

附件：

项目编号	GGJG2020B015		思政专项项目	是 ☐ 否 ☒
项目类型	重点项目 ☐	一般项目 ☒	自选项目 ☐	

广东工程职业技术学院

教育

教

学改革与研究

项目研究报告

项 目

称

： 贝叶斯模型在课程过程性考核与结果性考核有机结合中的研究与实践——以《软件测试技术》为例

项 目 负 责 人：

黄轶文

项 目 组 成 员：

朱珍、卢世军、刘越涛、陈华荣、苏进廷

联 系 电 话：

18011982151

填 表 日 期：

2023 年 4 月 6 日

目录

一、项目名称

贝叶斯模型在课程过程性考核与结果性考核有机结合中的研究与实践——以《软件测试技术》为例。

二、研究的问题

随着大数据技术的发展以及教学过程网络化、数字化，为“过程性考核与结果性考

核”有机结合研究提供了数据基础，在国内也有一些学者进行了基于贝叶斯网络的“过程性考核与结果性考核”有机结合初步研究，例如郭富强设计了一个包括学生知识水平、学习偏差程度、学生认知能力、学习情绪的智能教学系统学生模型，并通过模型来调整教学行为。杨卉等提出了一个两层动态学生模型，通过跟踪和综合评价相结合的方法获取学习者的综合特征，并采用广义模糊综合评判方法对学习者的二级特征进行评价。闫汉原，申麟等人提出了基于“态度”的知识追踪模型，在标准贝叶斯知识追踪模型的基础上，根据学生答题的持续时间，新增了态度节点，从而提高模型的预测准确率，通过使用智能在线教育系统的数据集，证明了预测准确度的提升。

贝叶斯知识跟踪模型有机结合“过程性考核与结果性考核”，具有一定的普适性，可以广泛用于各种以学习为本质的教育活动中，但是也正是由于这种普适性，无法适配具体的教学实践，造成比较大的误差，在实际中难以取得良好的效果。基于这种情况，本研究基于高等职业院校《软件测试技术》课程研究，以期得到具有推广价值的经验。

三、研究的背景

习近平总书记在全国教育大会上明确提出，“要扭转不科学的教育评价导向”。陈宝生部长在新时代全国高等学校本科教育工作会议上的讲话也提出“要以学生为中心办教育，以学生的学习结果为中心评价教育，以学生学到了什么、学会了什么评判教育的成效”。教育部发布的《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》，要求严把考试和毕业出口关，完善过程性考核与结果性考核有机结合的学业考评制度。因此，将过程性考核与结果性考核有机结合，边考边学、融考于学、以考促学，改变传统单一的考试模式，把考试真正作为学习的一种手段，建立和完善立体的、全面的、科学的学业考评体系是一项迫在眉睫的教学改革任务。

过程性考核与结果性考核是辩证统一的关系，两者既有区别也有联系。过程性考核侧重在教学过程中，采用分阶段、多形式的考核，例如课外作业、上机操作、面试、单元测验等等方式。结果性考核侧重在一个学习周期，用考试方式确定这门课程的最终得分，并作为是否授予学分的依据。

在具体的教学实践中如何将过程性考核与结果性考核有机结合，国家有关部门没有具体的标准化工作指引，教师们往往要结合教学理念、经验、课程特点、学生素质等各种因素综合考虑，在实际教学中采用不同的方法。同一门课程，不同授课教师的工作经验、责任心可能会对教学质量产生很大的影响。因此，研究和实践“过程性考核与结果性考核”有机结合的普遍性技术和工具，能有力地促进《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》的落地执行。

四、研究的目的

在智能教学系统(Intelligent Teaching System)中，由 Corbett 和 Anderson 提出的基于贝叶斯网络的学生知识点追踪模型 BKT (Bayesian knowledge tracing)，是“过

程性考核与结果性考核”有机结合的普遍性技术和工具。一方面：BKT 模型在教学中与过程性考核结合，能清晰地反映整个课程实施过程中学生的知识状态，教师可以根据学生的知识状态动态调整教学重点、时长、方式和节奏，能极大提高教与学的效率。另一方面：BKT 模型在课前就能预测学生的成绩，与结果性考核成绩进行全面的立体的分析，能指出教学过程中的薄弱项，为下一学期的教学改进提供参考建议，通过经验积累与教学优化迭代，不断提高教学质量。

五、研究的基础

本研究有一定的前期研究工作基础，得到专业教研室的高度重视和大力支持，本课题申请人有多年的从教经验，有坚实的软件、硬件支撑，具有可行性。

前期研究工作的支撑：团队在信息理论领域有深厚积累，在学生学习过程评估方面也有着丰富的经验。学校的支撑：学校领导高度重视和大力支持，同时，本课题申请人作为软件测试专业负责人，参与多项教科研项目，负责实训室建设、学术、科研工作等，可以保障课题的顺利开展。软件的支撑：拥有常用的软件开发平台，可以满足本课题研究的需要。硬件的支撑：目前专业建有省级校内实训基地、信息技术类省级公共实训中心，智能移动应用设备及软件测试实训、软件及大数据服务外包工作室、软件技术专业协同创新创业实训基地及企业工作站、软件技术专业软件仿真训练营等软硬件研究条件。

本课题邀请了两位资深的企业专家参与，为课题的研究和实践提供强有力的支持和协助，为课题的顺利展开提供了保证。

六、研究的方法与过程

按照“调查筛选、课题论证、制定方案、研究事件、交流总结、申请结题”的程序进行。通过文献研究、专家访谈对 BKT 跟踪考核进行充分学习，组织课题组内部全体成员对课题的可行性、现实意义、研究价值等充分讨论和评估，共同探讨工作方案，落实人员分工、时间安排、阶段性成果、里程碑等重要事项。

本课题的研究方法主要有文献研究法、实证研究法以及访谈法。文献研究法：搜集相关文献资料，深入了解知识追踪领域相关的理论基础以及研究现状。实证研究法：在现有的教学环境中，使用 BKT 跟踪考核模型对学生进行测试。访谈法：组织与教师的交流会议，根据反馈意见，做进一步的迭代优化。

第一阶段：2020 年 11—2021 年 3 月 准备阶段。收集国内外 BKT 跟踪考核模型研究成果，访谈国内有关领域的专家学者，听取他们的观点和建议，总结 BKT 跟踪考核模型优点和应用情况。

第二阶段：2021 年 4—7 月 知识点分解、建模。整体思路：按照教学大纲和教材，对软件测试课程进行知识点的分解，并征求其他院校教师的意见，汇集后进行讨论。

第三阶段：2021 年 8—12 月 测试阶段。根据知识点设计作业和测试，多轮作业和测试只给出题目的对错，而不直接给出分数。根据 BKT 跟踪考核模型计算学生知识状

态、学时预测与成绩预测，并将数据存入相应的数据库中。

第四阶段：2022 年 1—6 月 验证阶段。得出学生的知识掌握水平，最终生成个性化的诊断报告和直观的图表表示。在此阶段进行总结，组织研讨会进行广泛的讨论。

第五阶段：2022 年 7—9 月 成果发布阶段。将研究成果以论文方式凝固。

七、研究的创新

本研究在《软件测试技术》课程中引入贝叶斯知识跟踪模型，有机结合“过程性考核与结果性考核”，彻底改变原有的“结果性考核”为中心的模式。课前、课中、课后三个阶段基本流程都围绕知识点展开，“过程性考核与结果性考核”穿插整个教学过程。根据知识点设计教学内容、作业及过程性考核，动态跟踪学生的知识点掌握程度，及时调整教学重点、时长、方式和节奏。比对学生的预测成绩与结果性考核成绩，进行全面的立体的分析，输出教学过程中

的薄弱项分析报告以及下一学期的教学改进建议。

本研究把教学过程分解为一个渐进的过程。基本流程围绕知识点展开，包括教师设计知识点，根据知识点设计教学内容、作业及测试。学生自主学习教学内容后，进行作业或测试，保存结果数据。BKT 计算器通过数据采集，参照知识点对学生知识状态进行更新，同时也可以进行学时预测和成绩预测，最后将结果反馈给教师和学生。

贝叶斯知识跟踪模式主要通过知识状态的更新和学生成绩的预测两方面应用，具有一定的普适性，可以广泛用于各种以学习为本质的教育活动中，但是也正是由于这种普适性，无法适配具体的教学实践，造成比较大的误差，在实际中难以取得良好的效果。基于这种情况，本研究基于高等职业院校《软件测试技术》课程研究，以期得到具有推广价值的经验。

本研究是国内首次利用贝叶斯知识跟踪模式有机结合“过程性考核与结果性考核”，应用到具体专业课程的研究与实践。

八、研究的成果

本研究是国内首次利用贝叶斯知识跟踪模式有机结合“过程性考核与结果性考核”，应用到具体专业课程的研究与实践。成果具体内容包括：

1、发明专利：一种基于记忆网络模型的学生学习过程评估的方法和系统

本发明专利公开了一种基于记忆网络模型的学生学习过程评估方法及系统，涉及教育监控技术，针对现有技术中技术点关联性不足已经对学习情况描述不准确的问题提出本方案。主要通过利用静态的键矩阵以及动态的值矩阵分别进行知识点基本数据和掌握程度的数字化表示。挖掘不同知识点之间的关系，并直接评估学生对每个知识点的掌握程度。优点在于，能自动发现知识点之间的联系，并能描述学生不断变化的知识状态，更加符合人类知识追踪的实际情况。

2、论文：《基于记忆网络模型的学习过程评估方法研究》

利用新型记忆网络模型 (Mode of Memory Network, MMN), 挖掘知识点之间的关系, 评估学生对知识点的掌握程度, 使用键静态矩阵存储知识点基本数据, 值动态矩阵存储和更新相应知识点的掌握程度, 并评估学生不断变化的知识状态, 具有实时性和客观性. 为教师统筹教学资源, 优化教学计划提供客观量化的依据.

3、软件著作权：“i 班学习管理平台”

软件功能包括：

利用 i 班学习管理平台对整个班级体制进行科学管理，对学生的个人信息各项数据的收集，对老师发布通知、作业、接龙更加规范、科学，更好地提升教师班级管理的效果和质量。

4、在线开放课程：《软件测试技术》

在“软件测试技术”在线开放课程中，本项目成果也有独特的应用：

学习过程评估方法在软件测试课程中的应用，帮助教师和学生了解学生在课程中的学习进展和成果，从而进一步改进和优化教学策略和方法。

作业：学生在课程中需要完成一定数量和难度的作业，包括编写代码、编写测试用例、分析测试结果等。教师可以通过批改作业来了解学生的学习情况，发现学生的薄弱环节并及时给予辅导。

测试：课程中需要进行多次测试，包括编写测试用例、执行测试用例、分析测试结果等。教师可以通过测试结果分析来了解学生对知识点的掌握情况，并发现学生存在的问题和不足。

小组讨论：课程中经常会组织小组讨论，让学生分享自己的实践经验和心得体会，并从其他同学那里学习到新的知识和技能。这有助于学生相互促进，提高自己的表达能力和团队合作能力。

项目作品：在课程结束时，教师通常会要求学生提交一些项目作品，包括编写代码、设计测试用例、制作测试工具等。这可以帮助学生将所学知识应用于实践，并锻炼团队合作和创新能力。

课堂表现：教师在课堂上会观察学生的表现，包括听课态度、回答问题的能力、与同学的交流等。这可以帮助教师了解学生的学习状态和困难，及时调整教学节奏和方法。

通过以上评估方法的应用，教师可以全面了解学生在课程中的学习情况和成果，从而有针对性地制定教学计划和改进教学方法，提高学生的学习效果和综合素质。

九、成果的应用

本次应用以软件测试知识域为例。2020 年 9 月份至 2021 年 1 月份，以软件测试专业 180 名同学为样本，应用 MMN 模型进行动态学习过程评估。将软件测试的知识点进行分类，以集合的方式划分为单元测试、集成测试、系统测试和验收测试四个知识域^[14]，如下表：

表 1 软件测试知识域

Tab.1 Software testing knowledge domain

知识域	描述
单元测试 C1	单元测试(模块测试)是编写一小段代码,用于检验被测代码的一个很小的、很明确的功能是否正确。一个单元测试是用于判断某个特定条件(或者场景)下某个特定函数的行为。执行单元测试,就是为了证明这段代码的行为和期望的一致
集成测试 C2	集成测试是单元测试的逻辑扩展。最简单的形式是:两个已经测试过的单元组合成一个组件,并且测试它们之间的接口。组件是指多个单元的集成聚合。单元组合成组件,组件又聚合成程序的更大部分。方法是测试片段的组合,并最终扩展进程,将模块与其他组的模块一起测试。最后,将构成进程的所有模块一起测试。
系统测试 C3	系统测试是将经过测试的子系统装配成一个完整系统来测试。检验系统是否确实能提供系统方案说明书中指定功能的有效方法。目的是对最终软件系统进行全面的测试,确保最终软件系统满足产品需求并且遵循系统设计。
验收测试 C4	验收测试是部署软件之前的最后一个测试操作。验收测试的目的是确保软件准备就绪,并且可以让最终用户将其用于执行软件的既定功能和任务。验收测试是向经集成测试后,按照设计把所有的模块组装成一个完整的软件系统,接口错误基本排除后,进一步验证软件的有效性,即软件的功能和性能与用户的期望符合度。

每月进行一次评估和数据收集, T1-T5 分别是 9 月 11 日、10 月 8 日、11 月 2 日、12 月 3 日和 1 月 4 日, 每次评估由一百道选择题组成, 题目涵盖了单元测试、集成测试、系统测试和验收测试四个知识域。为了便于及计算机处理, 答案用二进制 0 和 1 表示, 0 表示错误, 1 表示正确。同一个题目中可以包含单元测试、集成测试、系统测试和验收测试中的一个或多个知识点, 传统的考卷方式只能得到笼统的分数, 从分数高低可以判断学生的整体掌握情况, 但是无法分析其中某个知识点的掌握分布情况, 因此我们尝试使用 MMN 模型来对学生知识水平进行追踪和判断。

评估共 180 名学生, 使用 MMN 模型跟踪学生知识域的掌握程度。数据结果分为三个等级, P 在 0.5 以下的表示较差, 0.5-0.8 表示良好, 0.8 以上表示优秀。下表的数据为全部学生的中位数, 另外每个学生还建立单独数据表, 从学生的整体知识掌握情况、群体差异和个体差异等方面来分析和应用评估结果。

表 2 MMN 模型跟踪学生知识域掌握程度中位数

Tab.2 The model tracks the median of students' mastery of knowledge domain

C1	C2	C3	C4
----	----	----	----

T1	0.82	0.70	0.80	0.60
T2	0.83	0.69	0.87	0.50
T3	0.80	0.75	0.86	0.51
T4	0.90	0.77	0.80	0.70
T5	0.91	0.78	0.87	0.75

注：Cn 知识域，Tn 评估时间点

从整体来看，单元测试和系统测试这两个知识域，学生掌握程度良好，授课教师按照教学计划进行，对课时和课程内容没有做过多调整。成绩预测和实际考核结果呈收敛趋势。集成测试和验收测试这两个知识域，学生觉得最难，掌握程度一般。根据跟踪结果，授课教师具体分析了原因：集成测试需要编写测试脚本，要具备一定编程能力；验收测试需要结合实际项目，根据不用的客户要求制定测试方案，而学生缺乏项目实操经验，难以理解和灵活应对不同的项目要求。授课教师根据跟踪分析结果，及时采取了有关措施：与C语言课程的教师协商，调整授课计划，让学生及时具备初步的编程能力，能够编写基本的测试脚本；与合作企业联系，带领学生到真实的项目验收现场，听取客户的意见，协助部分工作，对项目验收流程有初步的认知。经过几个月的调整，集成测试和验收测试这两个知识域的掌握程度，逐步呈现上升的趋势。

每个学生根据自己的学习跟踪数据表，认识到个体知识点的掌握程度差异，根据自身条件有针对性地调整学习计划。学生的个体学习跟踪数据表进入信息化系统，用大数据方法为每个学生进行数据画像，突显学生特点。在学生求职和企业招聘过程中，数据画像也为企业全面深入了解学生的情况提供了更多的参考。