ISSN 2460-9056
http://socj.telkomuniversity.ac.id/ijoict/

Intl. Journal on ICT
Vol. 9, No. 1, Jun 2023. pp. 14-26
doi: doi.org/10.21108/ijoict.v9i1.698

Portfolio Optimization Based on Return Prediction and Semi Absolute Deviation (SAD)

Gharyni Nurkhair[1]*, Deni Saepudin[2], Aniq Atiqi[3]
[1,2,3]*School of Computing, Telkom University*
[12,3] *Telekomunikasi St., Terusan Buahbatu, Bandung, West Java, Indonesia 40257*

*[1]gharyninurkhair@student.telkomuniversity.ac.id, denisaepudin@telkomuniversity.ac.id, aniqatiqi@telkomuniversity.ac.id

Abstract

A portfolio is a collection of financial assets and investments managed by financial institutions, investment managers, or individuals. In investment activities, investors expect a minimum risk of loss in stock investments and, of course, the optimum stock portfolio weight to get maximum profit. Current stock price movements are difficult to predict; however, investors can monitor changes in the stock index value from time to time. These changes can be used as a measuring tool to compare the portfolio's performance. More knowledge is needed to make it easier for investors to monitor changes in the value of their stock index. This research has discussed how to build a portfolio based on stock datasets with the LQ45 index using return predictions from the artificial neural network (ANN) method with semi-absolute deviation (SAD). Furthermore, the portfolio is optimized by looking for weights that match it. After that, a comparison of portfolio performance was carried out using the Sharpe ratio (SR) method between the semi-absolute deviation (SAD) portfolio and the portfolio resulting from the formation of the equal weight (EW) portfolio. Portfolio performance with ANN prediction and SAD is better than equal weight portfolios in terms of mean return, standard deviation, and sharpe ratio for portfolios with few stocks, namely 2 and 3 stocks. In addition, a portfolio with a higher number of shares can make the portfolio value from the ANN close prediction algorithm process and the selection of weights based on SAD better than, portfolios with an equal weight of each list of stocks in the portfolio.

Keywords: Portfolio, Artificial Neural Network (ANN), Semi Absolute Deviation (SAD), Sharpe Ratio (SR)

I. INTRODUCTION

A portfolio is a grouping of financial assets and investments managed by financial institutions, investment managers, or individuals. The portfolio contains securities such as stocks, bonds, and mutual funds. The stocks selected must be following their risk profile and can also be monitored regularly.

In investment activities, investors expect a minimal risk of loss in stock investment and the right stock portfolio to get maximum profit. Current stock price movements are difficult to predict, but investors can monitor changes in the stock index's value from time to time. These changes can be used as a measuring tool to compare portfolio performance. More knowledge in this area is required to make it easier for investors to monitor changes in the value of their stock index.

This paper discusses building a portfolio using return predictions obtained from close price predictions through the Artificial Neural Network (ANN) algorithm. Furthermore, to make ANN predictions better, the method is updated by combining it with Semi Absolute Deviation (SAD) as an optimization variable to reduce

Received on 02 Feb 2023. Revised on 14 Feb 2023. Accepted and Published on 18 June 2023.

the size of portfolio risk. SAD can determine the proper weight for each stock in the portfolio and still maintain a high return from the portfolio using a linear programming solver. The train data set used is historical data on the LQ45 index stock with a series of data spanning eight years starting from April 30, 2012 - April 27, 2020. Furthermore, the test data used spans two years, or 106 weeks, starting from May 5, 2020 - May 5, 2022. The dataset used in this paper was obtained through the Yahoo finance website. The results of the optimization method are compared with the equal weight portfolio selection method through the indicators of mean return, standard deviation, and Sharpe ratio.

II. LITERATURE REVIEW

Many studies have implemented machine learning algorithms in portfolio selection. One of the most popular and widely used machine learning algorithms is the Artificial Neural Network (ANN). The ANN algorithm is implemented into several stock datasets and compared with other machine-learning and non-machine algorithms. Pe´rez-Rodrı´guez et al.[1] have conducted research on the Spanish Ibex-35 index stock regarding the prediction of stock returns using the Artificial Neural Network (ANN) model in the form of three models, namely MLP, JCN, and Elman networks. The three models were compared with the linear AR model, the ESTAR, and LSTAR smooth transition autoregressive models. Based on the results of research that has been carried out with the help of percentage filter ranges and trading costs, it is found that the Artificial Neural Network (ANN) model has gone one step further and represents an improvement in investment strategy. Not only that, but Selvamuthu et al. [2] (2019) also conducted research on stock returns using the Artificial Neural Network (ANN) model with an algorithm. The neural network is based on three different learning algorithms, namely, Levenberg-Marquardt (LM), Scaled Conjugate Gradient (SCG), and Bayesian Regularization (BR) for stocks in India. The study results show that the three algorithms sequentially have an accuracy value of 96.2%, 97.0%, and 98.9%. For prediction speed, SCG is the fastest but also compensated with a high error distance. These results show that the Artificial Neural Network (ANN) model is very good at predicting stock returns seen from its accuracy value.

In 2019, Nagaraj et al. [3] conducted a study using the Recurrent Neural Network (RNN) method with Long Short Term Memory (LSTM) to predict stock returns. The dataset used comes from the National Stock Exchange of India (NSE). Based on the results of this experiment, the results show that this model has the potential to outperform the accuracy of stock return predictions, with an MAE value of 0.1645 and an RMSE of 23.54.

Subsequent research was carried out with two classes of Artificial Neural Networks (ANN) combined with three classes of econometric predictors by Nametala et al. [4] with a stock dataset in Brazil. Analysis of this method with econometric predictors, ANN trend and the candlestick, Bovespa index, and Buy and Hold strategy) results in Shapiro-Wilk and Fligner-Killen evaluations obtaining satisfactory values, respectively, 0.85 and 0.75, which makes this method good.Stocks and Stock Portfolio.

Stocks are part of an interesting type of investment to choose from because they can get enormous profits compared to other types of investment [5]. In stock investment, there are 2 types of benefits that will be obtained, namely dividends and capital gains. The dividend is the distribution of profits to investors. Meanwhile, the capital gain is the difference from the selling price which is higher than the previous purchase price. In addition to getting profit, in stock investment, there is also a risk of loss that will be obtained by investors. To minimize this risk, investors must have a greater ability to see the latest stock movements. Meanwhile, a stock portfolio is a series of combinations of assets that are invested and held by stock owners. The purpose of doing this portfolio is to minimize the risk of loss in investment activities [6].

A. Stock returns

Stock returns are systematically represented as follows [7]:

$$r_{(n)} = \frac{S(n) - S(n-1)}{S(n-1)}, \quad (1)$$

Where:

r(n) : return value period n
S(n) : stock price period n
S(n - 1) : stock price period n-1

B. Return Portfolio Stock

The return value of stocks in the portfolio is called portfolio return. The portfolio return value can be formulated as follows[7]:

$$r_{pt} = \sum_{i=1}^{n} r_{in} W_{in}, \quad (2),$$

Where:

r_{pt} : portfolio returns
r_{in} : stock return to-i period to-n
W_{in} : the weight value of the i-th stocks

C. Artificial Neural Network (ANN)

An artificial neural network, or what is known as an "artificial neural network," is an artificial intelligence [8] method inspired by the human brain. To some extent, this artificial neural network is modeled after the structure of the human biological brain, and its structure and links can be used to solve machine learning and computer-based problems. The artificial neural network consists of abstract models that have relationships with neurons. Artificial neural networks have a structure and an operation to help simplify the problem. First of all, the abstract model of the neural network consists of neurons, where these neurons are usually referred to as nodes or units. These nodes can retrieve data from outside or from other neurons, which they can then transmit to other neurons or immediately issue as output, which is the final result.

This artificial neural network consists of three types of neurons: input neurons, hidden neurons, and output neurons. Input neurons function to receive data or information in the form of patterns or signals from the outside world. Hidden neurons are located between the input neurons and the output neurons. Hidden neurons function to map internal information patterns. Then, as a result, the output neurons carry out the delivery of information and signals that have been mapped to the hidden neurons to the outside world.

D. Multi-Layer Perceptron (MLP)

The multi-layer perceptron (MLP) was first introduced by S. Papert and M. Minsky in 1969. The Multi-Layer Perceptron (MLP) is an artificial neural network model, commonly called the Artificial Neural Network (ANN), which consists of 3 layers: one input layer, several hidden layers, and one output layer. Each layer's number of neurons and layers is a hyperparameter of the multi-layer perceptron (MLP) [9]. Each neuron in the hidden layer consists of an input, weight, and bias term. Each neuron also has a non-linear activation function that will provide a cumulative output based on the previous neuron.

Multi-Layer Perceptron (MLP) is often used to perform classification and regression. The multi-layer perceptron (MLP) will produce more accurate classification results because this model is a type of neural network modeling that has a much better weight value compared to other types of modeling [10].

E. Semi-Absolute Deviation (SAD)

The semi-absolute deviation (SAD) calculation is used to estimate how much loss an investor may experience [11].

F. Equal Weight Portfolio (EWP)

Equal Weight Portfolio (EWP) is an approach that is often used because the way the method works is easy and simple. The way this portfolio works is by giving equal weight to all stocks in the same portfolio. The formula is as follows:

$$EWP = \frac{1}{N}, \tag{3}$$

Where:
EWP : Equal Weight Portfolio
N : Number of stock data

G. Sharpe Ratio

To evaluate the formation of this portfolio, the Sharpe Ratio (SR) calculation method is used. The following is the calculation formula along with its explanation.

$$S_P = \frac{r_{pt(exp)}}{\sigma_P}, \tag{4}$$

Where:
$r_{pt(exp)}$: Expected return portfolio
σ_P : Standard deviation of portfolio return

III. RESEARCH METHOD

Forming a portfolio based on stock datasets with the LQ45 index using the Semi-Absolute Deviation (SAD) method with an Artificial Neural Network (ANN) After that, a portfolio performance comparison will be made using the Sharpe Ratio (SR) method between the Semi-Absolute Deviation (SAD) portfolio and the portfolio resulting from the formation of an Equal-Weight Portfolio (EWP). The dataset used is stock price data with the LQ45 index for 10 years, starting from May 8, 2012- May 8, 2022. Fig.1 show an overview of the system design flow.

A. Stock Dataset Input

The stock dataset used in this report is the LQ45 index stock with a span of 10 years, starting from May 8 2012 to May 8, 2022. This dataset is taken from the yahoo finance website.

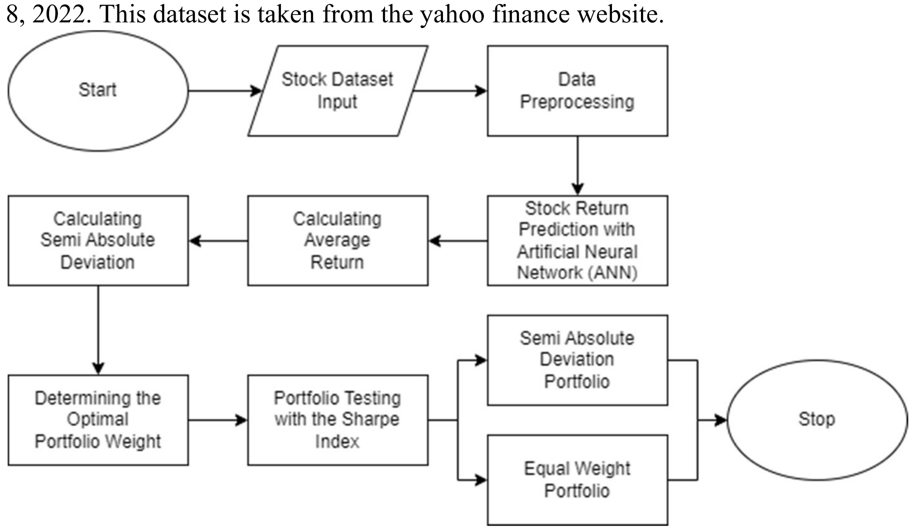

Fig 1. The system design

B. Data Preprocessing

Data preprocessing aims to produce quality data. Data preprocessing has a very large impact on data generalization performance [12]. The following is an explanation of the steps taken in the data preprocessing process. The data preprocessing process carried out in this study is data normalization. The data normalization process is carried out with the aim that the data distribution becomes normal. In each feature, there is often a difference between the maximum value and minimum value. Examples are the values 0.05 and 5000. When the data normalization process is carried out, the value scale will be scaled to a very low value.

C. Stock Return Prediction with Artificial Neural Network (ANN)

In predicting returns with an Artificial Neural Network (ANN), it is done using the Back Propagation (BP) algorithm with the following steps [13];

1. Setting up a training pattern
2. Setting up the Artificial Neural Network (ANN) model which consists of the number of input neurons, hidden neurons, and output neurons.
3. Initiate learning rate and momentum level.
4. Set the minimum error (E min).
5. Start training patterns one by one on the layer.
6. Error Backpropagation via the output layer and hidden layer.
7. Error Backprograte through the hidden layer by adjusting the weight of the stock portfolio.
8. Checking the error must be less than E min. If not, perform steps 5 through 8.
9. Obtain the return prediction results from the training pattern steps 5 to 8.

D. Calculating Average Return

After obtaining stock returns from the LQ45 stock portfolio using an Artificial Neural Network (ANN), the average stock return is calculated according to the number of stocks in the LQ45 stock index portfolio.

E. Calculating Semi-Absolute Deviation (SAD) as a Risk Measure

Following are the formulas and information used to perform Semi-Absolute Deviation (SAD) calculations[12].

$$w_t(X) = \frac{\left|\sum_{i=1}^{n}(r_i - r_{it})x_i\right| + \sum_{i=1}^{n}(r_i - r_{it})x_i}{2}, \quad (5)$$

Where:

w_t : Semi Absolute Deviation (SAD) t-period
X : portfolio-X
n : total assets
i : asset i
r_i : expected return asset i
r_{it} : expected return of assets to-i period t
x_i : the weight of each asset

Calculations to find out the total risk obtained using the Semi-Absolute Deviation (SAD) method. The following is the formula and information used to calculate the total risk.

$$R_{total}(X) = \frac{1}{T}\sum_{t=1}^{T} w_t(X)\ , \quad (6)$$

Where:

R_{total} : total risk

X : portfolio-X
T : calculated time period
w_t : Semi Absolute Deviation (SAD) t-period

F. Determining Portfolio Weights

To obtain the optimal portfolio weight, the formula below is used.

$$Min_{bobot} = \frac{1}{T}\sum_{t=1}^{T} d_t \quad , \tag{7}$$

Equation 9 above must fulfill the following:

$$d_t \geq 0 \; dan \; d_t \geq \sum_{i=1}^{n} (r_i - r_{it}) \, x_i \; , \tag{8}$$

So, it can be said that:

$$\sum_{i=1}^{n} r_i x_i \geq r_p \; , \tag{9}$$

Where:
T : calculated time period
d_t : decrease in the t-period Semi-Absolute Deviation (SAD).
r_i : return asset-i
r_{it} : return asset i-th period t
x_i : the weight of each asset
r_p : desired portfolio return

G. Portfolio Testing with the Sharpe Index

In the portfolio testing process, the aim is to compare portfolio performance using the Sharpe Ratio (SR) method between Artificial Neural Network (ANN), Semi-Absolute Deviation (SAD), and portfolios resulting from the formation of an equal weight portfolio (EWP). If the results obtained are of high value, the better portfolio performance. To evaluate the formation of this portfolio, the Sharpe Ratio (SR) calculation method is used.

IV. RESULTS AND DISCUSSION

A. Testing Scenario

In making close predictions using ANN, several steps are carried out as follows:

1. Train data span of 8 years or as much as 416 weeks starting from April 30, 2012 - April 27, 2020.
2. Test data with a span of 2 years or as much as 106 weeks starting from May 5, 2020 - May 5, 2022.
3. The ANN used has 3 hidden layers, the first to third input layers are sequentially 1024, 512, 256, epoch, and batch with a value of 100 and a validation split of 0.001.

Furthermore, the close results obtained with RMSE and MSE can be seen in Table 1:

TABLE 1.
RMSE AND MAE RESULTS

Stock	RMSE	MAE
BBCA	0.0959344	0.0165228
TLKM	0.0768817	**0.0116989**
BMRI	0.1080667	0.0169059
ASII	0.0843774	0.0149742
TPIA	0.0496641	0.0275145
UNVR	0.2254033	0.0169622
PGAS	**0.0303920**	0.0165210

Table 1 shows the RMSE and MAE values of each stock close prediction. RMSE and MAE are increasing well if it gets closer to zero. Stocks with the best close prediction evaluation seen from RMSE sequentially are PGAS, TPIA, TLKM, ASII, BBCA, BMRI, and UNVR. Meanwhile, if viewed from the MSE value, the best stock close predictions sequentially are TLKM, ASII, PGAS, BBCA, BMRI, UNVR, and TPIA.

In the test scenario, which is done with a combination of stocks based on market order LQ45 index stock portfolio capitalization. The combination of stocks is based on the market capitalization made contained in the following Table 2.

TABLE 2.
PORTFOLIO TEST LIST

Number of Stocks	Stock Name
2	BBCA, TLKM
3	BBCA,TLKM, UNVR
4	BBCA, TLKM, UNVR, ASII
5	BBCA, TLKM, UNVR, ASII, BMRI
6	BBCA, TLKM, UNVR, ASII, BMRI, TPIA
7	BBCA, TLKM, UNVR, ASII, BMRI, TPIA, PGAS

B. Test Result

Fig. 2 to Fig. 7 show a comparison of portfolio growth values.

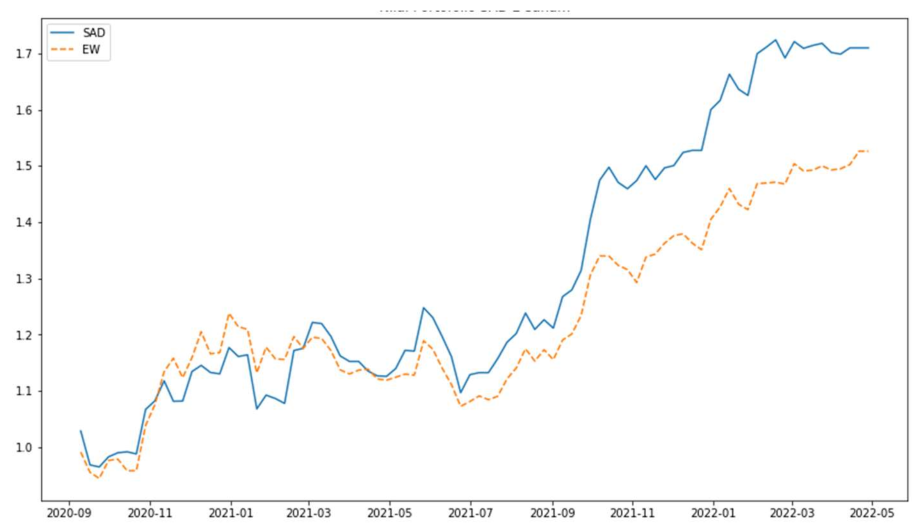

Fig. 2. Portfolio value of 2 stocks

Fig. 2, shows that the growth of the portfolio with predictions of close stock of ANN and weight selection through SAD results in higher portfolio growth compared to equal weight portfolio. For the movement of the value of the two portfolios that are built on 2 stocks with the two methods are relatively the same.

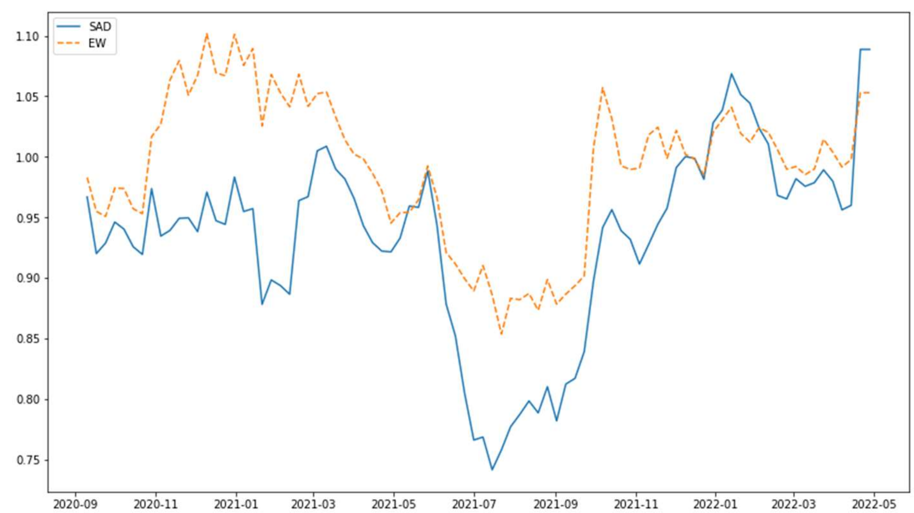

Fig. 3. Portfolio value of 3 stocks

Figure 3 shows that the growth of the portfolio with predictions of close stocks of ANN and weight selection through SAD resulted in portfolio growth which was still higher than with an equal weight portfolio. For movement in value, the portfolio with ANN and SAD is more extreme movement up or down compared to an equal weight portfolio.

Fig. 4. Portfolio value of 4 stocks

Fig. 4 shows that the growth of the portfolio with predictions of close stocks of ANN and weight selection through SAD resulted in portfolio growth which was still lower than with an equal weight portfolio. For movement in value, both portfolios with ANN and SAD as well as equal weight portfolios are relatively the same even though there are slightly different ones during the downward movement (more extreme SAD and ANN).

Fig. 5. Portfolio value of 5 stocks

Figure 5 shows that the growth of the portfolio with predictions of close stocks of ANN and the selection of weights through SAD results in relative lower than portfolio growth with an equal weight portfolio, but during the middle period it is better for equal weight portfolio growth. For the value movement, fine portfolios with ANN and SAD as well as equal weight portfolios are relatively the same although there are some differences during the downside movement in the middle period (more extreme SAD and ANN).

Fig. 6. Portfolio value of 6 stocks

Figure 6 shows that the growth of the portfolio with predictions of close stocks of ANN and weight selection through SAD results in lower portfolio growth compared to equal weight portfolio. For movement in value, portfolios with ANN and SAD are more extreme increase or decrease compared to an equal weight portfolio.

Fig. 7. Portfolio value of 7 stocks

Figure 7 shows that the growth of the portfolio with predictions of close stocks of ANN and selection of weights through SAD results in lower portfolio growth compared to equal weight portfolio. For the movement of the value of the two portfolios built by both methods is relatively the same.

Based on testing of making portfolios with various numbers of shares obtained different results that portfolios with a small number of shares can make a portfolio value from the algorithmic process created based on this report (prediction of close ANN and selection of weights based on SAD) is better than, portfolios with equal weight or equal weight of each list of stocks in the portfolio. Furthermore, for a higher number of stocks, the movement in the value of the portfolio will be more extreme, either when the value of a portfolio increases or decreases. The more stocks used in the portfolio can increase the profits derived from the portfolio. Portfolios with 2 and 3 shares can generate greater profits because stocks are selected based on a high level of capitalization and the performance of these stocks is indeed good.

TABLE 3.
PORTFOLIO PERFORMANCE EVALUATIONS

Portfolio	Methods	Evaluation		
		Mean Return	**Standard Deviation**	**Sharpe Ratio**
2 stock	**SAD**	0.0066303	0.0274756	0.2413170
	Equal Weight	0.0052275	0.0246640	0.2119517
3 stock	**SAD**	0.0015037	0.0323215	0.0465239
	Equal Weight	0.0009351	0.0261258	0.0357939
4 stock	**SAD**	-0.0043035	0.0305002	-0.1410989
	Equal Weight	0.0023229	0.0255951	0.09075789
5 stock	**SAD**	-0.0025826	0.0336443	-0.07676247
	Equal Weight	0.0031211	0.0254116	0.12282425
6 stock	**SAD**	0.0016616	0.0295322	0.05626617
	Equal Weight	0.0040960	0.0251034	0.16316772
7 stock	**SAD**	0.0025912	0.0402592	0.06436490
	Equal Weight	0.0044424	0.0279735	0.15880830

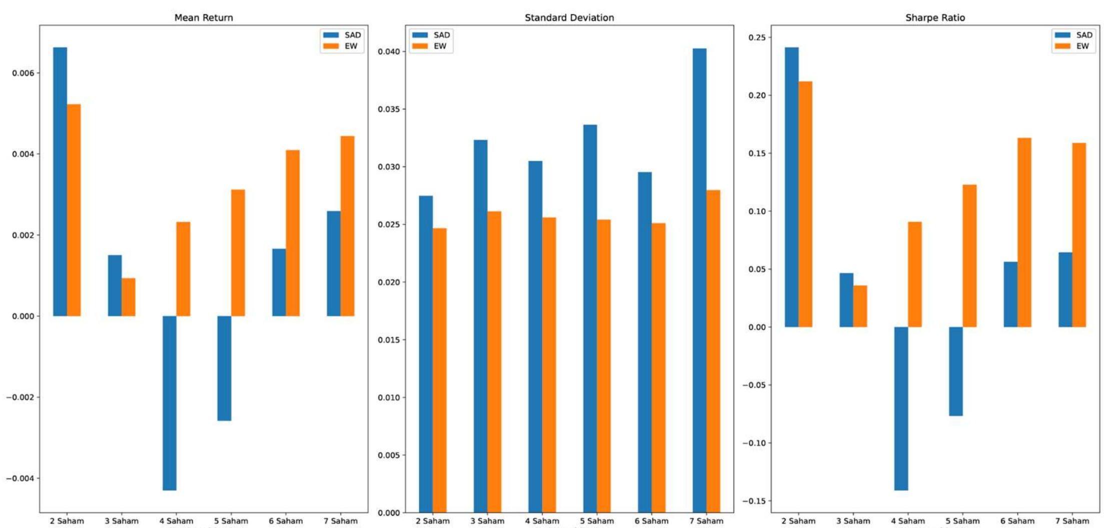

Fig. 8. Comparison of portfolio performance

Based on Table 3 and Fig. 8, overall testing of portfolio creation with various the number of stocks found that making a portfolio with SAD has a higher risk than Equal weight portfolio as seen from the standard deviation value which is larger. The less stock that is used in making a portfolio, affects portfolio performance. This can be seen from the mean side of higher returns and sharpe ratios than equal weight portfolios (portfolios with 2 and 3 stocks). Thing the opposite also applies, the more the amount in the portfolio, the mean return, and sharpe SAD portfolio ratio will be worse than the equal weight portfolio

V. CONCLUSION

From the results of the experiments conducted, several conclusions can be drawn as follows: The performance of the equal weight portfolio is better than the SAD portfolio, which can be seen based on the smaller standard deviation, which is 0.0246640 (the risks are also less). SAD portfolios that have the best performance are those with 2 stocks (number of stocks least) with a sharpe ratio value of 0.2413170 and also obtained the best mean return from testing in this portfolio is equal to 0.0066303.

VI. REFERENCES

[1] Pérez-Rodríguez, J. V., Torra, S., & Andrada-Félix, J. (2005). STAR and ANN models: Forecasting performance on the Spanish “Ibex-35” stock index. *Journal of Empirical Finance*, *12*(3), 490-509.

[2] Selvamuthu, D., Kumar, V., & Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. Financial Innovation, 5(1), 1-12.

[3] Naik, N., Mohan, B.R. (2019). Study of Stock Return Predictions Using Recurrent Neural Networks with LSTM. In: Macintyre, J., Iliadis, L., Maglogiannis, I., Jayne, C. (eds) Engineering Applications of Neural Networks. EANN 2019. Communications in Computer and Information Science, vol 1000. Springer, Cham.

[4] Nametala, C. A., Souza, J. V. D., Pimenta, A., & Carrano, E. G. (2022). Use of Econometric Predictors and Artificial Neural Networks for the Construction of Stock Market Investment Bots. *Computational Economics*, 1-31.

[5] Suyudi, M. A. D., Djamal, E. C., & Maspupah, A. (2019, August). Prediksi Harga stock menggunakan Metode Recurrent Neural Network. In Seminar Nasional Aplikasi Teknologi Informasi (SNATI).

[6] Saputro, R. N., & Badjra, I. B. (2016). Kinerja portofolio stock berdasarkan strategi investasi momentum pada industri manufaktur (Doctoral dissertation, Udayana University).

[7] Capinski, M., & Zastawniak, T. (2003). Mathematics for Finance: An Introduction to Financial Engineering and Springer.

[8] Mijwil, Maad & Esen, Adam & Alsaadi, Aysar. (2019). Overview of Neural Networks. 1. 2.

[9] Nurachim, R. I. (2019). Pemilihan Model Prediksi Indeks Harga stock Yang Dikembangkan Berdasarkan Algoritma Support Vector Machine (Svm) Atau Multilayer Perceptron (Mlp) Studi Kasus: stock Pt Telekomunikasi Indonesia Tbk. Jurnal Teknologi Informatika dan Komputer, 5(1), 29-35.

[10] Pérez-Rodríguez, J. V., Torra, S., & Andrada-Félix, J. (2005). STAR and ANN models: Forecasting performance on the Spanish “Ibex-35” stock index. *Journal of Empirical Finance*, *12*(3), 490-509.

[11] Ramchoun, H., Ghanou, Y., Ettouil, M., & Janati Idrissi, M. A. (2016). Multilayer perceptron: Architecture optimization and training.

[12] Qiu, M., Song, Y., & Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. Chaos, Solitons & Fractals, 85, 1-7.

[13] Qiu, M., Song, Y., & Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. Chaos, Solitons & Fractals, 85, 1-7.