第7章 可视化数据挖掘工具Rattle

教案

**课程名称：**R语言编程基础

**课程类别：**必修

**适用专业：**

**总学时：**48学时（其中理论18学时，实验30学时）

**总学分：x**学分

**本章学时**：6学时

# 材料清单

* 1. 《R语言编程基础》教材
  2. 配套PPT
  3. 引导性提问
  4. 探究性问题
  5. 拓展性问题及其资料
  6. R语言编程相关软件环境

# 教学目标与基本要求

### 教学目标

主要介绍一个用于数据挖掘的工具Rattle，阐述如何安装Rattle，使用Rattle的功能。紧接着使用Rattle完成数据挖掘建模的整个流程，导入CSV、ARFF等数据，对数据进行探索性分析如，构建常见聚类分析、关联规则等算法的模型，以及使用混淆矩阵、ROC图等对模型进行评估。

### 基本要求

1. 了解并安装Rattle工具。
2. 掌握使用Rattle工具导入数据。
3. 掌握使用Rattle工具探索数据的方法。
4. 掌握使用Rattle工具构建模型的方法。
5. 掌握使用Rattle工具估计模型的方法。

# 问题

### 引导性提问

引导性提问需要教师根据教材内容和学生实际水平，提出问题，启发引导学生去解决问题，提问，从而达到理解、掌握知识，发展各种能力和提高思想觉悟的目的。

* 1. Rattle主要功能有什么？
  2. Rattle有哪些优势？
  3. 有那些探索数据的方法？
  4. 常用的算法有哪些？
  5. 常用的模型评价方法？

### 探究性问题

探究性问题需要教师深入钻研教材的基础上精心设计，提问的角度或者在引导性提问的基础上，从重点、难点问题切入，进行插入式提问。或者是对引导式提问中尚未涉及但在课文中又是重要的问题加以设问。

1. Rattle可以导入什么数据？
2. Rattle中可以使用哪些数据探索方法？
3. Rattle可以构建哪些算法模型？
4. Rattle可以使用哪些模型评价方法？

### 拓展性问题

拓展性问题需要教师深刻理解教材的意义，学生的学习动态后，根据学生学习层次，提出切实可行的关乎实际的可操作问题。亦可以提供拓展资料供学生研习探讨，完成拓展性问题。

1. 数据挖掘建模的整个流程是怎么样的？
2. Rattle工具的实际应用？

# 主要知识点、重点与难点

### 主要知识点

1. 了解并安装Rattle工具
2. Rattle工具导入数据的方法
3. Rattle工具探索数据的方法
4. Rattle工具构建模型的方法
5. Rattle工具估计模型的方法

### 重点

1. 了解并安装Rattle工具
2. Rattle工具导入数据的方法
3. Rattle工具探索数据的方法

### 难点

1. Rattle工具构建模型的方法
2. Rattle工具估计模型的方法

# 教学过程设计

### 理论教学过程

1. 认识并安装Rattle
2. 使用Rattle功能
3. 导入CSV、ARFF、ODBC数据
4. R Dataset——导入其他数据源
5. 导入RData File数据集、Library数据
6. 数据总体概况
7. 数据分布探索
8. 相关性、主成分、交互图
9. 聚类分析、关联规则、决策树、随机森林
10. 混淆矩阵、风险图、ROC图及相关图表、模型得分数据集

### 实验教学过程

1. 安装Rattle
2. 在Rattle上导入数据、探索数据、构建模型、评估模型

# 教材与参考书

### 教材

林智章 张良均．R语言编程基础 [M]．北京：人民邮电出版社．2017．

### 参考资料

[1] 张良均、谢佳标、杨坦、肖刚．R语言与数据挖掘．北京：机械工业出版社．2016．