

工业大数据架构与应用课程多层次实验教学设计

覃海煥¹ 张威² 余翔¹ 芦立华¹ 计春雷¹

(1 上海电机学院电子信息学院 上海 201306; 2 上海朋禾智能科技有限公司研发部 上海 201210)

摘要: 工业 4.0 时代, 工业大数据是实现工业转型升级的基础和核心, 架构为解决其广泛应用所带来的挑战提供了基础。为培养提升学生的工业大数据架构与应用实践能力, 针对该课程内容覆盖面广、业务创新及架构设计层次较高较抽象、实验难以落地的问题, 探索并设计了多层次、递进式的实验教学内容, 涵盖了对学生自主学习能力、技术实践能力、分析解决问题与创新实践能力等的锻炼, 并以工业 4.0 的核心技术 CPS 环节的实验为例进行阐述。实施效果表明, 该设计内容满足教学要求, 能激发学生进行自主学习和实践, 使学生能够更清晰地理解工业大数据在制造业各重要环节的应用所面对的不同业务需求、架构设计、基本原理和关键技术。

关键词: 工业大数据; 多层次实验教学设计; 架构与应用; CPS

中图法分类号: TP311

Multi-level Experimental Teaching Design for Architecture and Application of Industrial Big Data Course

QIN Haihuan¹, ZHANG Wei², YU Xiang¹, LU Lihua¹, JI Chunlei¹

(1 School of Electronic Information Engineering, Shanghai Dianji University, Shanghai 201306, China)

(2 R&D Department, Shanghai Penghe Intelligence Technology Co., Ltd, Shanghai 201210, China)

Abstract: In the era of Industry 4.0, industrial big data is the foundation and core of industrial transformation and upgrading. The architecture provides the foundation for solving the challenges brought about by its wide application. The industrial big data architecture and application course covers a wide range of content and its business innovation and architecture design are relatively abstract, which make experiments difficult to design and conduct. In order to cultivate and improve students' practice ability of industrial big data architecture and application, multi-level and progressive experiment teaching contents have been explored and designed, which cover the training of students' initiative learning ability, technical practical ability, analytical problem-solving and innovative practical ability. The experiment of the core technology of Industry 4.0, CPS, is taken as an example to illustrate the design. The implementation effect shows that the design contents meet the teaching requirements. The design can stimulate students to carry out initiative learning and practice, and enable students to understand the different business needs, architecture design, basic principles and key technologies of industrial big data application in all important phases of the manufacturing industry more clearly.

Key words: industrial big data; multi-level experimental teaching design; architecture and application; CPS

近年来, 随着以美国工业互联网、德国工业 4.0 为代表的新一代工业革命的进行, 全球产业竞争格局正在发生重大调整, 工业转型升级势在必行。中国及大多数国家的工业企业都在积极探索转型升级方案, 如在生产过程中融入互联网、大数据、物联网、人工智能等新一代高新技术, 以期在未来的行业发展中抢得先机。工业大数据作为实现工业转型升级的基础和核心^[1], 企业界及学术界在其应用与实践、技术与架构、安全管理等方面已做了不少研究工作^[2-4]。为更好地衔接学校所学知识与企业所需的主流前沿技术与能力, 不少高校也陆续开设了工业大数据相关课程^[5], 并积极进行课程建设。

《工业大数据架构与应用》理论课程覆盖了产品生命周期与价值流、企业纵向层级、IT 价值链等内容, 知识体系庞杂, 且在各环节中通过工业大数据实现业务创新和转型升级的思路、解决方案和架构设计的相关内容层次较高, 相应实验较难设计并落地实践。为切实培养提升学生的工业大数据架构与应用实践能力, 实验内容中除了设计环节外, 还需结合业界当前主流技术增加技术实践环节, 形成多层次的实验教学内容^[6-8], 以锻炼学生各业务场景的关键技术应用与实践能力, 并借此加深其对业务场景的理解, 提高其解决方案和架构设计能力。

在此背景下, 课程组与上海朋禾智能科技有限公司合作, 选取

理论课程中重要的几个环节, 探索并设计了多层次、递进式的实验教学内容, 培养提升学生的自主学习能力、技术实践能力、分析解决问题和创新实践能力^[9]。并通过有趣的业务场景和企业主流、前沿的技术相结合, 激发学生的学习兴趣和实践自主性, 让学生能够在工业大数据和制造业各重要环节的应用中根据不同的业务需求选择恰当的基本组件和关键技术并完成创新解决方案和架构设计, 促使学生成功取得课程培养目标。

1 课程目标及建设思路

1.1 课程目标

我校数据科学与大数据技术专业的专业特色是培养智能制造背景下富有创新精神的新时代卓越应用型工业大数据人才。作为专业特色的核心课程之一, 工业大数据架构与应用课程需要支持的毕业要求指标点^[10]主要包括:

指标点 1 掌握专业基础知识, 并能用相关知识比较和综合不同的数据应用解决方案;

指标点 2 能运用本专业专业知识, 借助文献研究, 探索不同的数据应用解决方案, 通过分析获得有效结论;

指标点 3 能够针对不同行业应用需求, 设计体现行业应用特色的大数据应用系统, 并在设计中体现创新意识;

指标点 4 能够根据大数据应用需求,选择与使用合适的大数据平台、大数据开发工具和技术,对复杂数据工程问题进行分析、设计与实施。

相应地,课程组制定课程目标^[10]如下:

课程目标 1 掌握工业大数据相关的基础知识,并能选择适当的工业大数据应用解决方案;

课程目标 2 掌握工业大数据的典型应用场景、支持工业转型的典型方式、创新方式及分析方法,能设计具体业务场景的工业大数据应用解决方案;

课程目标 3 能运用 IT 价值链相关知识,完成工业大数据应用解决方案的具有新意的架构设计;

课程目标 4 掌握工业大数据解决方案关键环节的设计与实现方法。

课程目标与毕业要求指标点对应关系及其达成情况评价依据如表 1 所示。

表 1 课程目标与毕业要求指标点对应关系及达成评价依据

毕业要求指标点	课程目标	权重
指标点 1	课程目标 1	10
指标点 2	课程目标 2	15
指标点 3	课程目标 3	15
指标点 4	课程目标 4	60
合计 云数据库		100

1.2 实验教学建设思路

工业大数据架构与应用课程实验设计及开展的主要难点如下:

①知识体系庞杂,需选取适当的环节设计实验;②各环节基于工业大数据实现业务创新和转型升级的思路及其解决方案、架构设计层次较高,学生通常难以具备这样的知识和视野,需适当降低难度;③系统开发实现实验较难:系统复杂、实验环境配置复杂,因此需进行部分模拟和简化;④部分环节的实验有关联,需要连续实验,而普通机房不支持。为此,我们探索了多层次设计、校企合作和基于智慧云平台的建设思路。

1) 多层次设计

大部分实验都涵盖设计环节和技术实践环节,其中,设计环节一般为解决方案设计、架构设计或思考题;技术实践环节则包含基础实践内容和进阶实践内容,实验教学内容是多层次的。让学生至少能够在基础实践内容的基础上,部分完成进阶实验内容及设计环节,达成基本课程目标。

2) 校企合作

课题组与具备领先的边缘智能系统软件架构开发经验的工业智能解决方案服务商上海朋禾智能科技有限公司合作完成实验的多层次设计。首先,结合当前行业热点、企业项目和工业各重要环节,一起选取学生可能感兴趣的环节和业务场景;其次,一起完成设计环节的题目的要求和教学案例;再次,企业结合当前主流技术通过模拟或简化设计该工业环节或该业务场景的技术实践环节,然后课题组老师在此基础上将部分内容设计为进阶实践内容,根据内容难度及学生平均水平给出简单或详实的部分实现提示;最后,企业

梳理该实验所需使用的重要基础知识和基本原理,以帮助学生不仅知其然还能知其所以然,加深理解,完成知识储备。

此外,实验设计中借助行业热点、有趣的业务场景和企业前沿、主流的技术相结合,激发学生的学习兴趣和实践自主性,以达到锻炼学生的自主学习能力、技术实践能力、分析解决问题和创新实践能力的目标。

3) 智慧云平台

借助学院的基于 OpenStack 的智慧云平台,可以连续开展实验,也可提前将实验资料下载,让学生提前学习,且平台可保存课程模板,后续进行实验及实验环境的迭代优化也比较方便。

2 多层次实验教学内容设计

根据多层次、校企合作的实验教学内容建设思路,本文实验设计中最终选取的工业大数据架构中的维度、工业环节^[11]及相关业务场景如下:

(1) 生命周期与价值流维度的研发大数据、生产大数据和服务大数据环节,业务场景分别为电动汽车、数字孪生和工业机械设备的故障诊断;

(2) 企业纵向层级维度的信息物理系统(Cyber-Physical System, CPS)和互联网+时代的工业转型环节,业务场景分别为路况模拟及拥堵警示和白领西服个性化定制;

(3) IT 价值链维度的信息技术架构环节,业务场景为制造资源推荐或电影推荐;

(4) 综合维度的工业大数据治理环节,业务场景为工业大数据安全。

实验的多层次实验内容详见表 2,其中基础知识、基础实践内容和进阶实践内容、设计/思考环节分别培养锻炼学生的自主学习能力、技术实践能力和分析解决问题与创新实践能力;实验 1、2 和 5 使用 Node-RED 模拟实现软硬件间的数据流连接,并通过 Docker 对实验环境进行了封装简化。

表 2 多层次实验内容

序号	实验名称	实验内容			
		基础知识	基础实践内容	进阶实践内容	设计/思考环节
1	研发产品数据模型设计	MQTT 协议、设备平台联通基础架构等	模拟完成轻量级的设备数据采集、传输、存储和展示	理解数据发送、接收的代码	对工业产品(如电动汽车)进行数据建模
2	智能制造解决方案设计	智能制造发展三阶段、数字孪生、MQTT 等	模拟实现通过云端数字实体同步信息到物理实体的过程	按需修改车辆状态接收、改变的代码逻辑	给出一个制造环节的解决方案及其架构设计
3	机械设备故障诊断	机械设备故障诊断步骤、典型轴承故障	执行故障诊断代码得到故障图谱,分辨出故障类型	编程分析真实案例的数据,分辨故障类型	无
4	路况模拟及拥堵	CPS、TensorFlow、	完成图片识别示例代码的执行,	理解代码框架,增加拥堵	路况模拟及拥堵警示解

	堵警示	Fast R-CNN 等	理解其执行结果	堵判断逻辑	决方案及架构设计
5	红领西服个性化定制	个性化定制及其基础架构、柔性制造等	模拟完成客户体信息、推送及存档	按需增加几个个性化数据的采集	分析实现预期定制化效果需要解决的问题
6	猜你喜欢-推荐系统	推荐系统架构、相似度算法、协同过滤推荐	理解推荐流程, 修改代码实现特定的相似度计算及推荐	改用或加权其它距离函数, 比较推荐效果	无
7	工业大数据治理基础安全架构	PKI 安全总体架构、加密、哈希、签名算法等	完成对消息进行加密、哈希计算、签名及验证签名的过程	用一个函数完成原来分步执行的签名动作	设计企业和监管部门消息发送的基础安全方案

3 实验设计实例

本节以实验 4 为例, 阐述多层次实验内容的设计方法及该实验对学生能力的培养与锻炼。实验的业务场景如下: 给定一个城市里的十字路口, 要求设计出一个软硬件系统, 可以识别出是否有拥堵现象及拥堵程度, 若存在拥堵则通过一定方式通知给相关部门和人员, 例如道路交通监管部门和手机上安装了道路交通信息 APP 的用户等。然后设计实现该系统的拥堵判定算法, 给出是否拥堵的结论。

3.1 基础知识

本实验涉及的主要基础知识包括 CPS、TensorFlow 等, 简要介绍如下:

(1) CPS。广义上, CPS 是一个在环境感知的基础上, 深度融合了计算、通信和控制能力的可控可信可扩展的网络化物理设备系统, 它通过计算进程和物理进程相互影响的反馈循环实现深度融合和实时交互来增加或扩展新的功能, 以安全、可靠、高效和实时的方式监测或者控制一个物理实体^[12]。基于 CPS 的工业大数据基础架构^[11, 13]如图 1 所示。



图 1 基于 CPS 的工业大数据基础架构图

(2) TensorFlow: 谷歌推出的在工业界广受欢迎的深度学习框

架, 广泛应用于各类机器学习算法的编程实现^[14], 其对象检测模型可用来检测图片中的车辆数量。但该模型使用的是 Fast R-CNN (快速可重入卷积神经网络) 算法基于 Resnet 数据集训练出来的模型, 数据范围包括生活中的各种事物, 主要用于识别通用对象, 对车辆检测精度一般。实际应用中会用大量车辆数据训练出一个专用模型, 从而排除非目标对象, 大大提高检测速度和准确度。

3.2 设计/思考环节

实验内容: 进行路况模拟及拥堵警示解决方案及架构设计, 主要包括以下方面:

- (1) 路况感知以及数据采集方案设计;
- (2) 车辆信息数据处理及拥堵检测方案;
- (3) 路况信息可视化展示及拥堵警示设计;
- (4) 分析给出该解决方案的总体架构图。

设计思路:

(1) 数据采集方案可在合理位置和距离部署摄像头、传感器(如压力、距离传感器)等, 以适当频次采集数据; 或使用用户定位导航数据等;

(2) 设计数据存储、处理方法; 再根据车速、可接受的等待时间等设计出合理的拥堵判定算法;

(3) 将路况信息推送给相关系统和个人, 如将路况拥堵情况以红、橙、黄、绿等颜色显示在地图 app 上, 拥堵等待时间显示在主干道提示牌上等;

(4) 参考如图 1 所示的基础架构图, 选择适当的技术组件和应用, 设计上述方案的架构图。

3.3 技术实践内容

实验内容(2选1): (1) 编程实现你所设计的解决方案中的拥堵判定方法, 并测试给出结论;

(2) 理解并完成后面给出的基于图像识别的示例方案的可行性验证。

内容(2)的基础实践和进阶实践内容设计如下。

(1) 基础实践内容主要包括:

①完成实验依赖的安装部署, 包括 Tensorflow2 机器学习框架和 Tensorflow 模型库^[15];

②阅读并理解给出的基于 TensorFlow 的对象检测模型的车辆识别示例代码, 运行得到标注好车辆的图片, 如图 2 所示。

(2) 进阶实践内容: 理解示例代码框架, 在该代码文件中增加代码解读图片的识别结果, 并增加拥堵判断逻辑, 在屏幕上输出判定结果。实现提示:

①代码块 def run_detector(detector, path): 中, 语句 result = detector(converted_img) 会返回图像识别结果, 获取 result 中的内容的代码形如:

```
targets = result["detection_class_entities"].numpy()
targets_list = targets.tolist()
```

②用 filter 统计我们关心的对象, 包括 Taxi、Car、Bus、Land vehicle、Van 和 Truck 等, 可把车辆数量打印到屏幕上。代码形如:

```
cars = list(filter(lambda obj: obj == 'Car', targets_list))
```

权重及	实验内容平均分	课程目标达成度
-----	---------	---------

分值	基础 实践	进阶 实践	设计/ 思考	总分	课程 目标 1	课程 目标 2	课程 目标 3	课程 目标 4
占比	50	20	30	100	10	15	15	60
分值	43.4	13.4	17.4	74.2	0.859	0.553	0.607	0.839
百分制	86.8	67	58	74.2				

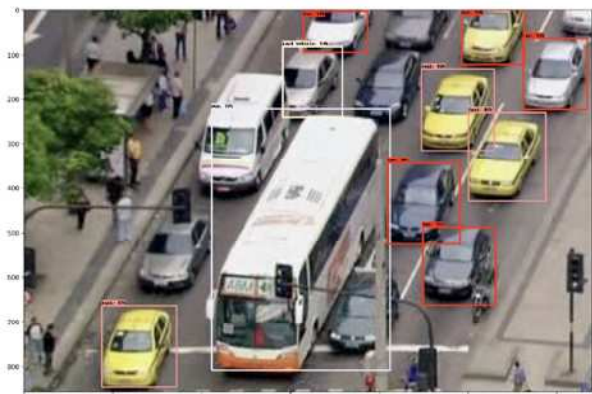


图2 图片基于 TensorFlow 的车辆标注结果

③根据统计到的车辆数量,判断路况是否拥堵,在屏幕上打印出结果。

3.4 培养提升的学生能力

本实验选取学生熟悉且有趣的路况拥堵警示场景和业界前沿主流的深度学习框架 TensorFlow,提升学生的学习兴趣和参与度,锻炼学生自主学习前沿技术的能力,并能将其与 CPS 进行结合,应用到业务场景的解决方案与架构的分析与设计,以及方案关键环节的验证实现中,从而锻炼了学生的技术应用实践能力和分析解决问题与创新实践能力。

4 教学效果分析

为使学生能更好地一起学习、实践、讨论解决方案及设计思路,实验采用 2-4 人一组的的方式进行,取得了较好的成效,多层次实验内容的平均分及课程目标的达成度如表 3 所示。

由表 3 可看出,实验成绩平均分为 74.2,达到了课程的预期目标,总体效果较好。其中基础实践和进阶实践内容成绩不错,其对应的课程目标 1 和课程目标 4 的达成度为 0.8 以上,效果很好;但设计/思考环节平均分仅为 58,其对应的课程目标 2 和课程目标 3 的达成度不够理想,原因在于工业领域场景的解决方案及架构设计对应用型本科学生来说难度较高且兴趣度也不高,在后续的教学改革实践中,需适当调整实验难度,并在实验过程监督及小组讨论引导等方面进一步进行探索与完善。

5 结语

通过在工业大数据架构与应用课程的实验教学中探索实践多层次实验内容设计,并借助行业热点及有趣的业务场景和企业前沿、主流的技术相结合,激发学生的学习兴趣,培养提升学生的自主学习能力、技术实践能力、分析解决问题和创新实践能力,以适应工业强国新变革时代对工业大数据创新应用型人才的新要求。

表 3 教学效果

权重及	实验内容平均分	课程目标达成度
-----	---------	---------

分值	基础 实践	进阶 实践	设计/ 思考	总分	课程目 标 1	课程目 标 2	课程 目标 3	课程目 标 4
占比	50	20	30	100	10	15	15	60
分值	43.4	13.4	17.4	74.2	0.859	0.553	0.607	0.839
百分制	86.8	67	58	74.2				

参考文献:

- [1] 李杰. 工业大数据—工业 4.0 时代的工业转型与价值创造[M]. 北京: 机械工业出版社, 2015.
- [2] 袁爱进, 岳滨楠, 闫鑫, 等. 工业大数据的应用与实践[J]. 大数据, 2017, 3(6):27-41.
- [3] 郑树泉, 覃海煊, 王倩. 工业大数据技术与架构[J]. 大数据, 2017, 3(4): 67-80
- [4] 张振山. 构建工业大数据安全管理体系[J]. 中国工业和信息化, 2021, (8): 26-30
- [5] 杨俊浩. 工业大数据在生产监控系统中的运用分析[J]. 无线互联科技, 2021, 18(19):88-89.
- [6] 苏兆品, 胡东辉, 张国富, 等. 信息论与编码课程内容的多层次设计[J]. 实验室研究与探索, 2021, 40(9):206-210.
- [7] 章喆, 吴谨. NB-IoT 实验室建设与实验课程设计[J]. 实验室研究与探索, 2021, 40(2): 252-255.
- [8] 崔晓龙, 张敏, 郭茜, 等. 基于 Hadoop 的大数据工程实践多层次教学设计[J]. 实验室研究与探索, 2021, 40(1): 157-161.
- [9] 覃海煊, 芦立华, 沈学东, 等. 软件工程专业学生多维能力培养探索与实践[J]. 福建电脑, 2021, 37(10): 118-121.
- [10] 中国工程教育专业认证协会. 工程教育认证标准, <https://www.cceaa.org.cn/gcjyzyrzh/rzcxjzb/gcjyrbz/tybz/599711/index.html>. 2017.11
- [11] 郑树泉, 宗宇伟, 董文生, 丁志刚. 工业大数据: 架构与应用[M].上海科学技术出版社, 2017.
- [12] 何积丰. Cyber-physical System [J]. 中国计算机学会通讯, 2010, 6 (1): 25-29.
- [13] 中国信息物理系统发展论坛. 信息物理系统白皮书. 中国电子技术标准化研究院, 2017.
- [14] Abadi, M., Barham, P., Chen, J., et al. (2016). Tensorflow: a system for large-scale machine learning. In OSDI (Vol. 16, pp. 265-283).
- [15] Tom Hope, Yehezkel S. Resheff, Itay Lieder. TensorFlow 学习指南: 深度学习系统构建详解[M]. 朱小虎, 李紫辉. 北京: 机械工业出版社, 2018.

基金项目: 上海电机学院电子信息学院工业大数据架构与应用课程建设(2021-1)

作者简介: 覃海煊(1978-), 女, 广西环江人, 博士, 讲师, 主要从事服务计算、工业大数据建模与分析方面的研究。