

**University of Warsaw**  
Faculty of Economic Sciences

**Krzysztof Wojdalski**

Nr albumu: 310284

# **Reinforcement Learning Portfolio Optimization for FX Trading**

**Praca magisterska  
na kierunku ECONOMICS**

The thesis written under the supervision of  
**dr Pawel Sakowski**

2025-05-03

Data

Podpis opiekuna

### **Statement of the Supervisor on Submission of the Thesis**

I hereby certify that the thesis submitted has been prepared under my supervision and I declare that it satisfies the requirements of submission in the proceedings for the award of a degree.

Date

Signature of the Supervisor:

### **Oświadczenie autora (autorów) pracy**

Data

Podpis autora (autorów) pracy

### **Statement of the Author(s) on Submission of the Thesis**

Aware of legal liability I certify that the thesis submitted has been prepared by myself and does not include information gathered contrary to the law.

I also declare that the thesis submitted has not been the subject of proceedings resulting in the award of a university degree.

Furthermore I certify that the submitted version of the thesis is identical with its attached electronic version.

Date

Signature of the Author(s) of the thesis



# Table of contents

- 1.1. asdad . . . . .
- 1.2. Introduction . . . . .
  - 1.2.1. Work Structure . . . . .
- 1.3. sdsa . . . . .
- 1.4. Statistical Arbitrage . . . . .
  - 1.4.1. asdsa . . . . .
- Bibliography . . . . .

## 1.1. asdad

## 1.2. Introduction

The conventional representation of trading floors as chaotic environments characterized by vocal trader interactions was historically accurate approximately three decades ago. However, the financial industry has since undergone substantial structural transformation.

The financial sector has consistently functioned as an early adopter of computational innovations. This technological integration represents a strategic imperative—competitive advantage through advanced technological implementation frequently correlates with enhanced financial performance metrics. The industry has systematically pioneered the deployment of state-of-the-art technologies, ranging from sophisticated Bloomberg terminal infrastructure in the 1990s to contemporary blockchain applications and ultra-low latency systems, all oriented toward maximizing operational efficiency<sup>1</sup>.

Trading entities, functioning as utility-maximizing economic agents, fundamentally aim to optimize market-derived profits. Multiple methodological approaches exist to achieve this objective, with varying degrees of complexity. Notably, Warren Buffett, a statistically significant outlier in investment performance metrics, continues to employ a passive buy-and-hold strategy.

In recent years, methodologies predicated on artificial intelligence and machine learning algorithms have demonstrated increasing significance. This phenomenon correlates with advancements in computational processing capacity, decreasing infrastructure expenditure requirements, and empirical recognition of cognitive biases in human decision-making processes (Arnold2017?). A growing consensus within the literature suggests that algorithmic systems should supplement or potentially supersede human involvement in decision-making and execution processes Turner (2015). The progressive automation of trading activities will likely

<sup>1</sup>The banking sector is frequently identified in empirical literature as a primary industry implementing blockchain technology.

continue its trajectory in subsequent temporal periods, contradicting the aforementioned stereotypical trading floor representation.

Although machine learning’s theoretical foundations date to the 1950s, its explicit implementation in trading contexts remains relatively limited. For example, in institutional foreign exchange trading, only entities with substantial capital resources and sophisticated quantitative infrastructure have developed effective machine learning trading systems Mosaic (2017).

This research investigates potential applications of machine learning—specifically reinforcement learning methodologies—in developing trading systems capable of generating statistically significant positive outcomes.

Currently, most systems documented in academic literature aim to maximize either absolute trading profits or risk-adjusted performance metrics. Despite numerous methodological approaches to create consistently profitable systems utilizing variables derived from financial econometrics, fundamental analysis, or machine learning algorithms, many have demonstrated suboptimal performance due to several factors, including:

- Frequent large drawdowns resulting in excessive performance volatility
- Prohibitive transaction costs rendering strategies economically unfeasible
- Excessive computational complexity, particularly problematic for high-frequency trading implementations

Even when empirical research demonstrates exceptional results, upon publication, any competitive advantage tends to diminish through market efficiency mechanisms. Strategies with persistent alpha-generating capabilities must remain proprietary to maintain their effectiveness.

### **1.2.1. Work Structure**

This thesis is structured to provide a comprehensive analytical framework for understanding reinforcement learning algorithms applied to trading contexts.

The first section provides a detailed examination of the intersection between artificial intelligence methodologies and financial markets, exploring the historical relationship between quantitative finance and computational science.

The second chapter presents a systematic review of selected literature from quantitative finance, examining both classical equilibrium models such as CAPM (the established paradigm in equity research) and contemporary approaches. This section evaluates the implicit advantages and limitations of various financial models, with particular emphasis on algorithmic trading methodologies employed in comparable research contexts.

The third section analyzes machine learning frameworks, providing a theoretical basis for why reinforcement learning may represent an optimal approach for certain trading applications. It presents a taxonomic comparison of major machine learning categories to elucidate their methodological distinctions, introduces key reinforcement learning concepts with illustrative examples, and addresses potential limitations and implementation challenges associated with these algorithms.

The fourth part details the experimental methodology, including research objectives, data characteristics, experimental design parameters, and empirical results. The primary objective was to develop trading agents capable of statistically outperforming established benchmarks

on risk-adjusted performance measures in the foreign exchange market—agents characterized by statistical robustness, adaptive learning capability, and consistent performance metrics. This chapter presents the mathematical formulations and procedural implementations leading to the empirical results, examining each component of the trading system.

The implemented algorithms utilize a dynamic optimization approach. Beyond a value function based on Differential Sharpe Ratio, the system incorporates various technical indicators such as Relative Strength Index to inform algorithmic decision-making processes. The methodology incorporates transaction cost models to simulate realistic trading conditions.

The value function integrates multiple statistical measures, including Sharpe and Differential Sharpe Ratios, to capture both risk and return dimensions. The algorithm outputs agent actions in the discrete action space  $\{-1,0,1\}$ . The final section of this chapter evaluates the reinforcement learning-based trading system against two benchmark methodologies:

- A buy-and-hold strategy (maintaining consistent long positions in selected currency pairs)
- Random action generation—producing stochastic values in the domain of  $\{-1,0,1\}$  to determine positions in underlying pairs. This benchmark excludes transaction costs, as such a strategy would incur prohibitive cumulative costs with position changes occurring in approximately two-thirds of states.

The concluding section presents a comparative analysis with similar research and proposes directions for future investigation, addressing research questions such as:

- What additional implementations could enhance performance metrics?
- What methodological limitations were encountered and how might they be addressed in subsequent research?

### 1.3. sdsa

## 1.4. Statistical Arbitrage

In financial markets the term arbitrage means the practice of exploiting riskless opportunities to make profit. In the purest form, if an asset is priced differently at two different places and a trader has access to both then he would have an arbitrage opportunity. In practice financial markets are designed in such a way that any deterministic arbitrage opportunities vanish, i.e., are being exploited by market participants. Still there are patterns in the market which seem to be persistent. Strategies which use statistical techniques to extract highly reliable patterns and then carry out trades to exploit them are called statistical arbitrage strategies.

Statistical arbitrage is a general term in which assets are put into baskets by market-based similarities. Once the relation between the co-moving financial instruments is found, any deviation from it can be exploited. If that relation is true then either one or both stocks are mispriced and a trader can take a long position on the undervalued one(s) and short position on the overvalued asset(s) to bet on their return to the prevalent relative price behaviour. Statistical arbitrage also fittingly resembles the fact that the strategy hedges risk from market movements. As the pair trade goes short and long on similar assets, the overall risk balance should cancel out and the strategy should be profitable without market risk exposure.

To determine if a security is over- or undervalued one can either use some theoretic approach to determine what the absolute price of the security should be or one could use the idea of relative pricing. In this approach it is only of interest how one security is priced in relation to another security. Now to form a simple mean-reverting pair trading strategy the only thing that remains is a definition of what constitutes a reliable price relationship between the asset prices. There are many ways to identify such security pairs. One approach would be to measure fundamentals (such as some accountancy variables) of the firms. The statistical approach, which also makes the whole concept easily applicable in a system is to use cointegration.

#### **1.4.1. asdsa**



# List of Figures

## Bibliography

- Mosic, Ranko. 2017. "Deep Reinforcement Learning Based Trading Application at JP Morgan Chase." *Medium*. <https://medium.com/@ranko.mosic/reinforcement-learning-based-trading-application-at-jp-morgan-chase-f829b8ec54f2>.
- Turner, Matt. 2015. "The robot revolution is coming for Wall Street traders." <http://www.businessinsider.com/robots-to-replace-wall-street-traders-2015-8?IR=T>.