

# 人大经济论坛统计软件培训中心

台湾辅仁大学

谢邦昌教授

## 站在云端的数据挖掘与商业智能分析课程

### 各位业界同仁、高校师生们:

当前数据挖掘应用主要集中在电信、零售、农业、网络、银行、电力、生物、天体、化工、医药等方面，并且应用领域还在不断地扩展。数据挖掘在市场研究和消费者数据分析等方面有极大的商业价值。而据 Gartner 的报告也指出，数据挖掘会成为未来 10 年内重要的技术之一。而数据挖掘，也已经开始成为一门独立的专业学科。

谢邦昌教授是数据挖掘界领军人物及世界知名统计学家，对数据挖掘的研究一直走在最前沿领域，并且实战经验丰富。国内最大的经管人士网络社区平台-人大经济论坛隆重推出数据挖掘现场培训班，详情如下：

### 一、 讲师介绍:

谢邦昌，湖南耒阳人，台湾大学生物统计学博士，台湾著名大学天主教辅仁大学统计信息学系教授。现任中华数据挖掘协会（Chung-hua Data Mining Society, CDMS）理事长，辅仁大学统计资讯学系教授，华通人商用信息有限公司高级顾问。中国人民大学应用统计科学研究中心学术委员会委员。中国人民大学统计学系 Data Mining 中心客座教授，上海财经大学统计学系客座教授。厦门大学计划统计学系客座教授。西南财经大学客座教授。他是数据挖掘界领军人物及世界知名统计学家。发表过近三百篇关于统计和数据挖掘的论文。出版了近五十余本相关专著。

### 二、 课程特色和培训内容

#### 1、 学员对象

- (1) 从事数据分析相关工作的各业界人员。
- (2) 从事数据挖掘和统计分析等相关领域教学的高校教师。
- (3) 有志于从事金融数据挖掘、商业数据挖掘分析工作和数据挖掘理论研究的学员。
- (4) 有志于学习数据挖掘技术和软件的社会各界人士。

#### 2、 课程目标

讲述云计算的观念及应用

◇ 本课程采用 MS Office Excel 2007(以上版本) Data Mining AddIn 及 R+EXCEL 实现数据挖掘功能，具有简单易学、功能强大、专业等特点。让您的数据挖掘分析变得如此简单、便捷！

- ◇ 本课程系统介绍数据挖掘理论和软件操作以及实际数据的案例分析，为学员奠定扎实的数据挖掘理论基础和实际的分析能力。课程中同时介绍 SAS/EM SPSS Clementine STATISTICA 的观念方法及操作
- ◇ 本课程理论结合实际，更偏重于实际应用。每一讲都提供了几个实际的数据挖掘案例分析，并详细讲解软件的操作和结果的解释。
- ◇ 本课程所提供的海量数据和实际案例都是数据挖掘中心与大型企业合作所作的课题资料。并提供了精心制作的长达 438 页的讲义。

### 3、培训教材

谢邦昌等编着.《Excel 2007 数据挖掘完全手册》. 北京：清华大学出版社. 2008

谢邦昌等编着.《数据挖掘 Clementine 应用实务》. 北京：机械工业出版社. 2008

谢邦昌等编着.《商务智能与数据挖掘 Microsoft SQLServer 应用》. 北京：机械工业出版社. 2008

授课时会同时提供讲课用幻灯片的讲义。

### 三、 课程大纲

#### -----第一天-----

#### 课程名称

#### 课程内容

数据挖掘介绍	云计算应用及前沿综述及介绍数据挖掘概念、功能、技术、步骤，尤其是 CRISP-DM 流程。
数据挖掘运用的理论和技巧	介绍数据挖掘(简单线性回归分析、多元回归分析、岭回归分析、Logistic 回归分析)、关联规则、聚类分析、判别分析、类神经网络分析、判定树分析、其他分析方法
数据挖掘与相关领域的关系	<ul style="list-style-type: none"> <li>◇ 数据挖掘与统计分析的不同</li> <li>◇ 数据挖掘与数据仓储的关系</li> <li>◇ 知识发现与数据挖掘的关系</li> <li>◇ OLAP 与数据挖掘的关系</li> <li>◇ 数据挖掘与机器学习的关系</li> <li>◇ 网络挖掘与数据挖掘的关系</li> </ul>
数据挖掘商业软件产品及其应用现状	<ul style="list-style-type: none"> <li>◇ 数据挖掘商业软件的分类</li> <li>◇ 主要软件的介绍 Data Mining AddIn 及 R+EXCEL SAS/EM SPSS Clementine STATISTICA 的介绍</li> <li>◇ 顾客关系管理</li> <li>◇ 数据挖掘的行业应用</li> </ul>

## -----第二天-----

课程名称	课程内容
准备	<ul style="list-style-type: none"> <li>◇ Data Mining 软件安装与设定数据挖掘加载项</li> <li>◇ 系统需求 开始安装 完成安装验证 组件设定 配置完成检查</li> </ul>
Excel Data Mining AddIn 数据挖掘入门	<ul style="list-style-type: none"> <li>◇ Excel Data Mining AddIn 数据挖掘功能介绍</li> <li>◇ 数据挖掘使用说明 (目录查询、开始功能、视频和教学)</li> <li>◇ 数据挖掘连接配置 (设定目前的连接、跟踪)</li> <li>◇ 数据准备 (浏览数据、清除数据、分割数据)</li> <li>◇ 数据建模</li> <li>◇ 精确度和验证 (准确性图表、分类矩阵、利润图)</li> <li>◇ 模型用法 (浏览功能、查询功能)</li> <li>◇ 模型管理 (重新命名挖掘模型、删除挖掘结构、清除挖掘结构、用原始数据处理挖掘结构、用新数据处理挖掘结构、导出挖掘结构、导入挖掘结构)</li> <li>◇ 数据建模: 开始进行数据挖掘步骤, 可以建立挖掘模型、预测分析等。数据建模的方法有分类、估计、群集、关联、预测等以及进阶。</li> </ul>
分类数据建模	<ul style="list-style-type: none"> <li>◇ 基本概念</li> <li>◇ 决策树模块贝叶斯概率分类 Logistic 回归 类神经网络 的建立</li> <li>◇ 决策树与判别函数比较</li> <li>◇ 计算方法 (确定预测精度的标准、选择分裂 (分层) 技术、定义停止分裂 (分层) 的时间点、选择适当大小的判定树)</li> </ul>

## -----第三天-----

课程名称	课程内容
聚类分析数据建模	<ul style="list-style-type: none"> <li>◇ 基本概念</li> <li>◇ 层次聚类分析的建立</li> <li>◇ 计算方法 (确定预测精度的标准)</li> <li>◇ 范例分析</li> </ul>
预测数据建模	<ul style="list-style-type: none"> <li>◇ 时间序列模型行基本概念 类神经网络 的建立</li> <li>◇ 计算方法 (确定预测精度的标准)</li> <li>◇ 范例分析</li> </ul>

- 
- |                 |  |
|-----------------|--|
| 关联规则            | <ul style="list-style-type: none"><li>◇ 基本概念</li><li>◇ 关联规则的种类</li><li>◇ 关联规则的算法: Apriori 算法</li><li>◇ 范例分析</li></ul>  |
| 时序聚类数据建模        | <ul style="list-style-type: none"><li>◇ 基本概念</li><li>◇ 相关研究和算法</li><li>◇ 范例分析</li></ul>  |
| 估计(推估)数据建模      | <ul style="list-style-type: none"><li>◇ 基本概念</li><li>◇ 决策树模块 线性回归 类 Logistic 回归 类神经网络 的建立</li><li>◇ 计算方法 (确定预测精度的标准)</li><li>◇ 范例分析</li></ul>  |
| 其他分析工具介绍        | <ul style="list-style-type: none"><li>◇ Data Mining AddIn 及 R+EXCEL 工具之介绍</li><li>◇ SAS/EM SPSS Clementine STATISTICA 的观念方法及操作</li><li>◇ 分析关键影响因素</li><li>◇ 检测类别</li><li>◇ 从示例填充</li><li>◇ 预测</li><li>◇ 突出显示异常值</li><li>◇ 应用场景分析 (目标查找、假设)</li></ul> |
| 上市公司投资价值分析的挖掘模型 | <ul style="list-style-type: none"><li>◇ 研究动机与目的</li><li>◇ 挖掘模型的构建</li><li>◇ 变量筛选</li><li>◇ 判定树模型</li><li>◇ 贝叶斯概率模型</li><li>◇ Logistic 回归模型</li><li>◇ 预测准确度比较</li></ul>   |
| 信用卡用户信用评测的挖掘模型  | <ul style="list-style-type: none"><li>◇ 研究背景</li><li>◇ 研究动机</li><li>◇ 研究目的</li></ul>   |
-

◇ 构建数据挖掘模型 (判定树分析、聚类分析、Logistic 回归)

市场营销 ◇ 研究动机与目的

与客户细 ◇ 研究方法与限制

分的挖掘 ◇ 数据分析

模型 ◇ 挖掘建模

讨论答疑 统一回答或者讨论之前学员的疑问

#### 四、 培训时间、地点：

培训时间：2013 年 4 月 29 日-5 月 1 日(三天)

培训地点：上海；南京东路培训中心

培训费用：3600 元(含发票)；差旅及食宿费用自理。全日制学生 2000 元。

授课安排：

(1) 授课方式：使用 Excel 2007(以上版本) Data Mining AddIn 及 R+EXCEL。中文多媒体互动式授课方式

(2) 授课时间：上午 9:00-12:00，下午 13:30-16:30(16:30-17:00 答疑)

**学院集体或多人一起报名学习有优惠，也可独立成班，可直接和我联系。**

#### 五、 报名流程

1. 登录人大经济论坛，提交报名信息。
2. 交费：<http://baoming.pinggu.org/paycenter.aspx>
3. 开课前一周发送培训教室路线图，培训现场领取发票

#### 六、 咨询联系方式

电话: (010)68456523 李老师

手机: 13811729406

Q Q: 1719155708

邮箱: [ivy@pinggu.org](mailto:ivy@pinggu.org)

人大经济论坛统计软件培训中心

北京国富如荷网络科技有限公司

二零一三年二月一日