**《机器学习实训》教学大纲**

|  |  |
| --- | --- |
| 课程代码 |  |
| 课程名称 | 机器学习实训 |
| 英文名称 | Machine learning training |
| 课程类别 | 集中实践课 |
| 课程性质 | 必修 |
| 学时 | 集中实践教学：2周 |
| 学分 | 2 |
| 开课学期 | 第三学期 |
| 开课单位 | 未来技术学院 |
| 适用专业 | 人工智能、数据科学与大数据技术 |
| 授课语言 | 英文授课 |
| 先修课程 | Python程序设计基础、机器学习 |
| 课程对毕业要求的支撑 | 本课程对学生达到如下毕业要求有如下贡献：  №3. 解决方案：能够设计针对人工智能/大数据领域相关复杂工程问题的解决方案，设计满足特定需求的系统、单元或流程，并能够在设计环节中体现创新意识，考虑社会、健康、安全、法律、文化及环境等因素。  №3.2 了解人工智能/大数据领域技术发展的现状与趋势，能够在方案设计中体现创新意识；  №9. 个人和团队：能够在计算机跨学科背景下的团队中承担个体、团队成员以及负责人的角色。  №9.1 能够在人工智能/大数据领域相关研究、开发和生产的团队中承担个体或成员角色，具有团队合作精神或意识；  №11. 项目管理：理解并掌握工程管理原理与经济决策方法，并能在多学科环境中应用。  №11.1理解并掌握工程管理原理与经济决策方法、能够识别人工智能/大数据领域相关工程项目管理与经济决策中的关键因素。  №11.2 能够将工程管理原理和经济决策方法运用于跨学科的复杂工程项目中。 |
| 课程培养学生的能力（教学目标） | 本课程的目的是练习学生的分析问题和编程的能力。  1）掌握机器学习相关模型、概念和有关算法，使用机器学习的思想，解决实际问题;  2）团队协作进行机器学习算法设计和程序编写，并且使用多种方式有效地调试和纠正错误;  3）通过撰写科学报告和发布陈述，对所学知识有更好的理解，同时培养学生的表达和推测能力。 |
| 课程简介 | 这门实践课程旨在培养学生对机器学习在相关领域解决实际问题的能力，基本要求包括以下几点：  1）每三四个学生可以组合在一起做项目。  2）每组的学生根据题目要求完成机器学习相关的算法编写和实际问题的解决，每个小组应提交建议、项目报告、相关策略和源代码;  3）技术：   1. 完成数据预处理，计算相关特征值。; 2. 验证特征有效性； 3. 利用机器学习算法，搭建相关模型; 4. 验证模型预测效果;   4）系统应设计和实施应当合理，高效和友好。  5）源代码应当有清楚的注释：对于类注释（详细描述的属性，行为，类继承和派生关系）;对于函数的注释（详细描述目标，函数参数，原理步骤/算法）。 |
| 教学内容 | 题目：本次课程提供恒生指数的日线数据与5分钟级数据，假设初始资金1000000元，投资日期为2020年11月17日至2023年11月17日。要求学生利用机器学习方法，实现指数增强策略。  要求：   1. 完成对所提供两种数据的预处理，自行查阅相关资料，挖掘相关因子。 2. 利用日线数据与相关因子，构建机器学习模型，自行设计训练集、测试集与验证集，编写程序，所使用的模型不设限，但要求结果需要对恒生指数有良好的预测效果，预测周期可以分为日级、周级以及月级。 3. 根据所提供的5分钟级别数据，构建模型，寻找日内最佳买卖时机信号，并根据日线预测结果与日内买卖信号编写具体交易策略，实现指数增强，获取超额收益。 4. 要求最终策略实现指数增强，并将最大回撤控制在合理范围内。 5. 需要将最终的投资结果做出可视化展示，并对策略与模型做出效果评估。 6. 在本次实验中，可以将指数当做股票进行交易，但必须遵守市场真实的交易规则，有能力的同学可以对所选指数所包含的个股搭建量化交易投资组合，实现指数增强策略。 7. 撰写课程报告时要明确表明所采用的具体算法、模型原理、训练过程、特征构建方法以及最终效果的验证。而且说明所构建的模型以及最终所采取的策略应用问题时有何优缺点？ |
| 实习方式 | 集体组织 |
| 实习地点 | 校内 |
| 考核要求 | 以考查形式，满分为100分，详情如下所示。 1.源代码60％  1）符合基本要求25%  2）指数增强效果10％ 3）所用策略结果的回撤情况10%  4）策略与模型的效果评估10％  5）附加功能5％  2. 答辩表现20％ 1）表达逻辑清晰度10％ 2）回答问题表现 10%  3．最终报告20％ |
| 实习注意事项 | 上课期间严格遵守实验室规定。 |
| 教材、指导书及教学参考书目 | Machine Learning, Ed. Zhou Zhihua, Tsinghua University Press, 2017 |
| 制定人及时间 | 徐进，2023-11-17 |

***“Advanced Language Programming Training”* Syllabus**

|  |  |
| --- | --- |
| Course Code |  |
| Course Title | Machine learning training |
| Course Category | Practice Course |
| Course Nature | Compulsory Course |
| Weeks | 2 |
| Credits | 2 |
| Semester | 2 |
| Institute | The School of Future Technology |
| Program Oriented | Artificial Intelligence, Data Science and Big Data Technology |
| Teaching Language | English |
| Prerequisites | Python programming basics, machine learning |
| Student Outcomes (Special Training Ability) | №3.Design / Development Solutions: An ability to design solutions for complex engineering problems and innovatively design systems, components or process that meet specific needs with societal, public health, safety, legal, cultural and environmental considerations.  №3.2 Being able to catch up with the current status and trends in AI/Big Data-related technological development and to demonstrate innovation in the solution design;  №9.Individual and Teams: An ability to function effectively as an individual, and as a member or leader in diverse teams and in multi-disciplinary settings.  №9.1 Being able to work well with team members in AI/Big Data -related research, development and production projects;  №11.Project Management: Demonstrate knowledge and understanding of engineering management principles and methods of economic decision-making, to function in multidisciplinary environments.  №11.1 Being able to understand and master management fundamental in engineering and economic decision-making methods, and to identify key factors in the managing and economic decision-making of AI/Big Data related projects.  №11.2 Being able to apply knowledge in engineering management and economics in complex interdisciplinary engineering projects. |
| Teaching Objectives | The purpose of this course is to practice students' analytical and programming skills.  1) Master machine learning related models, concepts and algorithms, and use machine learning ideas to solve practical problems;  2) Team work to design machine learning algorithms and program writing, and use a variety of ways to effectively debug and correct errors;  3) To write scientific report, and hence have a better understanding of the knowledge and to develop students' ability to express and speculative. |
| Course Description | This practical course aims to develop students' ability of machine learning to solve practical problems in related fields. The basic requirements include the following:  1) Every three-four students can be grouped together to do the project.  2) Students in each group should write algorithms related to machine learning and solve practical problems according to the requirements of the topic. Each group should submit suggestions, project reports, relevant strategies and source code.  3) Techniques: a) Complete data preprocessing and calculate relevant feature values.  b) Validate the effectiveness of features.  c) Build relevant models using machine learning algorithms.  d) Validate the predictive performance of the models.  4) The system should be designed and implemented reasonably, efficiently and friendly.  5) Source code should be clearly commented: comments for class (a detailed description of the properties, behaviors, class inheritance, and derived relationships); comment for functions (detailed description of objectives, function parameter, principle steps /algorithms). |
| Content | Title: Index Enhancement Strategy Implementation Using Machine Learning  Requirements:  a. Complete preprocessing of the provided daily and 5-minute interval data. Assume an initial capital of 1,000,000 yuan and the investment period from November 17, 2020, to November 17, 2023. Students are required to explore relevant factors through independent research.  b. Build machine learning models using daily data and related factors. Design training, testing, and validation sets, and write programs. The choice of models is open, but they must exhibit good predictive performance for the Hang Seng Index. The prediction periods can be daily, weekly, and monthly.  c. Use the provided 5-minute interval data to construct models, identify optimal intraday buying and selling signals, and develop specific trading strategies based on daily predictions and intraday signals. Aim to enhance the index and achieve excess returns.  d. Ensure that the final strategy achieves index enhancement and controls the maximum drawdown within a reasonable range.  e. Visualize the final investment results and evaluate the effectiveness of the strategy and models.  f. In this experiment, students can treat the index as a stock for trading, following real market trading rules. Those with the capability can build a quantitative trading portfolio based on the individual stocks included in the selected index, implementing an index enhancement strategy.  g. When writing the course report, clearly state the specific algorithms, model principles, training processes, feature construction methods, and the verification of the final results. Additionally, discuss the advantages and disadvantages of the constructed models and the applied strategies.  Practice period： 2 weeks |
| Internship Mode | collective organization |
| Internship Location | Campus |
| Assessment | It is evaluated with the defense. Full score is 100 points. The details are as the following.  1. Source code 60%  a. Meet the basic requirements by 25%  b. The index enhancement effect is 10%  c.10% of the retraction of the strategy results used  d. Evaluation of the effect of strategy and model 10%  e. Additional functions 5%  2. The defense performance is 20%  1) Expression logic clarity 10%  2) Answer questions with 10% performance  3. Final report 20% |
| Matters Needing Attention | Strictly comply with the requirements of the laboratory during class time |
| Teaching Materials and Reference Books | Machine Learning, Ed. Zhou Zhihua, Tsinghua University Press, 2017 |
| Prepared by Whom and When | Jin Xu, 2023 November. |