

高职大数据应用技术专业实践技能培养探究

沈 洋

(大连职业技术学院, 辽宁 大连 116035)

摘 要: 大数据技术对于国民经济生活的影响日趋明显,数据的商业价值也越来越重要,企业对于大数据分析与应用人才需求日益旺盛。通过对高职大数据应用技术专业的人才定位、知识结构与技能分析,以企业实际工作项目为载体设计实践学习内容,创新人才实践技能的培养内容和培养方式,使学生能够从事大数据相关领域的开发、运维与管理

关键词: 高职; 大数据应用技术; 实践技能

中图分类号: TP399 **文献标识码:** A

Research on Practical Skills Training of Big Data Application Technology Specialty in Higher Vocational Colleges

SHEN Yang

(Dalian Vocational Technology College, Dalian 116035, China)

Abstract: The influence of big data technology on national economic life is increasingly obvious and the commercial value of data is also increasingly important. As a result, enterprises' demand for talents in big data analysis and application is increasing. After the application of big data in higher vocational technology professionals positioning, knowledge structure and skill analysis, the practical learning content is designed with ongoing project as the carrier, innovating the contents and ways of cultivating talents' practical skills, and enabling students to engage in the development, operation, maintenance and management of big data related fields.

Keywords: higher vocational college; big data application technology; practical skills

1 引言(Introduction)

大数据是指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合,是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。

大数据及人工智能产业的发展对人才提出了新的需求,专业人才的培养是新一轮科技较量的基础,大数据人工智能人才培养也纳入各级各类的培养体系中。该专业是一个与行业契合度非常高的一个专业方向,不仅需要从业人员具有较强的逻辑思维能力,还需要计算机的编码开发能力以及对数据的分析处理能力,尤其对于高职院校的学生而言,数据的处理与分析能力是高职学生就业从事的主要岗位方向,因此在整个专业人才培养过程中,专业人才的培养一定要重视实践与实训教学,培养学生分析问题与解决问题的实践应用能力^[1]。

2 专业人才定位及能力需求分析(Talent orientation and capability demand analysis)

2.1 专业人才定位

高职大数据技术与应用专业人才培养一定要与本科院校培养的目标不一致,所面向的职业岗位应该是应用型技术人才所能从事的岗位,包括大数据应用开发、大数据运维管理、大数据处理三个层面为主。在现阶段高职院校开展的专业群建设中,大数据的相关课程也发挥着重要的作用,在计算机软件技术、网络技术以及电子商务技术等相关专业中,都涉及围绕大数据进行开发和数据处理相关知识的运用,因此培养大数据应用技术相关人才是高职院校信息技术专业群建设的重点建设方向。

2.2 专业人才知识与职业素养分析

在知识结构方面,应达到要求:(1)掌握数学方面的相关知识,具备抽象思维能力;(2)掌握至少一门程序设计语言,如Java、Python等,掌握数据库知识和分布并行计算至少一种方法;(3)了解机器学习,数据挖掘和神经网络算法;(4)懂

得不同行业的工作业务流程，能够挖掘数据的商业价值。

在职业素养方面，应达到如下要求：(1)具有工程思维，沟通交流能力和团队协作能力；(2)具有较强欲望的创新精神；(3)对数据要有较强的敏感性，对待事物要有较强的好奇心，善于跨界思考；(4)一定要能够严格保守数据信息安全，不能够损害企业和国家利益。

2.3 专业人才能力需求分析

随着国内多数高职院校大数据应用技术专业的开设，关于大数据应用技术专业的人才能力培养的体系逐步成熟，总体而言，专业的发展一定要面向大数据行业，对接大数据应用的不同产业，近几年该专业的人才能力培养发生以下变化：其一，人才需求从中高端人才向应用型人才过渡。大数据相关技术目前正处在落地应用的重要阶段，与大数据研发初期需要大量的中高端人才不同，在落地应用阶段则需要大量的应用型人才，这些应用型人才需要把大数据技术落地到广大的传统行业中。其二，大数据分析人才将是需求重点。在大数据落地应用的过程中，大数据分析将是人才需求的重点，因为数据分析是体现数据价值的重要途径，所以广大传统行业将首先会关注大数据分析领域。大数据分析领域的人才需求也会带动大数据运维和大数据开发领域的人才需求，当然大数据的发展也会带动物联网的发展。其三，大数据教育将结合行业特征。早期的大数据教育主要以培养大数据技术为主，包括大数据平台的搭建、大数据开发、算法设计、结果呈现等内容，未来随着大数据与传统行业的结合不断深入，大数据教育将进一步结合具体的行业特征。而具备行业背景知识的大数据人才将受到企业的欢迎，因为行业知识将是大数据落地应用的重要环节。其四，大数据与物联网等技术将进一步融合。随着5G通信标准的落地，未来的物联网、移动互联网、大数据、传统行业将深度融合，这些技术将作为产业互联网的重要组成部分共同服务于传统行业，所以对于大数据行业的从业者来说，应丰富自身的知识结构，了解物联网等相关技术^[2]。

基于以上分析，大数据应用技术方向人才需要具备本领域的理论性能力、实践性能力和应用性能力三个方向，其中理论性能力需求中要求学生能够掌握数据科学中的各种模型和算法的理解和运用，而这部分是高职学生的短板，在专业教学中，教师只需要按照教学计划中课程标准的要求完成理论内容的介绍，使学生能够掌握大数据的先进技术和理论，掌握大数据平台的理论技术基础即可，具体如图1所示。

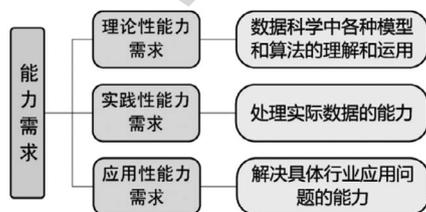


图1 大数据应用技术人才专业能力需求

Fig.1 Professional ability demand for the talents of big data application technology

在实践性能力和应用性能力需求方面，使学生能够处理相关工作场景中的实际数据，能够实现数据采集、数据存储、数据管理、数据分析与挖掘以及数据可视化的展示，能够完成不同行业应用的解决方案，并使得学生能够掌握与大数据分析的相关工具和平台。高职学生的普遍特点就是理论

基础薄弱，实践动手能力较强，因此在高职的大数据应用技术人才培养中，应该突出实践技能和应用操作处理能力的培养，重点培养学生的处理实际数据的能力和实际工作中解决不同行业应用实际问题的能力。

3 专业人才的实践技能培养探究(Research on the cultivation of practical skills of professional talents)

大数据应用技术人才应该属于新型的复合型人才，通过分析可知本专业要求具有较强的逻辑思维能力，能够掌握运用计算机处理数据的综合实践能力。但是作为高等职业院校人才培养目标而言，使本专业的学生应该具备能够完成实际工作中需求的各项技能是整个教学过程的重要手段，大数据应用技术领域本身就是运用计算机技术能够对数据进行综合性的操作和分析，在专业能力需求分析中，实践性能力也是联系理论性能力和应用性能力的重要纽带。因此作为高职大数据应用技术专业的人才能力培养的核心就应该是在掌握基本理论知识的前提下，基于企业真实的项目和案例对学生进行实践性能力和应用性能力的递进培养。

3.1 培养形式：校企合作共同培养

大数据应用技术专业的知识与能力素质培养前提一定是校企合作共同培养，因为大数据专业如果脱离了行业背景就失去了数据的价值与意义，因此人才的知识 and 能力素质培养要与行业内的资深企业合作，双方共同制定培养方案、师资培训计划、教学资源建设以及实践教学和实训。专业的产教融合应该也要为企业输送需求的技术人才，校内教师应该夯实学生的知识层面，完成专业基础课程内容的讲授，当然教师要在企业经过企业实践，受过专业培训和指导，在专业课阶段，应该是校内外教师团队共同讲授相关课程，给予学生指导，在实践课和实训中，应该由企业的老师以企业的实际案例进行项目化教学，通过这种教学组织形式，才能培养出学生坚实的基础知识和符合行业、企业要求的专业技术技能^[3]。

3.2 培养内容：以企业实际工作项目为载体设计实践学习内容

大数据应用技术专业的技能培养，一定要围绕企业实际项目为载体来设计实践教学内容，大数据实践项目需要来自工作一线的大量数据进行分析与操作，因此实践教学一定要基于实际项目来进行设计，只有这样才能让学生能够在实战中得到实践技能的提升和应用性能力的培养。

大数据应用技术人才的培养重在培养学生利用这项技术完成相关的实际案例的操作，因此在专业实践综合类实训课程的设计中，一定要基于实际应用的背景，围绕专业中的实践性能力需求和应用性能力需求开展实践类课程的设计。实践教学的项目选择一定是一个完整的项目过程，这个过程一定是能够覆盖行业内部整个工作流程，并且在整个实践项目的讲授过程中能够培养学生行业所需的核心技能。因此根据以上原则，选取“基于Hadoop的电商网站用户行为分析系统”这个综合性的实践项目可以作为大数据分析、处理与应用的实践项目，通过整个项目的教学与实训，使学生在实践技能方面得到提升。

项目依托的硬件基础设施为机房通用学生用机及相关网络设备，软件基础层选用开源的Linux操作系统。架构层采用Hadoop为核心，应用层分数据采集、数据存储、数据分析、数据可视化四个模块^[4]。

3.2.1 实践教学目标

本项目围绕的电商交易进行设计，生产消费过程中通过平台产生或记录了大量数据。这些数据如何分析从而得到和用户相关的有用信息。从大数据的实际应用出发，针对目前存在很多的电商网站存在相互竞争关系，解决如何精准定位目标客户的问题。为了提升网站的点击率和订阅量，采用行为分析和数据挖掘对网站进行分析，通多对网站操作的各种行为进行业务分析，从而提升网站的用户体验性，提高用户的转化率和广告效率等。通过这个教学案例，可以使培养学生掌握大数据技术中的数据采集、数据存储、数据管理、数据分析与挖掘以及数据可视化的实践技能。

3.2.2 实践教学项目应用价值

通过“基于Hadoop的电商网站用户行为分析系统”，进行用户行为分析系统可以很好地收集到原始数据，并且处理这些收集来的这些原始数据，把这些数据中有用的部分我们清洗提取处理，做成需要的用户行为的这种指标信息，然后分析，也可以形成用户画像，广告的精准投放等一些项目做一些基础的数据提供者，因此这个系统在培养学生实践与应用方面具有一定的研究价值^[5]。

3.2.3 实践平台功能分析

实践项目结构组成如表1所示。

表1 实践案例架构表

Tab.1 Architecture of practical cases

模块结构	功能介绍
数据采集模块	主要负责各种用户信息数据的收集工作
原始数据存储模块	主要负责把收集出来大量数据储存到可以做分析的系统平台上
原始数据清洗模块	主要负责原始数据的格式化清洗和查询出自己业务需要的数据格式
可用数据存储模块	主要负责把清除出来的需要用的数据放到数据仓库中使之可以用sql进行查询分析，然后得出每个业务需要的不同数据指标
数据分析展示模块	主要负责把分析出来的数据指标展示出来，进而分析用户行为的特征做出不同的策略

3.2.4 实践内容设计

(1)项目需求分析

①用户需求分析

在用户需求方面，需要针对用户对象的行为进行采集，完成一个用户行为的日志收集系统，记录用户在电商网站上的行为动作，这里可以用nginx服务器来做这个日志的收集系统。并将用户行为信息存储在平台中，进行数据的分析运算，清洗得出需求的数据。将清洗过的数据进行前端展示。

②系统需求分析

本系统下有四个模块：数据收集模块、数据存储模块、数据清洗处理模块、数据展示模块四大部分组成。

在数据收集模块中，要尽可能多的收集用户的行为信息，要有很多字段来解析这个人的点击流的行为，所以我们需要改收集模块的配置文件，然后收集很多我们需要的字段数据。

在数据存储模块中，我们收集来的数据肯定都是url的格式化的数据，那么我们的数据存储模块就要能很好存储格式化数据的数据库，收集每个用户的每个浏览网页的各种行为，因此收集的数据量非常大，需要能存储很大量数据的数据仓库，因此应结合这几点需求来设计我们的数据存储模块。

在数据清洗模块中，基于hadoop的大数据平台肯定要用

到hadoop里面的自带框架mapreduce，将原始数据清洗成能被数据厂库存储的表的格式。

在前端数据展示模块中，首先用户行为系统需要前端的展示，那么我们肯定需要这个系统来完成数据的可视化的展示。

(2)项目系统设计

本系统是一个Hadoop的离线数据分析的系统架构，主要用到的技术是Java+Mysql+JavaWeb+Linux+Hadoop/Hbase/Hive+Sqoop/Azkaban，然后通过数据的预处理，来实现数据的清洗。

在系统技术架构中，首先我们底层的就是Nginx作为服务器来收集日志然后加上Flume收集信息，再就是存储的设计就是用的hive的数据仓库加上Mysql数据库来完成数据存储模块，之后就是数据清洗，这就用到了Hadoop的Mapreduce计算模型来编写的一套数据清洗系统，最后就是前端可视化展示我们用的是ssm框架，加上大数据展示组件Echarts总体结构流程图框图。

(3)项目实现过程

①搭建服务器集群，规划集群中服务器功能。

②配置日志收集模块，实现日志文件的收集。

③应用Hdfs的Hive数据仓库来存储数据

④应用Hadoop平台下面的Mapreduce的计算模型来编写的数据处理程序，然后数据来源是有Hdfs提供，清洗出来的数据也是存储到hdfs，然后提交Maptask或Reducetask任务给集群计算。

⑤使用前端大数据可视化的图展示清洗出来的数据

本项目通过对电商交易网站的数据采集，充分运用Hadoop的Mapreduce集群分析模块，Hive和Hbase的数据存储模块，java web的数据展示模块，和Echarts大数据展示表格模块等大数据技术，挖掘出具有重要价值的信息，以此为企业提供有助于生产和营销的重要参考。

4 结论(Conclusion)

综上所述，在实践教学的方案设计中，应该以行业企业需求为主，企业实际项目为载体，能够通过学生的实践操作掌握专业的核心实践技能，本方案中采用的“于Hadoop的电商网站用户行为分析系统”完全贯穿了大数据应用技术的核心技能，能够使将课堂上学到的理论知识与实践技能运用到解决实际问题中去，因此，我们在大数据相关专业的实践教学项目的开发和设计过程中也一定要结合这种思路，使学生能够通过实践项目提升自己的专业技能，进而能够迅速的适应未来的工作岗位。

参考文献(References)

- [1] 罗恩韬,张少波,唐雅媛.新工科背景下产学研协同创新人才培养体系研究[J].山东工业技术,2018(10):236-238.
- [2] 田绪红,邝颖杰,肖磊,等.大数据应用人才的知识、能力、素质结构及其培养[J].计算机教育,2017(8):57-60.
- [3] 王崇刚.高职院校大数据技术与应用专业人才培养方案课程体系模块化设计的探索[J].广东蝉业,2017(12):91.
- [4] 傅伟,涂刚.高职院校大数据专业人才培养模式的构建[J].职教通讯,2017(32):13-15.
- [5] 李文强.基于Hadoop的电子商务大数据分析综合实训平台建设[J].现代工业经济和信息化,2018(6):36-39.

作者简介:

沈洋(1980-),男,硕士,副教授.研究领域:大数据技术与应用,信息安全.