

大数据导论本科通识课程建设*

杨先娣 熊素萍 张华 彭智勇

武汉大学计算机学院, 武汉 430072

摘要 根据“通识课程”建设理念, 引进校企协同育人的先进技术, 在教学过程中以学生为中心、采用翻转课堂, 并注重课程思政的融入, 从课程目标、课程内容设置、教学方法及考核方式等方面介绍面向本科通识教育的大数据导论课程建设情况, 为本科院校开展大数据导论通识课程提供参考。

关键字 大数据导论, 通识课程, 校企协同育人, 翻转课堂, 课程思政

Undergraduate General Education Course Construction of Introduction to Big Data

Xiandi Yang, Suping Xiong, Hua Zhang, Zhiyong Peng

School of Computer Science Wuhan University
Wuhan 430072, China
xiandiy@whu.edu.cn

Abstract—According to the construction concept of "general education course", the advanced technology of university-enterprise collaborative education is introduced. Student-centered teaching process and flipped classroom are adopted, meanwhile the integration of curriculum ideology and politics is emphasized. The construction of the Introduction to Big Data course for undergraduate general education is introduced from the aspects of course objectives, course content setting, teaching methods and assessment methods, so as to provides reference for undergraduate universities to carry out general courses on Introduction to Big Data.

Keywords—Introduction to Big Data, General Education, University-enterprise Collaborative Education, Flipped Classroom, Curriculum Ideology and Politics

1 引言

通识教育是高等教育人才培养的重要环节, 旨在培养学生独立思考能力及了解不同学科的常识, 进而能将不同的知识融会贯通, 最终目的是培养出一个完整意义上的人, 使学生在掌握专业技能的同时, 兼具科学素养、人文素养、批判性思维和创新思维, 以及良好的沟通表达技巧和团队合作能力。未来, 知识的更迭将更加迅速, 社会分工的日益细化要求有更加专业的知识, 同时, 各种复杂的问题需要更加全面的视野来进行分析和解决。因而, 实现通识教育与专业教育相融合, 成为众多高校推进高等教育高质量发展的必然选择。近些年来, 众多高校在通识教育的课程体系构建、课堂教学改革和制度创新等方面积极探索并有所建树, 但还需进一步深化、优化。

大数据既是一类客观数据, 也是一项技术革命, 更是一种思维理念。每个人都应或多或少地掌握大数据的概念、大数据的原理、大数据的应用, 特别是大数据的思维方式。学会用数据说话, 用数据分析的结果来证明“哪个更好, 哪个更坏”、并指导我们做出正确决策, 是每个大学生都应该学习掌握的。数据思维有利于培养一个健全的人, 作为理性的人类(实际上很多时候是非理性的), 我们不会仅仅满足于感情的宣泄, 而是希望通过数据掌握更多的事实, 进行理性思考。随着信息社会的发展, 大数据时代席卷而来, 大数据类的通识课程在全国高校陆续开设。

《大数据导论》是面向全校所有专业的本科生开设的通识课程, 这门课程不仅可以很好地拓展学生对数字世界的认知, 还有利于培养学生的跨学科视野以及终身学习的能力。同时, 本课程通过哲学的辩证眼光看待大数据带来的负面影响, 例如: 数据安全和隐私问题、大数据伦理问题及大数据交易问题, 引导学生形成正确的人生观和价值观。

*基金资助: 本文得到武汉大学“教育教学改革”建设引导专项子项目: “武大通识3.0”《大数据导论》(2020tsyb0034); 以及三个国家教育部产学合作协同育人项目(220801222122936, 202101215015, 202102342031)的资助。

2 课程建设情况

参照多所高校的大数据通识类课程的教学情况^[1-5],结合武汉大学“武大通识3.0”的实际需求,经过四年的课程调整与革新,《大数据导论》通识课程的建设已经趋于完善稳定。

2.1 课程目标

大数据时代已经全面开启,带来了信息技术发展的巨大变革,并深刻影响着社会生产和人民生活的方方面面。了解大数据概念、具备大数据思维,是新时代对人才的新要求,拥有大数据的理念,能够掌握数据和运用数据的人,才能在“一切都被记录,一切都被分析”的数据化时代生存和发展。作为本科通识课程,力求内容深入浅出,简单易懂,适合各院系本科生学习。课程旨在实现以下几个培养目标:

(1) 引导学生步入大数据时代,积极投身大数据的变革浪潮之中;

(2) 了解大数据概念,培养大数据思维,养成数据安全意识;

(3) 认识大数据伦理,努力使自己的行为符合大数据伦理规范要求;

(4) 熟悉大数据应用,探寻大数据与自己专业的应用结合点;

(5) 激发学生基于大数据的创新创业热情;

(6) 客观辩证地看待大数据的正面和负面影响,树立正确的人生观和价值观。

2.2 课程内容

作为一门通识课程,该课程的学习对象是全校所有专业的本科生,他们来自不同年级、不同专业,他们的知识结构和认知水平有明显差异。因此,在制订课程教学内容时,教师需要充分考虑学生的不同层次及不同学科背景的差异,将教学重点放在激发学生兴趣上,充分调动学生的学习积极性,打破学生原有的思维模式,拓展其数据思维能力并提升其创新能力。作为一门创新创业课程,我们和蚂蚁智信(杭州)信息技术有限公司合作开展校企协同育人,在课程中引入先进的国产分布式数据库OceanBase,将最新的产品和技术分享给学生,极大地提升了学生的学习积极性。

课程内容涉及广泛,首先从什么是大数据入手介绍基本概念及特征,并扩展到大数据时代的相关新兴技术的介绍,如云计算、物联网、人工智能、区块链和元宇宙,以及他们与大数据之间的密切关系,使学生们意识到学习大数据的重要性和必要性,同

时也开阔了学生的视野和见识;对大数据的存储、处理、分析及可视化三个方面的核心技术进行概要阐述,使得学生了解到大数据涉及的关键技术;引入国产自研分布式数据库OceanBase作为案例,并将其初级认证考试OBCA认证纳入课程考核体系,使同学们接触到最新的大数据产品并投入精力自学;通过翻转课堂进一步探讨了大数据时代的相关社会问题,包括大数据应用、大数据安全、大数据思维、大数据伦理、数据共享、数据开放、大数据交易与大数据治理8个主题。作为创新创业性质的通识课程,在相应的教学内容中引导学生将大数据技术及思维和本专业的应用相结合,鼓励学生开展相关的创新创业项目及竞赛。

2.3 教学设计

课程吸取往年的教学经验,在教学设计和方法上更加重视学生参与并主导课堂,以学生为核心,以问题为导向,启发学生思考,让学生成为课堂的主人;鼓励学生站上讲台,通过演讲、课堂讨论及辩论,融合不同的思想方法,产生思维的碰撞,激发学生的创新能力和哲学思辨精神,锻炼口才及团队组织能力;引入现代的信息化教学手段,通过QQ群引导学生线下学习交流、分享资源,借助学习通进行签到、课堂讨论和小组互评等,使课堂动起来,并对教学效果进行实时跟踪,及时对教学内容、教学方法进行修正。具体的教学设计有以下方面:

(1) 课程一共有11讲,大部分课次要求学生做到课前预习指定的课程阅读材料和讲义;

(2) 为保证教学效果,采用小班教学,班级人数控制在40人以内并将学生按照大数据应用、大数据安全、大数据思维、大数据伦理、数据共享、数据开放、大数据交易与大数据治理8个主题分成8个小组(3-5人/组),选取其中1人做组长。组长负责组织小组成员进行课堂讨论、组织组内成员进行翻转课堂讲课、分配结课项目任务等;

(3) 课程的前6讲主要由教师主讲,介绍大数据的概念、相关基础知识和特性;讲解目前最前沿的新技术及其和大数据的紧密联系;对大数据核心的三大技术——存储、处理、分析与可视化进行概要阐述;以先进的国产分布式数据库OceanBase为案例讲解大数据的具体应用。后5讲采用翻转课堂教学,由学生主导课堂,其中4讲里每一讲安排2个主题小组进行翻转课堂讲课,讨论并通过学习通APP打分互评,最后1讲是学生期末小论文的演讲,也要求打分互评;

(4) 在教学过程中安排两次讨论,一次随堂笔记,结课前安排一次学习通在线测试,最后一次大作

业要求小组分工合作完成一个主题报告或项目并提交结课论文。

以上这些都将有助于学生理解大数据导论的课程内容,同时也培养了学生的协同组织能力、哲学思辨能力、自主学习能力以及演讲与口才的锻炼,从而能更好地达到通识课程的教学目的。

2.4 教学安排

《大数据导论》通识课程教学周次共11周,每周一次课。第1周是2学时,后面10周都是一次课3学时,共计32学时,2学分。第11周最后一次课是小组汇报并提交结课论文,具体的周次和教学内容安排见表1。

表 1 大数据导论通识课程教学安排

课次	内容	要求
1	课程简介,什么是大数据?(大数据基础)	思考:谈谈你对大数据的认识以及你所在专业领域的大数据应用情况
2	为什么要学习大数据?(大数据和云计算、物联网、人工智能、区块链及元宇宙的关系)	讨论1:新兴技术对人类社会发展的影响
3	大数据存储	阅读资料;提交分组名单
4	大数据处理	阅读资料
5	国产分布式数据库OceanBase专题介绍	在线了解OB以及其他优秀的国产数据库
6	大数据分析	课堂同步作业:思维导图;布置期末大作业
7	大数据应用(1组)与大数据安全(2组)	OBCA认证
8	大数据思维(3组)与大数据伦理(4组)	讨论2:大学生应该具备什么样的大数据思维和大数据伦理?
9	数据共享(5组)与数据开放(6组)	阅读资料
10	大数据交易(7组)与大数据治理(8组)	测验:学习通在线测试
11	分组演讲答辩(撰写论文结课)	提交结课论文

3 教学实践

下面,我们以已经结束的2021-2022第2学期(春季)这一轮的教学实践情况对具体的教学实施及教学效果进行分析。

3.1 学生选课分析

大数据导论通识课程在2021-2022第2学期开课4个班,除去中期退课后的实际选课学生共计110人。武汉大学的四个校区除了四区(医学部)没有设课以外其他区均设课并顺利开课。在一区(文理学部)设两个班,分别有35人和25人,在二区(工学部)的班有35人,在三区(信息学部)的班有15人。按年级统计的选课人数如图1所示,大一、大二的学生是选课主力军,分别有61,38人,大三的选课学生有10人,大四的选课学生仅1人,这基本上符合通识课程以低年级学生为主的开课规律,高年级学生更多关注于专业课及实习实践课程。

按学院统计的选课人数如图2所示,选择大数据导论通识课程的学生来自于全校19所不同学院,充分体现了大数据导论通识课程面向全校培养交叉复合型知识人才的特性。从学院所属学科学部来看,选课学生主要集中在信息学部、理、工学部;人文、社科及

医学部的选课学生极少,这一方面是因为课程内容有技术性的硬核内容,人文、社科的同学有畏难情绪,另一方面是因为课程本身属于“社会科学与现代社会”模块,学校政策鼓励学生跨学科选课,这恰恰吸引了理、工、信息类学生,而医学部与其他学部距离较远,长途跋涉也阻碍了学生选课意愿,今后会在医学部开设课程。选课学生最多的前三位学院是网络安全学院、计算机学院和弘毅学堂,分别有24,18,12人,电气与自动化学院、数学与统计学院、电子信息学院、遥感信息工程学院、经济与管理学院选课学生随后,这些学院涉及的学科基本都是前沿科技,体现了大数据在前沿、热门领域的普遍关注度;另外,这些学院大都属于武汉大学的热门学院,高考录取分数较高,反映了相关专业学生的学习需求。

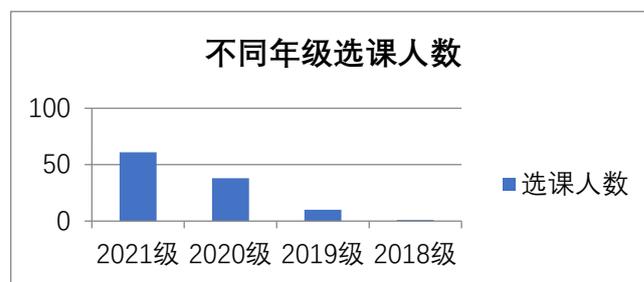


图 1 按年级统计的选课人数

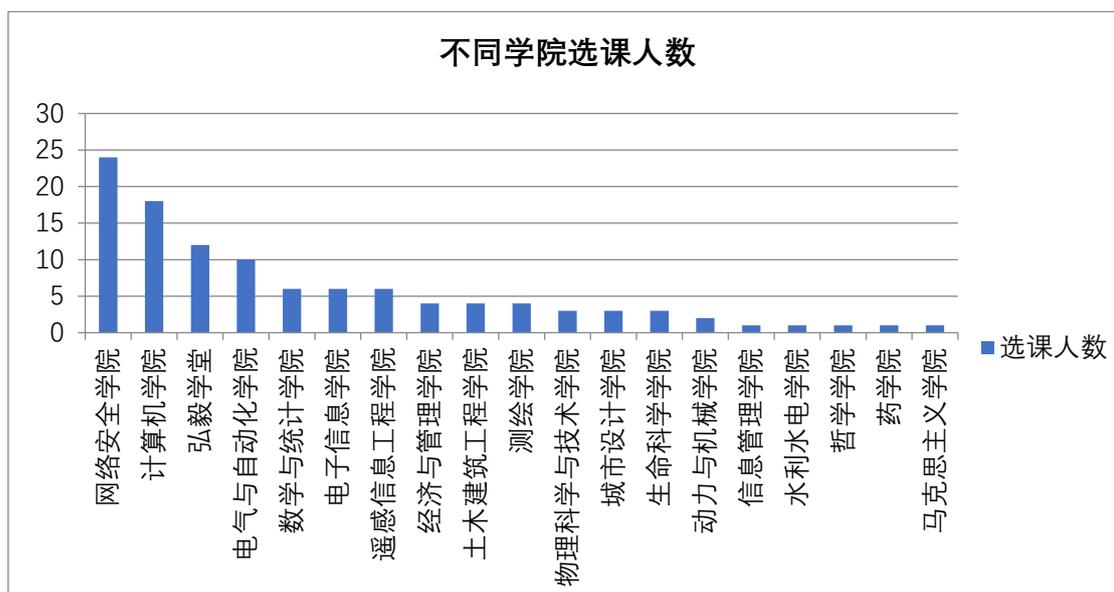


图 2 按学院统计的选课人数

3.2 课程考核及效果

通识课的学生来源和教学方法与专业课不尽相同,不能对所有学生采用同一考核评价标准。为了对每个学生作出客观、准确、全面的评价,本课程采取过程性考核方式,通过灵活多样的方式综合考查学生各方面的能力,考核内容主要由考勤、随堂笔记、课后作业、讨论、翻转课堂表现、在线测试、结课论文组成。

具体的考核方法:总评成绩=平时成绩*60%+期末大作业*40%。其中,平时成绩占到60分:考勤(5分)+讨论(10分)+随堂笔记(5分)+在线测试(15分)+翻转课堂(25分);期末大作业是小组分工合作完成并且演讲汇报,要求每位学生提交结课论文。另外,引入OBCA企业认证考试环节,鼓励学生业余时间自学分布式数据库的相关知识,在课程结束前通过OBCA认证考试的同学总成绩加5分,极大地提升了学生的学习热情,约有三分之一的同学能够通过OBCA认证考试加分。

表 2 大数据导论通识课程成绩分布

2021 -2022 学年第 2 学期成绩分布			
100-90分	87 人	89-80分	23 人
79-70分	0 人	69-60分	0 人
60分以下	0 人	不及格率	0%

课程成绩是教学效果的最好体现,表 2 是课程成绩分布情况,除了 6 人中期退课以外,所有的选课学生的成绩都 ≥ 80 分,79%的学生通过一学期的努力能够拿到 90 分以上的优异成绩。这个教学效果主要得益

于两个方面:首先,在第一次课上就明确了课程的教学内容及安排,尤其是严格的课程考核标准和详细的给分规则,这样就劝退了一部分对课程不感兴趣或者想来混学分的学生,留下来的学生都是认真学习且目标明确的同学;另外,武大宽松的退课政策也给了少部分同学反悔的机会,不至于到挂科的地步。

4 课程亮点

大数据导论通识课程自开设以来,不断吸收同行先进的教学方法和理念,同时收集热心同学们的建议并采纳。

近几年,通过对教学内容、教学方法及教学理念的不断调整,成为受到学生们普遍欢迎的高评价课程,主要教学亮点有以下方面:

(1) 课程采用了非常丰富的过程化考核方法,更加注重学生平时的表现,具体考核形式:总评成绩=平时成绩*60%+期末大作业*40%,其中,平时成绩的组成:考勤(5%)+讨论、随堂笔记及在线测试(30%)+翻转课堂(25%);期末大作业是小组分工合作完成并且演讲汇报,提交小论文;

(2) 作为创新创业课程与蚂蚁科技集团校企合作引入先进的 OceanBase 分布式数据库,为有兴趣深入学习相关技术的学生提供企业认证、数据库大赛等机会。课程中引入了 OBCA 企业认证考试环节,鼓励学生业余时间自学分布式数据库的相关知识,在课程结束前通过 OBCA 认证考试的同学总成绩加 5 分,极大地提升了学生的学习热情,一部分同学能够通过 OBCA 认证考试加分;

(3) 课程以学生为中心,力求及时将学生们感兴趣的新技术引进课堂,比如去年的元宇宙,今年备受关注的 ChatGPT。通过课堂讨论环节或者线上讨论专题启发学生思考;通过翻转课堂让每一位学生都有机会站上讲台主导课堂,不仅加强了学生的课堂参与度,也锻炼了演讲与口才,培养学生思辨能力;

(4) 立德树人是教育的根本任务^[6-8],在课程思政内容的设计方面,考虑到学生的接受程度,主要以视频故事或技术发展史等方式自然展开。例如:在介绍新兴技术的云计算时会讲述阿里云的创建者王坚院士的创业故事;在大数据存储介绍环节引入 OceanBase 分布式数据库的首席科学家阳振坤团队的奋斗历程等;另外,通过国产自研数据库的专题讨论让同学们了解到祖国优秀的数据库及其进展。这些与课程内容息息相关的素材生动形象,真实感人,随着课程进度自然衔接,同学们深受鼓舞,能够帮助学生塑造正确的人生观和价值观。

5 结束语

大数据导论本科通识课程已经在武汉大学经历了四轮的课程建设,目前的教学形式及考核方式等已经趋于成熟稳定,由于计算机新技术日新月异,今后,

我们也会及时引进大数据相关的新技术、新案例、新话题对课程内容及素材进行相应的更新和改进。希望能够为兄弟院校开展大数据导论通识课程提供参考和借鉴。

参考文献

- [1] 朝乐门. 数据科学导论的课程设计及教学改革[J]. 计算机科学, 2020, 47(7): 1-7.
- [2] 李志杰. 数据人才培养模式探索与实践[J]. 计算机教育, 2020(1): 60-63.
- [3] 李卫强, 王继强, 马先珍. 基于 OBE 的大数据课程体系建设[J]. 计算机教育, 2021(7): 139-142.
- [4] 易发胜, 赵丽琴, 胡德昆. “以学为中心”的大数据导论课程教学设计与实施[J]. 计算机教育, 2023(2): 177-181.
- [5] 林子雨. 大数据技术原理与应用课程建设经验分享[J]. 大数据, 2018, 4(6): 29-37.
- [6] 许家焯. 论课程思政实施中德育元素的挖掘[J]. 学科与课程建设思想理论教育, 2021(1): 70-74.
- [7] 张明慧, 郭欣. 大数据科学与技术专业课程群育人探索与实践[J]. 计算机技术与教育学报, 2022(10): 66-69.
- [8] 钱明珠, 关玉蓉. 教育元宇宙生态下计算机类课程思政教学路的探究[J]. 计算机技术与教育学报, 2022(10): 85-90.