

工信部教考中心《大数据治理工程师》简介

一、项目简介

大数据治理工程师是由工业和信息化部教育与考试中心推出的一套针对大数据治理领域专业人才的考核体系，涵盖了从基础到高级的全面技能要求，旨在培养能够在互联网、金融等众多行业中，运用大数据治理技术进行数据整合、处理、分析与应用的专业人才，以提升企业数据管理水平，挖掘数据价值，助力企业数字化转型和决策优化。

“大数据治理工程师”证书分为初级、中级、高级三个等级，每个等级对相关从业人员的知识和技能要求逐步提升，形成了一套完整的人才培养和评估体系。

二、能力标准

1、大数据治理工程师（初级）

掌握Linux操作系统的基本操作，具备一定的Python数据分析基础知识，了解数据分析的基本流程，能够进行简单的数据预处理和分析任务。

对ETL大数据整合与处理有初步认识，理解数据抽取、转换和加载的基本概念，能够在指导下完成简单的数据处理任务，如创建数据库连接、进行基础的数据处理操作等。

2、大数据治理工程师（中级）

熟练掌握Linux操作系统的进阶技能，能够编写简单的脚本实现自动化任务，提升工作效率。深入理解Python数据分析与挖掘技术，能够运用数据分析工具和技术解决实际业务问题。

熟练掌握 ETL 大数据整合与处理流程，能够独立完成数据处理任务。具备一定的数据存储和管理知识，了解数据仓库的基本概念，能够进行数据库的简单管理和维护。

3、大数据治理工程师（高级）

精通 Linux 操作系统，能够对系统进行优化配置，解决复杂的系统故障和性能问题，确保大数据处理环境的稳定运行。能够运用高级数据分析算法和技术进行深度数据挖掘，为企业提供数据驱动的决策建议。

熟练掌握 ETL 大数据整合与处理技术，能够设计和优化大型数据处理流程，处理海量数据的集成和转换任务。深入理解数据存储技术，能够构建高效的数据存储和管理系统。

掌握多种大数据实时处理框架，满足企业对实时数据处理的需求。在机器学习进阶知识领域，熟悉各类算法、集成学习和强化学习等技术，能够将这些技术应用于实际项目中，解决复杂的业务问题。

具备大数据治理项目的整体规划和管理能力，能够带领团队完成大数据治理项目的实施，确保项目目标的达成。

三、课程内容

阶段	课程	知识模块
基础知识	Linux操作系统基础	Linux概述与系统安装 Linux基本命令 Linux Vi 编辑器 Linux Shell 编程
	Python数据分析与挖掘技术	数据分析概述 NumPy数值计算基础 Pandas统计分析基础 Pandas数据预处理
	ETL大数据整合与处理	认识ETL 创建数据库连接 数据处理

		特征字段构建 记录集连接 数据输出 数据检查
大数据治理技术	数据存储技术	数据仓库 数据库管理系统与数据存储 查询事务管理与系统结构 分布式文件系统概述 环境配置 Hadoop集群配置与启动 Hadoop安全模式与HDFS操作 非关系型数据库概述 MongoDB部署模式与数据分片 系统部署单节点MongoDB Linux系统部署MongoDB副本集与分片集群 数据库管理工具
	Zookeeper分布式服务框架	大数据传输 分布式服务框架Zookeeper基础及安装 Znode创建与查看 权限设置与查看
	Storm大数据实时处理	大数据实时处理框架Storm简介及安装 Storm开发环境配置实战 Storm集群运行介绍及相关命令实战 Storm集成Hadoop API Storm集成Kafka
	Flink大数据实时处理	大数据实时处理框架Flink概念及应用场景 Flink数据模型/时间窗口及核心概念 Flink安装启动与监控 Flink基础API及核心数据结构 Flink开发环境配置及运行实例 Flink DataStream API编程 Flink与Kafka整合实例
机器学习进阶知识	分类与回归	回归分决策树算法介绍 泰坦尼克号生还者预测 人工神经网络介绍 单样本网络训练 全样本网络训练 K近邻算法 朴素贝叶斯 支持向量机算法介绍 支持向量机代码实现 市财政收入分析及预测析
	聚类分析	聚类分析概述密度与层次聚类 根据刷卡经纬度信息分析乘客上车站点与实现
	推荐算法	智能推荐系统简介 关联规则 协同过滤 FM&Ranking模型
	时间序列	时间序列原理与实现

		脑卒中发病环境因素分析及干预 金融服务机构资金流量预测
	集成学习与强化学习	集成学习概述与实现 强化学习概述与实现 网络入侵用户自动识别
职业技术考核	大数据治理工程师（高级）职业技术考试	

四、报考条件

初级：无要求，皆可报考。

中级：（满足其中一个条件即可）

- 1.获得初级证书。
- 2.年满18周岁，具备高中以上学历，工作年限满1年

高级：（满足其中一个条件即可）

- 1.获得中级证书。
- 2.年满20周岁以上，工作年限满2年
- 3.年满20周岁以上，具备大专及以上学历

五、考试方式

考试方式分为线下考试站点或线上考试系统统考两种方式，考试形式为上机答题，闭卷。

考试题型：

初级：120 分钟，单选题+多选题+判断题+操作题+简答题，上机答题。

中级：120 分钟，单选题+多选题+判断题+操作题+简答题，上机答题。

高级：120 分钟，单选题+多选题+判断题+操作题+简答题，上机答题。

六、成绩评分

考试最终成绩满分为100分，成绩80-100分为优秀；成绩60-79分为合格；成绩60分以下为不合格。

七、证书样本

学员经考核合格，由工业和信息化部教育与考试中心颁发大数据治理工程师职业技术证书，证书可登录国家工信部教育与考试中心官网查询。

