Stock Prediction using a Transfomer model and Portfolio Optimization with Mean-Variance model

INC27

CONTROL SYSTEMS AND INSTRUMENTATION ENGINEER PROGRAM

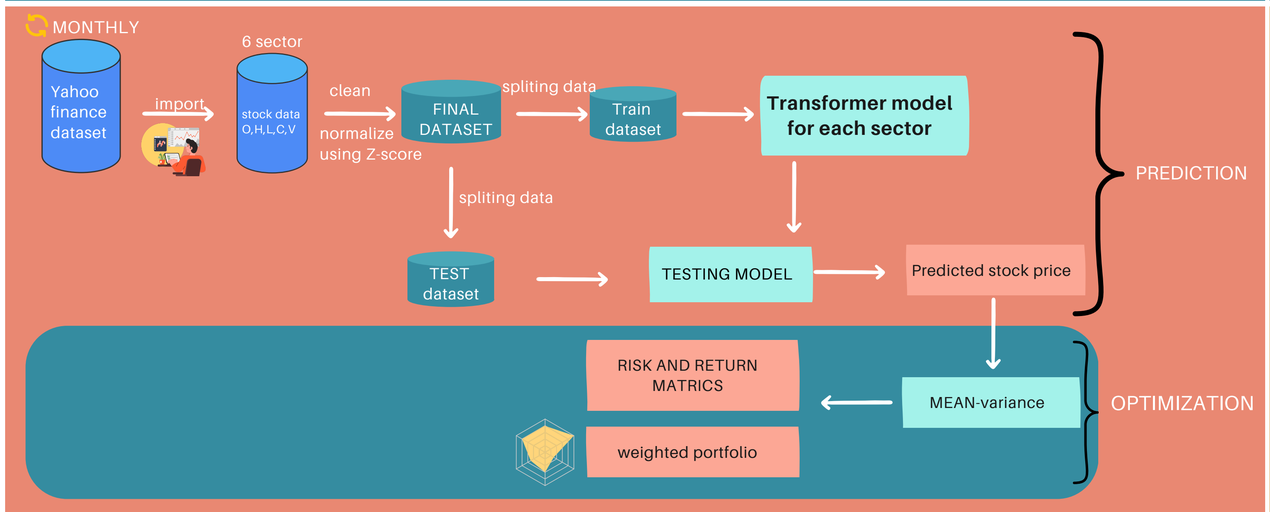
Boonyarit Kerdsuwan, Siraphop Noiamphaeng , Advisor: Dr. Kajornvut Ounjai, Co-Advisor: Dr. Teema Leangarun

Abstract

This project presents a deep learning-based framework for stock prediction and portfolio optimization using data from Thailand's SET50 index (2004–2024). An encoder-only Transformer model is employed to forecast future stock returns, utilizing key features such as open, high, low, close prices, and trading volume.The predicted returns are integrated into a mean-variance optimization (MVO) framework to construct dynamic portfolios. Experimental results across six major market sectors show that the optimized portfolios achieve higher returns, lower volatility, and better Sharpe ratios than traditional historical-data-based strategies. This approach demonstrates the effectiveness of combining Transformer models with financial optimization to enhance real-world investment decision-making.

Methods

System Overview



This system is designed to predict stock prices and optimize portfolios using a two-phase process: Prediction and Optimization.

1. Data Collection and Preprocessing (Input Stage)

- Source: Stock data is collected from Yahoo Finance.
- Imported Data: Includes stock data in 6 sectors, with key features like Open, High, Low, Close, and Volume.
  - 6 sectors include, Financial,Property,Resources,Services,Technology and Food industry.
- Cleaning & Normalization: The collected data is cleaned and then normalized using Z-score normalization to standardize the values.

2. Dataset Preparation

- This dataset is then split into two parts:
  - Train dataset: Used data from 2014-2023 to training.
  - Test dataset: Used data from 2024 for model evaluation.

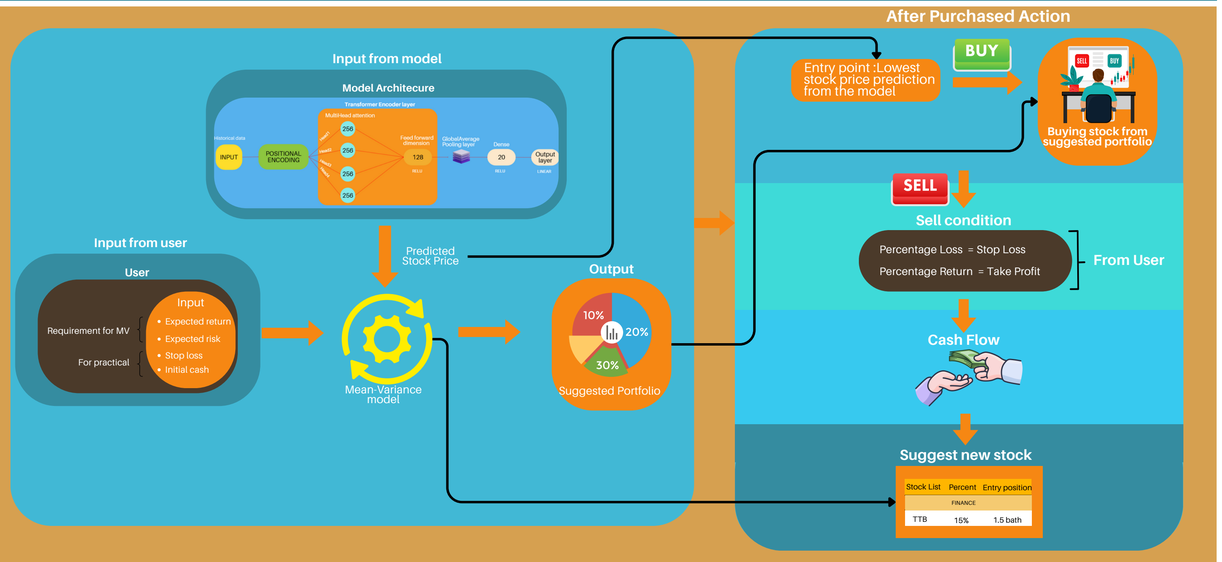
3. Prediction Phase

- A Transformer model is trained separately for each sector using the train dataset.
- The trained model is then evaluated using the test dataset, producing predicted stock prices.

4. Optimization Phase

- The predicted stock prices are input into a mean-variance optimization process.
- This optimization considers risk and return metrics to create a weighted portfolio that balances risk and expected return.

Back testing logic



This diagram illustrates the real-time backtesting loop of our Project. Forecasted stock prices from sector-specific Transformer models are combined with user-defined investment preferences

1. Set Your Investment Rules

- You choose your investment goals: Expected return, Expected risk, Stop-loss and initial investment.

2. Predict Future Prices

- The system uses AI (Transformer models) to forecast stock prices for the whole year

3. Build the optimal portfolio

- Using those predictions, it finds the best mix of stocks (using mean-variance optimization) to match your return and risk targets.

4. Buy Stocks at the Right Time

- It waits and buys a stock only when its actual price hits the model's predicted lowest point.

5. Sell When Conditions Are Met

- If a stock's price hits your highest preidiction or falls below your stop-loss, the system automatically sells it.

6. Re-investment

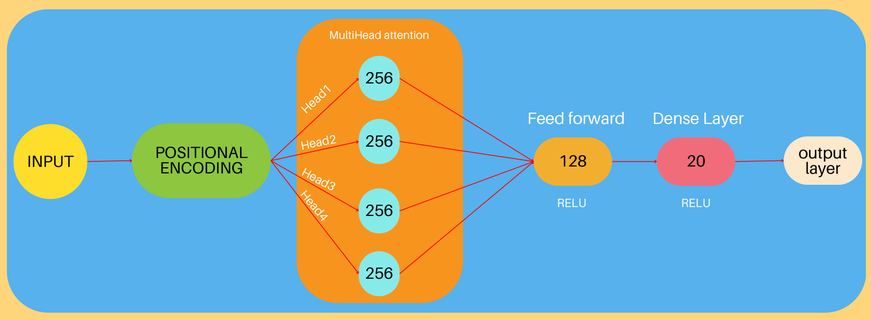
- The money from sold stocks is used to buy new stocks — based on the forecast and your original strategy.

7. Track Performance

- At the end, we use recording data to calculate total return, Sharpe ratio, and other performance metrics — to see how well this strategy worked.

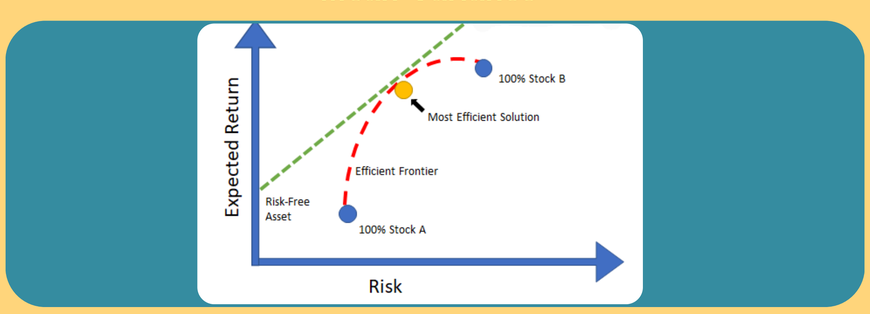
Model

Transfomer model



The heart of stock price prediction is the **'Transformer model'** we have developed. It learns from historical stock price data (e.g. high-low prices, trading volume) to find complex relationship patterns and predicts what each stock's price is likely to be in the future.

Mean-Variance



After we can predict the stock price, we will use the **'Mean-Variance'** principle to help decide which stocks to invest in and in what proportion, so that overall our investment portfolio has the highest chance of profit, without too high risk.

Results

Model Result

Financial and Resource sector had the lowest errors, while Technology sector had higher errors, reflecting volatility. Overall, the model generalizes well across sectors.

MODEL/SECTOR	MSE	MAE	RMSE
Agro & Food Industry	0.0129	0.083	0.1135
Financials	0.03	0.0473	0.0416
Property & Construction	0.0295	0.1307	0.1719
Resources	0.0284	0.045	0.00401
Services	0.0131	0.0817	0.1143
Technology	0.0896	0.2174	0.2993

Portfolio Result

The MV model using predicted values outperforms the model using historical values in terms of cumulative portfolio growth and risk-adjusted returns since 2014.

MV model with Predicted value

	Return	Sharpe ratio
2014	52.76%	3.31
2015	-2.66%	-0.25
2016	29.73%	1.61
2017	2.21%	0.03
2018	-11.69%	-0.91
2019	14.98%	0.68
2020	-14.6%	-0.57
2021	2.71%	0.06
2022	19.26%	1.28
2023	70.7%	1.05
2024	-6.43%	-0.38

Portfolio Value Average 14.36% per Year

VS

MV model with Historical value

	Return	Sharpe ratio
2014	26.63%	0.42
2015	-7.39%	-2.46
2016	25.54%	0.35
2017	10.70%	-1.11
2018	0.35%	-1.93
2019	22.76%	0.17
2020	-19.44%	-1.12
2021	34.83%	0.88
2022	6.82%	-1.52
2023	0.35%	-2.13
2024	7.34%	-1.37

Portfolio Value Average 9.86% per year

Conclusion

This project successfully demonstrates the integration of a Transformer-based deep learning model with mean-variance optimization for stock prediction and portfolio management. By applying an encoder-only Transformer to historical SET50 stock data, we were able to effectively forecast future stock returns using key market features. These predictions, when incorporated into a mean-variance optimization framework, enabled the construction of portfolios with superior performance in terms of return and sharpe ratio.The results highlight the strength of combining modern AI techniques with traditional financial models to enhance decision-making in dynamic market environments. This approach offers a scalable, data-driven strategy that can support more accurate investment planning and risk management in real-world applications.

Department of Control Systems and Instrumentation Engineering, Faculty of Engineering
King Mongkut's University of Technology Thonburi(KMUTT)

iNC AUTOMATION FOR ALL

PROJECT DAY 2025 FACULTY OF ENGINEERING