

大数据分析工程师（基础级）考试大纲

CBDA 大数据分析工程师考试大纲是 CBDA 命题组基于 CBDA 大数据分析工程师等级认证标准而设定的一套科学、详细、系统的考试大纲。考纲规定并明确了 CBDA 大数据分析工程师认证考试的具体范围、内容和知识点，考生可按照 CBDA 大数据分析工程师考试大纲进行相关知识的复习。

大数据数学基础（占比 30%）

- 1 微积分基础
- 2 概率论与数理统计
- 3 线性代数
- 4 数值计算基础
- 5 多元统计分析

数据分析基础（占比 20%）

- 1 数据分析的基本流程
- 2 数据的描述分析
- 3 数据的推断分析
- 4 数据的相关分析
- 5 数据的回归分析
- 6 数据的可视化分析

Excel 数据处理与分析（占比 10%）

- 1 Excel 函数及使用
- 2 Excel 数据管理
- 3 Excel 数据透视表
- 4 Excel 数据可视化图表分析

数据采集与预处理（占比 10%）

- 1 数据采集方法
- 2 数据取样方法
- 3 市场调研
- 4 数据预处理方法
- 5 数据存储方法

数据建模分析基础（占比 30%）

- 1 主成分分析法（占比 3%）、因子分析法（占比 2%）
- 2 系统聚类法（占比 3%）、K-Means 聚类法（占比 3%）
- 3 对应分析（占比 2%）、多维尺度分析（占比 2%）

- 4 预测性分析法 (多元线性回归 (占比 5%), 逻辑回归 (占比 5%))
- 5 时间序列 (占比 5%)

参考目录

- 1 大数据数学基础, , 人民邮电出版社, 2017
- 2 实用多元统计分析 (第 6 版), [美]约翰逊, [美]威克恩 著; 陆璇, 叶俊 译, 清华大学出版社
- 3 经济计量学, [美]斯托克 (Stock J.H.) 著; 王庆石 译, 东北财经大学出版社, 2005
- 4 数据库系统及应用 (第 3 版), 崔巍, 高等教育出版社, 2012

大数据分析工程师（专业级）考试大纲

CBDA 大数据分析工程师考试大纲是 CBDA 命题组基于 CBDA 大数据分析工程师等级认证标准而设定的一套科学、详细、系统的考试纲要。考纲规定并明确了 CBDA 大数据分析工程师认证考试的具体范围、内容和知识点，考生可按照 CBDA 大数据分析工程师考试大纲进行相关知识的复习。

数据挖掘基础理论（占比 10%）

- 1 数据挖掘的基本概念
- 2 数据挖掘任务
- 3 数据挖掘流程
- 4 数据挖掘的方法
- 5 数据挖掘的评估方法

数据预处理（占比 20%）

- 1 数据清理
- 2 数据集成
- 3 数据变换
- 4 数据归约
- 5 异常检测

数据库基础（占比 20%）

- 1 数据框概论
- 2 关系数据库基础
- 3 设计、创建与管理数据库
- 4 SQL 及关系型数据库基本概念
- 5 SQL 数据类型、运算符、函数
- 6 使用 SQL 语句操作数据库

中级数据分析方法（占比 20%）

- 1 实验设计
- 2 参数检验
- 3 多元回归分析
- 4 主成份分析/因子分析
- 5 K-means 聚类
- 6 随机森林(Random Forest)
- 7 最近邻方法 KNN
- 8 朴素贝叶斯分类(Naive Bayesian)
- 9 时间序列
- 10 常用的数据挖掘方法和思想

编程基础 (R/Python 二选一) (占比 30%)

R 语言基础

- 1 R 语言概述
- 2 数据对象与数据读写
- 3 数据集基本处理
- 4 函数与控制流
- 5 初级绘图函数
- 6 高级绘图函数 (lattice、ggplot2、交互绘图)
- 7 Rattle 可视化数据挖掘工具

Python 基础

- 1 Python 基础知识
- 2 Python 数据结构
- 3 程序流程控制语句
- 4 函数
- 5 文件基础
- 6 NumPy 数值计算基础
- 7 Matplotlib 数据可视化基础
- 8 Pandas 统计分析基础
- 9 使用 scikit-learn 构建模型

参考目录

- 1 R 语言编程基础, 林智章 张良均, 人民邮电出版社, 2017
- 2 Python 编程基础, 张健 张良均, 人民邮电出版社, 2017
- 3 Python 编程:从入门到实践, [美]埃里克·马瑟斯, 人民邮电出版社, 2016
- 4 数据挖掘导论, [美]Pang-Ning Tan, Michael Steinbach, Vipin Kumar 著, 译者: 范明 范宏建, 人民邮电出版社。
- 5 Data Mining: A Tutorial Based Primer, 作者: Roiger, Richard, Geatz, Michael, Addison-Wesley
- 6 统计学习方法, 李航, 清华大学出版社, 2012

大数据分析工程师（专家级）考试大纲

CBDA 大数据分析工程师考试大纲是 CBDA 命题组基于 CBDA 大数据分析工程师等级认证标准而设定的一套科学、详细、系统的考试纲要。考纲规定并明确了 CBDA 大数据分析工程师认证考试的具体范围、内容和知识点，考生可按照 CBDA 大数据分析工程师考试大纲进行相关知识的复习。

高级数据分析技术（占比 20%）

- 1 MySQL 理论及实战
- 2 R/Python 基础（二选一）
- 3 Linux & Ubuntu 基础

大数据项目管理（占比 15%）

- 1 大数据项目管理概述
- 2 基本职业素养
- 3 大数据技术基础
- 4 业务理解与沟通
- 5 建设项目团队
- 6 项目流程及拆解
- 7 项目时间管理
- 8 项目资源管理
- 9 项目质量管理
- 10 项目风险管理
- 11 项目沟通管理

大数据技术概论（占比 30%）

- 1 大数据概论
- 2 大数据采集及预处理
- 3 大数据分析
- 4 大数据可视化
- 5 Hadoop 概论
- 6 HDFS 和 Common 概论
- 7 MapReduce 概论
- 8 NoSQL 技术介绍
- 9 Spark 概论
- 10 深度学习概论
- 11 云计算与大数据

大数据分析建模（R、Python 二选一）（占比 35%）

- 1 回归分析(Regression Analysis)

- 2 决策树(Decision Tree)
- 3 集成学习提升方法
- 4 随机森林(Random Forest)
- 5 最近邻方法 KNN
- 6 朴素贝叶斯分类(Naive Bayesian)
- 7 支持向量机(Support Vector Machine)
- 8 人工神经网络(Artificial Neural Network)
- 9 极大似然估计和 EM 算法
- 10 聚类分析(Cluster Analysis)
- 11 关联规则(Apriori)
- 12 PCA 降维
- 13 智能推荐算法 (协同过滤)
- 14 特征工程与算法
- 15 模型的评估与选择

参考目录

- 1 R 语言商务数据分析实战, 韩宝国 张良均, 人民邮电出版社, 2017
- 2 Python 数据分析与应用, 黄红梅 张良均, 人民邮电出版社, 2017
- 3 利用 python 进行数据分析, 麦金尼 (Wes McKinney), 机械工业出版社, 2014
- 4 Hadoop 大数据开发基础, 余明辉 张良均, 人民邮电出版社, 2017
- 5 Spark 大数据技术与应用, 肖芳 张良均, 人民邮电出版社, 2017
- 6 机器学习, 周志华, 清华大学出版社, 2016
- 7 元数据-用数据的数据管理你的世界, [美] 杰弗里•波梅兰茨 (Jeffrey Pomerantz) 著; 李梁 译, 中信出版集团, 2017
- 8 架构之美 Till Adam 著; 王海鹏 / 蔡黄辉 / 徐锋译; 机械工业出版社, 2009