

中国民用航空飞行学院硕士研究生

同等学力加试科目

《大数据技术原理与应用》考试大纲

第一部分 考试说明

一、考试性质

《大数据技术原理与应用》是中国民用航空飞行学院理学院大数据技术与工程（专业学位）专业的硕士研究生同等学力加试科目之一。它的评价标准是高等学校、科研院所的优秀本科毕业生能达到及格或及格以上水平，以保证被录取者具有大数据技术原理与应用的基础理论知识和初步的应用能力。

二、考试内容范围

大数据的基本概念；大数据处理架构 Hadoop；分布式文件系统 HDFS；分布式数据库 HBase；NoSQL 数据库；云数据库；分布式并行编程模型 MapReduce；数据仓库 Hive；基于内存的大数据处理架构 Spark；流计算框架 Flink。

三、评价目标

掌握大数据技术原理与应用的相关概念、基本原理，具备在 Hadoop、HDFS、HBase、MapReduce、Hive、Spark 和 Flink 等平台的搭建及基本操作能力。

四、课程涉及的相关书目

1. 《大数据技术原理与应用（第三版）》，林子雨，人民邮电出版社，2021.1.
2. 《大数据技术原理与应用》，程秀峰，严中华，科学出版社，2022.9

第二部分 考查要点

1. 能够建立对大数据知识体系的轮廓性认识，了解大数据发展历程、基本

概念、主要影响、应用领域、关键技术、计算模式和产业发展，并了解云计算、物联网的概念及其与大数据之间的紧密关系。

2. Hadoop 的发展历史、重要特性和应用现状，Hadoop 项目结构及其各个组件，并熟练掌握 Hadoop 平台的安装和使用方法。

3. 分布式文件系统的基本概念、结构和设计需求，掌握 Hadoop 分布式文件系统 HDFS 的重要概念、体系结构、存储原理和读写过程，并熟练掌握分布式文件系统 HDFS 的使用方法。

4. 分布式数据库 HBase 的访问接口、数据模型、实现原理和运行机制，并熟练掌握 HBase 的使用方法。

5. NoSQL 数据库与传统的关系数据库的差异、NoSQL 数据库的四大类型以及 NoSQL 数据库的三大基石；基本掌握 Redis、MongoDB 等 NoSQL 数据库的使用方法。

6. 云数据库的概念、基本原理和代表性产品的使用方法。

7. 分布式编程框架 MapReduce 的基本原理和编程方法。

8. 数据仓库 Hive 的基本原理和使用方法。