Risk-adjusted Cryptocurrency Portfolio Management with Deep Reinforcement Learning
Master's Project Proposal

Barış Özakar, Emine Didem Durukan, Doğan Parlak
{baris.oezakar, eminedidem.durukan, dogan.parlak}@uzh.ch

February 10, 2022

1 Problem Definition & Related Work

Portfolio management is the decision making process of continuously reallocating an amount of fund into a number of different financial investment products, aiming to maximize the return while restraining the risk [1]. There are several approaches to the problem of portfolio management which are built with different combination of methodologies and algorithms. However, there is no framework which systematically compares those approaches. There is a number of publications reviewing the field. All of those do it on the level of ideas, however. Side-by-side comparisons are lacking.

Some of the existing deep machine-learning approaches attempt to predict price movements or trends [2]. In these approaches, the model outputs a predicted vector of asset prices for the next period. Then the trading agent can act upon this prediction. 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 [3].

Other approaches to the portfolio management problem directly optimize the policy, without explicitly predicting future prices. These model-free and fully machine-learning schemes to the algorithmic trading problem include Moody and Saffell (2001), Dempster and Leemans (2006), Cumming (2015) and the recent deep RL utilization by Deng et al. (2017) [4], [5]. These RL algorithms output discrete trading signals on a single asset.

Deep RL is lately drawing much attention due to its remarkable achievements in playing video games (Mnih et al., 2015) and board games (Silver et al., 2016) [6].

A general-purpose continuous action and state space deep RL framework, the actor-critic Deterministic Policy Gradient Algorithms, was recently introduced (Silver et al., 2014; Lillicrap et al., 2016) [6], [7]. The continuous output in these actor-critic algorithms is achieved by a neural-network approximating action policy function, and a second network is trained as the reward function estimator. Training two neural networks, however, is found to be difficult, and sometimes even unstable. The recent paper by Jiang et al. (2017) proposes an RL framework specially designed for the task of portfolio management by using the Ensemble of Identical Independent Evaluators (EIIE) topology [3].

The main difficulty arises because each publication focuses on its own data source, data preparation methodology, considers different time periods and have a different approach to back-testing. If one is to compare the methods, a fair comparison would involve comparing them on the actual performance. This is what the project is aiming to achieve.

2 Deliverables

The aim of the project is to design and implement a unified framework in order to compare methodologies for portfolio management. Furthermore, this comparison will be carried out on a larger sample of publications and results will be reported in an accompanying text.

3 Overview of the field

The optimization criteria considered in the literature contain average logarithmic return, Sharpe ratio, Sortino ratio and Mean-Variance portfolio optimization criterion.

The data sources that will be considered are high frequency cryptocurrency market data, technical indicators, sentiment data.

The function approximator families used are often CNN, RNN, and LSTM. A recent study that focuses on portfolio management, has used CNN that has two hidden layers: convolution layer and a fully-connected layer to solve the task [8]. In order to mitigate the problems that can arise when making the network deeper is CNN with Deep Residual Network [9]. The authors proposed a network topology that consists of CNN and Deep Residual Network.

Network topologies that have been mentioned previously depend on how we formulate the problem, the methodology that we want to optimize, and of course, the data.

Backtesting assesses the viability of a trading strategy by discovering how it would play out using historical data. It takes into account the transaction costs and other market frictions. In the literature, one usually assumes

- *Zero slippage*: The liquidity of all market assets is high enough that, each trade can be carried out immediately at the last price when a order is placed. A possible relaxation may include introducing a lag to the completion of order.

- *Zero market impact*: The capital invested by the software trading agent is so insignificant that is has no influence on the market.

Besides comparing the selected methodologies with their corresponding algorithms, we will also compare them with passive sector standard strategies including

- **the Uniform Buy and Hold(UBAH)** is a portfolio management approach simply equally spreading the total fund into the pre-selected assets and holding them without making any purchases or selling until the end.
- **Risk-parity Weighted Buy and Hold** is a portfolio management approach where the total fund assets are spread inversely with respect to the volatility of the asset.
- **Tangency Portfolio of the Efficient Frontier (Modern Portfolio Theory)** is a portfolio management approach by Markovitz (1952) [10]. The efficient frontier is a set of portfolios that maximize the expected return for a given level of risk. The tangency portfolio is the portfolio on the efficient frontier with the highest Sharpe ratio, tangent to the best capital allocation line (CAL).

4 Task Distribution

- **Baris Ozakar** Data collection, data pre-process, algorithm and criterion determination and implementation, project report and presentation
- **Didem Durukan** Data visualization and analysis, algorithm and criterion determination and implementation, result visualization, project presentation
- **Dogan Parlak** Data pre-process, algorithm and criterion determination and implementation, result analysis, project report

References

[1] R. A. Haugen, "Modern investment theory," *Prentice Hall*, 1986.

[2] N. G. P. J. B. Heaton and J. H. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry*, 2016. [Online]. Available: http://www.ssrn.com/abstract=2838013

[3] Z. Jiang, D. Xu, and J. Liang, "A deep reinforcement learning framework for the financial portfolio management problem," 2017.

[4] J. Moody and M. Saffell, "Learning to trade via direct reinforcement," *IEEE Transactions on Neural Networks*, vol. 12.

[5] M. Dempster and V. Leemans, “An automated fx trading system using adaptive reinforcement learning,” *Expert Systems with Applications*, vol. 30.

[6] D. S. V. Mnih, K. Kavukcuoglu, “Human- level control through deep reinforcement learning,” *Nature*, vol. 518.

[7] A. P. T. P. Lillicrap, J. J. Hunt, “Continuous control with deep reinforcement learning,,” *arXiv*, 2016.

[8] Z. Jiang and J. Liang, “Cryptocurrency portfolio management with deep reinforcement learning,” *CoRR*, vol. abs/1612.01277, 2016. [Online]. Available: http://arxiv.org/abs/1612.01277

[9] Z. Liang, H. Chen, J. Zhu, K. Jiang, and Y. Li, “Adversarial deep reinforcement learning in portfolio management,” *arXiv preprint arXiv:1808.09940*, 2018.

[10] H. M. Markowitz, “Portfolio selection: efficient diversification of investments,” *Yale university press*, vol. 16, 1968.