新一代信息技术专业群 大数据技术应用实训室升级-Hadoop大数据技术与应用课程资源包项目参数要求

特别提醒：（报价人在确认参数无误后请在最后一页签字盖章）

《Hadoop大数据技术与应用课程资源包》

Hadoop 环境配置

1. Hadoop的基础环境配置；
2. Hadoop的安装；
3. Hadoop的配置；
4. Hadoop的运行；
5. Hadoop相关的步骤与作业项
6. 连接Hadoop
7. 导入导出Hadoop集群数据
8. 执行HiveQL语句
9. 执行MapReduce
10. 执行Spark作业

## Hadoop 快速部署工具 Ambari

1. 基础环境准备
2. 设置权限
3. 测试连接
4. 安装jdk
5. 部署ambari-server
6. 安装ambari-server
7. 设置ambari-server
8. 启动ambari-server
9. 查看监听端口
10. 访问ambari server
11. 创建Hadoop集群

## 分布式文件系统HDFS

1. HDFS简介；
2. HDFS架构；
3. HDFS文件块；
4. HDFS常用操作

## 资源调度枉架 YARN

1. YARN架构；
2. YARN调度策略；
3. YARN常用操作

## 分布式计算框架 MapReduce

1. MapReduce原理；
2. MapReduce作业数据流；
3. Hadoop流处理；
4. MapReduce程序实现词频统计；
5. MapReduce程序的Reducer数量；
6. MapReduce程序的Combiner；
7. MapReduce程序实现数据连接

## 大数据数据仓库 Hive

使用Hive进行大数据处理，包括Hive的安装、配置和运行以及Hive的常用操作——数据定义、数据操纵和数据查询。

一、Hive简介

Hive概述；Hive的安装；Hive的运行

二、Hive数据定义操作

数据库操作；数据表基本操作；存储格式和行格式；数据类型；外部表

三、Hive数据操纵操作

数据导入；数据插入；数据导

四、Hive数据查询操作

基本查询；数据聚合；数据连接

## 分布式数据库HBase

1. HBase基础

1.1 、 HBase概述；1.2 、 HBase的安装；1.3 、 HBase的运行；

2、HBase数据定义操作

2.1 、 命名空间Namespace；2.2 、 DDL基本操作；

3、HBase数据操纵操作

3.1 、 插入数据；3.2 、 获取数据；3.3 、 删除数据

## 分布式日志采集工具Flume

1. Flume下载与安装
2. Flume架构学习
3. Flume配置
4. Flume特性
5. Flume事务
6. Flume数据传输
7. Flume选择器介绍
8. Flume副本
9. Flume负载均衡
10. Flume聚合
11. Flume监控
12. Flume自定义Source
13. Flume拦截器精讲
14. Flume高级聚合

## 分布式数据库与传统数据库数据相互转换工具 Sqoop

1. sqoop入门与精通
2. sqoop工具认识
3. Sqoop下载与安装
4. sqoop底层架构与原理
5. Sqoop命令与指令
6. sqoop数据导出应用
7. sqoop连接hdsf实现数据导入
8. sqoop导入主键表到hdfs
9. sqoop将无主键表导入到hdfs中
10. Sqoop连接mysql应用
11. sqoop与hdfs案例实战
12. mysql数据导入到hdfs索例实战
13. hdfs导入到mysql数据库案例实战
14. sqoop与hive高级应用
15. sqoop导出指定格式数据到mysql和hive中+sqoop数据增量更新导入
16. 定时任务操作sqoop导入导出数据
17. sqoop使用sql语句完成增量操作
18. sqoop的高级应用与高级优化

## 提供分布式协调 致性服务的ZooKeeper

1. Storm架构及各组件作用
2. Zookeeper核心功能
3. 安装配置Zookeeper集群
4. 安装配置Storm集群

## 课程总计配套

▲教学PPT不少于400页；教学视频不少于40个，总时长不低于600分钟；

▲课程资源包需兼容学校现有大数据技术专业实训平台。

报价人签字盖章