自定义数据字段、通过输入课程或学科的教学目标要求重新定义对学生知识学习的评价；系统在评价维度或指标设计上支持自定义，使系统可以根据学科需要对学生进行多维度的评价；系统基于规则或模板的评价报告部分，可以根据教师的教学风格和语言习惯自行设置和调整。

（3）探求共性的评价数据模型

本文研究的学生评价系统在考虑软件工程课程数据特征的前提下采用了混合数据评价模型。为提升模型通用性，需要学生评价系统能够采集和分析多学科不同类型的数据，探索具有共性的评价数据模型。

6.2.2 学生评价研究展望

学习评价研究上更趋向于个性化和因材施教，讲求对学生更具体的，针对性的评价，以更加全面客观地了解学生的学习情况，个性化程度更高，能够为教师教学提供针对性的参考依据。基于学生数据的混合式教学模式下的学生个性化评价研究有以下改进和展望：

（1）高度数据驱动以及深入借助人工智能技术

采用更加多元化的数据来源，在课程教学数据的基础上融入学生自我评价、学生活动、音视频、甚至是社交媒体数据等，以更加全面的评价学生的学习情况。在此基础上，通过自然语言处理、机器学习等技术，可以更加全面的获取学生的实际表现，可以实现将学生学情分析领域融入学生个性化评价，例如实时语言识别分析、课堂微表情识别、学生文本识别等。

（2）评价融合更多教育新理念

注重使用评价来引导学生自主发展，激发学习动机，紧扣成果导向教学，学生元认知诊断等新的教育理念，注重学生自我认知和自我管理能力的培养。同时，要求学生更加主动地参与评价过程，通过自我评价、目标设定和反思等方式，实现个性化发展。

参考文献

- [1] 许华琼,胡中锋.形成性评价及其反馈策略[J].教育测量与评价(理论版),2010,34(1):23-26.
- [2] 张立山,冯硕,李亭亭.面向课堂教学评价的形式化建模与智能计算[J].现代远程教育研究,2021,33(1):13-25.
- [3] 马志强,王雪娇,龙琴琴.基于同侪互评的在线学习评价研究综述[J].远程教育杂志,2014,32(4):86-92.
- [4] 翟苗,张睿,刘恒彪.高校混合式教学形成性评价指标研究[J].现代教育技术,2020,30(9):35-41.
- [5] 张晓刚.面向程序开发类课程的混合教学形成性评价体系构建[J].计算机教育,2023,No.339(03):212-215+220.
- [6] 曹国刚,蓝星,李文举等.形成性评价在软件工程专业课程中的实践研究[J].计算机教育,2017,270(6):40-44.
- [7] 刘雅.形成性评价与终结性评价联合应用实践与思考——以《计算机基础》课程为例[J].高教学刊,2019,119(23):70-72.
- [8] Zhang Yi et al. Classroom Quantitative Evaluation: A Method of Both Formative and Summative Evaluation[J]. Sustainability,2023,15(3):1783-1783.
- [9] Suaco P. Thea and Mangaliag D. Antonio and Gadgad M. Marilyn. Collaborative Summative Assessment: Means for Enduring Learning and Attainment of 21st Century Skills in the Online Platform[J]. Journal of Education and Learning, 2023, 12(1):118.
- [10] 牟智佳,李雨婷,彭晓玲.基于学习测评数据的个性化评价建模与工具设计研究[J].电化教育研究,2019,40(8):96-104+113.
- [11] 贾积有,乐惠骁,张誉月等.基于大数据挖掘的智能评测和辅导系统设计[J].中国电化教育,2023,43(3):112-119.
- [12] 刘进,高胜寒,林松月.从标准化到数据化:基于大数据的大学生学习评价研究方法论转向[J].现代教育技术,2022,32(9):31-37.
- [13] 马辛巳,于明鹤,王虹,张天成.基于学生画像的学生评价系统设计与实现[J].软件导刊,2023,22(1):20-28.
- [14] 周楠,周建设.基于深度学习的学生行为分析与教学效果评价[J].现代教育技术,2021,31(8):102-111.
- [15] Arbaugh J B . What Might Online Delivery Teach Us About Blended Management Education? Prior Perspectives and Future Directions[J]. Journal of Management Education,2014,38(6):784-817.
- [16] 张志红.学生参与式课堂评价:现实问题、理念引导及实施策略[J].当代教育科学,2021,(10):51-57.
- [17] 王楠,王淇.基于深度学习的学生课堂专注度测评方法[J].数据分析与知识发现,2022:1-16.
- [18] 魏勃颀,邹宏,左有游.基于 HMM 模型的网络课程学习过程多维评价研究[J].软件导刊,2020,19(5):247-251.
- [19] 郑勤华,孙洪涛,陈耀华,刘春萱.基于学习分析的在线学习测评建模与应用——综合建模方法研究[J].电化教育研究,2016,37(12):40-45.

- [20] Wu M,Zhao H,Yan X, et al. Student Achievement Analysis and Prediction Based on the Whole Learning Process[C]// 2020 15th International Conference on Computer Science & Education (ICCSE).IEEE,2020.
- [21] 牟智佳,武法提.教育大数据背景下学习结果预测研究的内容解析与设计取向[J].中国电化教育,2017(7):26-32.
- [22] 张玉欣.混合式教学模式下基于概率神经网络的学生学习评价模型研究[J].中国管理信息化,2023,26(1):217-220.
- [23] Yanfei Shen. The Construction of College English Blended Learning Assessment Model under the Guidance of CSE[J]. Frontiers in Educational Research,2022,5(13):80-84.
- [24] Li Deming et al. Evaluation of College Students' Classroom Learning Effect Based on the Neural Network Algorithm[J]. Mobile Information Systems,2022,1:1-8.
- [25] Na Liu. An evaluation model of distance learning effect based on MOOC theory[J]. International Journal of Continuing Engineering Education and Life-Long Learning,2022,32(2):129-142.
- [26] Christos Troussas and Konstantina Chrysafiadi and Maria Virvou. Personalized tutoring through a stereotype student model incorporating a hybrid learning style instrument[J]. Education and Information Technologies,2021,26:2295-2307.
- [27] Lwande Charles and Muchemi Lawrence and Oboko Robert. Identifying learning styles and cognitive traits in a learning management system[J]. Heliyon,2021,7(8):2-7.
- [28] Baherimoghadam Tahereh et al. The effect of learning style and general self-efficacy on satisfaction of e-Learning in dental students[J]. BMC Medical Education,2021,21(1):463.
- [29] 岳俊芳,孙道金.远程学习者二维满意度评价量表编制及其应用研究——以“网上人大”为例[J].中国电化教育,2016,(8):53-60+73.
- [30] 柴唤友,陈丽,郑勤华,王辞晓.学生综合评价研究新趋向:从综合素质、核心素养到综合素养[J].中国电化教育,2022,(3):36-43.
- [31] 张治,刘小龙,徐冰冰,陈雅云,吴永和.基于数字画像的综合素质评价:框架、指标、模型与应用[J].中国电化教育,2021,(8):25-33+41.
- [32] 王怀波,柴唤友,郭利明,刘紫荆,郑勤华.智能技术赋能学生综合素养评价:框架设计与实施路径[J].中国电化教育,2022,(8):16-23.
- [33] Alsariera Yazan A et al. Assessment and Evaluation of Different Machine Learning Algorithms for Predicting Student Performance.[J]. Computational intelligence and neuroscience,2022,1:4151487-4151487.
- [34] 陈曦,梅广,张金金,许维胜.融合知识图谱和协同过滤的学生成绩预测方法[J].计算机应用,2020,40(2):595-601.
- [35] 罗杨洋,韩锡斌.基于增量学习算法的混合课程学生成绩预测模型研究[J].电化教育研究,2021,42(7):83-90.
- [36] Mengash H A. Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems[J]. IEEE Access,2020,8:55462-55470.
- [37] Qingshan Deng and Zhili Wu. Students' Attention Assessment in eLearning based on Machine Learning[J]. IOP Conference Series: Earth and Environmental Science, 2018,199(3):032042-032043.

- [38] Zafari Mostafa et al. A Practical Model for the Evaluation of High School Student Performance Based on Machine Learning[J]. Applied Sciences, 2021,11(23):11534-11535.
- [39] VYGOTSKY L.Interaction between learning and development[J].Readinas on the Development of Children.1978,23(3):34-41.
- [40] 刘宁,余胜泉.基于最近发展区的精准教学研究[J].电化教育研究,2020,41(7):77-85.
- [41] 刘强.基于 OBE 理念的“软件工程”课程重塑[J].中国大学教学,2018,(10):25-31.
- [42] Yossy E H,Heryadi Y,Lukas. Comparison of Data Mining Classification Algorithms for Student Performance[C]// 2019 IEEE International Conference on Engineering, Technology and Education (TALE).2019:1-3.
- [43] Supianto A A,Dwitama A J,Hafis M. Decision Tree Usage for Student Graduation Classification: A Comparative Case Study in Faculty of Computer Science Brawijaya University[C]// 2018 International Conference on Sustainable Information Engineering and Technology (SIET). 2018:308-310.
- [44] 娄增辉,王红林,孙彩云.教育大数据背景下的学业发展分析模型仿真[J].计算机仿真,2022,39(12):306-310.
- [45] Bujang S,Selamat A,Ibrahim R,et al. Multiclass Prediction Model for Student Grade Prediction Using Machine Learning[J]. IEEE Access,2021,9:95608-95618.
- [46] Q. Jingliang, "A Grade Prediction Method Based on Long Short-Term Memory Network in Smart Campus," 2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China,2022:345-350.
- [47] 教育部高等学校教学指导委员会.普通高等学校本科专业类教学质量国家标准(上册)[M].高等教育出版社,2018:325-327.
- [48] 中华人民共和国教育部.教育部关于印发《高等学校课程思政建设指导纲要》的通知[Z].教高(2020)3号.
- [49] 核心素养研制组.《中国学生发展核心素养》[EB/OL].[2016-09-14]<http://edu.people.com.cn/n1/2016/0914/c1053-28714231.html>.
- [50] 倪晓莉主编,社会心理学[M],西安交通大学出版社,2007.09,第193页.
- [51] 苏贵.正向勉励:教师激励学生的关系理性与路径探讨[J].当代教育论坛,2022,312(06):86-94.
- [52] 张爱华.个性教育的发现激励原则[J].中国教育学刊,2001(04):28-30.
- [53] Barbara A.Soloman,Richard M.Felder.ndex of Leaing Styles Questionnaire [EB/OL].[1999-01]<http://wwwengr.nc-su.edu/learningstyles/ilsweb.html>.
- [54] ElHaddioui Ismail and Khaldi Mohamed. Learning Style and Behavior Analysis: A Study on the Learning Management System Manhali[J]. International Journal of Computer Applications,2012,56(4):9-15.
- [55] 黄雨恒,郭菲,史静寰.大学生满意度调查能告诉我们什么[J].北京大学教育评论,2016,14(04):139-154+189.
- [56] Laurie A.Schreiner,Stephanie L...Juilerat,Samples[EB/OL].[2018-12-11]<https://www.rufalon.com/complete-enrolmentmanagement/student-success/student-satisfaction-assessment/student-satisfaction-inventory/samples>.