Bilkent University
Department of Computer Engineering

Senior Design Project
T2438
Para-Meter

Project Specification Document

22003598, Abdullah Samed Uslu, samed.uslu@ug.bilkent.edu.tr

22102566, Tuna Saygın, tuna.saygin@ug.bilkent.edu.tr

22102825, Sıla Özel, sila.ozel@ug.bilkent.edu.tr

22002756, Muti Kara, muti.kara@ug.bilkent.edu.tr

Selim Aksoy

Mert Bıçakçı, Atakan Erdem

21.11.2024

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfilment of the requirements of the Senior Design Project course CS491/2.

1. Introduction

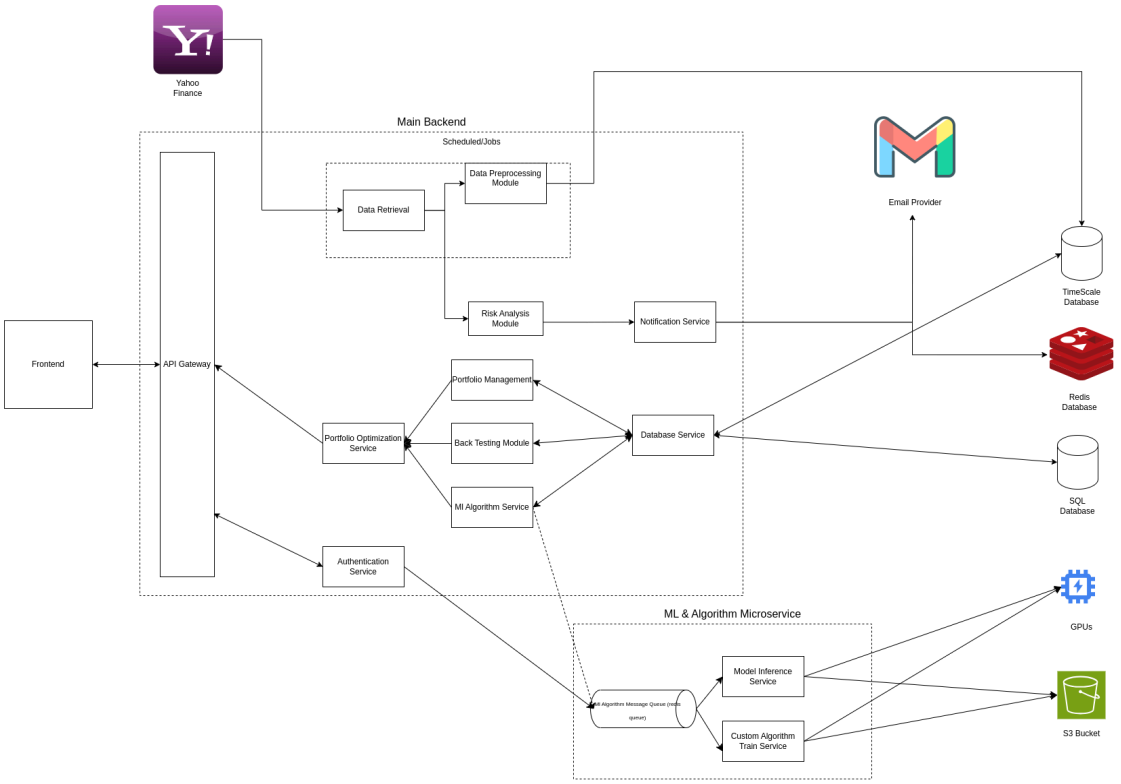
1.1. Description

The financial market has seen a growing demand for accessible, data-driven investment tools that enable users to optimize their portfolios with precision and flexibility. While effective for institutional investors with specialized knowledge, retail investors find traditional portfolio optimization solutions hard to use due to their technical complexity, expensive costs, and reliance on conventional analysis techniques.

This project aims to address these challenges by introducing the Para-Meter. By leveraging advanced machine learning algorithms, alternative data integration, and a user-friendly interface, the proposed solution enables institutional and retail investors to optimize their portfolios in an accessible, cost-effective, and customizable manner.

1.2. High-Level System Architecture & Components of Proposed Solution

The architecture of the portfolio optimization tool is designed to provide a modular, scalable, and customizable framework for users to manage their financial portfolios effectively. It enables users to utilize predefined algorithms or train their own machine-learning models on custom datasets while integrating core finance methodologies for portfolio optimization. The system architecture consists of multiple interacting components that work together to retrieve, process, and analyze data, perform optimizations, and provide actionable insights.



The solution's components are as follows:

- **The Frontend Module** provides an interface for users to interact with the systems. It allows users to create and manage portfolios and optimize them with highly customized algorithms.
- **API gateway** routes the frontend request to the appropriate API endpoint.
- **Authentication Service** manages user authentication and authorization, ensuring secure access to endpoints that connect microservices.
- **Portfolio Optimization Service** manages communication between backend services related to finance and centralizes workflow.
- **Data Retrieval Module** fetches daily stock data from Yahoo Finance to ensure models train with up-to-date information.
- **Data Preprocessing Module** prepares data by reducing noise and irrelevant features to improve training efficiency and model performance.
- **Risk Analysis Module** calculates risk metrics and sends alerts when user-defined thresholds are exceeded.
- **Notification Service** manages user notifications and handles related features like deletion and "Do Not Disturb" settings.
- **Email Provider** sends essential notifications to users via email.
- **Custom Algorithm Creation Service** gets the related parameters and custom dataset and sends a request to the GPU units to train the required algorithm and store it in the model storage.
- **Back Testing Module** tests the optimized portfolios from the historical stock data and returns the portfolio performance to the frontend via the API gateway.
- **ML Algorithm Service** generates future stock predictions for portfolio optimization service. This service also merges the future forecasts of multiple machine-learning algorithms. This service works closely with the ML & Algorithm Microservice.
- **ML & Algorithm Microservice** manages the model inferences and training part. This service separates GPU concerns from the main backend.
- **ML Algorithm Message Queue** is used for asynchronous training processes to increase the availability of the application. This part also verifies the request for whether or not the source is our backend.
- **Model Inference Service** manages model inferences and is used to make models more efficient.
- **Custom Algorithm Training Service** manages model fine-tuning tasks.
- **Portfolio Optimization Service** returns an optimized portfolio from the given portfolio and future predictions from the portfolio service.
- **Portfolio Management Service** handles portfolio addition, deletion, and update operations, as well as rebalancing the portfolio with the help of the Portfolio Optimization Service and other services.
- **Storages** are used to store essential data.
- **Database Service** is a service that manages databases and connects other backend services to databases.
- **GPU units** are used to train and infer the data.

1.3. Constraints

1.3.1. Implementation Constraints

- Accessing reliable and high-quality financial data is important for accurate modeling and optimization. However, many financial APIs have rate limits or limited historical data. Also, the data might be inaccurate or incomplete. These might hinder the accuracy of the results.
- Training machine learning models and running optimization algorithms require significant computational resources, including processing power and memory. The lack of access to high-performance hardware, such as GPUs or distributed computing systems, can slow down model training and testing, particularly when working with large datasets or performing hyperparameter tuning.
- The limited development timeline of 8 months can restrict the scope of development and testing. This limited duration requires careful prioritization of features and functionalities to ensure the delivery of a functional, high-quality system within the allotted time.

1.3.2. Economic Constraints

- The development of a machine learning-based portfolio optimization tool faces significant financial limitations. The initial budget for the project restricts access to advanced cloud infrastructure and premium financial datasets. Cloud services, which are essential for computation and storage, incur substantial costs, especially when utilizing high-performance configurations like GPUs or distributed systems. Similarly, high-quality financial datasets, often critical for accurate modeling, can be expensive and may exceed budgetary limits. These constraints could affect the accuracy and speed of the system's development.
- Operational costs are another concern, as the platform aims to remain affordable for retail investors who are sensitive to pricing. Maintaining low operational expenses is important to ensure the tool is accessible, which may necessitate adopting a freemium or tiered pricing model. Cost-effective development practices, such as using open-source tools and optimizing resource usage, will be critical to staying within budget.

1.3.3. Ethical Constraints

- A primary ethical concern is user data privacy. Compliance with data protection regulations such as GDPR (General Data Protection Regulation), CCPA (California Consumer Privacy Act), and KVKK (Personal Data Protection Law) is crucial since the tool handles sensitive financial and personal information.
- Avoiding algorithmic bias is one of the most important ethical constraints for Para-Meter. Machine learning models trained on financial data may unintentionally favor certain asset classes, industries, or demographic groups, leading to biased

recommendations. Ensuring fairness in portfolio suggestions requires diverse, high-quality training datasets, which can be hard to detect and obtain.

1.4. Professional and Ethical Issues

- Transparency in model behavior is crucial for Para-Meter. Machine learning algorithms often function as black boxes, which can deteriorate user trust. The platform must include explainable AI techniques, allowing users to understand how specific inputs, like financial indicators or alternative data, impact portfolio recommendations. This transparency builds trust and ensures that users can make informed decisions.
- Para-Meter must maintain impartiality to avoid conflicts of interest. For instance, it should not prioritize certain assets or datasets due to partnerships or financial incentives. Ensuring neutrality in data sourcing and recommendation processes is crucial for maintaining user trust and meeting professional standards.
- We acknowledge that the financial market is inherently unpredictable and volatile. While Para-Meter leverages state-of-the-art techniques to provide optimized portfolio suggestions, these recommendations are not guaranteed to lead to financial gains. Therefore, it is essential to emphasize that the ultimate responsibility for any financial decision lies with the user. Our application is intended to serve as a decision-support tool, not a substitute for professional financial advice. We explicitly disclaim any liability for financial losses resulting from acting on recommendations provided by Para-Meter. The performance of portfolios depends on numerous external factors, which are beyond the scope of Para-Meter. Users agree to accept full responsibility for their financial decisions by using it.

1.5. Standards

- **IEEE 830:** For the requirements specification of this project, we will use IEEE 830 standards.
- **UML 2.5.1:** For the UML diagrams like use case diagrams, activity diagrams, and class diagrams, UML 2.5.1 will be used.
- **OAuth2.0:** We will use the industry-standard authentication protocol OAuth2.0 to authenticate the users.
- **ISO 27002:** For security-related standards, we will use ISO standards to comply with many regulations regarding data security.

2. Design Requirements

2.1. Functional Requirements

User Features

- Users can create accounts, log in and manage profiles, store their preferences, save portfolios, and select features.
- A dashboard provides real-time visualization of portfolio metrics, asset allocations, and performance over time through graphs and charts.
- Users can compare different machine learning models side-by-side based on performance metrics, accuracy, and historical returns before selecting a model for portfolio optimization.

Customization

- Users can import alternative datasets (e.g., social sentiment, satellite data, economic indicators) to enhance portfolio predictions.
- Users can create custom features from raw data or choose predefined features like technical indicators, fundamental metrics, or sentiment scores.
- Users can choose from various machine learning models, including decision trees, SVMs (Support Vector Machine), neural networks, and ensemble methods.
- Users can set return goals, risk tolerance levels (using metrics like volatility and VaR), and investment horizons.
- Users can configure portfolios as long-only or long-short, set net and gross exposure limits, and manage leverage constraints.

Performance Analysis

- The platform uses cross-validation, out-of-sample testing, and rolling window testing for robust model validation.
- Integrated stress testing and Monte Carlo simulations assess portfolio performance under various market conditions.
- The platform provides detailed performance tracking, including metrics like Sharpe ratio, drawdown, turnover, and sector or factor exposures.

Notification and Alert System

- Users can set notifications and alerts for specific events, such as portfolio rebalancing thresholds, market movements, or changes in alternative data sources (e.g., sentiment analysis).

Optimization and Maintenance

- Models update periodically to incorporate the latest market data, maintaining relevance over time.
- The tool preprocesses data, including normalization and feature engineering, to ensure high-quality input for models.
- Automated tuning methods (e.g., grid search and Bayesian optimization) are available to optimize model performance.

2.2. Non-Functional Requirements

2.2.1. Usability

- The tool should have an intuitive UI, accessible to non-technical users, with easy-to-use options for model selection, customization, and analysis.

2.2.2. Reliability

- High availability to ensure consistent access and minimal downtime, particularly for real-time portfolio updates and rebalancing.
- Robust security protocols to protect user data and ensure compliance with data protection regulations.

2.2.3. Performance

- The platform should ensure low latency for portfolio rebalancing and simulations using parallel computing techniques.
- Models and calculations should be optimized to minimize processing time without sacrificing accuracy, even under heavy computational loads.

2.2.4. Supportability

- Modular code structure allows easier updates, model additions, and future feature integration.

2.2.5. Scalability

- Support for portfolios of varying sizes, with efficient handling of large datasets and concurrent requests.
- Support for large-scale data scraping and processing.

3. Feasibility Discussions

3.1. Market & Competitive Analysis

It has been identified that six companies serve users with portfolio optimization tools. The following table lists the available features of the applications from these six companies. The last row of the table is what we aim to serve to our users.

Competitors\Features	Algorithms Customizable	Custom Dataset Addition	Train Own Algorithm	No Need To Code	Back Testing	Explainability	Friendly UI/Simplicity
Magnus	Yes	No	No	Yes	Yes	Yes	Yes
Blackrock-Aladdin	No	No	No	Yes	Yes	No	No
InvestSuite iVar	No	No	No	Yes	Yes	No	Yes
SimCorp-Axioma	Yes	No	Yes	Yes	Yes	No	No
Gsphere-Porfolio ThinkTank	Yes	No	Yes	Yes	Yes	Yes	Yes
Talos	No	No	No	Yes	NA	No	Yes
Para-Meter	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Magnus is a Turkish-based company that offers AI-based portfolio recommendations to their users. They are trusted by Kuvveyt Türk, one of the banks in Turkey [1]. This depicts that they are prominent in the market. They offer a variety of different algorithms with customizable parameters; however, they don't offer any features like adding a custom dataset and training your own machine learning algorithms using the custom dataset [2].
Blackrock is a big finance company that offers a portfolio management system called Aladdin, which has extensive data retrieval techniques and sophisticated risk analysis tools to make better predictions [3], [4]. Even though AI is used to optimize portfolios, there is no explanation of the methodology that decreases the user's trust. Additionally, lots of parameters and values in the user interface make it complex and hard to use.

InvestSuite iVar also offers a portfolio management tool with advanced risk management calculations and has advanced backtesting simulation [5]. Their simplistic UI, not showing every parameter and metric on one page, makes their UI friendly [6]. However, they don't allow users to train their machine learning algorithms to predict the future values in the stock, which, in turn, will be used to optimize the portfolio.

SimCorp-Axioma also offers a portfolio management tool. What is different from their predecessors is that it allows users to use their 3rd party or their own models to optimize their portfolios. Even though they offer users that feature, they have no feature that allows users to train their own algorithm [7].

Gsphere Portfolio ThinkTank is also a portfolio management tool that allows users to select different algorithms and parameters while optimizing the portfolio. Additionally, they allow

users to rebalance, and they offer a variety of constraints while optimizing, such as each asset cannot be above a threshold percentage value [8]. Even though there is a feature that allows users to train their own algorithms, it has significant costs [9].

Talos is a portfolio construction tool with a simplistic and user-friendly UI, which allows users to optimize their portfolios and even rebalance them. The optimized portfolio can undergo several tests and comparisons with benchmarks [10].

3.2. Academic Analysis

The academic literature on portfolio management and optimization has evolved with the integration of machine learning algorithms into the area. However, traditional methods like Markowitz's Mean Variance Theorem still dominate the area. Therefore, academic research for Para-Meter was conducted to cover both classical techniques and modern portfolio optimization approaches, aiming to develop feasible solutions for our problem set. After research, the source set was created to gain particular knowledge on the area and reference during the project.

3.2.1. Books

- **Modern Portfolio Theory and Investment Analysis, 9th Edition**, E.J. Elton, M.J. Gruber, S.J Brown, W.N. Goetzmann
- **Portfolio Optimization**, M.J.Best.
- **Quantitative Portfolio Management**, M. Isichenko
- **Portfolio Management With Heuristic Optimization**, D. Maringer
- **Uncertain Portfolio Optimization**, Z. Qin

3.2.2. Articles

- **Portfolio Management Using Deep Reinforcement Learning,** A.A. Pawar, V.P. Muskawar, R. Tiku
- **Autonomous Sparse Mean-CVaR Portfolio Optimization,** Y. Lin, Y. Zhang, Z. Lai, C. Li
- **Covariance Matrix Analysis for Optimal Portfolio Selection,** L.H.S. Keith
- **Large-Scale Time-Varying Portfolio Optimisation Using Graph Attention Networks,** K. Korangi, C. Mues, C. Bravo
- **On Accelerating Large-Scale Robust Portfolio Optimization,** C.H. Hsieh, J. Lu
- **Derivatives Portfolio Optimization and Parameter Uncertainty,** A. Vorobets
- **Integrating Minds: An Ensemble Approach to Portfolio Optimization,** F. Shi, L. Shu, X. Gu, J. Qiu
- **Reinforcement Learning Applied to Dynamic Portfolio Management: A Real Market Data Application,** M. Mastrogiovanni
- **On the Average Sensitivity of Unconstrained Markowitz Optimization,** C. O'Cinneide

- **A Fuzzy Online Portfolio Selection Algorithm Based on Pattern Matching Approach,** M. Abdi, S.B. Ebrahimi, A.A. Najafi
- **Optimizing Portfolio with Two-Sided Transactions and Lending: A Reinforcement Learning Framework,** A. Habibnia, M. Soltanzade
- **Nonparametric Analysis of Financial Portfolio Performance,** L. Chrchye, B.D. Rock, D. Saelens

4. Glossary

AI: Artificial Intelligence
ML: Machine Learning
GPU: Graphics Processing Unit
Portfolio: A collection of financial investments such as stocks, bonds, commodities, closed-end funds, exchange-traded funds (ETFs), etc. [11].

5. References

[1] “Get Magnus,” Magnus, [Online]. Available: https://getmagnus.com/ [Accessed Nov. 21, 2024].

[2] Magnus, “Magnus Paneli Portföy Oluşturma”, Youtube, [Online]. https://www.youtube.com/watch?v=xkLZfYrgjNY [Accessed Nov. 21, 2024]

[3] Ryder the cat, “Aladdin® by BlackRock - software for portfolio management,” Aladdin, [Online]. Available: https://www.blackrock.com/aladdin [Accessed Nov. 21, 2024]

[4] “Blackrock Aladdin - Fixed Income Performance Analyst,” Youtube, [Online]. Available: https://www.youtube.com/watch?v=nFwJlw0Ut9k [Accessed Nov. 21, 2024]

[5] “iVaR | InvestSuite” InvestSuite, [Online]. Available: https://www.investsuite.com/investsuite-ivar [Accessed Nov. 21, 2024]

[6] Investsuite, “On-demand Portfolio Performance Reporting,” Youtube [Online]. Available: https://www.youtube.com/watch?v=Hwg1NDqXYn4 [Accessed Nov. 21, 2024]

[7] “Axioma Portfolio Optimizer for portfolio construction | SimCorp,” SimCorp, [Online]. Available: https://www.simcorp.com/solutions/point-solutions/axioma-portfolio-optimizer [Accessed: Nov. 21, 2024]

[8] Portfolio ThinkTank, “Optimizing Wall Street Bets (WSB) stock idea using Swaggystocks and Portfolio ThinkTank,” YouTube, [Online]. Available: https://www.youtube.com/watch?v=-q0qyV7zIxM [Accessed Nov. 21, 2024].

[9] “Custom Pre-Trained Deep Learning Predictions for Business,” Portfolio ThinkTank [Online]. Available: https://advisors.portfoliothinktank.com/product/custom-pre-trained-deep-learning-predictions/ [Accessed: Nov. 21, 2024]

[10]Talos | Portfolio Construction | Crypto Portfolio Management

[11] C.Tardi. (2024, June 6). “Financial Portfolio: What It Is, and How to Create and Manage One,” Investopedia, [Online]. Available: https://www.investopedia.com/terms/p/portfolio.asp. [Accessed: Nov. 21, 2024].