

农业大数据研究生培养模式探索

牟少敏, 温孚江, 宋长青

山东农业大学农业大数据研究中心, 山东 泰安 271018

摘要

大数据应用尚处于初级阶段, 农业大数据又是多学科的交叉融合产物, 相应人才的缺乏是其研究和应用的重要瓶颈。简要介绍了农业大数据的概念、处理技术及其应用。并结合农业大数据的特点, 针对目前人才严重缺乏的现状, 提出了农业大数据研究生培养的重要性, 并在此基础上, 探讨了培养方案的要素以及人才培养模式。

关键词

农业大数据; 机器学习; 大数据处理技术; 研究生培养模式

中图分类号: G643

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2016006

Training mode of graduate students majored in agricultural big data

MU Shaomin, WEN Fujiang, SONG Changqing

Agricultural Big Data Research Center, Shandong Agricultural University, Taian 271018, China

Abstract

Big data application is still in its primary stage, and agricultural big data is the integrated products of many different subjects, so the insufficiency of big data specialists or experts is an important bottle-neck of big data application. The concept of agricultural big data, the processing technology and its application were briefly introduced. Combining the characteristics of the agricultural big data, aiming at the situation of agricultural big data talent lack, the importance of graduate students majored in agriculture big data was discussed. On this basis, the essence of training scheme and the training mode of graduate students were discussed.

Key words

agricultural big data, machine learning, big data processing technology, training mode of graduate students

1 引言

自2008年国际综合性期刊《Nature》发表有关大数据(big data)的专刊以来,大数据的采集、存储、挖掘、处理、分析和利用,已成为我国各行业和科学研究关注的焦点。全球IT、互联网巨头都已经意识到数据的重要性,包括EMC、IBM、惠普、微软和阿里巴巴等大型IT企业都纷纷通过收购和建立相应的大数据研发机构等手段实现技术整合。人们进入了一个“数据为王”的时代。

我国政府、高等院校、科研院所和各个行业都认识到信息科技时代大数据技术和应用的重要性,其发展将成为我国应对经济全球化、加快经济结构调整、培育新的经济增长点的催化剂。目前,我国不仅对数据共享、数据处理和数据安全等迫切需要解决的问题进行了顶层设计,而且对大数据的技术参考模型、数据交易平台交易数据描述、数据能力成熟度评价模型、数据质量评价指标等有关标准的制定也在加紧进行中,其必将为大数据的可持续发展提供强有力的保障。

农业大数据指的是来源广泛、类型多样、结构复杂、具有潜在价值,并难以应用通常方法处理和分析的数据集合^[1],具有农业地域性、季节性、多样性、周期性、准公益性等特点。目前,农业大数据在我国已经取得了初步的应用。利用地理信息系统(geographic information system, GIS)、遥感和机器学习(machine learning, ML)等技术,对小麦和玉米等农作物病虫害的数据进行分析,并对其发生程度进行预测预警。农产品检测信息已成为消费者追溯产品最直观、最可信的信息^[2]。利用物联网和数据库等技术,实现农

产品检测数据的整合,进行农产品质量安全溯源。采用O2O(online to offline)电商模式,使互联网成为线下交易的前台,将农产品的线上交易与线下服务相结合,大大节省了农产品的交易成本。在我国经济进入新常态的背景下,互联网+农村金融也受到了人们的广泛关注。农业大数据可以当好政府的智囊,为企业提供支撑,为学科提升和转型提供平台,为提高管理水平提供手段^[3]。

由以上分析可以看出,大数据广阔的应用前景需要大量的大数据人才。2018年美国需要44万~49万名大数据深入分析人才,存在14万~19万名的人才缺口;需要400万名具备基于大数据分析进行决策的经理和分析师,这一类人才缺口在2018年将达到150万人^[4]。而我国系统和专业培养农业大数据人才的机构和单位基本没有,从事农业大数据的人才更是非常缺乏,甚至是农业大数据发展的一个短板。

目前,我国农业大数据人才的主要来源是农业信息化的人才,农业信息化人才本身就存在数量严重不足的问题。同时,农业信息化人才主要是来自农学专业、计算机类专业和经济管理类专业等。这些专业培养人才的专业针对性强,知识面不广,而既具有农业基本知识又掌握大数据技术的人才则是少之又少,如果不加快培养复合型的人才,将会制约农业大数据的发展。

我国农业大数据的研究和应用可以视为一个非常巨大和复杂的系统工程。农业生产过程中,多年积累的数据是许多领域无法比拟的,它具有大量、高维、动态、不完全(缺值)、不确定(数据中的系统或随机噪声)和稀疏性(很少甚至没有有用的记录)等特性。只有将机器学习、深度学习和自然语言的理解等人工智能技

术、云计算技术、Hadoop等大数据处理技术和农业领域进行深度融合,农业大数据才能产生巨大的效益。因此,未来农业大数据的人才需求必然呈现出爆发性增长,农业大数据研究生人才必然会成为炙手可热的人才,如何突破目前国内农业大数据领域人才需求的瓶颈以及对人才培养战略的积极布局,加速其培养,显得尤为重要。

近年来,山东农业大学在农业大数据的研究和应用方面取得了一定的成绩,在经过深入调研和基础上,率先将农业大数据人才培养纳入了研究生培养体系,设立了大数据科学与技术研究生研究方向,下面重点探讨一下农业大数据研究生人才培养的新模式。

2 农业大数据研究生人才培养必要性、可行性及其模式

2.1 农业大数据研究生人才培养的必要性与可行性

汪洋副总理指出“用发展新理念大力推进农业现代化”,而农业现代化的发展需要物联网、大数据和云计算等技术,需要大量的农业大数据人才,因此农业大数据研究生人才的培养是一项非常必要和迫切的工作。

大数据时代赋予了农业大数据研究生人才培养特殊意义,是非单一学科所能完成的,需要多元化的培养体系,需要跨界和多学科的深度融合。山东农业大学是一所多学科的农业院校,既有农业相关学科,也有计算机科学与技术与信息科学等学科,在农业大数据研究生人才培养方面具有跨界和多学科深度融合的优势和基础。

目前,国内设置大数据研究生人才培养的状况如下:2014年中国人民大学、北京大学等5所大学联合培养大数据分析研究生,首批招生50名,5所高校还将共同组建大数据分析研究生培养协同创新平台。香港中文大学设立数据科学与商业统计(data science & business statistics)研究生课程,纽约大学上海分校设立商业数据分析科学(master of science in business analytics)研究生课程。北京航空航天大学、清华大学等高校也启动了大数据项目,培养相关人才。上海复旦大学自设了数据科学本科专业。

但是,针对农业大数据的研究生人才培养还未见报道,山东农业大学的农业大数据研究生第一期实验班2016年开始招生,计划招生人数30~50人。

2.2 农业大数据研究生人才培养模式

国务院促进大数据发展行动纲要指出:“鼓励高校设立数据科学和数据工程相关专业,重点培养专业化数据工程师等大数据专业人才。鼓励采取跨校联合培养等方式开展跨学科大数据复合型人才培养。”这为农业大数据研究生人才培养模式指明了方向。针对农业大数据的特点,使计算机科学与技术、信息与计算科学、统计学与经济金融、生命科学、资源与环境、农学、园艺科学与工程和植物保护等众多学科领域相互渗透,深度交叉,建立跨学科、跨领域的师资队伍和研发团队,开展农业大数据研究生人才培养工作。

研究生培养模式指的是在一定的现代教育理论、教育思想指导下,按照特定的培养目标和人才规格,以相对稳定的教学内容和课程体系、管理制度和评估方式,实施研究生教育的过程。下面结合农业大数

据研究生人才培养模式的特点,以山东农业大学的农业大数据研究生人才培养模式为例,从以下几个方面探讨如何进行农业大数据研究生人才培养。

2.2.1 农业大数据方向培养目标

针对农业大数据的特点,在充分调研的基础上,确定了农业大数据方向培养目标。即从数据管理、数据分析、数据应用3个主要应用层面出发,依托农业学科的优势,强化学科交叉,着重培养学生数据的收集、融合、管理能力,挖掘农业大数据背后价值的的能力,成为未来农业大数据等领域具有丰富实践经验和能力的实用型大数据人才。

2.2.2 师资队伍

农业大数据研究生人才培养的关键问题之一就是师资队伍的建设。农业大数据的跨学科、跨领域的特点,决定了其师资队伍建设需要打破以往传统的构建模式,需要新思路和新模式。

农业大数据研究生人才培养主要依托信息科学与工程专业的师资力量,同时根据人才培养的需要,联合相关学科师资力量和领域专家学者力量共同培养。如山东农业大学在涉农、经济管理、计算数学等学科建设方面具有雄厚的学科基础,能够为数据科学与应用学科起到良好的支撑作用。

研究生的培养实行双导师负责制和导师组相结合的模式。为了强调学科交叉课程学习,完善研究生多学科的知识结构与农业大数据的思维方式,采用了双导师制。双导师1名来自本校,另外1名来自其他国内知名大学或国内知名企业。本校导师是研究生培养和管理的第一责任人。由于农业大数据应用会渗透到农业产业链的产

前、产中和产后的整个过程,这还需要构建以双导师为中心,多学科导师相辅助的集体指导小组,结合实际科研课题或工程项目进行培养。

2.2.3 人才培养

为了进一步加强人才培养力度,重视科研能力和方法的训练,还采用与名校、名企联合培养模式。因材施教,强调自学和对话式、讨论式教学,发挥导师和研究生双方的积极性和创造性。教学方式也有别于传统的课堂讲授,采取案例教学和讨论班的方式,或者以团队的形式共同完成一个项目。同时,聘请国内外知名数据科学和农业等领域的专家教授定期进行讲座,不断扩大对外交流,通过创新培养机制,保障农业大数据研究生人才培养质量。

2.2.4 科学研究

依托山东农业大学信息科学与工程学院、山东农业大数据研究中心和山东省农业大数据产业技术创新战略联盟,在农业大数据的科学研究上已经完成了以下工作:利用物联网技术进行农业数据(如温度、湿度等)的采集;建立了分布式的存储数据库HBase和基于Hadoop的农业数据分析平台。将采集到的农业相关数据(农业病虫害等结构化数据以及图像、文字等非结构化数据)存储到分布式数据库中。借助于Hadoop大数据处理平台将支持向量机等主要的机器学习算法进行并行化数据分析,实现农业数据从存储到分析的并行化,极大地提高了农业数据处理的吞吐效率和分析准确率,为农业数据的收集存储分析提供了一种新的处理模式。

充分发挥先期已经取得的数据技术与应用科研优势,结合国家“十三五”规划,从数据采集、数据存储、数据分析与处理技术、数据的可视化等几个方面入手,加速培养适应农业信息化各方面需要的高层次、多层次的复合型研究生人才。充分发挥山东农业大学农业科学优势及生命科学特色,充分发挥大数据技术与应用技术的强大功能及优势,进一步拓宽大数据技术在农业上的应用领域,提高农业数据收集整理分析和决策的科学化。

2.2.5 课程体系与实践教学

科学合理的研究生课程体系对保障研究生培养质量具有重要作用。紧紧围绕研究生培养目标,从体现农业大数据的特色出发,构建开放式的研究生课程体系。课程体系主要包括核心课程和专业学位课程。课程学习时间为1年。确定了“农业大数据分析技术基础”、“数据挖掘”、“机器学习”、“农业信息科学”、“非结构化大数据分析”和“农业大数据案例分析”等必修课程。每门课程不是由一位老师独立完成,而是由一个课程组各有特长的多位老师共同进行讲授。

实践教学是培养农业大数据创新人才非常重要的一个环节,是合作培养模式的具体体现。研究生在修满专业规定的学分后,以实际项目为背景,由校企双方发挥各自的优势,共同指导农业大数据课题的研究。实习基地是学生实习和社会实践的重要场所,建立跨学科的研究中心和实验室,并以如泰山神农智谷农业大数据产业园区作为农业大数据研究生人才的校外实习基地,计划逐步增加更多的实习基地,使得农业大数据研究生的培养更加贴合学科前沿的发展,更好地满足国民经济建设需求。

3 结束语

大数据时代背景下,如何培养农业大数据的复合型研究生人才无论是在国内还是在国外都没有先例,数据科学和技术的研究也还处于起步阶段。山东农业大学率先在农业大数据研究生人才培养方面进行了相关的探索和研究。农业大数据研究生人才的培养必将是一项长期而艰巨的工作,仍需要不断地创新、探索和实践。随着大数据时代的到来和研究生培养模式的不断创新和完善,一定能够培养出高质量的农业大数据研究生人才,服务于我国农业现代化的建设。

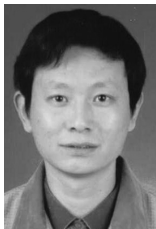
参考文献:

- [1] 赵璞. 浅谈农业大数据[EB/OL].[2015-01-11]. <http://www.caas.net.cn/zt/kxjd1/250773.shtml>.
ZHAO P. Discussion on agricultural big data[EB/OL].[2015-01-11]. <http://www.caas.net.cn/zt/kxjd1/250773.shtml>.
- [2] 钱建平, 张保岩, 邢斌, 等. 集成实时快速检测信息的蔬菜追溯系统改进与应用[J]. 农业工程学报, 2015(4): 306-311.
QIAN J P, ZHANG B Y, XING B, et al. Improvement and application of vegetable traceability system with integrated real-time rapid detection of information[J]. Transactions of the Chinese Society of Agricultural Engineering, 2015(4): 306-311.
- [3] 温孚江. 农业大数据研究的战略意义与协同机制[J]. 高等农业教育, 2013(11):3-6.
WEN F J. Strategic significance and synergetic mechanism of agricultural big data research[J]. Higher Agricultural Education, 2013(11): 3-6.
- [4] 何海地. 美国大数据专业硕士研究生教育

的背景、现状、特色与启示[J]. 图书与情报, 2014(2): 48-56.
HE H D. The status and suggestions of

American great universities with master's programs in big data analytics[J]. Library & Information, 2014(2):48-56.

作者简介



牟少敏(1964-),男,博士,山东农业大学教授,主要研究方向为大数据、机器学习和模式识别。



温孚江(1955-),男,现任山东农业大学校长、教授,农业大数据创新战略联盟理事长,全国人民代表大会常务委员会委员。早年留学美国,并获得博士学位。主要从事植物保护研究和宏观农业研究工作。发表论文210余篇,专著5部。最近一本专著《大数据农业》由中国农业出版社于2015年9月出版。目前主要从事农业大数据应用研究工作,是我国农业大数据研究主要发起人之一。



宋长青(1963-),男,山东农业大学教授,农业大数据研究中心常务副主任,主要从事农业大数据采集、入库、平台建设、示范基地等研究与示范工作。

收稿日期: 2015-12-15

基金项目: 国家科技支撑计划基金资助项目(No.2013BAD15B02); 公益性行业(农业)科研专项经费基金资助项目(No.201303107); 中国农业科学院科技创新工程基金资助项目(No.CAAS-ASTIP-201X-AII-04)

Foundation Items: The National Key Technology R&D Program (No.2013BAD15B02), Public Sector (Agriculture) Scientific Research Funding Program (No.201303107), Scientific and Technical Innovation Project of CAAS (No.CAAS-ASTIP-201X-AII-04)