附件1

**课程内容安排表**

|  |  |  |
| --- | --- | --- |
| **时 间** | **课 程 主 题** | **主 要 内 容** |
| 第一天上午9:00-12:00 | 云计算与大数据分析基础 | * 云计算和大数据的概念
* 大数据分析的要点
* 云计算与大数据国内外发展现状
* 大数据分析的应用和发展趋势
* 大数据分析的典型商业模式
 |
| Hadoop的起源与核心架构介绍 | * 传统大规模系统与Hadoop的比较
* Hadoop应用案例解析
* Hadoop 核心架构
* Hadoop与传统分布式环境的区别
 |
| Hadoop 2.0介绍 | * Hadoop 2.0介绍
 |
| 第一天下午2:00-：5:00 | Hadoop实验环境搭建 | * 实验手册
* 实验环境说明
* 实验机器安装操作系统配置
 |
| 第二天上午9:00-12:00 | Map-Reduce原理、体系架构与工作机制 | * MapReduce流程
* Hadoop JobTracker 剖析
* Hadoop TaskTracker 剖析
* Hadoop 任务提交流程剖析
 |
| Map-Reduce Streaming 流式计算 | * Map-Reduce Streaming 介绍
* Hadoop Streaming 和 Java MapReduce Api 差异
* Python 与 mrjob
 |
| 第二天下午2:00-：5:00 | Map-Reduce编程实战，日志分析 | * 剖析MapReduce日志分析程序
 |
| Hadoop实验 | * Map-Reduce word count实验
* mrjob 实验
* Map-Reduce日志分析
 |
| 第三天上午9:00-12:00 | Hadoop数据仓库-Hive | * Hive基础
* Hive的作用和原理说明
* Hive与传统数据库的对接使用
* Hadoop/Hive仓库数据数据流
 |
| HBase体系架构介绍 | * Hadoop 分布式数据仓库-HBase
* HBase概念与架构
* HBase核心知识点
* HBase安装、部署
* HBase配置优化综述
 |
| 第三天下午2:00-：5:00 | Hadoop实验 | * Hive 安装、部署
* HBase安装、部署
 |
| 第四天上午9:00-12:00 | YARN、Pig、ZooKeeper、Oozie介绍 | * Hadoop2.0资源调度器-YARN
* 并行计算数据流语言-Pig
* 分布式管理组件-Zookeeper
* 工作流管理- Ozzie
 |
| Hadoop发行版本介绍 | * Hadoop发行版本介绍
 |
| 第四天下午2:00-：5:00 | Hadoop 不同环境分析 | * Hadoop 单机版、伪分布式、分布式环境介绍及差异比较
* Hadoop分布式环境安装部署要点
 |
| Hadoop实验 | * Hadoop分布式环境安装、部署
 |
| 总结与答疑 |  |

附件2

**云计算大数据分析平台Hadoop最佳实践高级培训班报名回执表**

经研究，选派下列人员参加学习：

|  |  |
| --- | --- |
| 单位名称 |  |
| 详细地址 |  |
| E－mail |  | 邮政编码 |  |
| 联系人 |  | 电话 |  | 传真 |  |
| 姓名 | 性别 | 部门 | 职务 | 手机 | 身份证号码 | 参培地点 | 是否住宿 |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| 您想了解或希望解决的主要问题: |

**注：本表复印有效**

传真：64700872邮箱： hyb\_trainingcentre@126.com 王斌