

数据挖掘与商务分析

授课教师:

肖升生

办公室: 科研实验大楼数字经济系 837 室

E-mail: xiao.shengsheng@shufe.edu.cn

联系电话: 021-65904410-837

课程类别:

通识选修课

课程对象:

全校本科生

课程安排说明:

2024 年 9 月 9 日—2025 年 1 月 6 日

每周一晚上: 18:00—19:40

教室: 科研实验大楼 603

授课形式: 线下教学

教学课时数:

$2 \times 18 = 36$ 课时

教学学分:

2 学分

课件网址:

上财教学网 Canvas 平台 (<https://canvas.shufe.edu.cn/>)

在线学习网址:

<https://www.xueyinonline.com/detail/215844957>

教材和参考书目:

指定教材:

- 《数据挖掘导论(完整版)》, Pang-Ning Tan, 人民邮电出版社, ISBN: 9787115241009;

参考书目:

- 《商务智能方法与应用(第2版)》, 刘红岩, 清华大学出版社, ISBN: 9787302558101;

建议先修课程:

概率论与数理统计、高等数学、编程基础

课程教学目的

海量细粒度的数据客观记录了商业管理的各个环节, 信息技术的迅猛发展使得管理实践中对以人工智能、数据挖掘为代表的数据处理和分析技术产生了迫切需求。《数据挖掘与商业分析》是一门从商业数据中发现知识及固有规律的综合课程。课程以目前主流的数据分析语言作为基础, 结合商业管理实例, 重点培养学生使用数据分析技术及工具发现、分析和解决商业管理实践问题的能力。

思政育人元素

课程开展过程中将采用事件分析法、历史沿革法以及实验实践法, 在技术伦理、数据隐

私保护、信息素养、实践创新等方面对同学们展开思政教育。

课程考核形式

将各章节内容适时举行课堂讨论、随堂测验，安排个人平时作业和 1 次期末团队项目报告。其中，团队作业以不超过 3 人一组协作完成并需要进行项目汇报，锻炼学生运用所学知识解决领域问题的能力。课程最终考核方法如下：

平时成绩（含课堂参与、作业等）	45%
期中测验	15%
期末团队实践项目报告	40%

学术诚实

涉及学生的学术不诚实问题主要包括作弊；抄袭；伪造或不当使用在校学习成绩；未经老师允许获取、利用考试材料；使用其他课程汇报材料作为本门课程期末汇报项目材料等。对于学术不诚实的最低惩罚是考试或者作业给予 0 分。其它的惩罚包括报告学校相关部门并按照规定进行处理。

教学要点¹

一、 课程导论

- 1. 数据类型与价值使用
- 2. 数据挖掘、AI 大模型与商业智能
数据挖掘与隐私保护
- 3. 跨行业的数据挖掘流程
- 4. 课程内容与设计
- 5. 课程学习材料

二、 数据预处理

- 1. 数据预处理概述
数据获取的技术伦理
- 2. 数据清理
- 3. 数据抽样
- 4. 数据规约
- 5. 数据变换
- 6. 数据预处理案例

三、 分类分析与商务实践

- 1. 分类分析基本概念
- 2. 决策树分类模型
- 3. 朴素贝叶斯模型

¹ 注：标红部分为课程思政育人元素。

4. 逻辑回归模型
5. 分类模型评估
6. 商务案例分析

四、 关联规则分析与商务实践

1. 关联规则分析基本概念
2. 频繁项集生成
3. 关联规则生成
4. 关联规则分析拓展
5. 商务案例分析

五、 聚类分析与商务实践

1. 聚类分析基本概念
2. 聚类分析基本流程
3. 聚类分析主要方法
4. 聚类分析效果评估
5. 商务案例分析

六、 模型应用开发与部署

1. 模型部署工具介绍
2. 交互式功能开发
3. 模型的部署与分享

七、 自然语言处理与商务实践

1. 文本挖掘、自然语言处理（NLP）基本概念
2. 文本数据预处理
3. 文本表示-向量空间模型
4. 文本表示-主题模型&词向量模型
5. 大语言模型及其最新进展
大语言模型中的伦理问题
6. 商务案例分析

八、 时空数据挖掘与商务实践

1. 时空数据挖掘应用场景及基本概念
2. 时空数据管理
3. 时空数据可视化
4. 时空数据挖掘算法
5. 商务案例分析

九、 社交网络分析与商务实践

1. 社交网络分析基本概念

2. 图论基础
3. 社群识别
4. 影响力度量
5. 商务案例分析

十、推荐系统与商务实践

1. 推荐系统基本概念
2. 基于内容的推荐
3. 基于近邻的协同过滤
4. 基于矩阵分解的协同过滤
5. 商务案例分析

教学进度安排

序号	内 容	课时数	进度
1	课程导论	2	第 1 周
2	数据预处理	2	第 2 周
3	分类分析与商务实践	4	第 3、4 周
4	关联分析与商务实践	4	第 5、6 周
5	聚类分析与商务实践	4	第 7、8 周
6	模型应用开发与部署 期中测验	2	第 9 周
7	自然语言处理与商务实践	4	第 10、11 周
8	时空数据挖掘与商务实践	2	第 12 周
9	社交网络分析与商务实践	2	第 13 周
10	推荐系统与商务实践	4	第 14、15 周
11	商务智能实践项目/案例报告	2	第 16 周
12	期末考试周	4	第 17、18 周
合计		36	

注：第 16 周为学生进行期末团队项目课堂展示