《大数据商务分析与应用》教学大纲

课程编号：071133B

课程类型：□通识教育必修课 □通识教育选修课

□学科基础课 □专业核心课

☑专业提升课 □专业拓展课

总 学 时：48 讲课学时：32 实验（上机）学时：16

学　　分：3

考试类型：□考试 ☑考查

适用对象：电子商务（商务智能）

□是 ☑否 适合作为其他专业学生的个性化选修课

先修课程：数据库原理，统计学

一、教学目标

商务数据特别是大数据的挖掘、分析和利用，已经成为当前大型企业之间竞争的重要工具和手段。本课程主要学习大数据商务分析的相关原理、方法和技术，根据实际业务需求，基于一种大数据商务分析开发工具如Python对相应的业务数据进行处理、分析和应用。

1、课程专业教学目标

通过本课程学习，使学生了解大数据的采集、传输、处理、分析和应用的相关技术，掌握Python的基础语言知识、大数据分析和应用技术，能够基于Python进行数据采集、处理、分析和应用开发工作。鉴于商务数据分析已经成为电子商务运营和网络营销中重要的工具和手段，因此该课程是电子商务专业培养方案的重要组成部分，是先修课程“统计学”和“商务智能与数据挖掘”的实战延续，并可为同期进行的专业必修课“电子商务运营管理”打下良好技术基础。

2、课程思政目标

本课程思政教育的总体目标是把马克思主义立场观点方法的教育与科学精神的培养结合起来，提高学生正确认识问题、分析问题和解决问题的能力。具体目标包括：（1）强化工程伦理教育。教育引导学生深刻理解并自觉实践数据科学伦理，增强职业责任感，激发学生科技报国的家国情怀和使命担当；（2）科学思维方法的训练。培养学生探索未知、追求真理、勇攀科学高峰的责任感和使命感；（3）社会主义道德。教育引导学生深刻理解信息安全与数据滥用违法带来的道德和法律问题，培养遵纪守法、爱岗敬业、诚实守信的职业品格和行为习惯。

二、教学内容及其与毕业要求的对应关系

1、教学内容

本课程主要教学内容是在学生掌握基本的统计学、数据挖掘等基本原理和算法后，学习基于Python编程语言的基本语法、一般应用开发方法和常用的数据分析和数据挖掘方法等。具体包括：Python基础知识，Python数据分析工具，数据探索，数据预处理，挖掘建模，案例分析等。

2、教学方法和手段

根据教学目标，拟采用的教学方法有：课堂讲解基本概念和核心知识，讲授和讨论相结合领会知识要点，案例教学训练解决问题的能力，最后借助Python开发环境软件让学生进行上机操作和具体实践。

3、实践教学环节要求

根据教学进度和要求布置相应的小作业，通过上机实践。每一章都有对应的上机内容，如Python开发环境配置与入门程序开发、各章节开发例程等。建议自学时安装和配置这些软件进行自学，完成书上案例的内容。通过上机学习本课程的理论，掌握基于Python语言和工具的数据分析方法。

4、学习要求

为有效学习本课程，要求学生首先具备数据库、面向对象程序设计与开发等方面的基本知识，熟练掌握面向数理统计与数据挖掘等方面的完整专业知识。

自学时应该熟读大纲，提纲挈领地掌握Python数据分析的内容，随后按照大纲熟读教材，并通过课后思考和上机实践进行多角度和多层次的反复学习。

5、与毕业要求的关系

随着大数据时代的到来，大型企业、政府和机构逐渐积累了大量的业务数据，基于互联网技术和大数据分析手段挖掘商务和业务应用的价值，逐渐成为电子商务运作和企业竞争的必要组成部分，也是未来发展的必然趋势。因此，学习和掌握大数据的商务分析方法、技术及相关应用，不仅是当前电子商务产业发展的需要，还是适应未来大型组织运营和竞争的基本要求，也是电子商务专业培养的基本要求。

三、各教学环节学时分配

**教学课时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 章节内容 | 讲课 | 实验 | 其他 | 合计 |
| 第一章 | Python基础知识 | 8 | 4 |  | 12 |
| 第二章 | Python数据分析工具 | 8 | 4 |  | 12 |
| 第三章 | 数据探索 | 4 | 2 |  | 6 |
| 第四章 | 数据预处理 | 4 | 2 |  | 6 |
| 第五章 | 挖掘建模 | 6 | 2 |  | 8 |
| 第六章 | 案例分析 | 2 | 2 |  | 4 |
| **合计** | | **32** | **16** |  | **48** |

四、教学内容

**1、第一章 Python基础知识**

第一节 大数据与商务分析概述

第二节 Python基础语法和数据类型

第三节 文件系统和数据库访问

教学内容：概要介绍Python开发工具与安装配置，数据类型与数据组织方法，基本流程语句和使用方法，常用类和库的方法和接口，文件系统和数据库开发等。

重点和难点：Python数据类型与数据组织方法，常用类和库的方法和接口。

考核要求：使学生了解常见的Python集成开发工具和资源，深入理解Python数据类型与数据组织及数据转换方法，能够应用Python基础语言和类库进行基本的文字处理和数据存取操作。

实验一：Python开发环境搭建与入门程序开发。

实验二：学生信息的存储与读取处理（基于文件或数据库）。

课程思政切入点：提出大数据和数据科学所带来的伦理问题，培养学生树立职业道德和职业伦理理念。

**2、第二章 Python数据分析工具**

第一节 Numpy基础

第二节 Pandas基础

第三节 Matplotlib数据可视化基础

第四节 科学计算与统计基础：Scipy与StatsModels

教学内容：常用数据分析工具的介绍与练习，包括Numpy，Scipy，Matplotlib，Pandas，StatsModels等。

重点和难点：Numpy、Matplotlib和Pandas类库的常用接口形式、数据组织格式及常用绘图工具的使用和开发。

考核要求：了解Python常用数据分析工具的种类、安装配置与功能定位，掌握Numpy、Matplotlib和Pandas类库的编程接口与使用方式，能够应用上述类库进行初步的数据统计。

实验三：给定样本数据，练习使用Numpy和Pandas进行基本的统计分析。

实验四：给定样本数据，练习使用Matplotlib进行基本的统计绘图。

**3、第三章 数据探索**

第一节 统计与概率基础

第二节 假设检验与推断

第三节 数据特征分析与数据探索

教学内容：统计与概率基础，假设检验与推断，数据特征分析，Python主要数据探索函数。

重点和难点：Python主要数据探索函数的使用与结果分析。

考核要求：了解数据质量分析和数据特征分析常用方法，熟练掌握Python主要数据探索函数的使用方法，理解数据探索分析结果。

实验五：电子商务网站用户行为的数据探索分析。

课程思政切入点：基于Python进行数据探索时需要注意的问题、细节和规范，引导学生养成良好的职业品格和行为习惯，培养学生的社会主义工匠精神。

**4、第四章 数据预处理**

第一节 数据清洗

第二节 数据集成与转换

第三节 数据规约

教学内容：数据清洗，数据集成，数据变换，数据规约。

重点和难点：Python主要数据预处理函数的使用。

考核要求：了解数据清洗、数据集成、数据变换和数据规约的基本概念和原理，掌握Python主要数据预处理函数的使用。

实验六：应用系统运行或电子商务消费者行为数据的预处理练习，重点练习数据集成与数据转换。

**5、第五章 挖掘建模**

第一节 分类问题

第二节 聚类问题

第三节 关联分析

第四节 时间序列分析与预测

教学内容：分类与预测，聚类分析，关联规则，时序分析。

重点和难点：基于Python的分类与预测及关联规则的相关算法和工具使用。

考核要求：了解Python的分类分析、聚类分析、时间序列分析的相关算法，掌握基于Python的分类与预测及关联规则的类包工具使用，理解数据分析结果。

实验七：给出样本数据，练习使用分类与预测及关联规则的数据分析工具，并对分析结果进行解释。

课程思政切入点：给出电子商务运营或政务管理等实际场景，提出基于数据分析的判断和决策问题，训练学生科学思维方法，提高学生正确认识问题、分析问题和解决问题的能力。

**6、第六章 案例分析**

第一节 课程总结

第二节 案例剖析

教学内容：综合利用已学数据分析和数据挖掘知识，运用Python对商业案例数据进行分析和预测。

重点和难点：数据分析各阶段工具的综合利用与结果解释。

考核要求：能够综合运用Python提供的各类分析工具，对给定的商业案例数据进行必要的分析，并能够理解分析结果。

实验八：电商产品评论数据的综合数据分析。

五、考核方式、成绩评定

本课程采用平时、期末作业相结合的考核方式，尤其注重实践环节的考核。各部分所占比重为平时成绩占50%，期末大作业成绩占50%。

平时成绩包括考勤成绩（10%）和上机作业成绩（40）。

期末考核采用大作业形式，指定某个具体业务场景和相应业务数据，要求学生能够对数据进行处理、编码和分析，最后提供相应的数据分析方案和分析报告。考核方式注重知识应用能力和解决问题能力的考核，需要学生提交自己开发的基于Python的数据分析方案及分析报告。

六、主要参考书及其他内容

[1] 赫苏斯·罗格尔-萨拉查（Jesús Rogel-Salazar） 著，白皓 刘江一 上官明乔 刁娟译 译．数据科学与分析：Python语言实现．北京：机械工业出版社．2019.04

[2] Fabio Nelli.Python Data Analytics(2nd Edition).美国加州：Apress.2018.10.

执笔人：付东普 教研室主任： 系教学主任审核签名：