BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF INFORMATION TECHNOLOGY
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF INTELLIGENT SYSTEMS
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

REINFORCEMENT LEARNING FOR AUTOMATED STOCK PORTFOLIO ALLOCATION

VYUŽITÍ ZPĚTNOVAZEBNÉHO UČENÍ PRO AUTOMATICKOU ALOKACI AKCIOVÉHO PORTFOLIA

BACHELOR'S THESIS
BAKALÁŘSKÁ PRÁCE

AUTHOR **ZDENĚK LAPEŠ**
AUTOR PRÁCE

SUPERVISOR **doc. RNDr. MILAN ČEŠKA, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2023

Bachelor's Thesis Assignment

148202

Institut: Department of Intelligent Systems (UITS)
Student: **Lapeš Zdeněk**
Programme: Information Technology
Specialization: Information Technology
Title: **Reinforcement Learning for Automated Stock Portfolio Allocation**
Category: Artificial Intelligence
Academic year: 2022/23

Assignment:

1. Study the state-of-the-art methods for automated stock portfolio allocation. Focus on the methods based on reinforcement learning and planning in Markov Decision Processes.
2. Experimentally evaluate selected open access tools for automated portfolio allocation including e.g. FinRL-Meta and identify their weak points.
3. Propose and implement improvements of a selected method/tool allowing to mitigate theseweak points.
4. Using suitable benchmarks and datasets, perform a detailed experimental evaluation of the implemented improvements with the focus on the portfolio allocation returns.

Literature:
Rao A., Jelvis T., Foundations of Reinforcement Learning with Applications in Finance. 1st Edition, Taylor & Francis 2022
* Li, Xinyi and Li, Yinchuan and Zhan, Yuancheng and Liu, Xiao-Yang, Optimistic Bull or Pessimistic Bear: Adaptive Deep Reinforcement Learning for Stock Portfolio Allocation, In ICML 2019.
* Liu X.-Y. Rui J. Gao J. aj.: FinRL-Meta: A Universe of Near-Real Market Environments for Data-Driven Deep Reinforcement Learning in Quantitative Finance. Workshop on Data Centric AI 35th Conference on Neural Information Processing Systems at NeurIPS 2021.
* Mao Guan and Xiao-Yang Liu. 2021. Explainable Deep Reinforcement Learning for Portfolio Management: An Empirical Approach. In ICAIF 2021.

Requirements for the semestral defence:
Items 1, 2, and partially 3.

Detailed formal requirements can be found at https://www.fit.vut.cz/study/theses/

Supervisor: **Češka Milan, doc. RNDr., Ph.D.**
Head of Department: Hanáček Petr, doc. Dr. Ing.
Beginning of work: 1.11.2022
Submission deadline: 10.5.2023
Approval date: 3.11.2022

Abstract

Portfolio allocation is about selecting a set of assets to invest in. The goal is to maximize the expected return for a given level of risk for an investing horizon selected by us. In former times, this the process is usually done manually by an expert investor. Nowadays, there are many portfolio allocation methods that are not successful in the real world or can be improved and the potential of current technologies is not fully explored in the financial market. To solve the problem of portfolio allocation, I propose a methods based on reinforcement learning. The agent will learn the exact strategies of selecting assets and their weight for the portfolio based on which the human expert would select it, like fundamental analysis and the health of the company. The thesis deals with the problem of portfolio allocation using reinforcement learning, which helps in selecting the best assets and their importance for the portfolio.

Abstrakt

Alokace portfolia se zaměřuje na výběr souboru aktiv, do kterých investujete. Cílem je maximalizovat očekávaný výnos pro danou míru rizika pro námi zvolený investiční horizont. V dřívějších dobách tohle proces je obvykle prováděn ručně zkušeným investorem. V dnešní době existuje mnoho metod alokace portfolia, které nejsou úspěšné v reálném světě nebo je lze zlepšit a potenciál současných technologií není na finančním trhu plně prozkoumán. Pro řešení problému alokace portfolia navrhuji metody založené na posilovacím učení. Agent se naučí přesné strategie výběru aktiv a jejich váhu pro portfolio, na základě kterého by je lidský expert vybíral, jako je fundamentální analýza a zdraví společnosti. Diplomová práce se zabývá problémem alokace portfolia pomocí posilovacího učení, které pomáhá při výběru nejlepších aktiv a jejich důležitosti pro portfolio.

Keywords

artificial intelligence, AI, reinforcement learning, stock portfolio allocation, modern portfolio theory, Q-learning, neural networks, stock market

Klíčová slova

umělá inteligence, AI, posilované učení, alokace akciového portfolia, moderni teorie portfolia, Q-learning, neuronové sítě, akciový trh

Reference

LAPEŠ, Zdeněk. *Reinforcement Learning for Automated Stock Portfolio Allocation.* Brno, 2023. Bachelor's thesis. Brno University of Technology, Faculty of Information Technology. Supervisor doc. RNDr. Milan Češka, Ph.D.

Reinforcement Learning for Automated Stock Portfolio Allocation

Declaration

I hereby declare that this Bachelor's thesis was prepared as an original work by the author under the supervision of Mr. Milan Češka, Ph.D. I have listed all the literary sources, publications and other sources, which were used during the preparation of this thesis.

....
Zdeněk Lapeš
January 25, 2023

Acknowledgements

I would like to thank my supervisor, Mr. Milan Češka, for his guidance and support during the preparation of this thesis. I would also like to thank my family and friends for their support.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Portfolio allocation problem is to spread appropriate finite cash budget into financial instruments e.g. stocks [7]. Portfolio allocation has been studied for a long time. Under the financial instruments we can imagine:

1. **Stocks:** Stocks represent ownership in a company and can provide a source of income through dividends and capital appreciation.
2. **Bonds:** Bonds are debt securities that pay a fixed or variable rate of interest to bondholders. They can provide a more stable source of income compared to stocks.
3. **Mutual Funds:** Mutual funds are investment vehicles that pool money from multiple investors to buy a diversified portfolio of stocks, bonds, or other securities.
4. **Options:** Options are financial derivatives that give the holder the right, but not the obligation, to buy or sell an underlying asset at a certain price and time.
5. **Futures:** Futures are financial derivatives that obligate the holder to buy or sell an underlying asset at a certain price and time in the future.
6. **Currency:** Currency trading, also known as forex, involves buying and selling different currencies in the foreign exchange market.
7. **Commodities:** Commodities are raw materials and primary products that can be bought and sold, such as gold, oil, and agricultural products.
8. **Derivatives:** Derivatives, such as options and futures, are financial products that derive their value from an underlying asset.
9. **Real Estate Investment Trusts (REITs):** REITs are investment vehicles that own and operate income-producing real estate properties.
10. **Exchange-Traded Funds (ETFs):** ETFs are investment funds that are traded on stock exchanges, similar to stocks. They provide exposure to a diversified portfolio of assets, such as stocks, bonds, or commodities.

11. **Certificates of Deposit (CDs):** CDs are deposit accounts with banks or other financial institutions that pay a fixed rate of interest. They are considered a low-risk investment option.

12. **Savings Accounts:** Savings accounts are deposit accounts with banks or other financial institutions that pay a low rate of interest. They are considered a low-risk investment option.

The goal is to allocate assets instead of cash and thanks to constantly valorize the initial capital on a selected investing horizon. It is important to think about the minimization of risk while maximizing the expected return on the investment. Portfolio management is the most important task for financial institutions, like banks, hedge funds, and asset management companies. There exist various existing portfolio management strategies, such as the Markowitz mean-variance portfolio optimization proposed by Hanry Markowitz [10], the Black-Litterman model [3] and the CAPM model [1].

The earlier approaches based on predictive models (unsupervised learning methods) are not too appropriate for portfolio management, because the stock market is stochastic, volatile, quickly changing, and uncertain environment. Predictive models are not flexible enough to adapt to the changing environment like the stock market and are not able to predict future market conditions accurately. The current strategies and models are limited by the fact that they are based on historical data which includes a huge noise.

So, the most recent state-of-the-art portfolio management strategies are based on machine learning techniques, such as deep reinforcement learning (DRL), which shows promising results in the financial domain, such as stock trading and portfolio management tasks. The benefits of RL have been used in many different fields, such as robotics, and games, and a lot of people are interested in applying RL to the financial domain.

Reinforcement Learning is a subfield of machine learning, concerned with how an agent can learn to act in an environment to maximize the cumulative reward [7]. In the last decade, RL has become popular, because of its ability to learn difficult tasks in a variety of domains without knowing the environment model [5]. The agent is trained in a particular environment, which is modeled using MDPs.

MDPs (Markov Decision Processes) is a mathematical framework for modeling decision-making (actions) in specific situations (called states) [4, p. 93]. MDPs are the next level of MRP (Markov Reward Processes) [4, p. 84] and it is the next level of MP (Markov Processes) [4, p. 59]. In the real world, the transition matrix of the MDP model for a particular environment is unknown, so the agent must learn it from experience, and here comes into play the RL algorithms [5]. RL algorithms are divided into two categories: model-based[9] and model-free[8]. These algorithms have been studied on how much a change in input would influence the output and what contributes to the decision-making process of the RL agents [2]. This thesis experiments with multiple RL algorithms, such as DDPG, TD3, SAC, PPO, A2C, ..., and compare their performance in the same environment.

Deep Reinforcement Learning (DRL) was a breakthrough in training RL agents, which is a subfield of RL using deep neural networks to approximate the large and complex state-action spaces and help to understand the stochastic environments[6].

1.2 Problem Statement

Reinforcement learning has advantages, such as flexibility, adaptability, and utilization of various information like e.q. experience gained from the environment under certain conditions. The RL agents have huge potential to perform pretty well, via follow these factors:

- **State space**
  The actions computed by the model should change only the portfolio managed by the agent, but not the conditions of the financial market itself.
- **Action space**
  Action space should be designed so that the agent weights the assets in the portfolio. How much of the budget should be allocated to each asset and should be this asset in the portfolio? These design decisions are crucial for the performance of the agent. It is really difficult to find the optimal policy for the portfolio allocation because the agent has to choose between multiple assets with various differences in information about the assets. Also, actions should be considered profitable and safe in the long term, which means that the agent has to make decisions based on long-term rewards.
- **Reward function**
  The reward should reflect the agent's performance in the environment. Is the current portfolio value increasing or decreasing after the agent takes actions proposed by the policy?
- **Model Overfitting**
  The model should be able to generalize the environment and not overfit the training data. The dataset should be large enough to generalize the environment and split into the training and testing dataset.

1.3 Limitations

The current solutions, such as from FinRL [2], consider the technical indicators and covariances of the assets as the state space. This puts the agent at a disadvantage because the agent has to learn the relationship between the technical indicators and the covariances of the assets and the portfolio value. However, the covariance matrix is great for portfolio optimization, but I suppose it is not the best choice for the state space because it puts a huge amount of noise and complexity into it.

I suppose it should be possible to solve the portfolio allocation task in a simpler way, using fundamental data about the company, which should result in better agent performance.

The environment should be designed in a way that the agent can learn the relationship between the state space and the portfolio value according to the lowest level of complexity. Otherwise, the agent will not be able to learn the best policy for the portfolio allocation task in an efficient way.

1.4 Aim of the Thesis

This thesis focus on the experimental comparison of existing portfolio allocation agents, including testing and benchmarking of the portfolio allocation agents on real-world data. The agents will be experimentally changed for different input data, which should be able to improve the agent's performance, considering the factors mentioned in the previous section.

The objectives of this thesis are:

- Experimental evaluation and benchmark of the existing portfolio allocation agents.
- Create a suitable dataset for the portfolio allocation problem with a finite set of stocks. Dataset consists of financial data, such as fundamental and technical analysis data.
- Reimplement and Train RL agents with a finite cash budget on the datasets.
- Evaluate the performance of the RL agents by comparing them with the baseline portfolio management strategies (mean-variance optimization and baseline indexes such as DJI30 etc...) and compare the performance of the RL agents with the existing RL agents including the FinRL-Meta Portfolio Allocation agent.

Using the programming language Python and open-source libraries such as NumPy, Pandas, Stable Baselines3, OpenAI gym, etc... I will implement the RL agents and evaluate their performance on the created dataset.

This thesis aims to discover if RL can learn optimal portfolio management that yields the same or even better results than the current state-of-the-art portfolio management strategies focusing on the objectives mentioned above, such as Data Engineering, Data analysis, and Data visualization.

Bibliography

[1] Fama, E. F. and French, K. R. The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives* [https://mba.tuck.dartmouth.edu/bespeneckbo/default/AFA611-Eckbo%20web%20site/AFA611-S6B-FamaFrench-CAPM-JEP04.pdf]. American Economic Association. august 2004, vol. 18, no. 3, p. 25–46. DOI: 10.1257/0895330042162430.

[2] Guan, M. and Liu, X.-Y. *Explainable Deep Reinforcement Learning for Portfolio Management: An Empirical Approach.* arXiv, 2021. DOI: 10.48550/ARXIV.2111.03995. Available at: https://arxiv.org/abs/2111.03995.

[3] He, G. and Litterman, R. The Intuition Behind Black-Litterman Model Portfolios. *SSRN Electronic Journal* [https://papers.ssrn.com/sol3/papers.cfm?abstract_id=334304]. Elsevier BV. 2002. DOI: 10.2139/ssrn.334304.

[4] i=A., f. given and i=T., f. given. *Foundations of Reinforcement Learning with Applications in Finance* [https://stanford.edu/~ashlearn/RLForFinanceBook/book.pdf]. Upper Saddle River, NJ, United States: Prentice Hall, 2022 [cit. 2022-11-20]. ISBN 1032124121.

[5] i=R.S., f. given and i=A.G., f. given. *Reinforcement Learning, second edition: An Introduction (Adaptive Computation and Machine Learning series)* [http://www.incompleteideas.net/book/the-book-2nd.html]. Second editionth ed. Bradford Books, november 2018 [cit. 2022-11-20]. ISBN 0262039249.

[6] Li, X., Li, Y., Zhan, Y. and Liu, X.-Y. *Optimistic Bull or Pessimistic Bear: Adaptive Deep Reinforcement Learning for Stock Portfolio Allocation.* arXiv, 2019. DOI: 10.48550/ARXIV.1907.01503. Available at: https://arxiv.org/abs/1907.01503.

[7] Oshingbesan, A., Ajiboye, E., Kamashazi, P. and Mbaka, T. *Model-Free Reinforcement Learning for Asset Allocation.* arXiv, 2022. DOI: 10.48550/ARXIV.2209.10458. Available at: https://arxiv.org/abs/2209.10458.

[8] Pong, V., Gu, S., Dalal, M. and Levine, S. *Temporal Difference Models: Model-Free Deep RL for Model-Based Control.* arXiv, 2018. DOI: 10.48550/ARXIV.1802.09081. Available at: https://arxiv.org/abs/1802.09081.

[9] Yu, P., Lee, J. S., Kulyatin, I., Shi, Z. and Dasgupta, S. *Model-based Deep Reinforcement Learning for Dynamic Portfolio Optimization.* arXiv, 2019. DOI: 10.48550/ARXIV.1901.08740. Available at: https://arxiv.org/abs/1901.08740.

[10] Širůček, M. and Křen, L. Application of Markowitz Portfolio Theory by Building Optimal Portfolio on the US Stock Market. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. Mendel University Press. september 2015, vol. 63, no. 4, p. 1375–1386. DOI: 10.11118/actaun201563041375. Available at: http://dx.doi.org/10.11118/actaun201563041375.