The TSC Thesis Symposium

Thesis Title:

AI for Generating Investor Views in Black–Litterman Portfolio Optimization

Registration No.: 20065

AI for Generating Investor Views in Black–Litterman Portfolio Optimization

ABSTRACT

When it comes to financial prediction, artificial intelligence (AI) outperforms conventional economic models with huge and complicated datasets. Active investor uses the Black-Litterman (BL) model to incorporate investor views into market equilibrium returns. However, there hasn't been much debate on generating investor views that based on artificial intelligence models in BL portfolio optimization.

The objectives of this research are to proposes a new methodology to implement BL portfolio models integrated with quantitative opinions generating from AI algorithms. For the first model, we used AI models with past 4 years' fundamental financial factors as inputs to predict the next quarter's return and find model outperforms with daily and quarterly financial factors in predictive accuracy. For the second model, we design a novel BL_AI hybrid models tested on portfolios of Yuanta/P-shares Taiwan Top 50 ETF's stocks, in general, the results reveal better portfolio performance with Sharpe ratios.

The contribution of this study has been to confirm BL_AI framework outperforms comparable portfolio models with a greater Shape ratio in general. The practitioner implication of this study is that we provide fundamental investors to build portfolio in which artificial intelligence is used to provide investment views in a volatile market environment to enhances portfolio performance.

Keywords: Artificial Intelligence, Sliding Window, Black–Litterman, Investor View, Fundamental Factors

1. Introduction

A deep learning algorithm with time series data is a valuable tool for stock forecasting when there is nonlinear behavior (Lee & Yoo, 2020). Since the traditional technique uses hand-engineering to develop an algorithm to recognize the features, the result is usually time-consuming. On the contrary, deep learning can utilize the vast pool of data and automatically learn these features from the data, which results exceed the traditional techniques (Zhou, 2019).

What we know about Markowitz introduces a rigorous mathematical framework for asset allocation (Markowitz, 1952). This framework, which is based on a mean-variance optimization technique, only evaluates previous prices to determine returns and volatility, making it less predictable. Following Markowitz's thinking, Black–Litterman (BL) model combines market equilibrium with investors' expectations in the portfolio. Black and Litterman (1992), which overcomes the problem of unintuitive, highly-concentrated portfolios, input sensitivity, and estimation error maximization, also reacts more quickly to changes in the economic cycle and adjusts the asset allocation more rapidly (Bessler, Opfer, & Wolff, 2017).

However, it is unclear if using a misspecified asset pricing model can provide valuable insights of investor views. This indicates a need to understand the process of generating the view vector considering the complex and dynamic nature of financial markets. Artificial intelligence (AI) can be employed to model complex patterns and forecasting problems to produce investor views to enhances the prediction performance in volatility market.

This paper aim to employ different forecasting methodologies to produce investor views in a more accurate or unbiased way by using recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU) algorithm to construct a data-driven portfolio wit future return prediction.

Two important stages went into the develop of the BL_AI portfolio framework, which can be listed as: (1) using financial factor as feature for the moving window neural network model and (2) using of BL model to build portfolio in which artificial intelligence is used to provide investment views.

The paper is organized as follows: Section 2 provides the related previous research on financial forecasting methodologies of AI algorithm and asset allocation technology. Section 3 discusses research design architecture and research method. Section 4 summarizes the experimental results. Chapter 5 discusses the findings results and future research directions.

2. Literature Review

This chapter mainly be positioned between two research streams: forecasting model base on artificial intelligence algorithm and hybrid approaches of asset allocation technology.

2.1 Stock Market Forecasting Base on Artificial Intelligence

To date, several studies have investigated stock market forecasting using quantality technology. Coen, Gomme, and Kendall (1969) are aware of the crucial reliance on one- or two-time units, lag features effects, and stock returns (sometimes up to four in quarterly data to indicate seasonal effects). Stocks that are safe, profitable, growing, and well managed should command a higher price from an investor, therefore on average, high-quality stocks do cost more (Asness, Frazzini, & Pedersen, 2019). Because financial time series are chaotic, noisy, and non-stationary, forecasting them has become difficult (Wang, Huang, & Wang, 2012). According to Cavalcante, Brasileiro, Souza, Nobrega, and Oliveira (2016), learning mechanisms tend to generate the best results with data that is very noisy, non-stationary, and high dimensional, which is a significant improvement in predicting future market movements. In Table 1, examples of research are briefly discussed and summarized with the methods applied and the performance of the model.

Table 1 Relative works on Stock Market Forecasting

Art	Data Set	Period	Feature Set	Method	Performance Criteria

Chatzis. et al.(2018)	Several finance datasets	1996-2017	Index data, Bond yield (10 year), Exchange rates	RF, SVM, NN, XGBoost, DNN	AUC=0.78, Geometric Mean=0.64, Likelihood ratio=0.61, Balanced accuracy=0.68
Fu et al. (2018)	Chinese stock data	2012-2013	Technical, Fundamental data	Logistic Regression, RF, DNN	AUC= 0.97, Accuracy=0.92, Precision=0.92, Recall=0.92
Lee and Yoo (2020)	S&P500 stock data	1997-2016	Price data	RNN, LSTM, GRU	Accuracy=0.60, Monthly return= 0.02
Jiang and Liang (2017)	Cryptocurrencies	2014-2017	Price data	CNN, RNN, LSTM	Annualized return = 47.15% Sharpe ratio =0.08
Alberg and Lipton (2017)	US stock data	1970-2017	Fundamental data	MLP, RNN	MSE = 0.47, Annualized return = 17.10% Sharpe ratio = 0.68

2.2 Asset Allocation and Portfolio Management

One of the key concerns with investments is portfolio management. Investors construct portfolios to get the best risk-reward ratio with variety investing strategies. Portfolio using AI models to optimize portfolio weights can increased Sharpe ratio (Zhang, Zohren, & Roberts, 2020). Yaman and Dalkılıç (2021) used a hybrid approach based on the AI model and the BL algorithm to address the portfolio optimization problem, improving the performance of the portfolio. Table 2 examples of research are briefly discussed and summarize the outcomes with various asset allocation techniques.

Table 2 Relative works on Portfolio Management

Art	**Data Set**	**Sliding Window**	**Feature Set**	**Algorithm**	**Performance Criteria**

Alexander and Svetlana (2013)	US equities, US fixed interest, US real estate commodities	-	Fundamental, Sentiment, Technical	Random Forest	Sharpe ratio= 2.20
Quintana and Isasi (2013)	Stock, Bonds	History=120D, Future=1D	Stock price	NSGA-I, GDE3, SMPSO, SPEA2	Return=2.17 Risk=0.02
Creamer (2015)	US stock market	-	News sentiment, Accounting, Earnings forecast	BL_ML	Sharpe ratio = 6.56
Simos, Mourtas, and Katsikis (2021)	MNI140, MRG17, AMGN and NFLX	-	Stock price	BL_neuronet	R Square =0.89 MAE =139.04 RMSE =176.29

In all the studies reviewed here, investor view generation in BL modeling through the use of stock price forecasts based on artificial intelligence technology can achieve better portfolio performance.

3. Methodology

In this section, we provide the novel BL_AI algorithm to build investment portfolio.

3.1 Research Design

The stages of the proposed methodology are summarized in Figure 2 below.

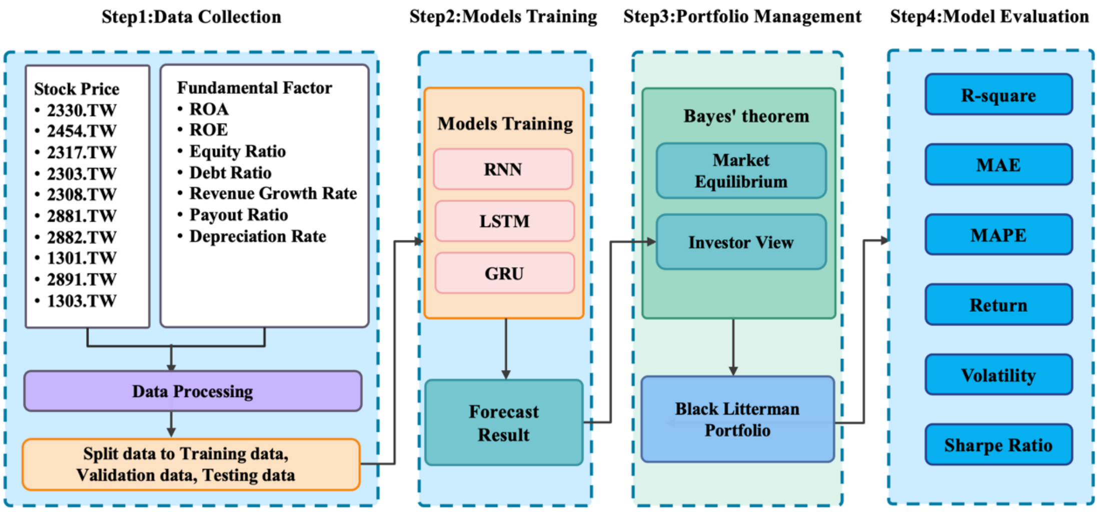


Figure 1 Research Framework Flow Chart

Step 1 Data Collection and Data Pre-processing

Yuanta/P-shares Taiwan Top 50 ETF's (0050.TW) price data were collected from YAHOO Finance. The Taiwan Economic Journal was used to obtain basic financial data. Before splitting the data into training and testing samples, the data were preprocessed for normalization and missing value imputation using the sliding window technique.

Step 2 Training stock price prediction model

Fundamental financial data and price data were used as features in prediction models based on RNN, LSTM, GRU to forecast stock prices.

Step 3 Portfolio Management

Using a hybrid approach based on the AI model to generating investor's view in BL model to address the portfolio optimization problem.

Step 4 Model Evaluation

(1) Artificial intelligence model was evaluated by R-squared, MAE, and MAPE to determine whether the model performance is satisfactory.

(2) BL_AI models was evaluated by Return, Standard deviation, and Sharpe ratio to measure a portfolio's performance.

3.2 Data Collection

The proposed model is applied in Taiwan stock markets: Yuanta/P-shares Taiwan Top 50 ETF.

3.2.1 Market Data

The asset universe consists of the top 10 stocks in Yuanta/P-shares Taiwan Top 50 ETF (0050.TW). We use data from January 2011 to December 2021 from Yahoo Finance. The dataset includes closing price as attributes for the following companies: Taiwan Semiconductor Manufacturing Company Limited (2330.TW), MediaTek Inc. (2454.TW), Hon Hai Precision Industry Co., Ltd. (2317.TW), United Microelectronics Corporation (2303.TW), Delta Electronics, Inc. (2308.TW), Fubon Financial Holding Co., Ltd. (2881.TW), Cathay Financial Holding Co., Ltd. (2882.TW), Formosa Plastics Corporation (1301.TW), CTBC Financial Holding Co., Ltd. (2891.TW), Nan Ya Plastics Corporation (1303.TW).

3.2.2 Financial Factor Data

An approach to figuring out a stock's intrinsic worth is through fundamental analysis. From the Taiwan Economic Journal database, we use the following financial factors: ROA (quarterly report), ROE (quarterly report), equity ratio (quarterly report), debt ratio (quarterly report), revenue growth rate (quarterly report), depreciation rate (quarterly report), and payout ratio (annual report).

All the data available including daily stock price and financial factor data.

3.3 Data Preprocessing

We proposed data preprocessing consists of two main stages: data normalization and split data to training set and test set.

3.3.1 Data normalization

During data processing, the missing value is removed after converting the various frequencies of time series data to a daily frequency. Data processing should be dealt with before data normalization. When data is rescaled using the Min-Max Normalization method, the feature's minimum value is converted to a 0 and its maximum value to a 1.

3.3.2 Split data to training set and test set

Using the rolling window method, the sample data is divided into training, validation, and testing sets (Selvin, Vinayakumar, Gopalakrishnan, Menon, & Soman, 2017). From 2011/01/01 to 2022/12/31, we chose training period to be the previous 4 years and the validation period to be the following 30 days. We established the 3-month sliding window as the warehouse transfer rule in accordance with finance sector custom. Fig. 2 depicts the prediction process in detail.

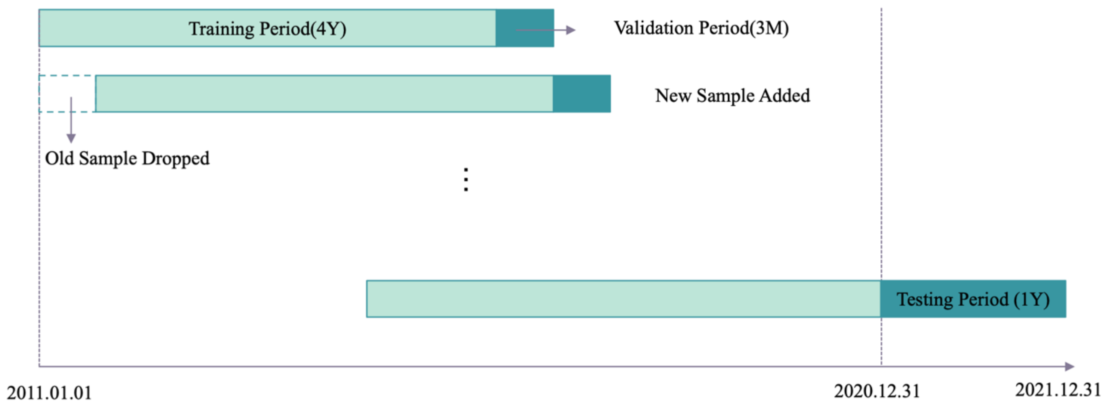


Figure 2 Scrolling Moving Sliding Window

3.4 Building Deep Learning Model

Deep learning models like RNN, LSTM, GRU were developed to forecast stock price.

3.5 Portfolio Management

In this study, we analyze different types of ensembles algorithms in portfolio optimization framework. The asset allocation benchmark are equally weighted portfolios (1/N) and Markowitz portfolios (Markowitz). Artificial intelligence including RNN, LSTM, GRU models is built to forecast stock price, which will be used as the inputs for the AI portfolio optimization framework. Black-Litterman model combine the views generated by artificial intelligence algorithm with the current evolution of the market (BL_AI). All asset allocation models are summarized in Table 3.

Table 3 Overview of asset allocation models

Number	Model	Abbreviation
Asset allocation benchmark		
1	1/N weight equally with N assets	1/N
2	Minimum-variance portfolio	MV
Portfolio optimization based on AI's prediction		AI
3	Portfolio optimization based on RNN framework prediction	RNN
4	Portfolio optimization based on LSTM framework prediction	LSTM
5	Portfolio optimization based on GRU framework prediction	GRU
Black–Litterman with AI series model's prediction as investor's view		BL_AI
6	Black–Litterman with RNN framework prediction as investor's view	BL_ RNN
7	Black–Litterman with LSTM framework prediction as investor's view	BL_ LSTM
8	Black–Litterman with GRU framework prediction as investor's view	BL_ GRU

3.5.1 Naive diversification rules

We create an equally weighted benchmark portfolio using a buy-and-hold approach without any rebalancing.

3.5.2 Mean-variance (MV) optimized portfolio

Investor optimizes a trade-off between risk and return (Markowitz, 1952). This framework assumes normally distribute which μ is the vector of expected return estimates, $\boldsymbol{\omega}$ is the weight of each asset, and δ is the coefficient of risk aversion.

$$\boldsymbol{max\ U = \omega'\mu - \frac{\delta}{2}\,\omega' \sum \omega}$$

3.5.3 Black–Litterman portfolio

The BL model combines two sources of information: implied returns referred to as Π and subjective return also referred to as views expressed in the vector Q. The combined return estimates are written as

$$\boldsymbol{\mu_{BL}} = [(\boldsymbol{\tau}\Sigma)^{-1} + \boldsymbol{P}'\boldsymbol{\Omega}^{-1}\boldsymbol{P}]^{-1}[(\boldsymbol{\tau}\Sigma)^{-1}\boldsymbol{\Pi} + \boldsymbol{P}'\boldsymbol{\Omega}^{-1}\boldsymbol{Q}]$$

in which τ is a scalar that measures the reliability of implied return estimates, P is a matrix that contains the information for which asset a subjective return estimate, Ω is a diagonal covariance matrix that measures the investor opinions uncertainty., Π is the implied equilibrium return vector, and Q expressed the investor's views of expected returns.

3.6 Evaluation Method

We first demonstrate the accuracy measurements with R-squared, MAE, MAPE to judge the model predictive performance. Then, we reveal the portfolio performance with log return, standard deviation and Sharpe ratio of each model.

3.6.1 Model Evaluation

(1) Coefficient of determination ($\boldsymbol{R^2}$)

Coefficient of determination is proportion of the variation of the dependent variable that can be explained by the independent variable, which is the explanatory power of the model.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$SS_{res} = \sum_{i=1}^{n}[f_i - y_i]^2,\ SS_{tot} = \sum_{i=1}^{n}[y_i - \bar{y}]^2,$$

n=number of data, y=true value, f=predicted value

(2) Mean Absolute Error (MAE)

Mean absolute error is a measure of errors between true value and predicted value.

$$\text{MAE} = \frac{1}{N} * \sum_{i=1}^{n}[f_i - y_i]$$

N=number of data, y=true value, f=predicted value

(3) Mean absolute percentage error (MAPE)

Mean absolute percentage error is a measure of errors between true value and predicted value as a percentage.

$$\text{MAPE} = \frac{1}{N} * \frac{\sum_{i=1}^{n} [f_i - y_i]}{y_i}$$

N=number of data, y=true value, f=predicted value

3.6.2 Portfolio Performance

(1) Log Return

Log return comprises any change in value of the investment.

$$R_i = \log{(P_t - P_{t-1})} = \log{(P_t)} - \log{(P_{t-1})}$$

P_t=stock price of the current day; P_{t-1}=stock price of the previous day

(2) Standard Deviation

Standard deviation is a measure of the amount of variation of a set of values.

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}$$

x_i=daily rate of return, μ=average rate of return

(3) Sharpe Ratio

The ratio is the average return earned in excess of the risk-free rate per unit of volatility.

$$\text{SR} = \frac{R_i - R_f}{\sigma_i}$$

R_i=rate of return, R_f=risk-free rate, σ_i=standard deviation

4. Results of the empirical application

To support our methodology for portfolio optimization, we split our experiments into two sections. In the first experiment, rolling window AI models will be created to forecast stock price. By comparing the models with various frequency attributes, we can determine which model is more accurate at forecasting. In the second experiment, we will change the weights of the chosen stocks to improve the performance of the portfolio. We present a hybrid framework to management portfolio using the BL_AI model, which combining with financial forecasting techniques and market equilibrium. We also back tested by comparing the results with those from other portfolio optimization frameworks.

4.1. Experiment 1: Rolling model with deep learning algorithm

In this study, we assess the efficacy of the forecasting model using several frequency features. The closing price of the stock is the daily frequency data. The following features are reported on a quarterly basis: ROA, ROE, Equity ratio, Debt ratio, Revenue growth rate, Depreciation rate. The Payout Ratio is annual frequency statistics.

4.1.1 Experiment 1: Rolling RNN Model Result

Table 4 presents the experimental data on RNN forecasting models with various frequency features. What stands out in the table is average R-square for the RNN model with daily feature is 87.02 percent, and the average MAPE is 2.59 percent. Average R-square dropped to 86.93 percent in the RNN model with daily and quarterly features, but average MAPE rose to 2.61 percent. The average R-square dropped to 62.60 percent and the average MAPE rose to 5.77 percent in the RNN model with daily, quarterly, and annual features.

Table 4 Rolling RNN Model with different frequency feature

Panel A: Rolling RNN Model with daily feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	98.51%	6.67	3.20%	94.21%	9.44	2.84%

2454	97.49%	12.49	4.23%	95.55%	15.67	3.39%
2317	97.55%	2.27	2.39%	77.32%	2.40	3.01%
2303	97.22%	0.28	1.98%	95.77%	0.42	2.37%
2308	93.00%	4.52	3.05%	92.52%	4.93	3.16%
2881	93.78%	0.98	2.28%	71.80%	0.99	2.58%
2882	96.65%	0.82	1.83%	81.54%	0.67	1.70%
1301	94.49%	2.33	2.46%	91.24%	1.97	2.21%
2891	94.59%	0.39	2.06%	82.77%	0.49	2.39%
1303	95.34%	1.30	1.95%	87.48%	1.41	2.28%
Average	95.86%	3.20	2.54%	**87.02%**	3.84	**2.59%**

Panel B: Rolling RNN Model with daily and quarterly feature

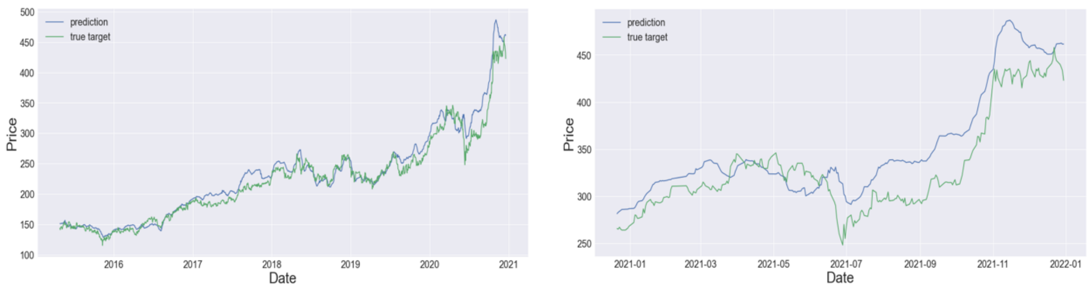
	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	98.70%	5.87	2.72%	94.66%	9.13	2.79%
2454	98.64%	8.37	2.68%	96.86%	13.23	2.89%
2317	98.28%	1.98	2.16%	80.95%	2.49	3.16%
2303	95.57%	0.38	2.71%	93.56%	0.53	3.03%
2308	91.91%	4.95	3.18%	92.28%	4.89	3.19%
2881	94.45%	0.93	2.16%	78.28%	0.87	2.27%
2882	95.29%	0.96	2.12%	81.58%	0.62	1.58%
1301	97.34%	1.64	1.90%	93.23%	1.50	1.80%
2891	91.17%	0.51	2.75%	70.60%	0.65	3.19%
1303	94.79%	1.40	2.06%	87.28%	1.39	2.25%
Average	95.61%	2.70	2.44%	86.93%	3.53	2.61%

Panel C: Rolling RNN Model with daily, quarterly and annually feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE

2330	94.22%	11.87	5.04%	66.51%	26.48	8.16%
2454	80.68%	31.58	9.51%	74.31%	47.02	11.19%
2317	92.09%	4.32	4.71%	40.76%	3.99	5.18%
2303	56.00%	1.16	7.81%	14.13%	2.09	11.41%
2308	86.04%	6.06	4.17%	81.17%	8.33	5.71%
2881	83.72%	1.57	3.51%	78.04%	0.88	2.23%
2882	94.34%	1.06	2.31%	80.63%	0.61	1.54%
1301	92.16%	2.68	2.95%	66.09%	4.00	4.58%
2891	86.60%	0.61	3.18%	48.28%	0.88	4.24%
1303	87.52%	2.19	3.16%	76.08%	2.19	3.43%
Average	85.34%	6.31	4.64%	62.60%	9.65	5.77%

In Fig. 3, the forecasting ability of the RNN models on 2330.TW are depicted.



(a) Training set output of 2330.tw (b)Testing set output of 2330.tw

Figure 3 The actual and forecasting prices of RNN models

4.1.2 Experiment 1: Rolling LSTM Model Result

Table 5 presents the experimental data on LSTM forecasting models with various frequency features. What stands out in the table is the average R-square for the LSTM model with daily feature is 76.57 percent, and the average MAPE is 3.49 percent. Average R-square dropped to 64.79 percent in the LSTM model with daily and quarterly features, and the average MAPE rose to 4.12 percent. The average R-square is reduced to 58.32 percent and the average MAPE is raised to 5.94 percent in the LSTM model with daily, quarterly, and annual features.

Table 5 Rolling LSTM Model with different frequency feature

Panel A: Rolling LSTM Model with daily feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	96.18%	10.96	4.99%	84.18%	17.37	5.36%
2454	98.24%	9.32	2.85%	95.39%	16.77	3.61%
2317	97.61%	2.27	2.49%	74.86%	2.65	3.40%
2303	78.60%	0.52	3.34%	36.50%	1.35	6.97%
2308	91.65%	4.92	3.35%	91.73%	5.02	3.25%
2881	90.12%	1.22	2.80%	70.39%	0.87	2.33%
2882	94.86%	1.00	2.20%	80.42%	0.62	1.57%
1301	94.85%	2.20	2.53%	80.30%	2.40	2.89%
2891	91.86%	0.46	2.44%	69.12%	0.64	3.15%
1303	94.25%	1.40	2.09%	82.80%	1.44	2.36%
Average	92.82%	3.43	2.91%	**76.57%**	4.91	**3.49%**

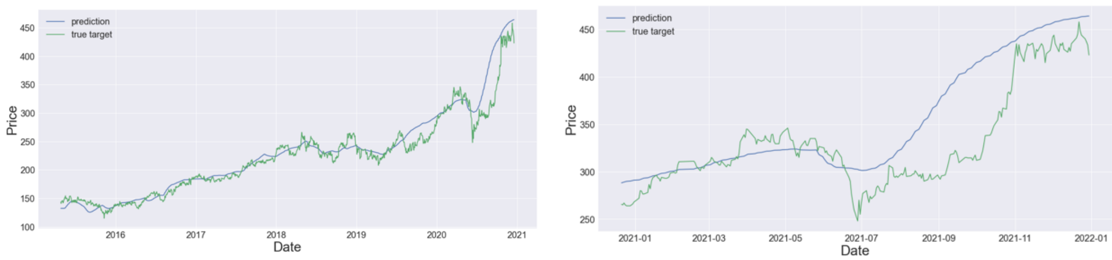
Panel B: Rolling LSTM Model with daily and quarterly feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	94.74%	12.31	5.47%	76.05%	21.13	6.55%
2454	95.75%	17.23	5.45%	91.32%	25.96	5.66%
2317	97.24%	2.43	2.52%	81.97%	2.13	2.74%
2303	90.67%	0.45	2.97%	79.09%	0.87	4.71%
2308	86.74%	6.34	4.33%	91.20%	5.47	3.68%
2881	73.50%	2.15	5.05%	-61.87%	2.74	7.06%
2882	94.95%	0.97	2.25%	60.30%	0.97	2.47%
1301	95.64%	2.11	2.48%	83.91%	2.29	2.76%
2891	88.05%	0.58	3.18%	66.83%	0.57	2.95%
1303	93.50%	1.46	2.20%	79.10%	1.59	2.63%
Average	91.08%	4.60	3.59%	64.79%	6.37	4.12%

Panel C: Rolling LSTM Model with daily, quarterly and annually feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	91.44%	12.46	5.29%	42.64%	28.82	8.97%
2454	85.72%	28.55	8.99%	78.51%	41.44	9.84%
2317	92.44%	4.32	4.39%	48.08%	3.79	4.72%
2303	52.67%	1.27	8.89%	16.92%	2.18	12.30%
2308	79.14%	7.75	5.31%	73.75%	9.31	6.61%
2881	76.23%	2.00	4.53%	63.81%	1.17	2.99%
2882	91.73%	1.30	2.81%	71.84%	0.79	1.98%
1301	90.94%	2.86	3.21%	75.43%	3.08	3.65%
2891	83.35%	0.66	3.50%	46.45%	0.89	4.41%
1303	83.72%	2.44	3.55%	65.81%	2.49	3.92%
Average	82.74%	6.36	5.05%	58.32%	9.39	5.94%

In Fig. 4, the forecasting ability of the LSTM models on 2330.TW are depicted.



(a) Training set output of 2330.tw (b) Testing set output of 2330.tw

Figure 4 The actual and forecasting prices of LSTM models

4.1.3 Experiment 1: Rolling GRU Model Result

Table 6 presents the experimental data on GRU forecasting models with various frequency features. What stands out in the table is the average R-square for an RNN model with a daily feature is 84.11 percent, and the average MAPE is 3.55 percent. Average R-square rose to 88.77 percent in the GRU model with daily and quarterly features, whereas average MAPE fell to 2.58 percent. The average R-square dropped to 75.03 percent in the GRU model with daily, quarterly, and annual features, but the average MAPE rose to 4.39 percent.

Table 6 Rolling GRU Model with different frequency feature

Panel A: Rolling GRU Model with daily feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	97.34%	9.37	4.28%	89.04%	14.80	4.53%
2454	97.33%	11.59	3.22%	92.11%	25.44	5.36%
2317	95.58%	3.18	3.03%	92.94%	1.30	1.64%
2303	68.11%	1.29	9.15%	65.04%	1.62	9.65%
2308	87.08%	6.46	4.05%	89.48%	5.50	3.32%
2881	93.77%	1.00	2.37%	70.48%	1.03	2.70%
2882	96.56%	0.83	1.89%	76.55%	0.77	1.97%
1301	97.40%	1.56	1.76%	93.32%	1.48	1.78%
2891	95.00%	0.37	2.00%	83.66%	0.48	2.32%
1303	92.40%	1.81	2.52%	88.51%	1.41	2.20%
Average	92.06%	3.75	3.43%	84.11%	5.38	3.55%

Panel B: Rolling GRU Model with daily and quarterly feature

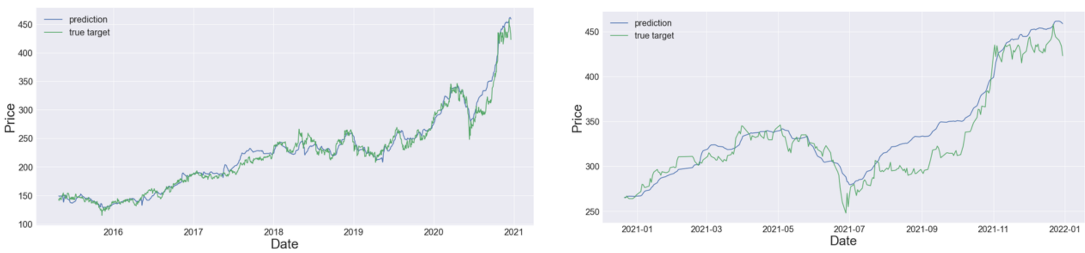
	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	99.34%	4.01	1.87%	97.14%	6.10	1.86%
2454	96.13%	18.29	6.50%	96.82%	15.43	3.53%
2317	96.87%	2.57	2.42%	94.97%	1.05	1.32%
2303	89.35%	0.68	4.73%	86.28%	0.94	5.51%
2308	93.52%	4.33	2.98%	93.00%	4.91	3.20%
2881	92.37%	1.08	2.42%	86.49%	0.60	1.59%

2882	95.80%	0.91	1.99%	82.20%	0.60	1.53%
1301	97.36%	1.61	1.86%	92.02%	1.58	1.91%
2891	91.48%	0.51	2.54%	69.38%	0.71	3.36%
1303	95.62%	1.28	1.88%	89.37%	1.22	2.00%
Average	94.78%	3.53	2.92%	**88.77%**	3.31	**2.58%**

Panel C: Rolling GRU Model with daily, quarterly and annually feature

	Train			Test		
Stock_ID	R^2	MAE	MAPE	R^2	MAE	MAPE
2330	97.35%	8.13	3.60%	88.24%	14.17	4.40%
2454	92.58%	21.28	6.79%	88.63%	31.00	7.32%
2317	93.08%	4.19	4.27%	56.80%	3.71	4.62%
2303	70.15%	1.06	7.95%	63.75%	1.50	9.29%
2308	85.22%	6.73	4.38%	92.47%	5.51	3.65%
2881	82.02%	1.63	3.67%	70.68%	0.90	2.36%
2882	91.71%	1.34	2.97%	73.49%	0.76	1.94%
1301	92.14%	2.67	2.91%	77.30%	3.23	3.66%
2891	91.75%	0.48	2.51%	65.69%	0.67	3.26%
1303	88.19%	2.16	3.17%	73.20%	2.17	3.44%
Average	88.42%	4.97	4.22%	75.03%	6.36	4.39%

In Fig. 5, the forecasting ability of the GRU models on 2330.TW are depicted.



(a) Training set output of 2330.tw (b) Testing set output of 2330.tw

Figure 5 The actual and forecasting prices of GRU models

4.1.4 Rolling Deep Learning Model Summary

What is interesting about the data in this table 7 is that GRU model employing intraday and quarterly features had the best R-squared performance by comparing the average R-squared evaluation of individual stocks with various frequencies features.

Table 7 R-squared evaluation with different frequencies features

	Train			Test		
Frequency	RNN	LSTM	GRU	RNN	LSTM	GRU
Daily	**95.86%**	**92.82%**	92.06%	**87.02%**	**76.57%**	84.11%
Daily+ Quarterly	95.61%	91.08%	**94.78%**	86.93%	64.79%	**88.77%**
Daily+ Quarterly+ Annually	85.34%	82.74%	88.42%	62.60%	58.32%	75.03%

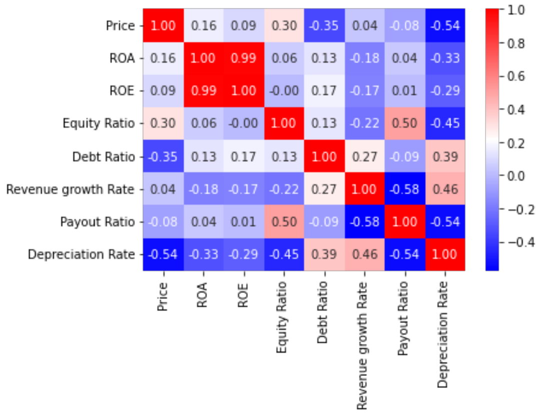
Some unanticipated finding in Table 8 revealed that firms utilizing AI models to forecast stock prices, such as Hon Hai Precision Industry Co., Ltd (2317.tw) and United Microelectronics Co., Ltd (2303.tw), tended to have R-squared values below 50%.

Table 8 Average R-squared evaluation of each stock

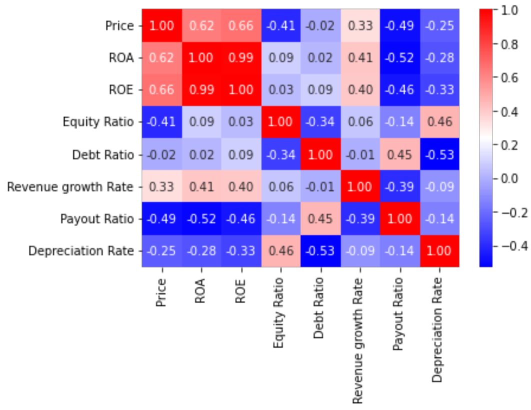
	2330	2454	2317	2303	2308	2881	2882	1301	2891	1303
Train	94.33%	86.33%	92.54%	59.60%	83.47%	80.66%	92.59%	91.75%	87.23%	86.48%
Test	65.80%	80.48%	**48.55%**	**31.60%**	82.46%	70.84%	75.32%	72.94%	53.48%	71.70%

Simple correlation analysis was used to determine the association between stock price and financial factors. In Fig.6(a) there is a clear correlational analysis of Hon Hai Precision Industry Co., Ltd.'s (2317.TW). The ROA, ROE, and revenue growth rate, which are indicators of the company's profitability and competitiveness, have a weakly positive link with the stock price. Depreciation rate, which measures the operating hazards of the company, has a moderately negative link with stock price.

In Fig.6(b) correlational analysis of United Microelectronics Co., Ltd.'s (2303.TW) demonstrate that the stock price and ROA and ROE, which measure the company's profitability, have a high positive correlation. The stock price and revenue growth rate, which measures competitiveness, have a moderate association.



(a) 2317.TW correlation analysis



(b) 2303.TW correlation analysis

Figure 6 correlational analysis on 2317.TW and 2303.TW

The summary MAPE performance of the RNN, LSTM, and GRU models is compared in Table 9. It is demonstrated that all models' prediction accuracy is comparable, with the GRU model with daily and quarterly features having a better level of predictability than the other two models.

Table 9 MAPE evaluation with different frequencies features

	Train			Test		
Frequency	RNN	LSTM	GRU	RNN	LSTM	GRU
Daily	2.54%	**2.91%**	3.43%	**2.59%**	**3.49%**	3.55%
Daily+ Quarterly	**2.44%**	3.59%	**2.92%**	2.61%	4.12%	**2.58%**
Daily+ Quarterly+ Annually	4.64%	5.05%	4.22%	5.77%	5.94%	4.39%

4.2 Experiment 2：Portfolio performance of asset allocation model

In this study, we use various optimization models to examine the performance of testing data. Equally weighted models (1/N) and Markowitz models (MV) serve as the benchmark portfolio for asset allocation. On the other hand, the portfolio optimization based on AI model's prediction used anticipated returns from RNN, LSTM, GRU algorithm as inputs. An artificial intelligence-based Black-Litterman model (BL_AI) combines market evolution and investor viewpoints generating from RNN, LSTM, GRU algorithm.

4.2.1 Experiment 2: Comparison of portfolio performance

In order to evaluate the performance of portfolios, we compare the top ten stocks (Top10), the bottom ten stocks (Down10), the top ten and bottom ten stocks (Top10+Down10), and all index constituent stocks (All) of Yuanta Taiwan 50ETF as investment targets (Targets), using benchmark model, AI model, and BL_AI models, respectively. The preliminary analysis of portfolio performance is presented in Tables 10 to 13.

Table 10 Comparison of portfolio performance with the Top10 stocks

Targets : Top10

Panel A: Annualized Portfolio Return

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	107.54%	107.54%	107.54%	7.78%	7.78%	7.78%	107.54%	40.02%
Q2	6.20%	58.41%	6.20%	0.00%	-4.20%	0.00%	58.41%	6.20%
Q3	33.57%	33.57%	33.67%	-16.67%	-16.67%	-16.67%	35.97%	0.86%
Q4	38.84%	32.45%	44.20%	5.21%	5.21%	5.21%	75.81%	14.04%
Average	46.54%	57.99%	47.90%	-0.92%	-1.97%	-0.92%	69.43%	15.28%

Panel B: Annualized Standard Deviation

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	41.10%	41.10%	41.10%	47.18%	47.18%	47.18%	41.10%	22.07%
Q2	24.72%	39.72%	24.72%	0.00%	22.61%	0.00%	39.72%	24.72%
Q3	28.92%	28.92%	29.13%	32.29%	32.29%	32.29%	40.84%	17.95%
Q4	12.99%	12.69%	14.58%	31.58%	31.58%	31.58%	33.13%	12.32%
Average	26.93%	30.61%	27.38%	27.76%	33.41%	27.76%	38.70%	19.27%

Panel C: Annualized Sharpe Ratio

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	2.62	2.62	2.62	0.16	2.71	0.00	2.62	1.81

Q2	0.25	1.47	0.25	-0.61	-1.02	0.00	1.47	0.25
Q3	1.16	1.16	1.16	-0.52	-4.15	-4.15	0.88	0.05
Q4	2.99	2.56	3.03	0.16	0.00	0.00	2.29	1.14
Average	1.75	**1.95**	1.76	-0.20	-0.61	-1.04	1.81	0.81

Table 11 Comparison of portfolio performance with the Down10 stocks

Targets : Down10

Panel A: Annualized Portfolio Return

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	269.96%	269.96%	269.96%	269.96%	269.96%	-20.99%	269.96%	38.83%
Q2	294.44%	294.44%	294.44%	294.44%	294.44%	294.44%	294.44%	41.24%
Q3	7.08%	48.38%	48.38%	-1.79%	48.38%	-3.15%	48.38%	-0.20%
Q4	23.19%	24.83%	23.22%	72.21%	72.21%	72.21%	72.21%	22.23%
Average	148.67%	159.40%	159.00%	158.70%	171.25%	85.63%	171.25%	25.52%

Panel B: Annualized Standard Deviation

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	58.69%	58.69%	58.69%	58.69%	58.69%	57.83%	58.69%	21.23%
Q2	63.37%	63.37%	63.37%	63.37%	63.37%	63.37%	63.37%	21.93%
Q3	5.02%	33.03%	33.03%	38.41%	33.03%	65.13%	33.03%	18.82%
Q4	8.39%	8.36%	8.48%	49.12%	49.12%	49.12%	49.12%	15.07%
Average	33.87%	40.86%	40.89%	52.40%	51.05%	58.86%	51.05%	19.26%

Panel C: Annualized Sharpe Ratio

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	4.60	4.60	4.60	4.60	4.60	-0.36	4.60	1.83
Q2	4.65	4.65	4.65	4.65	4.65	4.65	4.65	1.88
Q3	1.41	1.46	1.46	-0.05	1.46	-0.05	1.46	-0.01
Q4	2.77	2.97	2.74	1.47	1.47	1.47	1.47	1.48
Average	3.36	**3.42**	3.36	2.67	3.05	1.43	3.05	1.29

Table 12 Comparison of portfolio performance with the Top10+Down10 stocks

Targets : Top10 + Down10

Panel A: Annualized Portfolio Return

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	269.96%	269.96%	269.96%	7.78%	269.96%	7.78%	269.96%	90.92%
Q2	294.44%	294.44%	294.44%	294.44%	294.44%	294.44%	294.44%	46.97%
Q3	48.38%	48.38%	47.17%	-22.16%	-3.15%	-3.15%	48.38%	-0.52%
Q4	38.55%	47.73%	32.85%	72.21%	5.21%	5.21%	78.32%	38.70%
Average	162.83%	165.13%	161.11%	88.07%	141.62%	76.07%	172.77%	44.02%

Panel B: Annualized Standard Deviation

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	58.69%	58.69%	58.69%	47.18%	58.69%	47.18%	58.69%	40.59%
Q2	63.37%	63.37%	63.37%	63.37%	63.37%	63.37%	63.37%	43.74%
Q3	33.03%	33.03%	29.51%	24.53%	65.13%	65.13%	33.03%	33.95%
Q4	11.34%	12.89%	10.64%	49.12%	31.58%	31.58%	26.48%	23.92%
Average	41.60%	41.99%	40.55%	46.05%	54.69%	51.81%	45.39%	35.55%

Panel C: Annualized Sharpe Ratio

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	4.60	4.60	4.60	0.16	4.60	0.16	4.60	2.24
Q2	4.65	4.65	4.65	4.65	4.65	4.65	4.65	1.07
Q3	1.46	1.46	1.60	-0.90	-0.05	-0.05	1.46	-0.02
Q4	3.40	3.70	3.09	1.47	0.16	0.16	2.96	1.62
Average	3.53	**3.60**	3.48	1.34	2.34	1.23	3.42	1.23

Table 13 Comparison of portfolio performance with the All stocks

Targets : All

Panel A: Annualized Portfolio Return

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	269.96%	258.63%	261.00%	0.80%	269.96%	120.92%	269.96%	297.56%
Q2	2593.60%	2593.60%	2593.60%	152.27%	294.44%	294.44%	2593.60%	256.33%
Q3	45.50%	45.48%	45.54%	-3.15%	-3.15%	-3.15%	48.38%	-60.20%
Q4	204.27%	204.27%	204.27%	31.44%	72.21%	72.21%	204.27%	206.61%
Average	778.33%	775.49%	776.10%	45.34%	158.37%	121.11%	779.05%	175.08%

Panel B: Annualized Standard Deviation

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	58.69%	50.73%	52.10%	7.43%	8.14%	8.14%	58.69%	93.22%
Q2	99.06%	99.06%	99.06%	8.37%	8.37%	8.37%	99.06%	116.90%
Q3	24.04%	24.08%	23.95%	4.15%	6.28%	6.28%	33.03%	83.19%
Q4	54.07%	54.07%	54.07%	0.00%	0.00%	0.00%	54.07%	58.10%
Average	58.97%	56.99%	57.30%	4.99%	5.70%	5.70%	61.21%	87.85%

Panel C: Annualized Sharpe Ratio

Model	BL_AI series models			AI series models			Benchmark	
	BL_RNN	BL_LSTM	BL_GRU	RNN	LSTM	GRU	MV	1/N
Q1	4.60	5.10	5.01	0.11	33.15	14.85	4.60	3.19
Q2	26.18	26.18	26.18	18.18	35.16	35.16	26.18	2.19
Q3	1.89	1.89	1.90	-0.76	-0.50	-0.50	1.46	-0.72
Q4	3.78	3.78	3.78	0.00	0.00	0.00	3.78	3.56
Average	9.11	9.24	9.22	4.38	**16.95**	12.38	9.01	2.05

4.2.2 BL_AI Series model summary

Comparing benchmark models and pure AI series models with BL_AI series models, the results of this study show that portfolio built by BL_AI model with forecasting return as investor views generated better performance. Further analysis showed that the BL_LSTM model attains higher levels of Sharpe Ratios in general. Figure 7 reveals the Sharpe Ratio comparison of all asset allocation models.

Rather remarkable result is that portfolio performance reveals significant differences with diverse investment targets. This result is somewhat counterintuitive, investing in all constituent stocks of Yuanta Taiwan 50ETF produced the highest Sharpe ratio, whereas portfolios with top ten stocks as investment targets had the worst Sharpe ratio

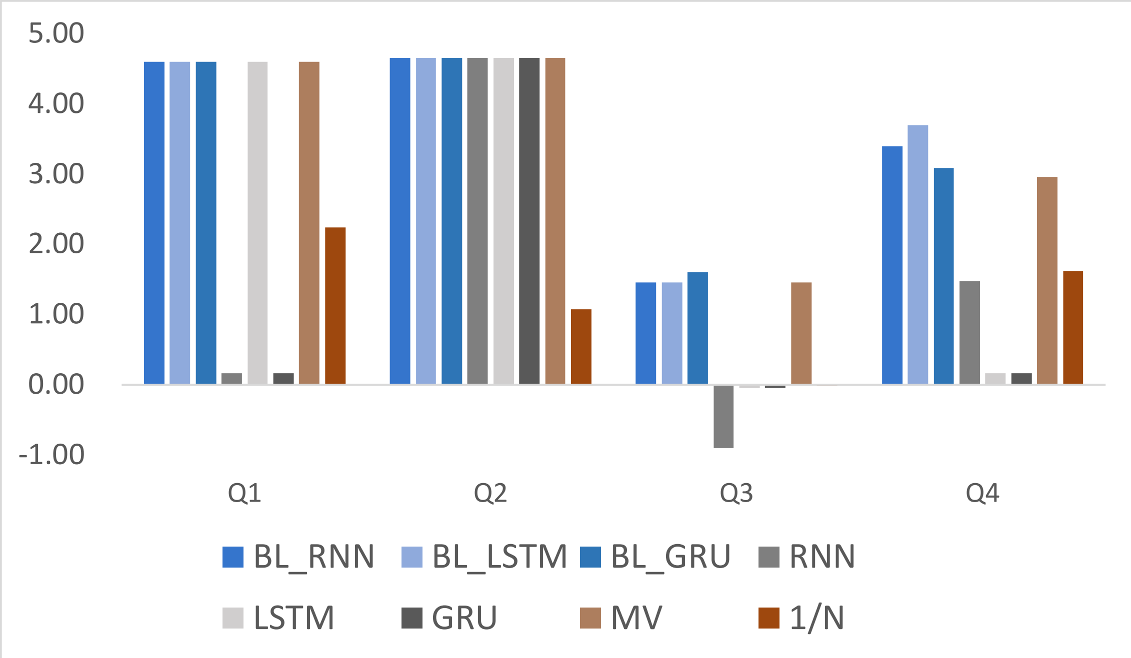


Figure 7 Comparison of portfolio performance with the Top10+Down10 stocks

4.3 Analysis and Discussion

Prior studies evaluate the outcomes of financial factors as features in AI models. We found that the average R-square of each rolling AI model with ROA, ROE, Equity ratio, Debt ratio, Revenue growth rate, Dividend payout ratio, and Depreciation rate as features is more than 60%, these results further support the idea of Alberg and Lipton (2017) who observing the price reflects the intrinsic value of the company. Another important finding was that combining daily frequency feature data and quarterly frequency feature data to forecast the daily stock price resulted in higher MAPE performances for all models.

The results of this study indicate that portfolio developed by the BL_AI model, achieve better risk-adjusted performance than pure AI models, it seem to be consistent with Ozbayoglu, Gudelek, and Sezer (2020) research which found that hybrid models are more preferred over the native or standalone models in most studies. Another significant discovery was that all BL_AI series model yield better Sharpe ratio than benchmark series models. This result may be explained by 1/N portfolios with strict investment weight policies, which cannot be dynamically adjusted when the market environment changes. In the Markowitz portfolio, the investments are overconcentrated, resulting in both high profitability and high volatility. In general, these findings indicate that the BL_AI series model attains higher levels of Sharpe ratio compared to other series models. The highest Sharpe ratio among the hybrid BL_AI series model is BL_LSTM, which supports evidence that the LSTM is more effective prediction approach at modeling time series data.

A note of caution is due here since other features need be considered in order to improve the model's accuracy. These findings also may be somewhat limited by using selected AI models to generate investor view. Further studies, which take these variables into account, will need to be undertaken.

5. Conclusions

This paper has argued that employ AI forecasting methodologies to produce investor views in BL portfolio models attains higher levels of Sharpe ratios.

The findings from this study make several contributions to the current literature. First, these experiments confirmed that AI model performed better with intraday and quarterly frequency features, which provided a better understanding of the impact of lagging features. Second, this study has been to confirm BL_AI framework outperforms comparable portfolio models with a greater Shape ratio, which indicates that the framework is still informative. A practitioner implication is that AI models provide fundamental investors to evaluation of corporate value based on the firm's financial factors, BL_AI models generating investor views from AI model in a volatile market environment can enhances portfolio performance.

Further research may focus on applying the framework on other features datasets and explore different aspects of AI models which consequently different investor's views with various asset allocation strategies.

6. References

Alberg, J., & Lipton, Z. C. (2017). Improving factor-based quantitative investing by forecasting company fundamentals. *arXiv preprint arXiv:1711.04837*.

Alexander, D., & Svetlana, D. (2013). Application of Ensemble learning for views generation in Meucci Portfolio Optimization Framework. *Review of Business and Economics Studies*(1), 100-110.

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies, 24*(1), 34-112.

Bessler, W., Opfer, H., & Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naïve diversification approaches. *The European Journal of Finance, 23*(1), 1-30.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial analysts journal, 48*(5), 28-43.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., & Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications, 55*, 194-211.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications, 112*, 353-371.

Coen, P., Gomme, E., & Kendall, M. G. (1969). Lagged relationships in economic forecasting. *Journal of the Royal Statistical Society. Series A (General), 132*(2), 133-163.

Creamer, G. G. (2015). Can a corporate network and news sentiment improve portfolio optimization using the Black–Litterman model? *Quantitative Finance, 15*(8), 1405-1416.

Fu, X., Du, J., Guo, Y., Liu, M., Dong, T., & Duan, X. (2018). A machine learning framework for stock selection. *arXiv preprint arXiv:1806.01743*.

Jiang, Z., & Liang, J. (2017). *Cryptocurrency portfolio management with deep reinforcement learning.* Paper presented at the 2017 Intelligent Systems Conference (IntelliSys).

Lee, S. I., & Yoo, S. J. (2020). Threshold-based portfolio: the role of the threshold and its applications. *The Journal of Supercomputing, 76*(10), 8040-8057.

Markowitz, H. (1952). The utility of wealth. *Journal of political Economy, 60*(2), 151-158.

Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. *Applied Soft Computing, 93*, 106384.

Quintana, D., & Isasi, P. (2013). Multiobjective algorithms with resampling for portfolio optimization. *Computing and Informatics, 32*(4), 777-796.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V. K., & Soman, K. (2017). *Stock price prediction using LSTM, RNN and CNN-sliding window model.* Paper presented at the 2017 international conference on advances in computing, communications and informatics (icacci).

Simos, T. E., Mourtas, S. D., & Katsikis, V. N. (2021). Time-varying Black–Litterman portfolio optimization using a bio-inspired approach and neuronets. *Applied Soft Computing, 112*, 107767.

Wang, B., Huang, H., & Wang, X. (2012). A novel text mining approach to financial time series forecasting. *Neurocomputing, 83*, 136-145.

Yaman, I., & Dalkılıç, T. E. (2021). A hybrid approach to cardinality constraint portfolio selection problem based on nonlinear neural network and genetic algorithm. *Expert Systems with Applications, 169*, 114517.

Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep learning for portfolio optimization. *The Journal of Financial Data Science, 2*(4), 8-20.

Zhou, B. (2019). Deep learning and the cross-section of stock returns: Neural networks combining price and fundamental information. *Available at SSRN 3179281*.