**OPTION I**

- idea and improvements are listed after the articles -

**1) An intelligent financial portfolio trading strategy using deep q-learning**

[**https://arxiv.org/abs/1907.03665**](https://arxiv.org/abs/1907.03665)

**Park, H., Sim, M. K., and Choi, D. G (2020)**

**-** published**: Expert Systems with Applications 2020**

- uses RL: DDPG

- deep Q-learning

- datasets 2010-2016: 3 US-ETFs: S&P500 Index, Russell 1000, iShares Microcap, 3 Korean Portfolio (KOR-IDX)

- results: 2017: Cumulative Return: 12.634% (not much) on the US portfolio – not unexpected considering the slow growth of ETFs

- code: some implementation: <https://github.com/Jogima-cyber/portfolio-manager>

**2) Deep reinforcement learning for portfolio management of markets with a dynamic number of assets**

[**https://arxiv.org/pdf/2012.13773.pdf**](https://arxiv.org/pdf/2012.13773.pdf)

**Betancourt, C.; Chen, W.H.**

**-** published **in Expert Syst. Appl. 2021**

- includes shorting (aka negative weights)

- uses RL: DDPG (deep Q-learning)

- added negative weights (allowed shorting)

- added arbitrage mechanism

- explainable model

- datasets: 4 portfolios: several stocks from the constituents of CSI300 Index (ETF)

- results: back-testing 1 year return 2020-2021: 40%, 18%, 35%, 11%

- no code provided

# **3) A Selective Portfolio Management Algorithm with Off-Policy**

**Reinforcement Learning Using Dirichlet Distribution**

# <https://www.mdpi.com/2075-1680/11/12/664/pdf>

# - published: **Axioms 2022**

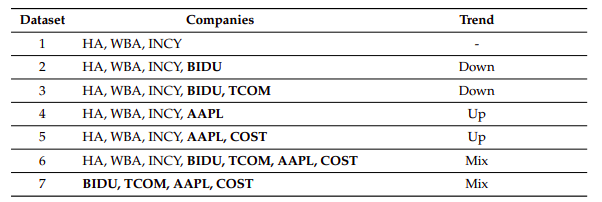
- references all articles above

- page 3: discusses the application of DRL in learning trading strategies

- scalable

- RL model: DDT-A

- datasets (stocks with different growth trends):



- results (2020-2022) – annual return for mid-risk level (investment portfolio preference by the user):

|  |  |
| --- | --- |
| Dataset | Cumulative Return with medium risk (2 years) |
| 1 | 24% |
| 2 | 12% |
| 3 | 7% |
| 4 | 70% |
| 5 | 30% |
| 6 | 11% |
| 7 | 70% |

- no code provided

**OPTION I idea:**

Use a similar set-up with 2) and apply:

- apply conservative q-learning

- add negative weights (shorting allowed)

- a more representative dataset: 7 datasets mixed stocks with up and down trends

- concatenate language features encoded from news/company prompts (like in Option II)

**OPTION II**

- idea and improvements are listed after the articles -

# **1) Reinforcement-Learning based Portfolio Management with Augmented Asset Movement Prediction States**

<https://paperswithcode.com/paper/reinforcement-learning-based-portfolio>

- published in [Journal of Financial Data Science 2020](https://paperswithcode.com/conference/journal-of-financial-data-science-2020-3)

- Dataset:

- Bitcoin

- high tech stocks: 9 companies which have the most news, 2006-2013

- uses stock prices as the internal features for the original state s\* + price predictions and news sources as external information to be augmented to the final state s

- the classifier used has 65.08% training and 60.10% testing accuracy

- HAN as an encoder to obtain a 100-dimensional embedding vector of stock movement prediction for each news. The training/testing accuracy is 64.10%/60.99%

- RL part: Deterministic Policy Gradient (DPG) -  off-policy learner

- achieves better results than the normal state-of-the-art RL model, which only uses stock prices

- the exploitation of diverse information can reduce the impact of environment uncertainty

- results (255 days testing) : 1.75 portfolio value increase for HighTech portfolio

- SARL outperforms the state-of-the-art method (DPM) by 140.9% on the Bitcoin dataset and 15.7% on the HighTech dataset

Why to use RL

“

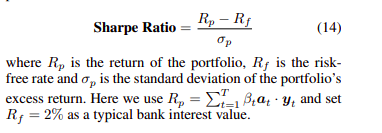
With history prices of all assets as its input, a neural network can output a predicted vector of asset prices for the next period. Then the trading agent can act upon this prediction. This idea is straightforward to implement, because it is a supervised learning, or more specifically a regression problem. The performance of these price-prediction-based algorithms, however, highly depends on the degree of prediction accuracy, but it turns out that future market prices are difficult to predict. Furthermore, price predictions are not market actions, converting them into actions requires additional layer of logic. If this layer is a hand-coded, then the whole approach is not fully machine learning, and thus is not very extensible or adaptable. For example, it is difficult for a prediction-based network to consider transaction cost as a risk factor.

“

(Jiang, Xu, and Liang 2017)

- the article provides a good explanation on page 2 why RL would be a good idea to use for Portfolio Management

**Sharpe Ratio explained**: • Sharpe Ratio The Sharpe Ratio (Sharpe and Pnces 1964) is the average return earned in excess of the risk-free rate per unit of volatility or total risk. It is used to compare the portfolio’s return to its risk and is defined as



- the RL implementation compares with DPM is the state-of-art standard RL method in Portofolio Management: **references** A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem (Jiang, Xu, and Liang 2017): <https://arxiv.org/abs/1706.10059> (FinRL implementation)

# **2) A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem**

(“Liang et al”) [Zhengyao Jiang](https://arxiv.org/search/q-fin?searchtype=author&query=Jiang%2C+Z), [Dixing Xu](https://arxiv.org/search/q-fin?searchtype=author&query=Xu%2C+D), [Jinjun Liang](https://arxiv.org/search/q-fin?searchtype=author&query=Liang%2C+J)

<https://arxiv.org/abs/1706.10059>

- not published

- uses DDPG

- code: <https://github.com/kftam1994/Robo_Advisor>

**OPTION II idea:**

use the experimental set-up from the first article 1) and consider improvements for the price prediction and language encoders (trying to get higher accuracy than 60%):

Improvements:

- dataset: – portfolios with mixed trends; average on certain years

- language encoding

- use company prompts

- use a better NLP model, e.g.:

Trade the Event: Corporate Events Detection for News-Based Event-Driven Trading

<https://paperswithcode.com/paper/trade-the-event-corporate-events-detection>

- published in [Findings (ACL) 2021](https://paperswithcode.com/conference/findings-acl-2021-8)

- disadvantages: need to train using DTU servers

- price prediction

Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction

<https://arxiv.org/pdf/2204.02623v1.pdf>

Advantages:

- understandable paper (I did something similar with wind prediction)

- understandable code

**Other RL strategies for PM/ Asset Allocation**

**-** experimental set-ups might differ -

# **Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy**

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3690996>

- already revised, uses FinRL

- ACM International Conference on AI in Finance.

**Adversarial Deep Reinforcement Learning in Portfolio Management**

Liang et al <https://arxiv.org/abs/1808.09940>

- not published

- uses Deep Deterministic Policy Gradient (DDPG),

- code: <https://github.com/kftam1994/Robo_Advisor>

# **Asset Allocation: From Markowitz to Deep Reinforcement Learning**

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4148379&fbclid=IwAR3ysuaN5T7GUWNWeEFHCGdK2RwBX2wn5w5Kh935dJHzp5VBEC7fr2o6dJs>

- not published yet

- adds improvements to the DDPG

- understandable code: <https://github.com/ricarddurall/benchmarking-strategies-for-asset-allocation>