

大数据工程化处理与应用
人工智能前端设备应用

职业技能等级证书

2021年重点工作

王欣欣

有关试点工作事项 说明



01

对应专业

学校类型	对应相关专业
中等职业学校	大数据技术应用、计算机应用、软件与信息服务、移动应用技术与服务、计算机网络技术、网站建设与管理、客户信息服务、通信运营服务、电子商务等专业。
高等职业学校	大数据技术、云计算技术与应用、计算机应用技术、计算机网络技术、软件技术、统计与大数据分析、人工智能技术服务等专业。
应用型本科学校	大数据工程技术、计算机应用工程、信息安全与管理、软件工程、数据科学与大数据技术、大数据管理与应用、数据计算及应用、计算机科学与技术、信息管理与信息系统、应用统计学、信息资源管理等专业。

02

试点院校条件

1

具备办学许可的法人单位，已开设大数据工程化处理与应用职业技能等级证书所对应的相关专业，且该专业近3年连续招生。有3年以上相关专业学历教育与职业培训经验。

2

配备具有相应培训能力的专兼职师资团队，团队成员不少于5人，具有相应工作领域项目经验的行业企业兼职教师占比不少于30%，“双师型”授课教师不少于50%；专业带头人具有中级或高级职称。

3

围绕大数据工程化处理与应用职业技能等级证书相对应的专业，已开发了较为成熟的课程体系和专业教学资源。

02

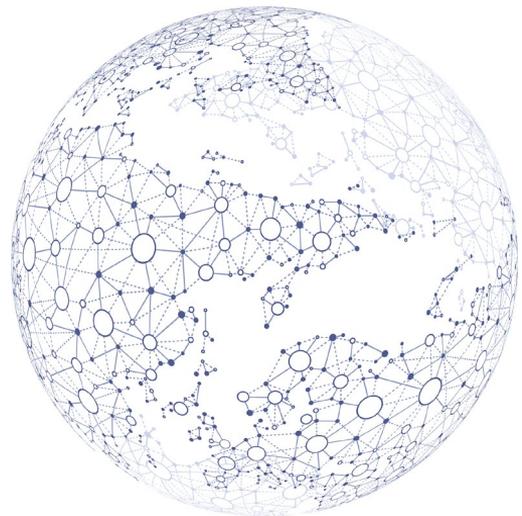
试点院校条件

4

具有大数据工程化处理与应用职业技能等级证书对应的培训和考核环境，可同时满足40人进行理论学习和实践操作。培训和考核环境包含稳定成熟且符合本职业技能等级证书考核要求的相关实训设备。

5

组织机构完善，具有满足大数据工程化处理与应用职业技能等级证书培训需求的管理团队，有固定办公场所。团队负责人能够充分调动资源，提供培训所需的保障条件。



人工智能前端设备应用职业技能等级证书 有关试点工作事项说明

01

对应专业

学校类型	对应相关专业
中等职业学校	电子信息技术、软件与信息服务、物联网技术应用、移动应用技术与服务、服务机器人装配与维护、工业机器人技术应用等专业。
高等职业学校	人工智能技术应用、软件技术、物联网应用技术、智能互联网络技术、智能产品开发与应用、计算机应用技术、嵌入式技术应用、工业机器人技术、智能控制技术等专业。
应用型本科学校	人工智能工程技术、计算机科学与技术、机器人工程、物联网工程技术、智能制造工程技术、电子与计算机工程、电子信息科学与技术、软件工程技术等专业。

02

试点院校条件

1

具备办学许可的法人单位，已开设人工智能前端设备应用职业技能等级证书所对应的相关专业，且该专业近3年连续招生。有3年以上相关专业学历教育与职业培训经验。

2

配备具有相应培训能力的专兼职师资团队，团队成员不少于5人，具有相应工作领域项目经验的行业企业兼职教师占比不少于30%，“双师型”授课教师不少于50%；专业带头人具有中级或高级职称。

3

围绕人工智能前端设备应用职业技能等级证书相对应的专业，已开发了较为成熟的课程体系和专业教学资源。

02

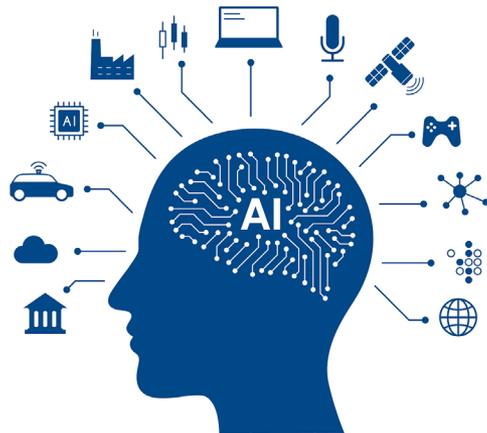
试点院校条件

4

具有人工智能前端设备应用职业技能等级证书对应的专业理论和实践教学场地，可同时满足30人进行理论学习和实践操作。

5

组织机构完善，具有满足人工智能前端设备应用职业技能等级证书培训需求的管理团队，有固定办公场所。团队负责人具备充分调动资源，提供培训保障的能力。



师资培训方案



培训目标

针对大数据工程化处理与应用、人工智能前端设备应用职业技能等级证书试点院校的教师，培养一支数量足够、结构合理、素质优良的专业知识丰富、实践能力强的高技能“双师型”教师队伍，打造能够满足教学与培训需求的教学创新团队。

师资团队建设

创建创新教学团队

建设专兼职师资团队，与行业企业加强合作，探索教产岗位互通、专任教师互聘机制，形成模块化的创新教学团队。

开展师资培训和交流

将职业技能等级证书有关师资培训纳入教师素质提高计划项目。定期参加师资培训和交流，提高教师实施教学、培训和考核评价能力。



02 培训内容

1

积极支持与参与教育行政部门组织开展的有关国培、省培师资培训项目。

2

根据试点工作需要，由培训评价组织分批次分区域分级别开展培训。培训采用线上线下相结合模式。

3

自行开展的师资培训由行业、企业及院校专家组成的培训专家队伍进行授课，主要授课形式包括现场讲授、案例分析与实践、研讨交流、考核等。



培训主题	培训内容	初级	中级	高级	培训形式
工程化数据采集	网络数据采集	✓	✓	✓	专家授课 + 现场实操
	离线数据采集	✓	✓	✓	
	作业调度	✓	✓	✓	
	实时数据采集			✓	
工程化数据处理	数据清洗	✓	✓	✓	
	数据计算	✓	✓	✓	
	数据拍摄	✓	✓	✓	
工程化数据应用	OLAP系统应用	✓	✓	✓	
	搜索系统应用	✓	✓	✓	
	报表系统应用	✓	✓	✓	
结业考试		✓	✓	✓	

培训内容	初级	中级	高级	培训形式
智能前端设备检测	✓	✓	✓	专家授课 + 现场实操
智能前端设备安装	✓	✓	✓	
智能前端设备调试	✓	✓	✓	
应用系统安装与配置	✓	✓	✓	
服务器和数据库安装配置	✓	✓	✓	
边缘端程序与模型部署	✓	✓	✓	
数数据采集与清洗	✓	✓	✓	
数据标注与转换	✓	✓	✓	
应用配置和调优		✓	✓	
模型微调训练与优化		✓	✓	
应用性能分析与监控			✓	
结业考试	✓	✓	✓	

考核方案、 考核站 点要求



大数据工程化处理与应用、人工智能前端设备应用 职业技能等级证书考核方案

01 考核方式

40% 理论考核

理论考试

实操考试

60% 实操考试

大数据工程化处理与应用

理论考核：上机考试，考试时长1小时，共25道题，满分100分，权重40%。

实操考试：上机考试考试时长2小时，共3-4道题，满分100分，权重60%。

人工智能前端设备应用

理论考核：考试时长60分钟，包含单选、多项选择题、判断题，共20道题，满分100分，权重40%。

实操考试：考试时长：初级120分钟，中高级150分钟，满分100分，权重60%

理论部分（样例）

大数据工程化处理与应用（初级）

使用可视化工具进行网络数据采集、离线数据采集、作业调度、数据清洗、数据计算、数据派生、OLAP系统应用、搜索系统应用、报表系统应用工作的能力。

大数据工程化处理与应用（中级）

编写脚本进行网络数据采集、离线数据采集、作业调度、数据清洗、数据计算、数据派生、OLAP系统应用、搜索系统应用、报表系统应用工作的能力。

大数据工程化处理与应用（高级）

使用框架进行网络数据采集、离线数据采集、作业调度、实时数据采集、数据清洗、数据计算、数据派生、OLAP系统应用、搜索系统应用、报表系统应用工作的能力。

1. Hive创建数据库的语句是（ ）。

- A、CREATE DATABASE
- B、SHOW DATABASES
- C、CREATE FUNCATIONS
- D、DESC DATABASE

实操部分（样例）

(1)数据观察：分别访问MySQL和HDFS，确定数据字段格式。（5分）

- ① 查询气象观测表的字段格式，将代码片段及代码执行结果保存到实验报告中。（2分）
- ② 查询气象站点信息的前5行，将代码片段及代码执行结果保存到实验报告的相应位置。（3分）

理论部分（样例）

人工智能前端设备应用（初级）

根据系统连接拓扑图和实施方案，完成智能前端设备的安装与配置，应用系统部署，数据采集与标注等工作。

人工智能前端设备应用（中级）

根据项目需求和目标场景，完成智能前端设备选型与调试，应用系统联调，目标场景数据采集与数据标注，使用工具软件进行应用配置，场景化模型微调训练等工作。

人工智能前端设备应用（高级）

根据项目需求和目标场景，完成系统连接拓扑和部署方案设计，边缘端模型部署优化，数据采集、清洗和增强，数据标注和转换，应用配置优化、场景化模型微调训练、应用性能分析与监控等工作。

1.人工智能产业链三层架构类型不包括（ ）

- A.基础层
- B.技术层
- C.应用层
- D.网络层

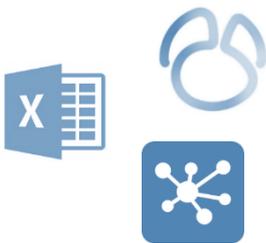
实操部分（样例）

(1)前端设备安装与调试：请根据给出的智慧校园人员车辆智慧管理场景的网络设备拓扑图，以及人工智能前端设备应用实训平台设备，按照安装连接要求，安装搭建至人工智能模块化工位，对设备进行正确接线，并对人工智能边缘网关、人脸门禁一体机、智能摄像头等设备进行正确的配置。

初级

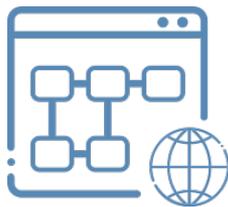


图形化工具



...

中级

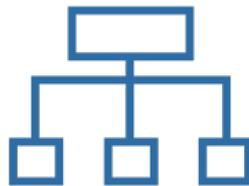


脚本化工具

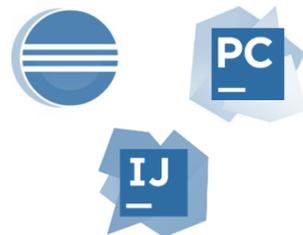


...

高级



项目化工具



初级

执行

设备安装部署

.....

应用系统部署

.....

数据采集与标注

.....

中级

分析与执行

根据需求选型与安装
设备

.....

根据需求进行系统部
署与联调

.....

根据场景进行数据采
集与标注

.....

模型微调训练

.....

高级

设计与优化

实施方案设计与实施

.....

应用部署方案设计与
优化

.....

数据采集和标注方案
设计

.....

场景化模型微调训练
与优化

.....

1 前端设备
安装与调试2 应用
系统部署3 数据
采集与标注4 场景化
应用与优化



1+X大数据工程化处理与应用、人工智能前端设备应用职业技能等级证书理论考试试卷满分100分，权重40%；
实操考试试卷满分100分，权重60%。
理论考试和实操考试合格标准为两项成绩相加大于等于60分，可以获得相应级别的职业技能等级证书。

1 考核场地要求

- ☑ 场地要求有理论考核场地、实操考核场地，及保密室。理论考核场地要求不少于40人。
- ☑ 考核场地建有过程监控系统，确保能够实施考核全过程音频、视频信息采集与存储。



2 考务管理人员要求

- ☑ 应设有健全的组织管理机构，能满足考核工作需要的专兼职管理人员（5人以上）。
- ☑ 按每40人标准考场，安排2位监考人员。



大数据工程化处理与应用职业技能等级证书 考核站点要求

设备名称	具体要求	建议数量
计算机	CPU: i7; 内存: 8GB; 硬盘: 500GB ; 显示器: 1920*1080; USB2.0接口:至少1个。	40台
服务器	CPU: 16核32线程; 内存: 256GB; 硬盘: 8TB; 网卡: 10/100/1000MB/s; 操作系统: Centos7.6。	3台
实验环境 (初级)	可支持本证书初级考证的实验环境应至少集成以下组件: 数据库: MySQL; 离线数据采集工具: Sqoop; 文件系统: HDFS; 数仓工具: Hive; 非结构化数据库: HBase; ETL工具: Kettle; 联机事务分析系统: Kylin; 可视化报表工具: Kibana; 作业调度工具: Azkaban; 备注: 需安装谷歌浏览器80以下版本或火狐浏览器。	1套
实验环境 (中级)	可支持本证书中级考证的实验环境在初级环境基础上应增加集成以下组件: 网络数据采集依赖: Requestium; 基础编辑语言: Python; 数据检索系统: Elasticsearch; 备注: 需安装谷歌浏览器80以下版本或火狐浏览器。	1套
实验环境 (高级)	可支持本证书中级考证的实验环境在中级环境基础上应增加集成以下组件: 日志采集工具: Flume; 消息队列: Kafka; 数据计算框架: Spark; SQL查询引擎: Presto; 备注: 需安装谷歌浏览器80以下版本或火狐浏览器。	1套
网络环境	带宽: 100/1000Mbps; 端口: 40个。	1台

1 考核场地要求

- ☑ 场地要求有理论考核场地、实操考核场地，以及保密室。
- ☑ 考核场地建有过程监控系统，确保能够实施考核全过程音频、视频信息采集与存储。



2 考务管理人员要求

- ☑ 应设有健全的组织管理机构，能满足考核工作需要的专兼职管理工作人员（5人以上）。
- ☑ 按每33人标准考场，安排2位监考人员。



人工智能前端设备应用职业技能等级证书 考核站点要求（实验设备配置）

设备名称	具体要求
计算机	操作系统：Windows10；内存：8GB以上 硬盘：128GB以上；CPU：i7 宽带：10Mbps以上（外网+局域网）；主机接口：至少1个串口，4个USB接口
人工智能前端设备 应用实操设备	<ol style="list-style-type: none">(1) 人工智能实训工位，数量：1(2) AI边缘网关，数量：1(3) 人脸门禁，数量：1(4) 枪型网络摄像机，数量：1(5) 传感器，数量：1(6) USB摄像头，数量：1(7) 人体红外、温湿度、火焰传感器，数量各1(8) 继电器控制器模块，数量：1(9) 全向麦克风，数量：1(10) 声光报警器，数量：1(11) 灯泡、风扇、智能柜子锁，数量各1(12) 路由器、交换机，数量各1

人工智能前端设备应用职业技能等级证书 考核站点要求（实验环境）

名称	具体要求
人工智能实验环境（初级）	人工智能前端设备应用实验环境（初级）应至少集成以下组件： 数据标注软件：Colabeler标注工具；终端工具：XShell 5、Xftp 5；Python版本：Python3.5,数据库：MySQL；Web服务器：Nginx；浏览器：谷歌浏览器；
人工智能实验环境（中级）	人工智能前端设备应用实验环境（中级）应至少集成以下组件： 数据标注软件：Colabeler标注工具、Labelme、Parrrt；终端工具：XShell 5、Xftp 5；Python版本：Python3.5数据库：MySQL；Web服务器：Nginx；浏览器：谷歌浏览器；开源python发行版本工具包：Anaconda；Web交互式平台：Jupyter Notebook；
人工智能实验环境（高级）	人工智能前端设备应用实验环境（高级）应至少集成以下组件： 数据标注软件：Colabeler标注工具、Labelme、Parrrt；终端工具：XShell 5、Xftp 5；Python版本：Python3.5,数据库：MySQL；Web服务器：Nginx；浏览器：谷歌浏览器；开源python发行版本工具包：Anaconda；Web交互式平台：Jupyter Notebook；Python 代码编辑器：Pycharm Community；

配套资源开发说明



1

实训教材准备情况

大数据工程化处理与应用职业技能等级标准将对应其初级、中级、高级的要求，配套开发3本实训课程教材。目前对应中级的教材已开发完成，提报出版社走出版流程中。初级教材预计于2021年6月-9月开发，10月可进入出版流程。高级教材预计于2021年底开始开发。

2

教辅准备情况

教学辅助资源围绕配套大数据工程化处理与应用职业技能等级的初中高级实训教材分别开发。教学辅助资源分免费与付费两个层级。免费资源将与教材绑定提供给客户院校，资源类型包括：教学大纲、教学课件、项目实验手册、实验源代码、实验所需环境安装软件或插件。付费资源将在免费资源的基础上，增加教学讲解视频、精品教学课件、实验操作视频、知识点讲解动画等类型的精品资源。中级免费资源可跟据证书运营情况陆续发布，付费资源预计发布时间为21年8月。初级免费资源预计21年9月陆续发布，付费资源预计于21年年底发布。

01

信息化资源的类型

慕课资源

教辅资源



解决
学生
问题



解决
老师
问题



TensorFlow由Google开源，是一款使用 C++ 语言开发的开源数学计算软件，使用数据流图（Data Flow Graph）的形式进行计算。图中的节点代表数学运算，而图中的线条表示多维数据数组（tensor）之间的交互。

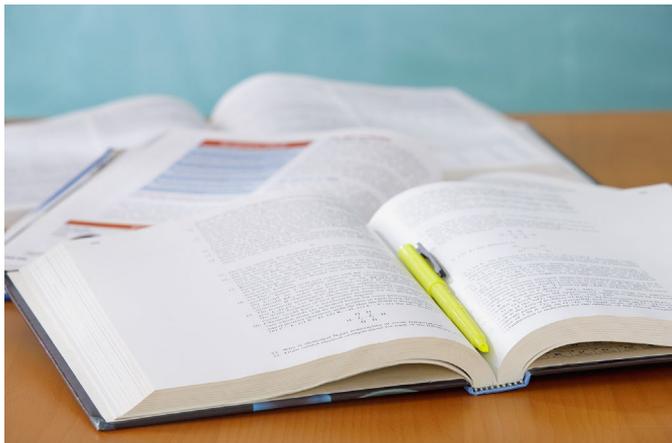
TensorFlow灵活的架构可以部署在一个或多个 CPU、GPU 的台式及服务器中，或者使用单一的 API 应用在移动设备中。

由于其语言太过于底层，目前有很多基于 TensorFlow 的第三方抽象库将 TensorFlow 的函数进行封装，使其变得简洁，目前比较有名的几个是 Keras, Tflearn, tflsim, 以及 TensorLayer。



- ☑ 完整的一堂课，可直接替代线下教师
- ☑ 对学生学习主动性要求高，更适合成人或普教学生
- ☑ 无法培养职业教育学生的动手能力
- ☑ 无法提升教师水平，发挥教师的能动性

- ☑ 完整的一堂课，可直接替代线下教师
- ☑ 对学生学习主动性要求高，更适合成人或普教学生
- ☑ 无法培养职业教育学生的动手能力
- ☑ 无法提升教师水平，发挥教师的能动性



人工智能前端设备应用职业技能等级标准将对应其初级、中级、高级的要求，配套开发3本实训课程教材。目前对应的中级教材预计在6月底开发完成，并提报出版社走出版流程。初级教材预计于2021年7月开始开发，10月进入出版流程。高级教材预计于2021年底开始进行开发。

信息化精品教辅资源样例展示

PPT：带有教学设计

教案：带有教学设计，详细并与PPT对应

引导视频：帮助教师说明真实工作场景应用

任务描述与要求视频：帮助教师说明实训设备与实际应用的关联性

理论动画：有助于教师讲解重点、难点

实操视频：帮助教师减轻实训指导1对多的负担，学生可根据视频完成实训项目



有经验的教师



资深教师



新教师

快速上手（60分）

教辅资源与教材完全匹配且类型丰富，保证新教师快速入门

赋能教师（80分）

教辅资源作为教师教学的素材赋能教师，教师逐步形成自己特有的教学套路

发挥教魂（100分）

教师不断优化自己的教学设计，沉淀优秀案例分享交流，发挥出真正的教学灵魂

持续满足教师培养需求
有效规范教师教学内容
不断优化教师教学效果

感谢聆听

THANKS FOR YOUR WATCHING