

中国现代教育装备杂志社

“大数据”系列课程（暑期）教学研修班 邀请函

各高校教务处、计算机学院、软件学院（系）、相关学院负责人：

近年来，大数据受到我国政府的高度重视和大力支持。2015年8月31日，国务院印发《促进大数据发展行动纲要》（国发〔2015〕50号），标志着将大数据纳入国家战略层面。据统计，未来3-5年我国需要大数据人才180万，但目前只有30万此类人才，缺口达150万。培养出足够的、合格的数据人才，对我国在未来掌握大数据的核心价值起着至关重要的作用。

为满足社会对大数据人才的需求，2016年2月，教育部公布新增“数据科学与大数据技术”（专业代码：080910T）本科专业，首批仅北京大学、对外经济贸易大学和中南大学获批，2017年3月公示中国人民大学等32家新增“数据科学与大数据技术”专业获批。目前全国已经有35所高校新增该专业。2016年9月教育部研究决定，正式批准“高职”《大数据技术与应用》专业作为《普通高等学校高等职业教育（专科）专业目录2016年增补专业》（专业代码：610215）。

国内外高校纷纷设立了大数据专业或在相关专业开设了大数据课程，经调研发现，大数据师资是目前制约大数据教育发展的主要因素。

为此，在中国高等教育学会指导下，学会下属《中国现代教育装备》杂志社（以下简称“杂志社”）定于2017年8月13至18日在青海西宁举办教学高级研修班。

特邀请各高等院校数据科学相关专业、计算机科学技术、网络工程、软件工程、信息工程、信息管理、统计学专业、数学专业、经济管理专业、电子商务、市场营销专业等科研、教学带头人，骨干教师、博士生、硕士生参加本次研修班学习，探讨大数据专业建设、教学经验交流等。

一、课程目标

1. 老师参加完培训后，可以直接开设专业课程。
2. 本次培训理论与实践相结合，提升动手实战能力。
3. 在大数据和数据科学领域找到新的研究课题。
4. 掌握大数据可视化；
5. 掌握数据科学核心理论、产品开发及典型实践；
6. 探讨“大数据”系列课程的专业建设；

二、培训形式

本次培训班采取**网络课程+面授**形式。

网课时间：学员自报名之日起，先开通网络学习课程，课程开通时间为半年。

面授时间：2017年8月14日至18日（五天）

三、培训时间、地点及费用

培训时间：2017年8月13日至8月18日，8月13日全天报到。

培训地点：东湖宾馆（青海省西宁市同仁路2号东湖宾馆）

前台电话：0971-7661666

培训费：4980元/人（培训费、资料费、证书等费用）。食宿统一安排，费用自理。

缴费方式：可提前转账、现场刷卡或支付现金。

四、相关说明：

1. 为了保证教学质量，本次培训班人数名额有限，请抓紧报名。
凡参加本次培训的学员均可获得杂志社合作伙伴——“全国高校大数据教育联盟”（以下简称“联盟”）即将出版的数据科学与大数据技术专业相关教材、讲义、视频等资料。同时对大数据感兴趣的老师、学生均可免费参加联盟“大数据公益人才培养工程”。
2. 本次培训的具体工作由优选创新（北京）科技有限公司承担，
培训费发票由优选创新（北京）科技有限公司开具。
3. 学费缴纳形式：支持现场 POS 机刷卡、如汇款请务必在备注栏里注明 “西宁培训班”，汇款后请及时告知会务组。

开户名称：优选创新（北京）科技有限公司

开户银行：中国民生银行北京上地支行

银行账号：603700218

五、培训证书：

对经学校有关部门推荐参加培训的教师，考评合格后获得由工业和信息化部教育与考试中心颁发的《大数据分析师》教师认证证书（见附件四），请提前准备好**白底二寸相片 3 张**（纸质版相片）。所有的证书将由联盟提供真伪查验。

对参加培训获得证书的教师，所在学校应承认其接受培训的经历，计入继续教育学时。

六、研修班联系人及联系方式

研修班联系人：刘老师 联系电话：13651136269 （010）82210085

张老师 联系电话：13601155010

E-Mail: contact@bigdataedu.org

请各单位积极组织和推荐相关课程骨干教师参加会议。参加培训的教师请培训班开始前一周将回执表(见附件三)发给组委会联系人。

七、其他

提示：请各位参会代表自带笔记本电脑。

附件：1、课程内容及安排

2、主讲专家介绍

3、回执表

4、证书样本

《中国现代教育装备》杂志社

2017年7月1日

附件 1

课程内容及安排

1、网络学习课程

| 课程名称 | 课时 | 课程内容介绍 |
|----------------------|----|-----------------------------------------------------------------------------------------------------------------------------|
| Linux 操作系统 | 10 | Linux 操作系统介绍、文件与目录管理、用户以及用户组管理、磁盘管理、文本编辑工具 vim、文档的压缩与打包、软件安装、Shell 脚本等内容。 |
| Hadoop 环境搭建及工作原理 | 10 | 讲解 Hadoop 生态系统, 包括操作与开发; 详细讲解 HDFS 和 Map-Reduce 的功能及作用; 了解 MapReduce 原理、运行流程、压缩数据处理、作业调度、计算器等环节。 |
| HDFS 程序开发实践 | 6 | 讲解 Hadoop 文件系统 HDFSJAVAAPI 的使用。如何使用 HDFSJavaAPI, 读写文件、读写目录、以及对文件进行压缩处理等。 |
| MapReduce 原理与开发实践 | 6 | 本课程针对 HadoopMapReduce 开发进行讲解。课程以案例为基础, 以及如何使用 MapReduce 进行数据统计, 去重, 排序, Map 端 Join, Reduce 端 Join 等关联操作, 掌握 MapReduce 处理过程。 |
| Hive 数据仓库的应用 | 12 | Hive 是基于 Hadoop 构建的数据仓库分析系统, 通过学习掌握 Hive 的函数、Hive 数据的加载、Hive 的 DDL 操作、自定义函数 (UDF) 等内容, 达到使用 Hive 进行查询、汇总、分析数据的能力。 |
| 分布式数据库 HBase 工作原理及应用 | 6 | HBase 是 Hadoop 生态系统中的重要一员, 主要用于海量结构化数据存储; 通过学习对 HBase 表设计、表操作、数据操作、JavaAPI 等内容, 掌握对 HBase 系统的开发及使用。 |
| 数据迁移工具 Sqoop 操作 | 4 | Sqoop 是关系型数据库和 Hadoop 生态系统之间进行数据转换的主要工具; 通过学习将 mysql 中的数据导入到 hdfs 中、将数据导入到 HBase 中、定义导入导出任务, 掌握对数据迁移的能力。 |
| 分布式日志框架 Flume 配置与应用 | 4 | Flume 对海量日志进行采集、聚合和传输的主流大数据工具; 课程内容包含 Flume 应用场景、FlumeNG、FlumeOG、Flume 的核心组件、Flume 的架构、Flume 的 source、sink 配置说明等。 |

| | | |
|------------------------|---|-------------------------------------------------------------------------------------------------------------|
| 流式数据采集 Kafka 的部署与操作 | 4 | Kafka 是分布式的消息队列，广泛应用于实时数据处理。学习内容包含 Kafka 的体系结构、安装模式及安装部署、Topic、Producer、Consumer、发布订阅消息以及 Kafka JAVA 开发等。 |
| Spark 环境搭建及工作原理 | 6 | Spark 是一款高性能的分布式计算框架，比 MapReduce 计算快百倍；本课程内容全面涵盖了 Spark 生态系统、Spark 与 Hadoop 对比、开发环境搭建、RDD、编程模型、Web 监控等内容。 |
| SparkStreaming 原理与应用优化 | 4 | SparkStreaming 是用户结合流式、批处理和交互式查询应用的实时计算框架；本课程内容详细讲解原理与特点、适用场景、容错、性能优化和内存优化等。 |
| SparkSQL 的原理以及操作 | 4 | SparkSQL 的出现，使得 SQL-on-Hadoop 的性能相对于 Hive 有了显著的提高。达到 Spark 兼容 Hive 的功能。本课程详细讲解特点、运行架构、数据源、数据缓存、DataFrame 等。 |

2、面授课程安排

| 数据科学与数据可视化课程（青海西宁） | | |
|--------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 时间 | 课程模块 | 课程内容 |
| 第 1 天 8 月 14 日 | 数据科学核心理论与产品开发典型实践 | <ul style="list-style-type: none"> 数据科学的核心理论（数据科学中的新理念、理论、方法、技术、工具、应用、机器学习以及如何跟踪数据科学的最新动态和如何继续学习数据科学） 数据科学的典型实践（最佳实践解读、典型算法及应用案例、基于 R 的数据科学实战、基于 SparkR 的数据科学实战） 数据产品开发（数据产品开发在数据科学中的重要地位；数据产品的特征；数据柔术；数据战略；数据能力；数据治理；开发技术；应用案例；如何继续学习 |
| 第 2 天 8 月 15 日 | | |
| 第 3 天 8 月 16 日 | | |
| 第 4 天 8 月 17 日 | 数据可视化 | <ul style="list-style-type: none"> 数据可视化之美（数据与图形、数据可视化的运用、传统的数据分析图表、数据可视化工具） 数据可视化方法（整理数据源、直方图、圆饼图、散点图） 数据引导可视化设计（可视化对认知的帮助、掌握可视化设计组件、分析数据，指导视觉探索、空间数据的可视化） 数据可视化入门（Tableau、创建视图、仪 |
| 第 5 天 8 月 18 日 | | |

| | | |
|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | <p>表板与故事)</p> <ul style="list-style-type: none"> • 数据管理与计算 (认识数据、数据架构、数据连接、数据加载、数据维护、高级数据操作) • 可视化分析 (条形图与直方图、饼图、压力图与突显表、树地图、气泡图与圆视图、标靶图、甘特图、地图分析) • 预测分析 (预测分析的可视化、预测、背景图像、趋势线) • 分享与发布 (导出和发布数据、导出图像文件、保存和发布) |
|--|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

附件 2

主讲专家



朝乐门

我国首部系统阐述数据科学理论、方法、技术与工具的专著——《数据科学》的作者。博士、中国人民大学信息资源管理学院副教授，硕士生导师。中国计算机学会信息系统专业委员会委员、国际知识管理协会正式成员。主持完成国家自然科学基金、国家社会科学基金等重要科学研究项目 10 余项。参与完成核高基、国家自然科学基金重点项目等国家重大科研项目 10 余项。获得北京市中青年骨干教师称号、中国人民大学优秀博士论文奖等多种奖励 30 余项。



周苏

浙江大学城市学院教授、《大数据可视化》作者、国家科技部认证的二级创新工程师、国家科技部一级创新培训师、浙江省高校优秀青年教师。长期从事计算机等专业教学工作，担任大数据技术与应用、大数据可视化（技术）、创新思维与创新方法、操作系统原理、计算机基础、计算机导论、软件工程、软件测试技术、软件体系结构、多媒体技术、项目管理、应用统计学、C 语言程序设计、面向对象程序设计等课程的教学，对高等院校计算机学科教学有深刻理解。

附件 3

“大数据”教学研修班（暑期）回执表

| | | | | | |
|--------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|--|--------|--|
| 姓名 | | 性别 | | 职务/职称 | |
| 单位名称 | | | | 手机 | |
| 发票抬头 | | | | 办公电话 | |
| 发票快递地址 | | | | E-mail | |
| 报到时间 | (第一期) <input type="checkbox"/> 8月12日 <input type="checkbox"/> 8月13日 | | | | |
| 房间预定 | 8月 <input type="checkbox"/> 12日 <input type="checkbox"/> 13日 <input type="checkbox"/> 14日 <input type="checkbox"/> 15日 <input type="checkbox"/> 16日 <input type="checkbox"/> 17日 <input type="checkbox"/> 18日 | | | | |
| | <input type="checkbox"/> 单间大床房 <input type="checkbox"/> 标间单住 <input type="checkbox"/> 标间双住 | | | | |

提示：

1. 请在回执表中将“□”替换为“☑”选中相关信息。
2. 请参加《数据科学与数据可视化》（青海西宁）培训班的老师于**2017年8月6日**前，通过电子邮件的方式回复至研修班联系人。

咨询联系电话：刘老师 13651136269 张老师 13601155010

电子邮件：contact@BigDataEdu.org

附件4 证书样本

