



上海理工大学  
UNIVERSITY OF SHANGHAI FOR SCIENCE AND TECHNOLOGY

## 本科课程教学大纲

数据挖掘实验

制定日期：2023 年 5 月 15 日

## 一、课程基本信息

课程名称	数据挖掘实验					
	<i>Machine Learning and Artificial Intelligence</i>					
课程代码	19101210		开课单位	医学信息工程		
课程负责人	周雷		课程类别	实验类		
课程性质	专业课程		学分	2	学时	32
学时分配	理论	0	实验	16	上机	0
学习负荷	16+16					
教学团队	周雷, 林勇, 尹梓名					
授课语言	中文					
适用专业	医学信息工程, 生物医学工程等					
前修课程	高等数学, python 程序设计					
后续支撑	毕业设计					
课程思政设计	<p>设计思路：通过介绍分析我国在人工智能发展领域上的现状、成就、规划和瓶颈等，说明我国人工智能行业的快速进步和领先地位，增进民族自豪感和自信心；通过课程理论和实验的专业知识教学，培养学生求真务实的科学精神；同时也实事求是的分析我国人工智能行业发展过程中所遇到的困境，鼓励同学们认真学习，毕业后积极投身相关行业，爱岗敬业为国家人工智能行业的发展做出贡献。思政育人目标：培养学生的爱国精神、民族自豪、求真的科学精神和敬业精神。在课堂建设中充分体现“四个意识、四个自信、两个维护”。</p>					
课程简介						

**课程定位:** 本课程的目的是让学生通过代码实现, 深入理解机器学习算法的基本原理、掌握数据分析的基础流程, 主要培养学生利用机器学习技术解决实际问题的能力; 培养学生应用逻辑回归、朴素贝叶斯、支持向量机等技术进行数据分析的能力, 并形成良好的编程习惯和团队合作精神;

**课程内容:** 本课程侧重数据分析方法的练习, 主要培养学生利用机器学习和人工智能等技术解决实际问题的能力; 培养学生应用逻辑回归、朴素贝叶斯、支持向量机等技术进行数据分析的能力, 并形成良好的编程习惯和团队合作精神; 以医学信息应用为背景, 培养学生利用利用机器学习和人工智能等技术解决实际问题的能力;

**核心学习成效:** 练习基本的机器学习算法原理; 培养编程习惯和团队合作精神; 以医学信息应用为背景, 掌握利用利用机器学习和云计算等技术解决实际问题的能力;

**教学方法:** 本门课程以实践教学为主, 主要结合课程实践进行教学。

## 二、课程目标

目标	课程目标	支撑毕业 要求指标点	毕业要求
1	编写决策树、逻辑回归、朴素贝叶斯和 k-means 和 PCA 等基础算法	1. 1, 1. 2, 2. 1, 2. 2, 2. 3, 3. 1, 3. 2, 3. 4, 4. 2, 5. 1, 5. 2, 6. 1, 8. 2, 8. 3, 9. 1, 9. 2, 9. 3, 10. 1, 11. 2, 12. 2	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12
2	深入理解神经网络的基本原理和练习使用	1. 1, 1. 2, 2. 1, 2. 2, 2. 3, 3. 1, 3. 2, 3. 4, 4. 2, 5. 1, 5. 2, 6. 1, 8. 2, 8. 3, 9	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12

	tensorflow、pytorch 等深度学习平台	. 1, 9. 2, 9. 3, 10. 1, 11. 2, 12. 2	
3	熟练掌握使用机器学习算法进行数据分析的方法；练习可用于数据分析的开源软件	5. 1, 5. 2, 9. 1, 9. 2, 9. 3, 10. 1, 11. 2, 12. 2	5, 9, 10, 12

### 三、教学内容

教学模块	教学内容	学生学习 预期成果	教学方式	支撑的 课程目标
一 决策 树算 法练 习	<p>1. 教学内容： 决策树算法练习</p> <p>2. 教学重点： 决策树算法的 python 实现</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政： 1) 富强、爱国、敬业——人工智能的发展现状和成就 2) 求真务实的科学精神和爱岗敬业精神——专业知识的学习</p>	编写完整的决策树实验代码，并在西瓜数据集等数据库上进行分析	<p>教师：</p> <p>线上：案例视频</p> <p>线下：课堂研讨，采用案例教学；国内外典型的人工智能应用介绍等</p> <p>思政教学：</p> <p>线上自行学习</p> <p>引导学生阅读新一代人工智能发展规划，理解人工智能的迅速发展对人类社会生活和世界改变所带来的重要影响。理解国家在为抢抓人工智能发展的重大战略机遇，构筑我国人工智能发展的先发优势，加快建设创新型国家和世界科技强国，所做的布局 and 规划。培养爱</p>	目标 1

			<p>国精神。</p> <p>学生：</p> <p>撰写报告：中国人工智能行业的发展与创新，体现富强、爱国、敬业、文明、和谐</p>	
二 朴素 贝叶 斯算 法练 习	<p>1. 教学内容： 朴素贝叶斯算法练习</p> <p>2. 教学重点： 朴素贝叶斯算法的python实现</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政：无</p>	<p>编写完整的朴素贝叶斯实验代码，并在西瓜数据集等数据库上进行分析</p>	<p>教师：周雷</p> <p>线上：无</p> <p>线下：理论教学，操作示例与课堂研讨</p>	目标 1
三 逻辑 回归 算法 练习	<p>1. 教学内容： 逻辑回归算法练习</p> <p>2. 教学重点： 逻辑回归算法的python实现</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政：无</p>	<p>1. 掌握逻辑回归算法的原理，会进行随机梯度下降算法的推导</p> <p>2. 学会用python实现逻辑回归算法，并进行数据分析</p>	<p>教师：周雷</p> <p>线上：无</p> <p>线下：理论教学，操作示例与课堂研讨</p>	目标 1
四 K-	<p>1. 教学内容： K-means 算法练习</p>	<p>编写完整的 k-means 代码，并在西瓜数据集和医学影像等数据</p>	<p>教师：周雷</p> <p>线上：无</p>	目标 1

means 算法练习	<p>2. 教学重点： K-means 算法的 python 和 matlab 代码实现</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政：无</p>	库上进行分析	线下：理论教学，操作示例与课堂研讨	
五 PCA 算法的介绍	<p>1. 教学内容： PCA 算法练习</p> <p>2. 教学重点： PCA 算法的 python 和 matlab 代码实现</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政：无</p>	编写完整的PCA代码，并在人脸识别数据库上进行分析	<p>教师：周雷</p> <p>线上：无</p> <p>线下：理论教学，操作示例与课堂研讨</p>	目标 1
六 综合练习	<p>1. 教学内容</p> <p>以电子病历等医学信息数据的分析为背景，讲解进行数据分析的完整流程，包括数据收集、数据清理、算法设计、模型训练和模型部署等，并对比决策树算法、逻辑回归算法和朴素贝叶斯算法在相关应用上的性能优缺点。</p>	练习使用机器学习算法进行完整的数据分析流程练习	<p>教师：周雷</p> <p>线上：无</p> <p>线下：理论教学，操作示例与课堂研讨</p>	目标 3

	<p>2. 教学重点</p> <p>数据分析完整流程练习</p> <p>3. 教学难点</p> <p>实现从数据采集到数据分析的完整流程</p> <p>4. 课程思政：无</p>			
七 神经网络和深度学习练习	<p>1. 教学内容： 深度学习算法练习</p> <p>2. 教学重点： 使用 pytorch 和 tensorflow 平台，进行深度学习网络搭建和训练</p> <p>3. 教学难点： 完成数据预处理，编写代码</p> <p>4. 课程思政：无</p>	编写完整的语义分割代码，并在医学影像分割数据库上进行分析	<p>教师：周雷</p> <p>线上：无</p> <p>线下：理论教学，操作示例与课堂研讨</p>	目标 2

#### 四、教材与学习资源

课程网站	
课程教材	[1] 机器学习(第一版)，周志华，清华大学出版社，2016 年



参考书目	[1] 机器学习实战 [Machine learning in action], Peter Harrington 著, 李锐, 李鹏, 曲亚东 等译, 人民邮电出版社, 2013 年
教学条件	使用线上线下混合教学的方式, 线上教学继续使用超星平台, 学生可观看相关的课程视频。

## 五、教学进程安排

序号	教学内容	课内学时	课外学时	课外学习内容
1	决策树算法练习	2	2	完成布置的作业
2	朴素贝叶斯算法练习	2	2	完成布置的作业
3	逻辑回归算法练习	2	2	完成布置的作业
4	K-means 算法练习	2	2	完成布置的作业
5	PCA 算法的介绍	2	2	完成布置的作业
6	综合练习	2	2	完成布置的作业
7	神经网络和深度学习练习	4	4	完成布置的作业
	合计	16	16	

## 六、课程考核

课程目标	考核要点	考核与评价方式及成绩比例 (%)					成绩比例 (100%)
		过程考核				期末 考试	
		作业	报告	设计	自测		
1	对理论推导的掌握程度					40%	40%
2	对算法实现的理解程度			10%		20%	30%
3	对数据分析流程的掌握程度			10%			10%
4	对前沿热点的掌握程度		10%				10%
5							
合计			10%	20%		60%	100%
<b>期末考试资格</b>							
无故缺课 3 次及以下且作业完成超过二分之一							
<b>期末考试形式</b>							
<input type="checkbox"/> 闭卷笔试 <input checked="" type="checkbox"/> 开卷/半开卷 <input type="checkbox"/> 小论文 <input type="checkbox"/> 报告 <input type="checkbox"/> 口试 <input type="checkbox"/> 作品 <input type="checkbox"/> 口笔试兼用 <input type="checkbox"/> 上机 <input type="checkbox"/> 技能操作 <input type="checkbox"/> 其他 (请注明) _____ (必填)							

附件：各类考核评分标准表

大纲制定：应填写课程负责人

大纲审核：应填写专业负责人或教研室主任等

制定单位：应填写课程归属单位（敲章）

制定日期：2023年5月 日

附件：各类考核评分标准表

机器学习与人工智能评分标准

课程目标	评分标准				权重 (%)
	90-100	80-89	60-79	0-59	
编写决策树、逻辑回归、朴素贝叶斯和 k-means 和 PCA 等基础算法	能够准确的运用所学机器学习算法系统性地进行分析	能够运用所学编写相关算法代码并应用	能够推导核心机器学习算法	了解核心机器学习算法的原理	60
深入理解神经网络的基本原理和练习使用 tensorflow、pytorch 等深度学习平台	能使用深度学习平台进行医学影像等数据分析	能够推导神经网络反向传播训练过程，并运用所学的基础知识编写相关算法代码并应用	能正确安装和使用 tensorflow、pytorch 等深度学习平台	了解神经网络算法和深度学习的原理	20
熟练掌握使用机器学习算法进行数据分析的方法；练习可用于数据分析的开源软件	能使用经典机器学习算法或深度学习熟练进行数据分析	能搭建和实现数据分析完整的流程	能设计数据分析的方案	了解利用机器学习进行数据分析的基本流程	20