Research Proposal

Integrating Relational and Time Series Features for Enhanced Stock Price Forecasting and Portfolio Optimization: A Case Study on the Iran Stock Market

August 27, 2023

Contents

1 Abstract

This research paper aims to explore the interplay between stock price forecasting and portfolio optimization techniques, with the objective of developing a comprehensive framework for efficient investment decision-making. The study begins by examining various approaches to stock price forecasting in the context of the Iran stock market. The performance and accuracy of each method will be critically evaluated based on historical data from a diverse set of securities.

We hope that by leveraging the combination of relational features extracted using Graph Neural Network (GNN) methods and time series features obtained from transformer-based models , we would enhances prediction accuracy .

In the second phase of the research, the focus shifts towards portfolio optimization techniques that help investors allocate their assets optimally. Modern Portfolio Theory, Markowitz's mean-variance optimization, and more recent approaches such as Black-Litterman and risk-parity strategies will be explored. The goal is to identify the most effective methodologies for constructing well-diversified portfolios that balance risk and return.

Furthermore, this paper delves into the integration of stock price forecasting techniques with portfolio optimization models. By incorporating accurate short-term and long-term price predictions into the optimization process, the aim is to improve the overall performance and risk-adjusted returns of investment portfolios. Various portfolio rebalancing and trading strategies based on the forecasted prices will also be investigated.

To evaluate the effectiveness of the implemented framework, extensive experiments and analysis are conducted using real-world data from the Iran stock market. Performance metrics, including prediction accuracy, risk-adjusted returns, and portfolio evaluation measures, are employed to assess the framework's performance compared to traditional methods and baseline models.

Keywords

Forecasting ,Deep learning , transformer, GNN, Optimization, Markowitz's mean-variance optimization

2 Introduction

our research consist of two main part

- Price forecasting
- Portfolio optimizationhow to select the number of items for allocating to assets in different market

Price forecasting predicts price movements for assets like stocks, commodities, currencies, and cryptocurrencies. It involves analyzing historical data, trends, and factors that impact prices. Techniques include technical analysis (examining price patterns and indicators) and fundamental analysis (evaluating asset value based on financial statements, industry analysis, etc.). Price forecasting supports decision-making in financial markets, but it isn't foolproof due to unpredictable market forces.

The purpose of stock price prediction is to explore the development law of stock market so as to provide a scientific basis for stock investments. As the stock price volatility is caused by many factors, it is difficult to grasp the uncertainty of these factors affecting stock prices

Relational features capture interdependencies and correlations between stocks from different markets. GNN techniques are used to encode and incorporate these features into the forecasting process, capturing local and global interactions.

Transformer-based models extract time series features from historical stock price data, capturing complex temporal patterns and dependencies. These models encode sequential information, providing valuable features reflecting stock price dynamics.

Integrating relational and time series features creates a comprehensive forecasting framework. By combining GNN-derived relational features and transformer-based time series features, the approach captures both structural and temporal aspects of the stock market. This fusion enhances prediction accuracy and provides a comprehensive view of market dynamics.

In addition, an alternative approach can be employed to forecast stock prices utilizing cutting-edge reinforcement learning techniques, which have gained prominence in recent times. The application of reinforcement learning in the context of stock price prediction for a particular stock is justified by its reliance on minimal historical data and a Markovian process. We can leverage the Q-learning algorithm, a model-free approach, to effectively utilize the available data. By maximizing the expected value of the total reward across subsequent steps, beginning from the current state, Q-learning enables the determination of an optimal policy.

Portfolio optimization : how to select the number of items for allocating to assets in different market

The first mathematical model for portfolio selection was presented by Markowitz [4] that is evaluating the mean and variance of investments A basic and important assumption of Markowitz mathematical model is that the investor knows the exact expected return but in reality, the true expected is not occurred and a lot of factors can change the results

In spite of the fundamentality of the Markowitz mean–variance model in optimal portfolio selection problem, always there have been critics because of being quadratic and making use of covariance matrix in the risk objective of this

Stephen Boyd [1] consider a basic model of multi-period trading, which can be used to evaluate the performance of a trading strategy based on markowitz model

3 Literature Review

Harry Markowitz's seminal paper 's called "Portfolio Selection" was published in 1952. It introduced the concept of Modern Portfolio Theory (MPT) and won Markowitz the Nobel Prize in Economics. It revolutionized investment theory by emphasizing the importance of diversification and balancing risk and return in constructing investment portfolios.[4]

Stephen Boyd [1] describe a framework for single-period optimization, where the trades in each period are found by solving a convex optimization problem that trades off expected return, risk, transaction cost and holding cost such as the borrowing cost for shorting assets. then describe a multi-period version of the trading method, where optimization is used to plan a sequence of trades, with only the first one executed, using estimates of future quantities that are unknown when the trades are chosen. The single period method traces back to Markowitz

4 Methodology

Price Forecasting

In this research , introduces the application of a recently introduced deep learning model — the Transformer model, to predict the future price of stocks.

Transformers, a revolutionary deep learning model, have gained significant attention and popularity in the field of natural language processing (NLP). Introduced in the seminal paper "Attention Is All You Need" by

Vaswani et al [5]. in 2017, Transformers have revolutionized the way we approach various NLP tasks, such as machine translation, text generation, sentiment analysis, and also time series analysis such as stock price prediction and more.

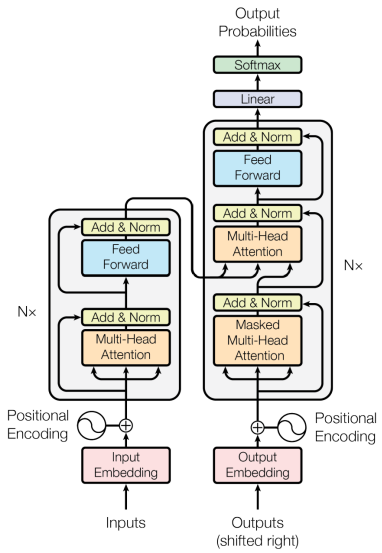

Figure 1. The Transformer - model architecture.

Our model is designed with an encoder component but does not include a decoder. This design is similar to well-known encoder models such as the BERT architecture [2].

The existing literature extensively explores the time series analysis of stocks . However, there is a noticeable research gap regarding the influence of stock correlation on the accuracy of stock price forecasting.

Therefore, in this research we are going to use relationship information which is extracted with a graph convolutional network (GCN) similar to the method presented by Cheng Zhao [3] and at the end integrate both the time series and relationship information to forecast the stock price.

Portfolio optimization

in order to have to solid optimization over a vast number of assets in the market we need to have good mathematical model that can represent the market

Parameters to be modeled:

- Period and time step: The time period in our model is arbitrary and could be daily, weekly, or one hour intervals
- holdings in an specific time interval : could be price value in a Fiat currency or just normalized vector since we only need to know what percentages an specific asset is in our Portfolio
- Trade vector : in each time step what changes happens to our holdings
- Transaction cost : how much each trade costs us
- Holding cost : for some assets we need to pay holding cost e.g shorting
- Assets return vector: in each time step how much change of value occurs in each asset (this is estimated is the previous section of research)
  bid - ask price mean is a good candidate
- Portfolio return : in each time step how much change of value in the whole porfolio
- Risk: we need to find a model that can model risk , Markowitz was first to do that but there are some limitation in his model we will gather an evaluate models that do better in practice
- Constraints: in every portfolio we face some constraints which are either forced by the market (e.g. you can't sell less than 500k toman in Iran stock market) or the constraints are trader strategy (e.g β neutrality)

Optimization:

after modeling the portfolio and market parameter we use optimization to see what the is best choice of **Trade vector** in each time step
we could use greedy algorithms in this part and evaluate the result
if our model parameter and constraint are convex we could use convex optimization to maximize Portfolio return over a period

5 Expected Results

The ultimate objective of this research is to provide practical insights into the symbiotic relationship between stock price forecasting and portfolio

optimization, enabling investors to make more informed investment decisions and achieve enhanced risk-adjusted returns in dynamic financial markets.

6 Challenges

the main data needed for this research can be gathered from Tehran Securities Exchange Technology Management Co tsetmc
other useful websites:

- codal
- rahvard365
- sahamyab

since we are using deep neural network we need a large data set for model training
because embedder models (e.g BERT and XLNET) are trained on the text corpus for NLP tasks, we can't use the pre-trained embedder models directly so the code needs to be written from scratch , training procedure also needs to be developed from zero
modeling some parameters discussed above might be a Challenge in Iran stock markets we need to pay careful attention on the structure of this market not only for portfolio optimization part but also for price forecasting
in Iran stock market there are a lot of noise that makes the market hard to predict we need to find a way to reduce the noise one solution might be to make time steps bigger e.g monthly or weekly instead of daily

References

[1] Stephen Boyd. Multi-period trading via convex optimization. *Foundations and Trends® in Optimization*, 3(1):1–76, 2017.

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[3] Qiang Liu, Xiaohui Yu, and Shouyang Li. Stock market analysis using time series relational models for stock price prediction. *Expert Systems with Applications*, 38(8):10683–10690, 2011.

[4] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.