

# Excel 数据挖掘大纲

## 第 1 章 Excel 如何用于数据挖掘

- 1.1 保证数据正确性
- 1.2 数据的导入
- 1.3 基础函数的掌握
- 1.4 数据透视表的掌握
- 1.5 分析工具库的加载
- 1.6 基础的统计概念和知识

## 第 2 章 数据挖掘概述

- 2.1 数据挖掘的发展历史
- 2.2 统计分析与数据挖掘的主要区别
- 2.3 数据挖掘的主要成熟技术以及在数据化运营中的主要应用  
决策树 神经网络 回归 关联规则 聚类 分类分析 主成分分析 假设检验 等介绍
- 2.4 分析师常见的错误观念和对治的管理策略

## 第 3 章 数据化运营中常见的数据分析项目类型

- 3.1 目标客户的特征分析
- 3.2 目标客户的预测（响应、分类）模型
- 3.3 运营群体的活跃度定义
- 3.4 用户路径分析
- 3.5 交叉销售模型
- 3.6 信息质量模型
- 3.7 服务保障模型
- 3.8 用户（买家、卖家）分层模型
- 3.9 卖家（买家）交易模型
- 3.10 信用风险模型
- 3.11 商品推荐模型
- 3.12 数据产品
- 3.13 决策支持

## 第 4 章 数据挖掘项目完整应用案例演示

- 4.1 项目背景和业务分析需求的提出
- 4.2 数据分析师参与需求讨论
- 4.3 制定需求分析框架和分析计划
- 4.4 抽取样本数据、熟悉数据、数据清洗和摸底
- 4.5 按计划初步搭建挖掘模型
- 4.6 与业务方讨论模型的初步结论，提出新的思路和模型优化方案
- 4.7 按优化方案重新抽取样本并建模，提炼结论并验证模型
- 4.8 完成分析报告和落地应用建议
- 4.9 制定具体的落地应用方案和评估方案
- 4.10 业务方实施落地应用方案并跟踪、评估效果
- 4.11 落地应用方案在实际效果评估后，不断修正完善
- 4.12 不同运营方案的评估、总结和反馈
- 4.13 项目应用后的总结和反思

## 第5章 常见的数据处理技巧

- 5.1 数据的抽取要正确反映业务需求
- 5.2 数据抽样
- 5.3 分析数据的规模有哪些具体的要求
- 5.4 如何处理缺失值和异常值
- 5.5 数据转换
- 5.6 筛选有效的输入变量
- 5.7 共线性问题

## 第6章 各种分析的典型应用和技术小窍门

- 6.1 聚类分析
  - 6.1.1 聚类分析的典型应用场景
  - 6.1.2 主要聚类算法的分类
  - 6.1.3 聚类分析在实践应用中的重点注意事项
  - 6.1.4 聚类分析的扩展应用
  - 6.1.5 聚类分析在实际应用中的优势和缺点
  - 6.1.6 聚类分析结果的评价体系和评价指标
  - 6.1.7 一个典型的聚类分析课题的案例分享
- 6.2 关联分析
  - 6.1.1 关联分析的典型应用场景
  - 6.1.2 主要关联算法的分类
  - 6.1.3 关联分析在实践应用中的重点注意事项
  - 6.1.4 关联分析的扩展应用
  - 6.1.5 关联分析在实际应用中的优势和缺点
  - 6.1.6 关联分析结果的评价体系和评价指标
  - 6.1.7 一个典型的关联分析课题的案例分享
- 6.3 回归分析
- 6.4 检验假设
- 6.5 分类分析