

CDA LEVEL III 数据科学家考试大纲

CERTIFIED DATA ANALYST LEVEL III EXAMINATION OUTLINE

一、总则

「CDA 数据分析师人才行业标准」是面向全行业数据分析及大数据相关岗位的一套科学化、专业化、正规化、系统化的人才技能准则。经管之家 CDA 数据分析师认证考试是评判「标准化人才」的唯一考核路径。CDA 考试大纲规定并明确了数据分析师认证考试的具体范围、内容和知识点，考生可按照大纲要求进行相关知识的学习，获取技能，成为专业人才。

二、考试形式与试卷结构

包括两个阶段，通过第一个阶段，才有资格参加第二个阶段考试。

第一阶段：150 分钟，客观题+主观题，闭卷，上机答题。

第二阶段：提供项目案例，1 个月内完成，开卷。截止日前，提交项目过程和结果，60 分钟，线上答辩面试。

考试成绩：分为 A、B、C、D 四个层次，A、B、C 为通过考试，D 为不通过。

三、知识要求

针对不同知识，掌握程度的要求分为【领会】、【熟知】、【应用】三个级别，考生应按照不同知识要求进行学习。

1. 领会：考生能够领会了解规定的知识点，并能够了解规定知识的内涵与外延，了解其内容要点和它们之间的区别与联系，并能做出正确的阐述、解释和说明。

2. 熟知：考生须掌握知识的要点，并能够正确理解和记忆相关理论方法，能够根据不同要求，做出逻辑严密的解释、说明和阐述。此部分为考试的重点部分。

3. 应用：考生须学会将知识点落地实践，并能够结合相关工具进行商业应用，能够根据具体要求，给出问题的具体实施流程和策略。

四、考试范围

◆ PART 1 计算机科学技术（占比 15%）

- a. 大数据的高级处理技术（占比 5%）
- b. 高性能编程计算（占比 6%）
- c. 常用机器学习框架（占比 4%）

◆ PART 2 大数据处理与架构设计（占比 15%）

- a. 大数据架构设计的方法论概述（占比 3%）
- b. 互联网场景的大数据解决方案设计（占比 5%）
- c. 大数据存储与计算的方案选型（占比 2%）
- d. 大数据指标系统与数据安全（占比 2%）
- e. 集群资源管理、调优（占比 3%）

◆ PART 3 机器学习（占比 25%）

- a. 特征选择与稀疏学习（占比 4%）
- b. 类别不平衡问题（占比 4%）
- c. 决策规则（占比 2%）
- d. 半监督学习（占比 1%）
- e. 强化学习（占比 2%）
- f. 文本挖掘（占比 4%）
- g. 社会网络分析（占比 4%）
- h. 区块链分析（占比 4%）

◆ PART 4 深度学习（占比 20%）

- a. 感知机与神经网络（占比 2%）
- b. 深度学习基础概念（占比 3%）
- c. 深度学习模型训练与优化（占比 3%）
- d. 深度学习神经网络-DNN/CNN/RNN/LSTM 神经网络（占比 6%）
- e. 生成式对抗网络（占比 2%）

- f. 深度学习在物体检测与定位上的应用（占比 1%）
- g. 深度学习在人脸识别上的应用（占比 1%）
- h. 深度学习在语音识别上的应用（占比 1%）
- i. 深度学习的未来发展趋势（占比 1%）

◆ **PART 5 数据治理（占比 15%）**

- a. 大数据治理概述、大数据建模（占比 3%）
- b. 元数据管理、数据体系建设（占比 3%）
- c. 大数据隐私、安全、立法（占比 3%）
- d. 大数据质量、热度（占比 3%）
- e. 大数据生命周期模型（占比 3%）

◆ **PART 6 项目管理（占比 10%）**

- a. 软件项目管理基础（占比 2%）
- b. 敏捷开发（占比 2%）
- c. 代码管理（占比 2%）
- d. 构建大数据团队（占比 2%）
- e. 项目管理相关知识及常用工具（占比 2%）

五、考试内容

PART 1 计算机科学技术

◆ 大数据的高级处理技术

1. 领会：Python、Java、Scala 等编程语言的特点和应用场景。
2. 熟知：Python、Java、Scala 对大数据的多线程编程，并行计算，及第三方常用类库等高级处理技术。

◆ 高性能编程计算

1. 领会：影响性能（运行时间及内存消耗）的因素，衡量性能的方法。
2. 熟知：加速运行的常用方法，使用编译代码加快运行速度，将数据处理交给数据库系统，并行计算（模型并行、数据并行、混合并行）提升运行速度，使用 GPU 加快运行速

度，减少内存使用的常用方法，使用有限的内存处理大型数据集。

3. 应用：搭建高性能计算环境及大数据处理的实作。

◆ 常用机器学习框架

1. 领会：Tensorflow 原理和系统架构、计算图、张量（Tensor）、会话、流（Flow）等基本元素，TensorBoard 实现方式，理解 Keras 的易用性、灵活性等特点。

2. 熟知：Tensorflow 前端系统和后端系统，构建和运行计算图，Keras 定义常见网络的方法及参数含义，Scikit-Learn、TFLearn 等算法库使用方法。

3. 应用：基于 Tensorflow 实现线性回归算法，并用 TensorBoard 记录图结构和各项运行指标；使用 Keras 实现 LeNet 网络的结构设计和训练，使用 Scikit-Learn 实现 DBSCAN 聚类，使用 TFLearn 实现 CNN 和 RNN 做分类并做预测。

PART 2 大数据处理及架构技术

◆ 大数据架构设计的方法论概述

1. 领会：大数据分层架构设计的思想，技术架构视图的概念及涉及范围，大数据处理框架选择，服务总线思想，基于大数据的机器学习架构，大数据架构发展趋势。

2. 熟知：分层架构设计的过程和内容、总体架构设计的工具和方法，通用大数据处理流程及主要环节（如采集、预处理、存储、处理、监控等）。

3. 应用：能运用架构设计的方法体系进行企业信息化架构设计的实现。

◆ 互联网场景的大数据解决方案设计

1. 领会：针对用户行为分析的架构设计。

2. 熟知：了解收集用户的准备埋点规范、实施步骤，了解数据流采集、计算和可视化，了解。

3. 应用：了解 OLAP 分析在企业级别应用的演进方式和工程效率提升。

◆ 大数据存储与计算的方案选型

1. 领会：海量存储、离线计算、在线计算、流式计算四种常见的大数据分析场景的区别与联系。

2. 知晓：HDFS、Hbase 等常用海量存储工具，MapReduce、Hive、Dremel、Drill、Impala 等离线计算工具，Kylin、Redis、MongoDB 等在线计算工具，Flink、Storm、Spark 等流式

计算工具，Zookeeper、Spark、Kafka 等常用大数据工具。

3. 应用：实时流和离线数据整合的架构设计。

◆ 大数据指标系统与数据安全

1. 领会：大数据资源管理通用架构，资源监控平台架构，集群安全管理，标准化异常处理流程，数据的安全体系介绍。指标如何在元数据进行定义、规范化和准入的平台化设计。

2. 熟知：数据脱敏动态和离线存储的安全设计方式。

3. 应用：数据的使用审计、追溯，用户的授权功能最少、时间最短的实践方式。指标应用的热度，指标动态 SQL 指导。

◆ 大数据处理性能调优、集群优化、实时计算

1. 领会：存储性能优化，实时计算优化，Lambda 架构思想，大数据组件化选型。

2. 熟知：YARN 和 Impala、Spark 的优化，缓存应用机制、资源硬件分配方案、资源动态调度等配置。

3. 应用：可基于 Hbase 实现数据的存储和查询方案设计。

PART 3 机器学习

◆ 特征选择与稀疏学习

1. 领会：特征工程的目标，特征的构造及压缩、特征的选择、及特征提取的基本原理和思想。

2. 熟知：不同特征构建、压缩及选择的方法，Pearson 相关系数、信息价值法(Information Value)，基尼指数(Gini Index)、信息增益法(Information Gain)，增益比例法(Gain Ratio)，压缩感知方法及应用，主成分分析（PCA）降维算法，SVD 降维算法。

3. 运用：能利用工具针对不同类型样本进行特征的构建、压缩及选择。

◆ 类别不平衡问题

1. 领会：不平衡数据定义，不平衡数据场景，传统学习方法在不平衡数据中的局限性，类别不平衡所造成的问题。

2. 熟知：类别不平衡问题的检测方法，过采样技术(Over-sampling)，欠采样技术(Under-sampling)，模型惩罚技术。熟知 EasyEnsemble 算法，BalanceCascade 算法，SMOTE 算法，Borderline-SMOTE 算法，ADASYN 算法，Ensemble 算法，并对各种算法进行评价。

3. 应用：能运用类别不平衡的处理技术，提升分类模型的分类效能。

◆ 决策规则

1. 领会：决策规则与决策树的关系，决策规则适用的场域。

2. 熟知：决策规则的优点及缺点，不同决策规则的算法，PRISM 算法，PART 算法，JRip 算法。

3. 应用：能运用决策规则，提升决策树的分类效能。

◆ 半监督学习

1. 领会：监督学习、无监督学习及半监督学习间的关系。半监督学习的基本思想。

2. 熟知：半监督学习的基本假设(平滑假设(Smoothness Assumption)、聚类假设(Cluster Assumption)、流形假设(Manifold Assumption))，半监督分类，半监督回归，半监督聚类，半监督降维。掌握基于 SVM 的半监督学习算法，基于核方法的半监督学习算法，EM 半监督学习算法。

3. 应用：能运用半监督学习，降低开发决策模型的成本。

◆ 强化学习

1. 领会：行为主义理论，强化学习基本原理，强化学习的实现过程和应用领域，在线策略、离线策略，马尔可夫决策过程的原理。

2. 熟知：值迭代求解、策略迭代求解、Q-learning 等，智能体、环境、状态、动作和反馈(reward)等基本概念。值函数求解：动态规划方法，蒙特卡罗方法，时间差分方法。策略函数、Q-函数的求解：DQN、A3C。

3. 应用：能运用强化学习，提升决策系统的效能。

◆ 文本挖掘

1. 领会：文本挖掘与数据挖掘的关系，文本挖掘的分析流程及相关应用。

2. 熟知：分词、词形归一化、词性标注、句法分析、语义分析、语境分析，文本特征提取与表示(表示模型：布尔模型、向量空间模型、概率模型、图空间模型。特征选择方法：TF-IDF、信息增益(IG)、互信息、LDA、Word2Vec、GloVe、向量空间模型等)，语言模型、N-Gram，知识图谱常用加工、存储、表示工具。

3. 应用：能够运用文本挖掘，进行文本分类、舆情分析、文本聚类、问答系统、自动文摘相关应用。

◆ 社会网络分析

1. 领会：社会网络的重要性及应用。
2. 熟知：社会网络表示方式，网络密度，网络节点进出程度，扩散分析，群组分析，社会相似性，分割群组，群组与群组成员描述（社群领袖分析）。
3. 应用：能够运用社会网络分析进行扩散分析/群组分析模型建置、选择模型及算法参数调整。

◆ 区块链分析

1. 领会：区块链的含义、来源、发展、分类、特征，及基础架构模型。
2. 熟知：区块链分析在智能合约、证券交易、电子商务、物联网、社交通讯上的运用方式。
3. 应用：能够运用数据挖掘与文本挖掘技术于区块链的应用分析上。

PART 4 深度学习

◆ 感知机与神经网络

1. 领会：感知机的网络结构和神经网络的神经元的原理，领会感知机的学习规则和神经网络训练；神经网络的训练分解；领会两者的权重和阈值概念，领会神经网络的算法原理。
2. 熟知：感知机的计算过程和神经网络的常见激活函数的原理，熟知神经网络迭代次数的控制，熟知神经网络过度学习的现象。
3. 应用：结合给定的数据集，完成对数据的输入和输出，并能控制其中的核函数，网络层数、隐藏神经元数量的优化，得到比较准确的预测结果，并完成对结果的评价。能绘出神经网络的精度折线图，并能控制好迭代和学习的精度防止过度学习。

◆ 深度学习基础概念

1. 熟知：基于梯度的学习、隐藏单元、激活函数、损失函数、反向传播、范数惩罚、多任务学习、稀疏表示、集成深度学习的原理，卷积神经网络和循环神经网络，深度学习中的结构化概率模型和生成模型。（随机）梯度下降算法、动量算法、优化策略和元算法。常见激活函数特性，损失函数，学习率优化算法，线性因子模型。
2. 应用：熟练掌握独立成分分析 ICA，数据的白化在深度学习中的应用，熟练掌握极大似然估计、贝叶斯估计、（无）监督学习的各种算法在深度学习中的应用，掌握 Bernoulli

输出分布的 Sigmoid (含 logistic) 单元, 掌握 BP 的计算方法, 熟练掌握深度学习的正则化技术, 掌握具有自适应学习速率的算法, 能用卷积神经网络、循环网络进行深度学习构建。

◆ 深度学习模型训练与优化

1. 熟知: 深层神经网络训练原理, 过拟合、欠拟合产生原因及解决方法, 数据增强原理, Google NIN 及 Inception 实现原理, 深度残差网络原理。特征工程、重采样等训练样本优化。激活函数选择、权重初始化、网络结构设计、学习率、正则化、batch/Epoch 大小设置、Dropout 等性能调优策略, 学习曲线、验证曲线、损失曲线表示意义, 模型训练的中间结果可视化方法, TensorBoard 训练过程可视化方法。

2. 应用: 能够结合不同的业务数据特点选择不同的网络结构并对模型超参进行设置。

◆ 深度学习神经网络-DNN/CNN/RNN/LSTM 神经网络

1. 熟知: 卷积神经网络、循环神经网络、长短期记忆网络的原理及结构特点, 卷积、池化的本质, 解码器-编码器设计思路及缺陷, 注意力模型的原理和特点。卷积算法, 池化算法, 通道、参数共享、稀疏连接、特征图、全连接等基本操作实现方法、Softmax、SVM 等输出层特性, 反卷积算法, RNN 常见分类, RNN 主要缺陷, 输入门、遗忘门、输出门控制过程, LSTM 存在问题及 GRU 特点。熟悉 LeNet, AlexNet, GoogLeNet 基本架构和原理。

2. 应用: 能够基于卷积神经网络的原理将其应用于非图片处理领域中, 将 LSTM 用于股票、天气、销售情况等具有时序特点的领域中。

◆ 生成式对抗网络

1. 领会: 生成式对抗网络理论思想和发展趋势。

2. 熟知: 自动编码器 AE、变分自编码器 VAE 原理, 生成模型与判别模型协作过程, GAN 的训练过程, 目前 GAN 存在的主要问题, DCGAN、Pix2pix、CGAN、WGAN 等常见生成式对抗网络结构及改进。

3. 应用: 基于 GAN 及其变种网络生成数字图片, 基于现有图片集训练生成式对抗网络并生成新图片。

◆ 深度学习在物体检测与定位上的应用

1. 领会: ROI、NMS、SS 等物体检测的基本概念, IoU、mAP 等评价标准, DPM、R-CNN、Fast R-CNN、Faster R-CNN、R-FCN、RPN、YOLO、SSD 等物体检测模型及其特点。ImageNet、COCO 等目前常见图像数据集。

2. 熟知：物体检测的实现原理，卡尔曼滤波、粒子滤波、mean-shift 等实现方法及 TLD (Tracking Learning Detection)、CSK 等目标跟踪算法。

3. 应用：能够基于 YOLO 框架实现物体检测与识别。

◆ 深度学习在人脸识别上的应用

1. 领会：人脸识别的主要分类（检测、识别），基于生物特征的识别原理，人脸识别的难点，人脸图像特点，人脸图像的构成要素（内部属性、光源等成像条件、摄像机参数等），传统人脸识别方法及局限性。

2. 熟知：人脸建模方法，矩形特征（Harr-like）等人脸特征表示方法，ASM 模型基本思路，人脸关键点模型（dlib）DeepID/DeepID2/DeepID2+、DeepFace、FaceNet 等基于深度学习的人脸识别模型特点及训练过程，基于几何特征的人脸识别技术原理，基于深度学习识别模型原理，活体检测的原理及常见方法。熟悉 VGGFace，Caffe-face 和 Lightened CNN。

3. 应用：基于人脸关键点模型实现人脸合成或替换，能够基于人脸识别实现人物图片的检测和比对。

◆ 深度学习在语音识别上的应用

1. 领会：频谱、基频、声响、响度、音高等声学物征，音素、元音、辅音、清音、浊音等基本名词，语音识别及合成的基本难题。

2. 熟知：基音周期、基音频率等语音信号参数，音频加窗、FFT 变换等音频信号处理方法。语音合成和语音识别的原理，音频质量评价方式和标准，MFCC、PLP 特征提取方法，心理声学模型，基于 GAN 理论实现语音生成的过程，端到端的深度学习语音识别框架。语音识别的算法：DNN-HMM，RNN-CTC，LSTM-DNN，FSMN。

3. 应用：能够基于现有的 LSTM 或端到端模型实现语音识别系统的基本框架搭建，能够设计基本的音频处理解决方案，并能确定模型的输入音频特征及损失函数中关键指标。

◆ 深度学习的未来发展趋势

1. 领会：深层网络模型在深度和宽度上的趋势，强人工智能，神经生物学研究方向，目前深度学习的瓶颈，对偶学习、迁移原理，梯度下降训练方法的改进，权值共享、量化、二进制神经网络等模型压缩方法，自动化模型调参（AutoML）原理，多任务学习技术。

PART 5 数据治理

◆ 大数据治理概述、大数据建模

1. 领会：数据治理的概念、框架、目标，数据治理能力成熟度模型（DMM）及分级实施。

2. 熟知：了解大数据数据数据建模的基本知识，建立数据仓库的分类，元数据维度表、数据分析主题事实表，并且掌握雪花模型、星型模型，通过工具建模，完成 OLAP 分析的数据流、可视化、数据的上卷下钻等分析，能在较短周期内快速交付、灵活支撑业务场景。

◆ 元数据管理、数据体系建设

1. 领会：元数据的概念、存储、管理，数据标准的概念。

2. 熟知：元数据的编码体系：语法编码和受控词汇表，都柏林核心元数据元素集，基于业务元数据、技术元数据、操作元数据间的关联关系构建元数据模型，了解元数据管理自动化，熟悉业务词库的建立和数据标准的实施。

3. 应用：建立企业级元数据体系建设，创建企业级元数据技术规范和实施指引，制定合适的元数据管理流程。

◆ 大数据隐私、安全、立法

1. 领会：个人可识别信息 PII，识别敏感的大数据，熟悉国外隐私相关法律要求（参阅《国际数据保护规则要览》）、以及国内大数据立法的历程和展望（参阅《中国大数据法治发展报告》）。

2. 熟知：对元数据库中的敏感大数据进行标记，隐私影响评估 PIA，监控特权用户对敏感大数据的访问，管理个人数据跨国界流动的情况，熟知欧盟《通用数据保护条例》（General Data Protection Regulation，简称 GDPR）中企业和个人的数据使用权限。

3. 应用：隐私保护措施的实施，包括从设计入手保护隐私；对数据采集、保留、处理的合理限制；获得用户的明示同意；对数据进行反识别；要求下游用户将数据以反识别的形式保存等。物理安全、系统安全、存储数据安全的实施，其中系统涉及技术包括 Hadoop、MapReduce、NoSQL 等；数据加密算法，如对称加密 DES，IDEA；非对称加密 RSA 算法、ECC 算法；散列算法；数字签名；数字证书等。

◆ 大数据质量、热度

1. 领会：数据剖析在大数据应用中的作用，数据处理理论，数据质量评测方法，数据质量管理平台基本功能及构架。

2. 熟知：数据剖析方法与原理，数据剖析的基本流程，剖析中的业务规则的应用，数据质量诊断的原理和方法，数据诊断流程，数据质量诊断报告及解读，数据治理中的业务规则定义及使用，企业业务规则库建立及使用，企业数据质量管理平台使用角色及职责权限设计，数据质量管理平台在不同的商业应用场景的定位及部署，常用质量管理工具，如帕累托图、鱼骨图、休哈特图、智能设备校准。

3. 应用：数据质量管理工具与 ETL、数据的使用热度、频次统计分析。

◆ 大数据数据库演进

1. 领会：大数据数据库存储计算的演进图谱，知道分类和主要代表产品。

2. 熟知：对传统的 RDBMS 到 NoSQL、NewSQL 演进的规则、产品特性和应用场景，并对部分具体实例有数据存取操作的熟知。

3. 应用：图数据库、KV 数据库、HTAP 数据库、分析型数据库 ClickHouse、数据采集方式。

◆ 大数据生命周期模型

1. 领会：大数据生命周期管理概念，对数据热度的理解（如热数据、温数据、冷数据），数据整合与主数据管理。

2. 熟知：对大数据生命各周期进行管理，如定义大数据范围、大数据采集、大数据存储、大数据整合、大数据呈现与使用、大数据分析与应用、大数据归档与销毁。

3. 应用：数据实时采集、抽取技术，大数据血缘关系分析，数据可视化技术。

PART 6 项目管理

◆ 软件项目管理基础

1. 领会：理解范围管理、时间管理、成本管理、质量管理等项目管理核心部分。理解项目周期中包含的启动、计划、执行、控制、结束五个关键过程。

2. 熟知：了解典型软件开发团队的人员组成与角色职责。了解典型软件开发工作的组成部分：需求分析、设计、实现、测试及维护。

3. 应用：能够对数据科学项目进行任务分解与分配，能够识别子任务的依赖关系，能识别项目中的风险点。能够利用甘特图把控项目进度与预期。

◆ 敏捷开发

1. 领会：理解敏捷式开发的关键思想，了解敏捷式开发相比于传统瀑布式开发模式的区别及其优势。

2. 熟知：理解 Scrum、看板、极限编程、精益编程等敏捷式开发方法，熟悉 Scrum 开发规则，熟悉 Scrum 团队成员与角色。

3. 应用：应用 Scrum 模式进行数据科学项目开发。

◆ 代码管理

1. 领会：理解版本控制系统的思想及其在软件开发中的重要作用，理解分布式版本控制系统的特点与优势。

2. 熟知：Git 版本控制系统的基本工作原理，仓库、分支、合并等 Git 概念，熟练掌握常用的 Git 命令。

3. 应用：安装并使用 Git 进行代码与文档管理，部署与应用 GitLab 以实现私有 Git 服务，在 Github 上使用和发布开源代码。

◆ 构建大数据团队

1. 领会：团队组织架构思想，项目管理的本质，将业务问题转化为分析问题方法论，外包与自研的优缺点。

2. 熟知：团队沟通技巧，成员分工与结果量化规则，人才培养方案，团队文化建设。

◆ 项目管理相关知识及常用工具

1. 领会：项目管理知识体系，项目计划制定原则，项目内容管理，质量管理、风控管理。

2. 熟知：项目管理体系，项目管理过程，目标可视化、要素标准化、绩效导向化等基本技巧。

3. 应用：将常用项目管理工具（甘特图、PERT 图、思维导图、HQQ 等）熟练应用到项目管理中。

六、推荐学习书目

[1] Micha Gorelick, Ian Ozsvald. Python 高性能编程[M]. 人民邮电出版社，2017.

[2] Aloysius Lim, William Tjhi. R 高性能编程[M]. 电子工业出版社，2015.

[3] Paul Chiusano, Rúnar Bjarnason. Scala 函数式编程[M]. 电子工业出版社，2016.

[4] 周志华. 机器学习[M]. 清华大学出版社，2016.

- [5] Ian Goodfellow, Yoshua Bengio. 深度学习. 人民邮电出版社, 2017.
- [6] 郑泽宇. TensorFlow 实战 Google 深度学习框架 (第 2 版) [M]. 电子工业出版社, 2018
- [7] 南森·马茨 (Nathan Marz), 詹姆斯·沃伦 (James Warren). 大数据系统构建 (可扩展实时数据系统构建原理与最佳实践) [M]. 机械工业出版社, 2017.
- [8] 约阿夫·戈尔德贝格. 基于深度学习的自然语言处理[M]. 机械工业出版社, 2018.
- [9] 迪潘扬·萨卡尔 (Dipanjan Sarkar). Python 文本分析[M]. 机械工业出版社, 2018.
- Maksim Tsvetov, Alexander Kouznetsov. 社会网络分析方法与实践[M]. 机械工业出版社, 2013.
- [10] 桑尼尔·索雷斯. 大数据治理[M]. 清华大学出版社, 2014.
- [11] 杰弗里·波梅兰茨. 元数据-用数据的数据管理你的世界[Metadata] [M]. 中信出版集团, 2017.
- [12] 艾伯特·拉斯洛·巴拉巴西, 巴拉巴西, 马慧. 爆发:大数据时代预见未来的新思维[M]. 中国人民大学出版社, 2012.
- [13] 阿里巴巴数据技术及产品部. 大数据之路:阿里巴巴大数据实践 [M]. 电子工业出版社, 2017.7
- [14] 欧盟. 《通用数据保护条例》(General Data Protection Regulation, 简称 GDPR), 2018.
- [15] Andrew Stellman, Jennifer Greene. 学习敏捷构建高效团队[M]. 人民邮电出版社, 2017.
- [16] Scott Chacon, Ben Straub. Pro Git 第二版[M]. 人民邮电出版社, 2017.
- [17] 电子版 <https://bingohuang.gitbooks.io/progit2/>
- [18] 科丽·科歌昂, 叙泽特·布莱克莫尔, 詹姆士·伍德. 项目管理精华[M]. 中国青年出版社, 2016.

CDA Institute

经管之家 CDA 数据分析研究院