《机器学习与量化投资》教学大纲

课程编号：2121493A

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：信息管理与信息系统专业

□是 ☑否 适合作为其他专业学生的个性化选修课

先修课程：程序设计基础与应用、数据结构、高等数学

一、教学目标

1、学习目标

《机器学习与量化投资》是信息管理与信息系统专业的一门专业核心课程。本课程是面向信息管理与信息系统开设的专业核心课。其教学重点是使学生掌握常见机器学习算法，包括算法的主要思想和基本步骤，并通过编程练习和典型应用实例加深了解；同时对机器学习的一般理论，如假设空间、计算学习理论，以及无监督学习和强化学习有所了解。量化投资部分介绍机器学习在量化投资方面的具体分析与应用，包括对未来股票价格的预测、大盘指数趋势分析等。

学生通过本门课程的学习，能够对机器学习和量化投资的内容有一个较为全面的了解和认识，更深刻地理解机器学习的实质内容，使学生具备前沿的计算机技术必要的专业知识。从而，为学生今后从事信息管理与信息系统与量化投资前沿研究，以及相关领域的科学研究做好理论和技术上的准备。

目标(1)： 通过对机器学习与量化投资基本概念、原理和基本方法的讲解，让学生理解并掌握机器学习和量化投资的基本技术。

目标(2)：培养学生利用机器学习方法，运用技能解决本专业和相关领域的实际问题的能力。

目标(3)：鼓励学生运用知识解决各自学科的实际问题，培养他们的独立科研的能力和理论联系实际的能力。

2、课程思政目标

《机器学习与量化投资》是大学本科生非常重要的一门课程，可以为学生后续专业课程的学习和以后工作奠定基础。应该将辩证唯物主义、政治思想道德、社会主义核心价值观、时代精神、家国情怀等元素融入课堂，贯穿整个课堂教学。

目标(1)：树立正确“三观”塑造良好人格。

目标(2)：培养学生民族自豪感和自尊心。

目标(3)：渗透社会主义核心价值观。

1. 教学内容及其与毕业要求的对应关系

本课程将以从入门者的角度，讲解Python 语言和sklearn 模块库内置的各种经典机器学习算法；介绍了股市外汇、比特币等实盘交易数据在金融量化方面的具体分析与应用，包括对未来股票价格的预测、大盘指数趋势分析等。通过教学和实践，培养学生运用数学工具和方法分析问题和从多角度运用数学工具解决问题的基本能力，培养学生利用模式识别方法，运用技能解决本专业和相关领域的实际问题的能力，训练学生的逻辑思维能力和想象力。

在教学中，鼓励学生运用算法知识解决各个学科的实际计算问题，培养学生初步的独立开展科研工作的能力和理论联系实践，解决实际问题的能力，同时，为后续课程以及将来的研究工作提供必要的基础。

此外，配合实验课程的教学，学生应理论联系实际，理论指导实践，通过规范地完成项目实践进一步巩固所学的相关书本知识，在知识、能力、素质上得到进一步的提高。

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 与机器学习 | 2 |  |  | 2 |
| 第二章 | 机器学习编程入门 | 2 |  |  |  |
| 第三章 | 金融数据的预处理 | 2 |  |  | 2 |
| 第四章 | 机器学习快速入门. | 2 | 2 |  | 4 |
| 第五章 | 模型验证优化 | 2 | 2 |  | 4 |
| 第六章 | 决策树 | 2 | 2 |  | 5 |
| 第七章 | 随机森林算法和极端随机树算法 | 2 | 2 |  | 4 |
| 第八章 | 机器学习算法模式 | 2 | 2 |  | 4 |
| 第九章 | 概率编程 | 2 | 2 |  | 4 |
| 第十章 | 实例算法 | 2 |  |  | 2 |
| 第十一章 | 正则化算法 | 2 |  |  | 2 |
| 第十二章 | 聚类分析代 | 2 | 2 |  | 4 |
| 第十三章 | 降维算法 | 2 |  |  | 2 |
| 第十四章 | 集成算法 | 2 |  |  | 2 |
| 第十五章 | 支持向量机 | 2 | 2 |  | 5 |
| 第十六章 | 人工神经网络算法 | 2 |  |  | 2 |
| **合计** |  | 32 | 16 |  | 48 |

四、教学内容

第1 章 Python 与机器学习

**课程思政切入点：**介绍我国在这门课程研究领域的成就让学生了解我国机器学习的发展历史和辉煌成就，从而增加民族自豪感和荣誉感，激发爱国热情。

**教学内容：**scikit-learn 模块库、开发环境搭建

**教学重点、难点：**开发平台搭建

**课程的考核要求：**重点模块版本测试

第2 章 机器学习编程入门

**教学内容：**经典机器学习算法、经典爱丽丝、机器学习数据集、线性回归算法

**教学重点、难点：**经典机器学习算法

**课程的考核要求：**机器学习算法流程.

第3 章 金融数据的预处理

**课程思政切入点：**讲授数据处理的重要性，要秉承实事求是的实验态度和价值观要求。不实的数据分析结果有违现实情况，通过这种方式获得的虚假信息阻碍我国经济发展，影响社会发展。

**教学内容：**至简归一法、股票池与Rebase、金融数据切割

**教学重点、难点：**金融数据切割

**课程的考核要求：**金融数据切割

第4 章 机器学习快速入门

**教学内容：**回归算法、LR 线性回归模型、常用评测指标、多项式回归.

**教学重点、难点：**回归算法、常用评测指标

**课程的考核要求：**常用评测指标

第5 章 模型验证优化

**课程思政切入点：**要求同学们潜心学习，掌握技能、用知识与能力像老一辈革命家那样为祖国效力。

**教学内容：**交叉验证评估器、交叉验证评分

**教学重点、难点：**交叉验证

**课程的考核要求：**交叉验证

第6 章 决策树

**教学内容：**决策树学习基础知识、决策树学习基本算法、决策树学习常见问题、决策树学习理解解释

**教学重点、难点：决**策树学习基本算法

**课程的考核要求：**决策树学习理解解释

第7 章 随机森林算法和极端随机树算法

**课程思政切入点：**鼓励学生利用本门课程的学习参与到统计实践中，关注民生民意，了解和挖掘社会热点问题，讨论是否可以解决社会中的短板问题，

**教学内容：**随机森林函数、决策树测试函数极端随机树算法

**教学重点、难点：**极端随机树回归算法

**课程的考核要求：**极端随机树函数

第8 章 机器学习算法模式

**教学内容：**学习模式、机器学习五大流派

**教学重点、难点：**经典机器学习算法

**课程的考核要求：**经典机器学习算法

第9 章 概率编程

**课程思政切入点：**调动学生的主观能动性，将课堂上的理论与实践结合，从大学生角度对国家社会建言献策，树立学生的主人翁精神。

**教学内容：**朴素贝叶斯的上证之旅、隐马尔可夫模型

**教学重点、难点：**上证朴素贝叶斯算法

**课程的考核要求：**HMM 算法与模型读取

第10 章 实例算法

**教学内容：**最近邻学习基础知识、最近邻学习基本思想、最近邻学习常见问题

**教学重点、难点：**最近邻学习基本思想、最近邻学习常见问题

**课程的考核要求：**最近邻学习基本思想

第11 章 正则化算法

**课程思政切入点：**通过典型案例，从辩证唯物主义的角度去理解算法思想。从而树立正确的方法论和世界观，渗透社会主义核心价值观。

**教学内容：**岭回归算法、套索回归算法、弹性网络算法、小角回归算法

**教学重点、难点：**套索回归算法应用、弹性网络算法应用

**课程的考核要求：**正则化算法应用

第12 章 聚类分析

**教学内容：**K 均值算法、BIRCH 算法

**教学重点、难点：**K 均值算法应用、BIRCH 算法应用

**课程的考核要求：**K 均值算法应用、BIRCH 算法应用

第13 章 降维算法

**课程思政切入点：**从算法机理分析学生勤奋努力学习对学习成就的重要性，实现对学生知识传授、价值引领、能力提升的培养目标，引导学生爱国敬业、积极进取。

**教学内容：**主成分分析、奇异值分解算法

**教学重点、难点：**主成分分析的应用、PCA 算法的上证戏法

**课程的考核要求：**主成分分析和奇异值分解算法

第14 章 集成算法

**教学内容：**sklearn 内置集成算法、装袋算法、AdaBoost 迭代算法

**教学重点、难点：**AdaBoost 迭代算法

**课程的考核要求：**算法应用

第15 章 支持向量机

**课程思政切入点：**通过典型案例为学生讲解科学家们的典型事迹，鼓励学生以严谨认真的态度对待学习、工作和生活。

**教学内容：**最大边缘超平面、线性支持向量机、非线性支持向量机

**教学重点、难点：**线性支持向量机、非线性支持向量机

**课程的考核要求：**线性支持向量机

第16 章 人工神经网络算法.

**课程思政切入点：**通过讲授算法发展历史，为学生讲解老一辈“为中华之崛起”而读书的光荣事迹，让学生明确学习志向。

**教学内容：**神经网络的定义、神经网络的发展历史、M-P神经元模型、单层感知机、多层前馈神经网络、深层神经网络

**教学重点、难点：**神经网络的正则化。

**课程的考核要求：**熟悉各种神经网络及其梯度下降的概念。

五、考核方式、成绩评定

课程考核方式是考试，考试时间是120分钟。

考试命题：课程考试的命题内容要从大纲的要求出发，围绕本课程的教学内容、知识点和教学要求，着重从知识、能力、素质三个方面对学生进行全面的考核，重点考核学生运用知识解决问题的能力，同时考察学生的综合素质。考核范围为除了最后一周教学的内容外，其他大纲确定的知识点都在考试范围之内。

最后成绩计算办法：期终考试成绩70%+ 平时成绩30%

六、主要参考书及其他内容

[1] 何海群，《Python机器学习与量化投资》，电子工业出版社，2018.12

执笔人：邱月 教研室主任：胡磊 系教学主任审核签名：