

课程教学进度计划表

一、基本信息

课程名称	大数据基础概论				
课程代码	2050545	课程序号	7020	课程学分/学时	3
授课教师	宋建虎	教师工号	23152	专/兼职	兼职
上课班级	软工 B21-1、2	班级人数	42	上课教室	计算中心 228 (单) 三教 216
答疑安排	周午中午三教 216				
课程号/课程网站	无				
选用教材	《大数据技术基础与实践》. 薛志东主编.人民邮电出版社.2020.6				
参考教材与资料	<p>《大数据技术原理与应用-概念、存储、处理、分析与应用》(第2版). 林子雨 编著.人民邮电出版社.2017.1</p> <p>《深入理解大数据-大数据处理与编程实践》.黄宜华 主编.机械工业出版社.2014.8</p> <p>《Hadoop 大数据开发实战》. 杨力 主编. 人民邮电出版社 2019.2</p>				

二、课程教学进度安排

课次	课时	教学内容	教学方式	作业
1	4	大数据技术概述	案例教学	
2	2	使用 VirtualBox 和 VMware Workstation 安装虚拟机	案例教学实验	练习与实践 1
3	4	Linux 常用命令、网络配置、Linux 集群配置	案例教学实验	练习与实践 2
4	2	Hadoop 集群安装、初始化以及日志查看	案例教学实验	
5	4	HDFS 简介、HDFS 基本命令、数据平衡	案例教学	

		优化, HDFS API 调用	实验	
6	2	MapReduce 分布式编程步骤、 MapReduce Shuffle 过程开发、 MapReduce 性能优化、Yarn 处理框架	案例教学	练习与实践 3
7	4	MapReduce 实战: 词频统计、绘制频度 分布	案例教学 实验	
8	2	Hive 安装配置、创建数据库数据表、 数据查询	案例教学	练习与实践 4
9	4	Hive 自定义函数运算、Hive 自定义函 数编程	案例教学 实验	练习与实践 5
10	2	Hbase 简介、安装、Shell 操作、客户 端 API、与 MapReduce 集成	案例教学 实验	
11	4	HBase 实战 公有云网盘系统管理	实验	
12	2	公有数据资源获取、使用网络爬虫获 取数据、使用 Flume 获取数据	案例教学	练习与实践 6
13	4	Spark 简介、Spark 快速部署、Spark 程序	案例教学	
14	2	Spark RDD 编程、Spark 生态系统	案例教学	
15	4	Hadoop 处理图像的问题和对策, HIPI 安装与部署, 使用 HIPI 进行图像处理		
16	2	大数据安全		

三、考核方式

总评构成	占比	考核方式
X1	大作业	60%
X2	实验报告	25%
X3	课堂表现	15%

任课教师:  (签名) 系主任审核:  (签名) 日期: 2024. 2. 20