

《数据分析(Python)》本科课程教学大纲

一、课程基本信息

课程名称	数据分析 (Python)				
	Data Analytics (Python)				
课程代码	2055059	课程学分		2	
课程学时	16	理论学时	16	实践学时	32
开课学院	信息技术学院	适用专业与年级		软件工程大三及以上	
课程类别与性质	专业选修	考核方式		考查	
选用教材	Python 数据分析与机器学习（微课视频版）			是否为马工程教材	否
先修课程	面对对象程序设计、概率论与数理统计				
课程简介	《数据分析(Python)》是软件工程和计算机类相关专业的专业课程。该课程介绍了 Python 语言的基础知识，数据分析和可视化基础，机器学习简介和数据预处理、模型评估方法等内容。本课程专为希望掌握数据分析技能的初学者和有一定编程基础的学生设计，通过使用 Python 这一强大的编程语言，学习如何处理、分析和可视化数据，为后续深入机器学习，人工智能学习奠定基础。该门课程的教学目标有三项，分别为掌握数据分析理论知识、培养数据分析技能、解决实际问题。为了达成这个三个教学目标，在理论课时上，会着重介绍 Python 语言的基础知识和数据分析的基础知识，可视化图表、数据挖掘模型、机器学习方法概述等相关知识。为了培养数据分析技能，在实践课时上会着重讲解如何使用 Matplotlib 进行图表绘制，运用 Scikit-learn 中的有监督和无监督机器学习模型库对数据进行分类，识别等，并能理解各种参数的含义、评价模型的性能优劣、分析结果的质量等。				
选课建议与学习要求	本课程为计算机类专业的选修课，是一门实践性较强的课程。要求先修课程包括编程基础类课程和概率与统计类课程，还需具备良好的自我学习和解决问题的能力，良好的团队协作和乐于分享的精神。能够在课堂分组的模式下，与小组成员合作完成课程中所布置的课后作业，实验报告，项目展示等等任务。				
大纲编写人	王晨啸		制/修订时间	2025 年 1 月	
专业负责人	朱丽娟		审定时间	2025 年 1 月	
学院负责人	靳桂娥		批准时间	2025 年 1 月	

二、课程目标与毕业要求

（一）课程目标

类型	序号	内容
知识目标	1	能够将所学的统计学相关知识应用于实际项目中，解决商业、科研或社会领域的数据分析问题，如市场分析、客户细分、预测建模等，并能够从数据中发现问题、提出假设并通过分析验证，形成基于数据的决策思维。
	2	了解监督学习、无监督学习和强化学习的区别，掌握常见的机器学习算法。能够使用 Python 语言构建、训练和评估机器学习模型。通过实际案例，完成从数据预处理到模型部署的完整机器学习项目流程。
技能目标	3	掌握 Pandas 库，Numpy 库和 Matplotlib 库的使用：做到能够使用这些库进行数据导入、数据清洗（处理缺失值、重复值）、数据筛选、分组聚合等操作和对数组的创建、索引、切片和基本数学运算以及绘制基本图表（如折线图、柱状图、散点图、饼图、热力图等），做到清晰、准确地传达数据洞察。
	4	能够使用 Scikit-learn 库构建、训练和评估机器学习模型（如线性回归、逻辑回归、决策树、K-Means 聚类等），掌握模型选择、交叉验证和性能评估指标（如准确率、召回率、F1 分数等）。
素养目标 (含课程思政目标)	5	培养学生理解数据隐私、数据安全和数据伦理的重要性，遵守相关法律法规和职业道德准则。

（二）课程支撑的毕业要求

L01 工程知识：能够将数学、自然科学、工程基础和专业知识用于解决软件系统中的复杂软件工程问题。②能够针对具体对象或问题，建立合适的数学模型，并能够编写计算机程序求解。
L05 使用现代工具：能够针对软件系统中的复杂软件工程问题，开发、选择与使用恰当的技术、资源和工具，对软件系统中的复杂工程问题进行预测与模拟，并能够理解其局限性。①了解软件开发过程中所涉及到的主要的软件开发平台、开发与管理工具的使用原理和方法，了解其差异和适用的领域，并理解其局限性。

（三）毕业要求与课程目标的关系

毕业要求	指标点	支撑度	课程目标	对指标点的贡献度
------	-----	-----	------	----------

L01	②	M	1 能够将所学的统计学相关知识应用于实际项目中，解决商业、科研或社会领域的数据分析问题，如市场分析、客户细分、预测建模等，并能够从数据中发现问题、提出假设并通过分析验证，形成基于数据的决策思维。	60%
			2 了解监督学习、无监督学习和强化学习的区别，掌握常见的机器学习算法。能够使用 Python 语言构建、训练和评估机器学习模型。通过实际案例，完成从数据预处理到模型部署的完整机器学习项目流程。	40%
L05	①	H	3 掌握 Pandas 库，NumPy 库和 Matplotlib 库的使用：做到能够使用这些库进行数据导入、数据清洗（处理缺失值、重复值）、数据筛选、分组聚合等操作和对数组的创建、索引、切片和基本数学运算以及绘制基本图表（如折线图、柱状图、散点图、饼图、热力图等），做到清晰、准确地传达数据洞察。	50%
			4 能够使用 Scikit-learn 库构建、训练和评估机器学习模型（如线性回归、逻辑回归、决策树、K-Means 聚类等），掌握模型选择、交叉验证和性能评估指标（如准确率、召回率、F1 分数等）。	30%
			5 培养学生理解数据隐私、数据安全和数据伦理的重要性，遵守相关法律法规和职业道德准则。	20%

三、课程内容与教学设计

（一）各教学单元预期学习成果与教学内容

第一单元：Python 语言基础 本单元重点学习 Python 基础，侧重于数据分析中会使用到的语法知识，为今后课程打好基础。知识点包括 Python 语言开发环境的配置、基础语法、函数、类与对象等。 第二单元：Numpy 数据处理基础 本单元主要学习使用 Numpy 库中对数组的操作。包括 Numpy 数据结构、数据的准备、数组的索引与切片、数组的基本运算、数组的排序、数组的组合、数组的分割等。 第三单元：Matplotlib 数据可视化基础 本单元主要学习如何使用 Matplotlib 库对数据进行可视化。需要掌握各种图标的绘制，包括折线图、柱状图、散点图、饼图、热力图、词云图等等，并且能够根据要求的样式进行参数调整。 第四单元：Pandas 数据处理与分析 本单元主要学习 Pandas 库对数据进行清洗和预处理。内容包括，数据的导入、缺省值、

异常值处理、类型转换、合并和连接、排序和排序、透视表等功能的使用。

第五单元：有监督学习方法基础

本单元主要学习有监督类机器学习方法的理论知识。介绍常见的模型，包括回归算法、分类模型的理论知识以及在 Scikit-learn 库中如何实现以及学习性能评估、参数调整等知识。

第六单元：无监督学习方法基础

本单元主要学习无监督类机器学习方法的理论知识。介绍常见的模型，包括聚类模型、关联规则等的理论知识以及在 Scikit-learn 库中如何实现以及学习性能评估、参数调整等知识。

第七单元：初识深度学习

本单元主要介绍卷积神经网络（CNN）的基本知识。通过识别图片中的数字这个经典案例，介绍 CNN 模型的搭建、训练和参数调优以及应用的过程中的理论知识并使用 TensorFlow 库进行实现。

（二）教学单元对课程目标的支撑关系

课程目标	1	2	3	4	5
教学单元					
第一单元：Python 语言基础		√	√	√	
第二单元：Numpy 数据处理基础	√				√
第三单元：Matplotlib 数据可视化基础	√		√		√
第四单元：Pandas 数据处理与分析	√		√		√
第五单元：有监督学习方法基础	√	√		√	√
第六单元：无监督学习方法基础	√	√		√	√
第七单元：初识深度学习	√	√		√	√

（三）课程教学方法与学时分配

教学单元	教与学方式	考核方式	学时分配		
			理 论	实 践	小 计
第一单元：Python 语言基础	理论授课	课后作业	2	2	4
第二单元：Numpy 数据处理基础	演示和上机操作	实验报告	2	2	4
第三单元：Matplotlib 数据可视化基础	演示和上机操作	实验报告	2	4	6
第四单元：Pandas 数据处理与分析	演示和上机操作	实验报告	2	2	4
第五单元：有监督学习方法基础	理论授课	实验报告、期末汇报	2	2	4
第六单元：无监督学习方法基础	理论授课	实验报告、期末汇报	2	2	4

第七单元：初识深度学习	理论授课	实验报告	4	2	6
合计			16	16	32

（四）课内实验项目与基本要求

序号	实验项目名称	目标要求与主要内容	实验时数	实验类型
1	Numpy 和 Pandas 库数据处理	能够掌握使用 Numpy 库和 Pandas 库中的方法，完成对数据的基本操作。	4	②
2	Matplotlib 数据可视化	能够使用 Matplotlib 库完成如折线图、柱状图、散点图、饼图、热力图等绘制，并对图表按要求完成客制化。	4	②
3	机器学习模型实践	能够使用 Scikit-learn 库实现使用有监督和无监督机器学习模型解决实际数据分析问题。	6	②
4	深度学习基础	能够使用 Tensorflow 库实现使用神经网络识别图片中的数字。	2	①

实验类型：①演示型 ②验证型 ③设计型 ④综合型

四、课程思政教学设计

该门课程思政建设应注重将专业知识与社会主义核心价值观、职业道德、社会责任等有机结合，引导学生树立正确的价值观、职业观和社会责任感。在教学设计方面，可以“大数据杀熟”等案例为切入点，引导学生思考数据滥用、隐私泄露等伦理问题，以及数据分析师应承担的社会责任。组织学生讨论数据收集、使用、存储等环节中的伦理规范，引导学生树立数据安全意识，遵守数据伦理准则，负责任地使用数据并且同时，鼓励学生思考如何利用数据分析技术传承和弘扬中华优秀传统文化，以及如何在数据分析工作中维护国家数据安全。

五、课程考核

总评构成	占比	考核方式	课程目标					合计
			1	2	3	4	5	
X1	40	期末汇报大作业	20	30	20	20	10	100
X2	35	平时作业	-	30	30	30	10	100
X3	25	平时课堂表现	30	20	20	10	20	100

六、其他需要说明的问题

--