

关于举办全国高等院校商务数据挖掘企业项目

教师高级研修班的通知(第二轮)

如何对庞大的企业数据源进行“数据深加工”，已成为企业信息处理的重要研究课题。在此背景下，商务数据分析人才也顺势成为当今社会最受欢迎的群体之一，而其中又以数据分析与挖掘人才的需求最为迫切。

但现实是一方面大量企业很难寻找到合适的数据分析人才，另一方面高等院校对于如何培养出符合企业需求的数据分析人才又深感困惑。因此加强高等院校与企业的合作，以企业真实项目和实际需求为基础，从项目中来到实际中去，对于增进高校教师对数据分析的理解，提升教学实战水平，培养出真正符合企业和社会需求的学生，有着极为重要的意义。

基于此，本期商务数据挖掘教师高级研修班将聘请多名具有丰富企业数据挖掘实战经验的专家及学者在介绍数据分析语言、数据分析工具、数据可视化处理等知识的基础上，重点介绍企业商务数据挖掘的典型项目并进行实训操作。本次培训通过多个企业商务数据挖掘与分析的真实案例实训，可以较大提升高校教师在企业数据分析实战方面的水平，对于高等院校教学水平的提升以及数据分析与应用等相关专业建设具有重要意义。

一、 培训对象

高等院校财经类、电子商务、计算机科学类、信息类等相关专业的专业负责人、教研室主任、专业责任教师及院校教学分管领导。

二、 培训内容

模块	内容
R 语言基础介绍	<ol style="list-style-type: none">1. 认识 R 语言2. 认识 R 的编译环境3. 使用 R 包4. 了解 R 包内置数据集
数据对象与读写	<ol style="list-style-type: none">1. 查看数据类型2. 判断数据结构3. 读写不同数据源的数据4. 实验操作：加载不同类型数据并写出

<p>常用数据分析</p>	<ol style="list-style-type: none"> 1. 数据质量分析 2. 数据特征分析 3. 统计特征函数、统计可视化
<p>常用数据管理</p>	<ol style="list-style-type: none"> 1. 数据清洗 2. 数据集成 3. 数据变换 4. 数据规约
<p>函数与控制流</p>	<ol style="list-style-type: none"> 1. 使用函数处理数据 2. 实现猜数游戏 3. 练习：用 R 实现猜字游戏 4. 练习：打印 100 到 200 间所有素数 5. 函数自定义 6. 练习：制作简易计数器
<p>数据可视化</p>	<ol style="list-style-type: none"> 1. 绘制基本图形 2. 图形主要参数设置 3. 组合图形 4. 高级绘图 lattice 包 5. 高级绘图 ggplot2 包 6. 交互式绘图工具
<p>可视化分析工具 Rattle</p>	<ol style="list-style-type: none"> 1. Rattle 简介与功能预览 2. 利用 Rattle 实现数据分析
<p>正则表达及 R 语言爬虫</p>	<ol style="list-style-type: none"> 1. 严格的字符匹配 2. 正则表达式的广义化 3. 例子：电话目录的整理 4. 字符串的 R 处理 5. stringr 组件 6. 其他实用函数 7. 练习：爬取网站信息
<p>了解数据分析</p>	<ol style="list-style-type: none"> 1. 数据分析的基本概念 2. 数据分析基本流程 3. 评估方法 4. 性能度量 5. 数据分析应用场景
<p>常用数据分析算法</p>	<ol style="list-style-type: none"> 1. 回归分析 2. 决策树 3. 贝叶斯

	<ul style="list-style-type: none"> 4. 神经网络 5. 集成学习 6. R 语言常用数据分析包
认识 Power BI	<ul style="list-style-type: none"> 1. 认识常用数据分析可视化软件 2. 了解 Power BI 的基本概念 3. 掌握 Power BI 安装 4. 熟悉 Power BI 基本界面
数据获取和整理	<ul style="list-style-type: none"> 1. Excel 数据、网页数据、SQL Server 数据获取 2. 合并文件与合并查询 3. 数据的基本整理 4. M 语言和 Power Query 介绍 5. 导入 Excel 表中的日期数据 6. 建立表间的关系 7. 透视列与逆透视列介绍 8. 交叉表和凭证表介绍 9. 数据泛化 10. 数据对比 11. 检测与处理重复值、缺失值、异常值 12. 有分隔符与无分隔符的单元格字符求和
DAX 建模、分析可视化及数据发布与共享	<ul style="list-style-type: none"> 1. DAX 语言、M 语言简介 2. 新建度量值、计算列 3. 新建表（日历表） 4. 日期的上探与下钻 5. 上下文操作 6. 图表元素的布局 7. 图表色彩的搭配 8. 图表美化的其它操作—插入图像、辅助线等 9. 对比分析：堆积柱形图、簇状柱形图、堆积条形图、簇状条形图、堆积面积图 10. 相关分析：折线图、散点图 11. 结构分析：饼图、环形图、着色地图 12. 趋势分析：折线图、折线和堆积柱形图（柏拉图） 13. 描述性分析：表、矩阵、雷达图、信息图、漏斗图、仪表图、KPI Indicator、子弹图 14. PBI online 版介绍 15. 工作报表的分享 16. 多人协作 17. 添加磁贴功能

	18. 使用自然语言问与答
案例一：商品零售购物篮分析	<ol style="list-style-type: none"> 1. 购物篮分析介绍 2. 商品热销情况分析 3. 零售数据预处理 4. Apriori 算法原理与实现 5. 实训：商品零售购物分析上机实验
案例二：电商产品评论数据情感分析	<ol style="list-style-type: none"> 1. 了解电商产品评论背景与现状分析 2. 实现电商评论情感分析的步骤与方法 3. 电商产品评论数据的采集与提取 4. 对电商产品评论数据进行冗余删除 5. 中文分词技术 jieba 中文分词库介绍 6. 对电商产品评论数据进行预处理 7. LDA 主题模型介绍 8. 基于 LDA 模型进行主题分析 9. 主题分析结果的分析与评价 10. 实训：电商产品评论数据情感分析上机实验
参观	

三、培训机构

主办单位：厦门城市职业学院

协办单位：厦门腾视科技有限公司、广州泰迪智能科技有限公司

四、培训证书

证书名称：全国工业和信息化人才培养工程职业技术证书

发证单位：工业和信息化部教育与考试中心

五、培训时间与地点

培训时间：2018 年 1 月 29 日—2 月 2 日

报到时间：2018 年 1 月 28 日全天

报到地点：厦门如是酒店会展店（地址：厦门市思明区前埔中路 319 号 电话：0592-2332888）

六、乘车线路

机场：从高崎国际机场到酒店，乘坐出租车约 38 元左右。

火车站：乘坐动车到厦门站的，可乘坐出租车到酒店，车费约 25 元。乘坐动车到厦门北站的，可先乘坐 BRT 快 6 线到前埔枢纽站下车，再乘坐出租车到酒店，车费约 10 元。

七、培训费用及缴费方式

1. 培训费：培训费 3200 元/人（含培训费、资料费、考试费及证书等）。食宿统一安排，费用自理。参加培训教师往返交通费自理。

2. 缴费方式：报到时以刷卡方式缴纳培训费；或者由学校提前通过电汇方式缴纳培训费，办妥电汇手续（汇款时请注明培训班名称：商务数据挖掘企业项目高级研修班），请提前办妥汇款手续以确保开班前到账，并在报到时出具银行电汇单复印件后开具发票。

户名：厦门城市职业学院

账号：35101553001059888888

开户银行：厦门建行槟榔支行

八、报名方式

参训老师请将报名回执表于 1 月 20 日前发邮箱:sjpxb@qq.com, 并请加入“数据挖掘高级研修班”QQ 群（群名片：校名+名字），群号：570050166。（咨询、交流与电子版文件、相关事项说明及交通住宿信息）。联系人：林老师 联系方式：电话：18206091172；QQ：2298543998



群名称：数据挖掘高级研修班
群号：570050166



附件 1:

报名回执表

姓名	单位名称	所在专业	性别	职称/职务	手机	电子邮箱/qq	房型	是否拼房 (标间)	入住 时间	离店 时间
							<input type="checkbox"/> 大床房 <input type="checkbox"/> 标间			
							<input type="checkbox"/> 大床房 <input type="checkbox"/> 标间			
							<input type="checkbox"/> 大床房 <input type="checkbox"/> 标间			

回执联系人: 林老师

电话: 18206091172

邮箱地址: sjpxb@qq.com

附件 2:

培训费开票信息表

学员姓名	单位名称	单位纳税号	开票内容（一般为培训费）

邮箱地址: sjpxb@qq.com

