

NATIONAL UNIVERSITY OF SINGAPORE

Master's of Computing (General-Track)



**Alpha Tree Search and Machine Learning
Approaches to Optimising Real Estate
Portfolios**

Leong Wei Ming

Supervisor: Professor Liu Li Li

Examiner: Professor Chin Wei Ngan

Department of Computer Science

Internal Capstone Project for AY2023/2024

ABSTRACT

An internal project about applying genetic algorithm to search for optimal alphas. State the major contribution:

DECLARATION

I hereby declare that this project report is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in this report.

This report has also not been submitted for any degree in any university previously.

ACKNOWLEDGEMENT

I would like to thank Professor Liu Lili for her guidance, support and encouragement throughout the course of this project. Working with Professor Lili has been a great learning experience, getting to learn much more about machine learning and its applications to finance in solving some challenges faced by industry practitioners.

Also many thanks to Professor Chin Wei Ngan for taking the time out to assess this project.

CONTENTS

1	Introduction	2
1.1	Problem Definition	2
1.2	Motivations	3
1.3	Major Contributions and Creativity	4
I	Background	1
2	Financial Terminology and Concepts	2
2.1	Key Terms	2
2.1.1	Stock	2
2.1.2	Real Estate Investment Trusts (REITs)	2
2.1.3	Portfolio	2
2.2	Evaluating Investments with Data	2
2.2.1	Technical Analysis with Price Volume Data	2
2.2.1.1	Profits and Loss (PnL)	2
2.2.1.2	Risk and Volatility	2
2.2.1.3	Sharpe Ratio	2
2.2.2	Fundamental Analysis with Financial Statements	2
2.2.2.1	Income Statement	2
2.2.2.2	Balance Sheet	2
2.2.2.3	Cash Flow Statement	2
2.2.3	Alpha Formulas	2
2.2.4	Other Methods of Analyses	2
3	Literature Review of Portfolio Optimisation Techniques	3
3.1	Optimal Portfolio Theory	3
3.2	Traditional Time-Series Analysis	3
3.3	Machine Learning Techniques	3
3.4	Comparison of Techniques	3
3.4.1	Scope	3
3.4.2	Profitability	3

3.4.3	Predictive Accuracy	3
II	Innovation	1
4	Datasets	2
4.1	Extended Fundamental Data Features	2
4.2	Feature Selection with Decision Trees	2
5	Methodology	3
5.1	Machine Learning for REITs Portfolio Optimisation	4
5.1.1	MLR / NN / LSTM Predictions with Extended Features	4
5.1.2	Trade Execution Logic	4
5.1.3	Performance Evaluation	4
5.2	Genetic Algorithm Search for Outperforming Alphas	4
5.2.1	Alpha Tree	4
5.2.2	Application of Genetic Algorithms to Alpha Trees . . .	4
5.2.2.1	Objective Function	4
5.2.2.2	Selection	4
5.2.2.3	Crossover	4
5.2.2.4	Mutation	4
5.2.3	Portfolio Allocation Using Alpha	4
5.2.4	Performance Evaluation	4
III	Experiments	1
6	Post-processed Financial Datasets	2
7	Machine Learning Results	3
7.1	Evaluating Stock Price Predictions	3
7.1.1	Multiple Linear Regression (MLR)	3
7.1.2	Neural Networks (NN)	3
7.1.3	Long-Short Term Memory (LSTM)	3
7.2	Trade Execution Results with Different Parameters	3
7.3	Overall Evaluation of Performances	3

8	Alpha Tree Search Results	4
8.1	Alphas Generated	4
8.1.1	Initial Set	4
8.1.2	Intermediate Alphas	4
8.1.3	Best Performing Alphas	4
8.2	Portfolio Allocation Results with Best Performing Alphas . . .	4
8.3	Overall Evaluation of Performance	4
9	Conclusion	5
9.1	Benchmarking Against Index Funds	5
9.2	Comparing Results with Literature Review	5
9.3	Key Findings	5
9.4	Major Contribution and Creativity	5
9.5	Future Work	5
9.5.1	More Operators and Features for Alphas	5

CHAPTER 1

INTRODUCTION

The saying goes that once a profitable trading formula has been discovered and traded on by enough people, its profits will be eroded away and it will cease to be profitable. Traders and investors appear to be playing a never-ending game of "Hide-and-Seek" in search of profitable trading strategies. Due to the evolving nature of the financial markets, traditional financial time-series forecasting models which are static in nature are becoming less effective than Machine Learning models and dynamic algorithms in identifying the best investments (Sheth & Shah, 2023). The success of Machine Learning has led to numerous research papers applying a myriad of Machine Learning techniques to predict stock prices (Obthong, Tantisantiwong, Jeamwatthanachai, & Wills, 2020). However, fewer papers apply these approaches in the Real Estate sector and these papers largely focus on future stock price prediction rather than portfolio allocation as a whole (Habbab & Kampouridis, 2024). Many of these papers also use price-volume data without fundamental financial data as inputs. In 2022, the global real estate sector was worth more than \$380 trillion and worth more than the global equity and bond markets combined (Tostevin & Rushton, 2023), with approximately 893 listed Real Estate Investment Trusts (REITs) (Nareit, 2024). Hence, addressing the research gaps in this sector is particularly valuable.

1.1 Problem Definition

The aim of this report is to develop an effective real estate portfolio optimisation model that maximises profits, minimises risks and outperforms the market indexes with a sharpe ratio of more than 2. A novel approach of applying Genetic Algorithms to search for outperforming Alpha formulas, as well as, three Machine Learning models, namely, Multiple Linear Regression (MLR), Neural Networks (NN) and Long-Short Term Memory (LSTM) are

used to construct models. A trading agent is also constructed to allocate funds to the portfolio based on the results from the models. The input data to the Machine Learning models will be extended beyond price-volume data to include fundamental stock data. The models will be trained on 10 years worth of real market data. Then, the performances of these approaches will be compared with traditional financial models and the market indexes.

1.2 Motivations

Interviews of traders and financial journalists have revealed that both technical and fundamental analysis are used for forecasting investments, for shorter and medium to long term investment horizons respectively (Oberlechner, 2001). Given that most investments in real estate are meant for the medium to long term, the lack of fundamental data inputs for REITs stock price predictions reveals a pressing need for fundamental data to be incorporated.

In the field of quantitative trading, quant firms seek to develop predictive algorithms for quantitative trading called "Alphas" that determine how capital is allocated to portfolios profitably (Tulchinsky, 2020). What began with investment experts manually constructing Alpha formulas has gradually been replaced with automated alpha mining techniques that employ Machine Learning algorithms (Wang et al., 2023). There appears to be an opportunity to apply the concept of quantitative finance Alphas for Real Estate portfolio allocation. In addition, researchers have been inspired by Charles Darwin's theory of evolution (Ruse, 1975) to develop a Genetic Algorithm and have applied it in finance (Aguilar-Rivera, Valenzuela-Rendón, & Rodríguez-Ortiz, 2015). Synthesising these two ideas, presents a new and innovative approach to use Genetic Algorithms to search for Alphas in the Real Estate sector.

Lastly, as most of past research is centered on price prediction, there is a

need to develop trade execution logic to allocate portfolios based on these predictions.

1.3 Major Contributions and Creativity

The first major contribution of this project is the analysis on REITs stock datasets with more than 150 features of technical and fundamental indicators. This expands the dimensionality of the input training data significantly beyond price-volume features to incorporate REIT company data from financial statements. The comprehensive dataset covers 150 out of the total worldwide population of 893 REITs and includes up to 20 years of historical data.

Creativity is demonstrated in the writing of a Genetic Search Algorithm that searches for outperforming REIT Alphas. It resulted in the second contribution of a trading model with superior returns that significantly outperforms the market indexes.

The third contribution is the implementation of a trading agent or trade execution logic that acts on individual REIT price predictions to allocate funds to selected REITS in a portfolio.

Overall, this research paper has expanded the body of knowledge on the application of Alphas and Machine Learning models in Optimising Real Estate Portfolios.

I. Background

CHAPTER 2

FINANCIAL TERMINOLOGY AND CONCEPTS

2.1 Key Terms

2.1.1 Stock

2.1.2 Real Estate Investment Trusts (REITs)

2.1.3 Portfolio

2.2 Evaluating Investments with Data

2.2.1 Technical Analysis with Price Volume Data

2.2.1.1 Profits and Loss (PnL)

2.2.1.2 Risk and Volatility

2.2.1.3 Sharpe Ratio

2.2.2 Fundamental Analysis with Financial Statements

2.2.2.1 Income Statement

2.2.2.2 Balance Sheet

2.2.2.3 Cash Flow Statement

2.2.3 Alpha Formulas

2.2.4 Other Methods of Analyses

CHAPTER 3

LITERATURE REVIEW OF PORTFOLIO OPTIMISATION TECHNIQUES

3.1 Optimal Portfolio Theory

3.2 Traditional Time-Series Analysis

3.3 Machine Learning Techniques

3.4 Comparison of Techniques

3.4.1 Scope

3.4.2 Profitability

3.4.3 Predictive Accuracy

II. Innovation

CHAPTER 4

DATASETS

4.1 Extended Fundamental Data Features

4.2 Feature Selection with Decision Trees

CHAPTER 5

METHODOLOGY

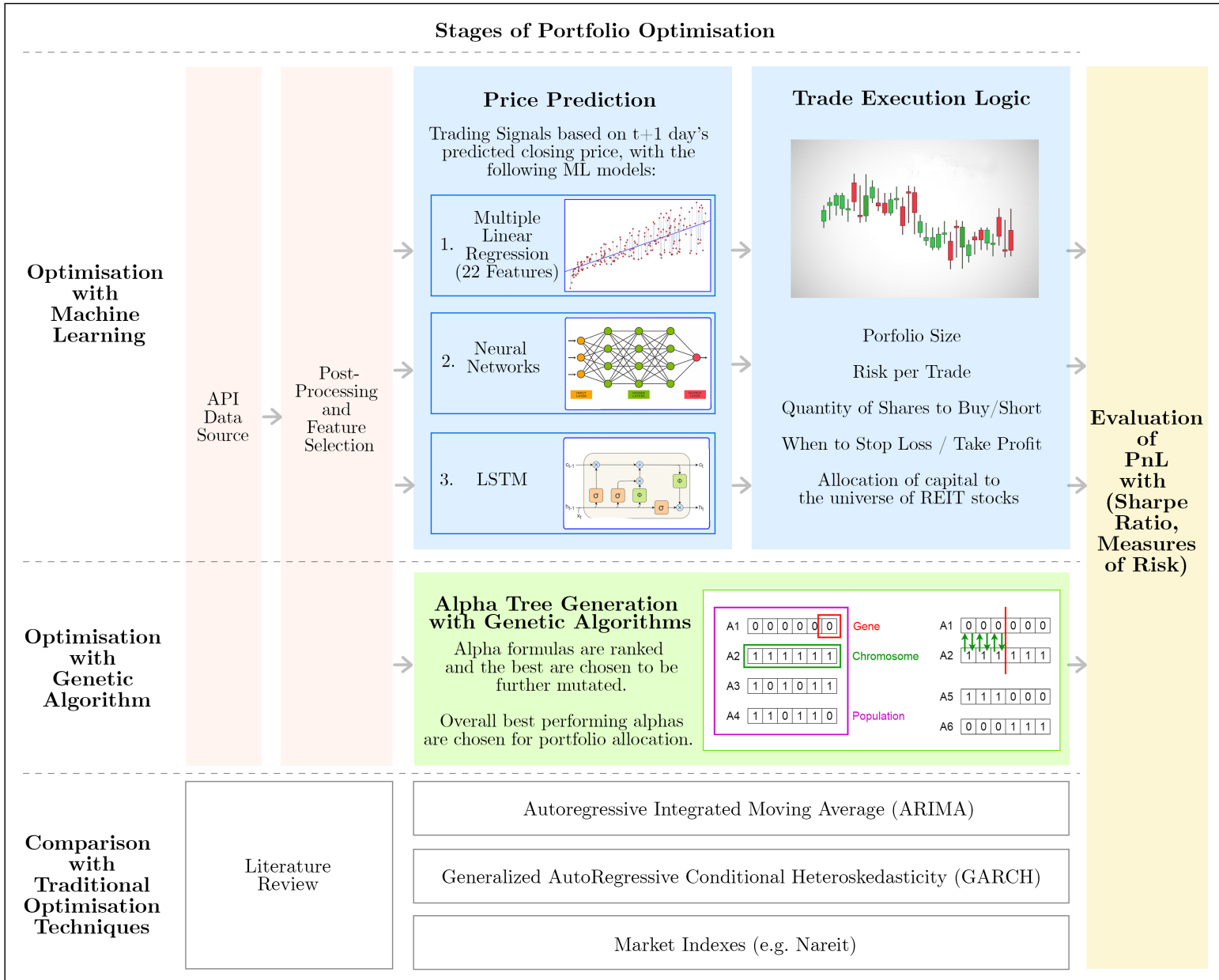


Figure 5.1: Write some caption here

5.1 Machine Learning for REITs Portfolio Optimisation

5.1.1 MLR / NN / LSTM Predictions with Extended Features

5.1.2 Trade Execution Logic

5.1.3 Performance Evaluation

5.2 Genetic Algorithm Search for Outperforming Alphas

5.2.1 Alpha Tree

5.2.2 Application of Genetic Algorithms to Alpha Trees

5.2.2.1 Objective Function

5.2.2.2 Selection

5.2.2.3 Crossover

5.2.2.4 Mutation

5.2.3 Portfolio Allocation Using Alpha

5.2.4 Performance Evaluation

III. Experiments

CHAPTER 6

POST-PROCESSED FINANCIAL DATASETS

CHAPTER 7

MACHINE LEARNING RESULTS

7.1 Evaluating Stock Price Predictions

7.1.1 Multiple Linear Regression (MLR)

7.1.2 Neural Networks (NN)

7.1.3 Long-Short Term Memory (LSTM)

7.2 Trade Execution Results with Different Parameters

7.3 Overall Evaluation of Performances

CHAPTER 8

ALPHA TREE SEARCH RESULTS

8.1 Alphas Generated

8.1.1 Initial Set

8.1.2 Intermediate Alphas

8.1.3 Best Performing Alphas

8.2 Portfolio Allocation Results with Best Performing Alphas

8.3 Overall Evaluation of Performance

CHAPTER 9

CONCLUSION

9.1 Benchmarking Against Index Funds

9.2 Comparing Results with Literature Review

9.3 Key Findings

9.4 Major Contribution and Creativity

9.5 Future Work

9.5.1 More Operators and Features for Alphas



Figure 9.1: The Universe

REFERENCES

- Aguilar-Rivera, A., Valenzuela-Rendón, M., & Rodríguez-Ortiz, J. (2015, November). Genetic algorithms and Darwinian approaches in financial applications: A survey. *Expert Systems with Applications*, 42, 7684–7697. doi: 10.1016/j.eswa.2015.06.001
- Habbab, F. Z., & Kampouridis, M. (2024). An in-depth investigation of five machine learning algorithms for optimizing mixed-asset portfolios including REITs. *Expert Systems with Applications*, 235, 121102. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0957417423016044> doi: <https://doi.org/10.1016/j.eswa.2023.121102>
- Nareit. (2024). *Global Real Estate Investment*. Retrieved 2024-04-08, from <https://www.reit.com/investing/global-real-estate-investment>
- Oberlechner, T. (2001, January). Importance of technical and fundamental analysis in the European foreign exchange market. *International Journal of Finance & Economics*, 6(1), 81–93. Retrieved 2024-04-08, from <https://onlinelibrary.wiley.com/doi/10.1002/ijfe.145> doi: 10.1002/ijfe.145
- Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G. (2020, May). A survey on machine learning for stock price prediction: algorithms and techniques. In (pp. 63–71). Retrieved 2024-04-08, from <https://eprints.soton.ac.uk/437785/> (Num Pages: 9) doi: 10.5220/0009340700630071
- Ruse, M. (1975). Charles Darwin’s Theory of Evolution: An Analysis. *Journal of the History of Biology*, 8(2), 219–241. Retrieved 2024-04-08, from <https://www.jstor.org/stable/4330635> (Publisher: Springer)
- Sheth, D., & Shah, M. (2023, February). Predicting stock market using machine learning: best and accurate way to know future

- stock prices. *International Journal of System Assurance Engineering and Management*, 14(1), 1–18. Retrieved 2024-04-07, from <https://link.springer.com/10.1007/s13198-022-01811-1> doi: 10.1007/s13198-022-01811-1
- Tostevin, P., & Rushton, C. (2023, September). *Total Value of Global Real Estate: Property remains the world’s biggest store of wealth*. Retrieved 2024-04-08, from <https://www.savills.com/impacts/market-trends/the-total-value-of-global-real-estate> (Section: Market trends)
- Tulchinsky, I. (Ed.). (2020). *Finding alphas: a quantitative approach to building trading strategies* (Second edition ed.). Chichester, West Sussex: Wiley.
- Wang, S., Yuan, H., Zhou, L., Ni, L. M., Shum, H.-Y., & Guo, J. (2023, July). *Alpha-GPT: Human-AI Interactive Alpha Mining for Quantitative Investment*. arXiv. Retrieved 2024-04-07, from <http://arxiv.org/abs/2308.00016> (arXiv:2308.00016 [cs, q-fin])