Jurnal Varian
Vol. 8, No. 1, October 2024, pp 113–124
e-ISSN: 2581-2017

Stock Price Index Prediction Using Random Forest Algorithm for Optimal Portfolio

Putri Humairah, Dina Agustina
Universitas Negeri Padang, Padang, Indonesia

Article Info

Article history:

Received : 07-19-2024
Revised : 09-02-2024
Accepted : 10-18-2024

Keywords:

Machine Learning;
Portfolio Optimal;
Random Forest.

ABSTRACT

With a majority Muslim population in Indonesia, Islamic capital markets such as the Jakarta Islamic Index (JII) are a relevant choice because the JII is an investment index that complies with Sharia principles. This research aims to predict stock prices in the JII using the Random Forest (RF) algorithm and form an optimal portfolio with the Mean-Variance Efficient Portfolio (MVEP) model. The data used is the daily closing price of JII stocks from April 2023 to March 2024, obtained from the Indonesia Stock Exchange and Yahoo Finance. The RF method is used to predict stock prices, with model performance evaluation using Mean Absolute Percentage Error (MAPE). The results showed that the application of ML with the RF algorithm in predicting stock prices produced very good predictions because the evaluation results using MAPE were in the 0%-10% range, namely a value of 2.522% for ACES shares; 1.222% for ICBP shares, and 0.760% for INDF shares. The optimal portfolio formed using MVEP produces a stock composition with a weight of 7.64% for ACES, 22.46% for ICBP, and 69.90% for INDF. The optimal portfolio's estimated expected return and risk are 0.0546% and 0.0103%.

Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: https://doi.org/10.30812/varian.v8i1.4276

Corresponding Author:

Dina Agustina,
Department of Mathematics, Universitas Negeri Padang, Indonesia,
Email: dinagustina@fmipa.unp.ac.id

How to Cite:
Humairah, P., & Agustina, D. (2024). Stock Price Index Prediction Using Random Forest Algorithm for Optimal Portfolio. *Jurnal Varian*, *8*(1), 113-124.

A. INTRODUCTION

Investment, in a broad sense, is the sacrifice of a number of current resources in the hope of obtaining future benefits (Desiyanti, 2017). Investments can be made, among others, by saving money, opening deposits, buying land and buildings, buying gold, participating in mutual funds, and securities such as stocks and bonds (Gunadi & Widyatama, 2021). Investing in stocks has the potential for high returns but also has a high risk because stock prices change from time to time. So, investors need to pay attention to stock price movements to predict future stock prices. With the prediction of stock prices, the future state of the capital market can be described and can be an important indicator for investors to make investment decisions (Hu et al., 2021).

This research involves the process of predicting stock prices and then forming an optimal portfolio. So, forecasting is carried out to predict stock prices before forming a portfolio. The solution is to use machine learning (ML) with the random forest (RF) algorithm to predict stock prices. The gap between this study and previous research is that while many studies have employed traditional methods, such as statistical models, to predict stock prices, there has been limited exploration of the RF algorithm in the context of forming an optimal portfolio after stock price predictions. This study aims to fill that gap by investigating the role of RF in stock price prediction and how these predictions can be used to construct an optimal portfolio. ML is a branch of artificial intelligence (AI) that focuses on learning from data (Cholissodin & Soebroto, 2019). This technique is a method for inferring data with a mathematical approach that aims to create a (mathematical) model that reflects data patterns (Putra, 2020). Meanwhile, RF is a model used in ML, especially in classification and regression, and part of ensemble learning methods that use (decision tree) DT as the base model. RF also creates diverse DT to reduce the correlation between the trees, helping to overcome the overfitting problem that often occurs with DT.

The purpose of this research is to involve the process of forecasting stock prices and then forming an optimal portfolio. So, forecasting is carried out to predict stock prices before forming a portfolio. The solution is to use machine learning (ML) with the random forest (RF) algorithm to predict stock prices. ML is a branch of artificial intelligence (AI) that focuses on learning from data (Cholissodin & Soebroto, 2019). This technique is a method for inferring data with a mathematical approach that aims to create a (mathematical) model that reflects data patterns (Putra, 2020). Meanwhile, RF is a model used in ML, especially in classification and regression, and part of ensemble learning methods that use (decision tree) DT as the base model. RF also creates diverse DT to reduce the correlation between the trees, helping to overcome the overfitting problem that often occurs with DT.

RF was chosen because it provides a higher level of accuracy in prediction compared to other models. This is because the randomness of RF makes it more resistant to outliers, which makes it easier to generalize the model. RF uses effective bagging techniques to avoid overfitting, which often occurs in machine learning models. In addition, it excels in handling large and complex datasets and efficient feature selection (Yin et al., 2023). Previous research, such as that conducted by Pinelis and Ruppert in 2022, shows that RF provides more accurate estimates in predicting stock returns. Research by Abraham et al. (2022) in 2022 showed a daily S&P 500 stock trend prediction accuracy of 80%, while Bastian et al. (2021) in 2021 showed that RF with technical analysis can minimize prediction errors with 84% accuracy.

Compared to previous research, The focus of this research is that this method is the first to apply RF to forecast JII stock prices and then form an optimal portfolio based on the prediction results to minimize risk and maximize return. MVEP is used to ensure efficient asset allocation for optimal portfolio formation. In investing, the main priority is not only maximizing profits but also reducing the risk of loss (Rivalno et al., 2019). Return and risk are positively correlated, so the higher the risk, the higher the expected return. Diversification, which is spreading investment in several different stocks and forming an optimal portfolio (Arisena et al., 2023) to minimize variance or investment risk with mean return (Rivalno et al., 2019).

B. RESEARCH METHOD

This type of research is applied research. Applied research aims to apply, test, and evaluate the ability of a theory to solve a practical problem (Abubakar, 2021). The type of data used in this research is secondary data. Secondary data is a source that does not directly provide data to data collectors, secondary data is collected by institutions or agencies and published to the public. The data source used is a list of daily closing prices recorded on the Jakarta Islamic Index (JII) stock index in the period April 1, 2023-March 31, 2024, from the Indonesia Stock Exchange Investment Gallery, Faculty of Economics, UNP and the official website of the Indonesia Stock Exchange (IDX), namely https://finance.yahoo.com/. Using daily data for one year helps identify patterns and trends in a shorter period of time. In addition, daily data provides more detailed information and includes daily fluctuations that are important for understanding stock price behavior, thereby increasing the accuracy of the prediction model (Ji et al., 2021). The work steps in this study are:

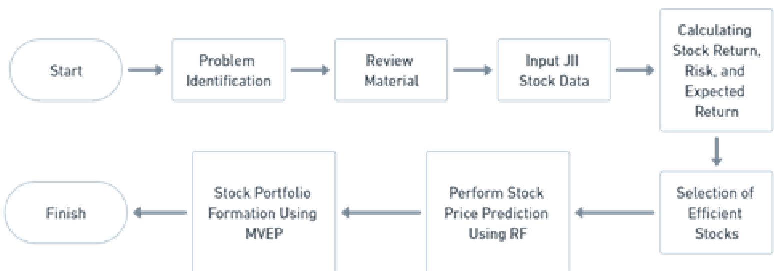


Figure 1. Research Steps

1. Calculating Return, Expected return, and Risk

Return is a reward for the investor's courage to take investment risks and the commitment of time and funds that investors have spent (Mercurio et al., 2020).

$$r_i = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \tag{1}$$

Where,

r_i = return of asset i

$P_{i,t}$ = stock price i at time t

$P_{i,t-1}$ = stock price i at time $t-1$

Expected return is the level of profit expected to be obtained by an investor as compensation for the risks that will arise from investment activities (Agustina et al., 2022).

$$\mu_i = E\left(r_i\right) = \frac{1}{N}\sum_{i=1}^{n} r_i \tag{2}$$

Where,

μ_i = $E(r_i)$ = expected return stock i

r_i = asset return i

N = number of returns in the observed period

The rate of return and risk are positively correlated. If the rate of return increases, the risk will also increase, otherwise if the rate of return decreases, the risk will also decrease (Desiyanti, 2017).

$$\sigma_i^2 = \frac{1}{N}\sum_{i=1}^{n} \left[r_i - E\left(r_i\right)\right]^2 \tag{3}$$

Where,

${\sigma_i}^2$ = variance of stock i

$E\left(r_i\right)$ = expected return of stock i

r_i = return aset i

N = number of returns in the observed period

2. Application of RF Algorithm

a. Decision Tree (DT)

DT is an ML algorithm that can perform regression and classification with a prediction model using a tree structure. DT is also a basic component of RF and is one of the most powerful ML algorithms available today (Géron, 2019). To form a DT structure, splitting is done to form branches. Measuring the quality of splitting in the case of classification using entropy and information gain. Entropy is an information-theoretic measure that can identify homogeneity and impurity in a data set, and information gain is used to determine optimal feature selection when building DT. While in the case of regression using mean square error (MSE) (Géron, 2019).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{4}$$

Where,

n = number of data points

Y_i = actual target value for data point i

$\hat{Y}_i$ = predicted target value for data point$i$

b. Lag Feature on Time Series Data

A sequence of observations made sequentially in time is called a time series (Box et al., 2016). Time series analysis is defined as the pattern of movement of variable values at certain time intervals, such as days, weeks, months, years, or other time units (Ardesfira et al., 2022). In time series analysis, lag is the time lag between data observations used to forecast future values. For example, if forecasting tomorrow's sales, sales data from the previous days will be used as input variables.

Using RF for time series forecasting is different from using simple regression. The previous values of the time series data, also known as lagged variables, serve as predictor variables. Therefore, the training time of the set will inevitably decrease if the number of predictor variables, specifically the selected lagged variables, is increased. The number of predictor variables used may decrease the amount of data obtained from the available knowledge of temporal dependence (Tyralis & Papacharalampous, 2017).

c. Random Forest (RF)

Random Forest (RF) is an ensemble learning algorithm that enhances accuracy and reduces overfitting by creating multiple decision trees through the bagging method, which involves sampling data to generate the trees (Agustina et al., 2024). Bagging stands for bootstrap aggregating, which consists of two stages. Bootstrapping is sampling from existing sample data, and aggregating is combining many conjectural values into one conjectural value. In the RF method, creating one tree at the time of training is independent of the previous tree. The classification DT is created based on the highest number of votes, and the regression DT is created by averaging each tree (Géron, 2019).

d. RF Evaluation with MAPE

To evaluate the performance of a statistical learning method on a given data set, it is necessary to know how well the predictions actually match the observed data (Jogunuri et al., 2024).

$$MAPE = \frac{1}{n}\sum_{i=1}^{N}\frac{Y_i - Y_N}{Y_i} \times 100\% \tag{5}$$

Where,

$MAPE$ = Mean Absolute Percentage Error
Y_N = Predicted value
Y_i = Actual value
n = Number of sample datasets

The lower the MAPE value, the better the model produced because the accuracy level will be high. The accuracy value is obtained based on the equation:

$$Accuracy = 100\% - MAPE \tag{6}$$

There is a range of MAPE values that can be used as a measurement of the ability of a prediction model (Fadhila & Zuliana, 2023):

Table 1. MAPE Value Range

MAPE Range	**Meaning**
$< 10\%$	The ability of the prediction model is very good
$10\% - 20\%$	The ability of the prediction model is good
$20\% - 50\%$	The ability of a decent prediction model
$> 50\%$	Poor prediction model ability

3. Optimal Portfolio Formation Using Mean-Variance Efficient Portfolio (MVEP)

A portfolio is a collection of several investments. The type of investment often used to form a portfolio is stock, so the portfolio in question is a collection of N stocks. The portfolio return is a linear combination of random variables from $r_1, r_2, \ldots, r_N$ with $(w_1, w_2, \ldots, w_N)$ or the sum of the return values of N stocks combined with the stock weight value $(w_1, w_2, \ldots, w_N)$

which is denoted by w_i. The sum of the weights denoted on each stock is equal to one, $\sum_{i=1}^{N} w_i = 1$. The portfolio return denoted (R_p) can be defined as (Logubayom & Victor, 2019):

$$R_p = w^T r \tag{7}$$

Where,
R_p = Portfolio return
w^T = Weight vector
r = Individual asset return vector

While the risk value of the stock portfolio can be calculated using the variance of the portfolio return (Wang et al., 2022).

$$Var\,(R_p) = w^T \Sigma w \tag{8}$$

Where,
$Var(R_p)$ = Portfolio return risk
w^T = Transpose of the weight vector
Σ = Covariance matrix
w = Weight vector

The mean-variance portfolio weighting is the portfolio that has the least risk among all possible portfolios, as indicated by the variance that can be formed at the expected profit level. The portfolio optimization problem minimizes the risk, and the constraint is that the weight equals one. Mathematically, it can be written with the following:

$$w = \frac{\Sigma^{-1} 1_{N\times 1}}{1_{N\times 1}^{T} \Sigma^{-1} 1_{N\times 1}} \tag{9}$$

Where,
w = Weight vector
Σ^{-1} = Inverse of covariance matrix
$1_{N\times 1}$ = Column vector with all elements equal to 1, and the length of the vector is N
$1_{N\times 1}^{T}$ = Transpose of vector $1_{N\times 1}$

C. RESULT AND DISCUSSION

The following is a Table 2 of differences between this research and previous research.

Table 2. Differences between this research and previous research

Differentiated Aspects	**This Research**	**Yin et al. (2023)**	**Pinelis & Ruppert (2022)**	**Abraham et al. (2022)**
Data	Jakarta Islamic Index (JII) stock data listed on the IDX for the period April 1, 2023-March 31, 2024	Four companies are listed on the US stock market, namely Integrated Electronics Corporation (INTC), Google (GOOG), Advanced Micro Devices Inc. (AMD), and Activision Blizzard Inc. (ATVI). The data period used is from January 3, 2015, to October 20, 2020.	Companies continuously listed on the NYSE, AMEX, or NASDAQ indices from 1927 to 2019	15 stocks listed in the Technology, Finance, and Healthcare sectors, taken from the period January 2, 2018 to June 30, 2019.
Prediction method	Random Forest	Random Forest	Prediction of return with random forest and volatility with Random Forest Model, Elastic Net	Genetic Algorithm (GA) based feature selection and classification using Random Forest (RF)

DOI: https://doi.org/10.30812/varian.v8i1.4276

Differentiated Aspects	This Research	Yin et al. (2023)	Pinelis & Ruppert (2022)	Abraham et al. (2022)
Model accuracy	ACES.JK 97.478% ICBP.JK 98.778% INDF.JK 99.24%	89%	Return prediction with random forest prediction accuracy of 64.52% and volatility estimation model accuracy of 79.84%	The model's accuracy in this study ranged from 55% for ABC stocks to 80% for EVTC stocks, which is the highest accuracy achieved by the model
Research focus	Predict stock prices using random forest then form an optimal portfolio with the mean-variance of the optimal portfolio	Make the Random Forest model more accurate and reliable for medium- and long-term stock prediction by applying model processing and optimization techniques	Develop and evaluate return prediction and volatility estimation methods using the Random Forest model in the context of portfolio allocation	Develop a stock price trend prediction model by utilizing Genetic Algorithm-based feature selection and classification using Random Forests

This study uses Jakarta Islamic Index (JII) stock data listed on the IDX for the period April 1, 2023-March 31, 2024, totaling 235 data. The expected return and risk values are attached in Table 4 to determine efficient stocks. The stocks listed in this research period are listed in Table 3.

Table 3. Stock Listed on JII for the period April 1, 2023-March 31, 2024

No.	Kode	Nama Stock
1	ACES.JK	Ace Hardware Indonesia Tbk.
2	ADRO.JK	Adaro Energy Indonesia Tbk.
3	AKRA.JK	AKR Corporindo Tbk.
4	ANTM.JK	Aneka Tambang Tbk.
5	BRIS.JK	Bank Syariah Indonesia Tbk.
6	BRMS.JK	Bumi Resources Minerals Tbk.
7	CPIN.JK	Charoen Pokphand Indonesia Tbk
8	EXCL.JK	XL Axiata Tbk.
9	ICBP.JK	Indofood CBP Sukses Makmur Tbk.
10	INCO.JK	Vale Indonesia Tbk.
11	INDF.JK	Indofood Sukses Makmur Tbk.
12	INKP.JK	Indah Kiat Pulp & Paper Tbk.
13	INTP.JK	Indocement Tunggal Prakarsa Tbk.
14	ITMG.JK	Indo Tambangraya Megah Tbk.
15	KLBF.JK	Kalbe Farma Tbk.
16	MIKA.JK	Mitra Keluarga Karyasehat Tbk.
17	PGAS.JK	Perusahaan Gas Negara Tbk.
18	PTBA.JK	Bukit Asam Tbk.
19	SMGR.JK	Semen Indonesia (Persero) Tbk.
20	TLKM.JK	Telkom Indonesia (Persero) Tbk.
21	TPIA.JK	Chandra Asri Petrochemical Tbk.
22	UNTR.JK	United Tractors Tbk.
23	UNVR.JK	Unilever Indonesia Tbk.

In Table 3, 23 JII stocks are consistently listed in this research period: April 1, 2023-March 31, 2024. The movement of 23 stock prices listed on the JII can be seen in Figure 2:

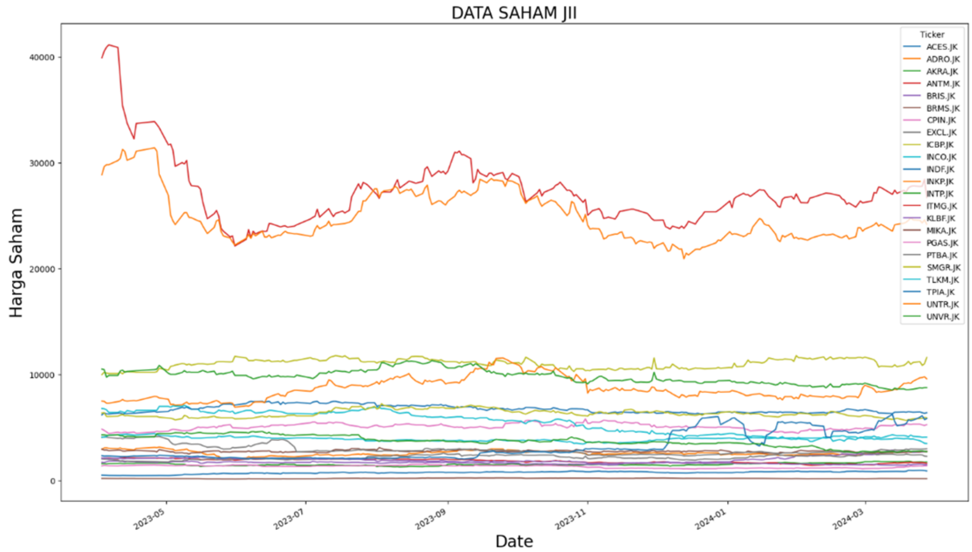


Figure 2. JII Stock Price Movement Chart

1. **Selection of Efficient Stock**

a. Calculating the daily stock return

Stock returns can determine how much return on investment is obtained from investing in a security. If a company's stock return is positive, the company is considered successful. Conversely, if the company's stock return is negative, the company is considered to have suffered a loss. Based on daily closing price data, Equation (1) (Rathi et al., 2024) can be used to calculate the return value of each stock with the help of Microsoft Excel software.

b. Calculating expected return and risk

The calculation of expected return and risk of 23 JII stocks can be calculated using Equations (2) and (3) (Rathi et al., 2024) assisted by using Microsoft Excel software. The expected return and risk values are presented in Table 4.

Table 4. Expected Return and Risk

No.	Stock	Expected Return	Variance
1	TPIA.JK	4.9478E-03	2.0634E-03
2	ACES.JK	2.9981E-03	8.3161E-04
3	BRIS.JK	2.2333E-03	4.5897E-04
4	INKP.JK	1.3859E-03	6.5275E-04
5	ICBP.JK	7.8906E-04	2.9045E-04
6	EXCL.JK	7.5987E-04	5.2241E-04
7	AKRA.JK	5.8896E-04	3.1559E-04
8	CPIN.JK	5.6052E-04	3.9409E-04
9	INDF.JK	1.9908E-04	1.2604E-04
10	PGAS.JK	1.4548E-04	2.5991E-04
11	BRMS.JK	-4.7195E-05	9.9578E-04
12	ADRO.JK	-1.2476E-04	4.7652E-04
13	ANTM.JK	-9.9717E-04	2.8981E-04
14	INCO.JK	-1.9457E-03	4.6519E-04
15	INTP.JK	-6.3565E-04	2.8691E-04
16	ITMG.JK	-1.5086E-03	4.0999E-04
17	KLBF.JK	-1.3247E-03	2.8888E-04
18	MIKA.JK	-8.3509E-05	4.0907E-04
19	PTBA.JK	-1.0429E-03	5.2359E-04
20	SMGR.JK	-1.8853E-04	2.5382E-04
21	TLKM.JK	-6.1235E-04	1.5907E-04
22	UNTR.JK	-5.9038E-04	3.2802E-04

No.	Stock	Expected Return	Variance
23	UNVR.JK	-1.8071E-03	3.3831E-04

Based on Table 4, the stocks have been sorted based on the highest to lowest expected return values. Of the 23 stocks, 10 stocks are obtained with positive expected return values TPIA, ACES, BRIS, INKP, ICBP, EXCL, AKRA, CPIN INDF, PGAS, meaning that the stock is experiencing an increase in return value. While BRMS, ADRO, ANTM, INCO, INTP, ITMG, KLBF, MIKA, PTBA, SMGR, TLKM, UNTR, UNVR show negative expected return values, meaning that the stock is experiencing a downward trend.

c. Selection of efficient stocks

Based on the positive expected return value in Table 4, not all positive expected return values are efficient stocks. Of the 10 stocks with positive expected return values, only 5 are efficient stocks. These efficient stocks have a high expected return value with a certain variance. The efficient stocks are shown in Table 5.

Table 5. Efficient Stocks

No	Stock	Expected Return	Variance
1	TPIA.JK	4.9478E-03	2.0634E-03
2	ACES.JK	2.9981E-03	8.3161E-04
3	BRIS.JK	2.2333E-03	4.5897E-04
4	ICBP.JK	7.8906E-04	2.9045E-04
5	INDF.JK	1.9908E-04	1.2604E-04

Table 5 shows 5 efficient stocks. Next, 3 optimal stocks will be selected to form a portfolio: ACES, JK, ICBP, JK, and INDF.JK.

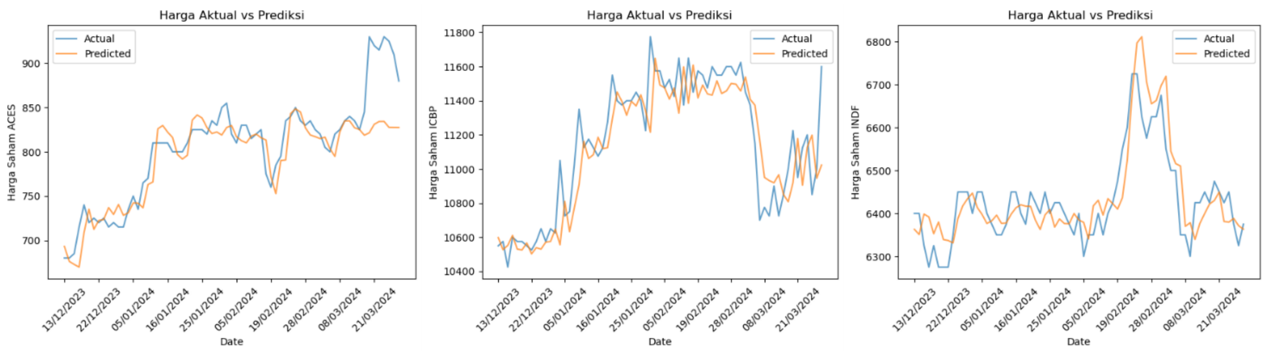
2. Predicting Stock Prices Using the RF Algorithm

a. Determining the Number of Steps (Lag Observations) Used as Features

Lag is the previous observation to predict the current value. This study will use 6 steps (6 lag observations). Create features and targets, where features are 6 observations and targets are stock prices.

b. Dividing Data into Training Data and Testing Data

The ratio of training and testing data is 7:3. Where training data starts from April 3, 2023, to December 11, 2023, and testing data from December 12, 2023, to March 28, 2024.



(a) Actual vs Predicted Price ACES Stock (b) Actual vs Predicted Price ICBP Stock (c) Actual vs Predicted Price INDF Stock

Figure 3. Actual vs Predicted Stock-Stock Price

Figure 3 shows a comparison chart of actual stock price data and predicted stock price data. Judging from the Figure, the predicted value generated by RF is almost close to the actual value, even though the predicted value is above or below the actual value.

DOI: https://doi.org/10.30812/varian.v8i1.4276

c. RF Evaluation Using MAPE

MAPE calculation is done using the price of each predicted stock and the price of each actual stock (Fadhila & Zuliana, 2023). MAPE results can be seen in Table 6.

Table 6. MAPE Value and Accuracy of Each Stock

No.	Stock	MAPE	Accuracy
1	ACES.JK	2.522%	97.478%
2	ICBP.JK	1.222%	98.778%
3	INDF.JK	0.760%	99.240%

Table 6 shows that the value of each RF model's MAPE for predicting average stock prices is in the 0%-10% range, which means that the model's ability to predict is very good.

3. Portfolio Formation Using MVEP

a. Calculating the variance-covariance matrix value

The variance-covariance matrix shows the statistical relationship between stocks. The values in the variance-covariance matrix indicate the extent to which changes in stock prices or returns of a stock are correlated with changes in other stocks. Positive covariance values indicate that the assets tend to coincide, while negative values indicate a tendency to move in opposite directions (Atta Mills et al., 2016). The variance-covariance matrix is shown in Table 7.

Table 7. Stock Variance-Covariance Matrix

Stock	ACES	ICBP	INDF
ACES	8.316E-04	3.206E-05	4.601E-05
ICBP	3.206E-05	2.905E-04	5.045E-05
INDF	4.601E-05	5.045E-05	1.260E-04

The matrix's main diagonal in Table 7 indicates the variance value of each stock. Where the variances of ACES, ICBP, and INDF stocks are 8.316E-04, 2.905E-04, and 1.260E-04, respectively. At the same time, the value outside the main diagonal shows the covariance value between the stocks.

b. Calculating the portfolio weight value (w)

The portfolio weight value can be determined from the variance-covariance matrix value and then calculated as the inverse of the variance-covariance matrix (Fadhila & Zuliana, 2023). The portfolio weight value of each stock will be shown in Table 8.

Table 8. Stock Portfolio Weight Value

No	Stock	Weight
1	ACES.JK	7.64%
2	ICBP.JK	22.46%
3	INDF.JK	69.90%

Table 8 shows the weight of each stock in the portfolio. The funds allocated to ACES stock amounted to 7.64%, to ICBP stock amounted to 22.46%, and to INDF stock amounted to 69.90%. This means that the largest funds allocated to the investment portfolio are in ICBP stock.

c. Calculating Portfolio Return and Risk

The expected return value and portfolio risk (Rathi et al., 2024) will be shown in Table 9.

Table 9. Expected Return and Risk Portfolio

Expected return portfolio	0.0546%
Portfolio risk	0.0103%

Table 9 shows that the portfolio of ACES, ICBP, and INDF stocks has an estimated expected return of 0.0546% with a risk level of 0.0000038. This shows the average expected return value of the portfolio of 0.0013598, but one must also be prepared to face a risk of 0.0103%.

D. CONCLUSION AND SUGGESTION

ML applications with the RF algorithm, it is found that the predictions produced are very good, with MAPE values in the 0%-10% range, namely 2.522% for ACES stocks, 1.222% for ICBP stocks, and 0.760% for INDF stocks, as well as the accuracy of 97.478%, 98.778%, and 99.24%, respectively. In addition, portfolio formation using the Mean-Variance Efficient Portfolio (MVEP) method shows an optimal composition with a weight of 7.64% ACES shares, 22.46% ICBP shares, and 69.90% INDF shares. The optimal stock portfolio's estimated expected return and risk are 0.0546% and 0.0103%, respectively.

Based on the conclusions of the above research results, further research can be developed by considering more comprehensive features, including external factors that can affect stock price movements, and comparing portfolios using other portfolio measurement methods.

REFERENCES

Abraham, R., Samad, M., Bakhach, A., El-Chaarani, H., Sardouk, A., Nemar, S., & Jaber, D. (2022). Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. *Journal of Risk and Financial Management*, *15*(5), 188. https://doi.org/10.3390/jrfm15050188

Abubakar, R. (2021). *Pengantar Metodologi Penelitian*. SUKA-Press UIN Sunan Kalijaga.

Agustina, D., Fitri, F., Zilrahmi, Winanda, R. S., & Sari, D. P. (2024). Utilization of machine learning for predictive maintenance in improving productivity in manufacturing industry. *AIP Conference Proceedings*, *3123*(1), 060001. https://doi.org/10.1063/5.0224329

Agustina, D., Sari, D. P., Winanda, R. S., Hilmi, M. R., & Fakhriyana, D. (2022). Comparison of Portfolio Mean-Variance Method with the Mean-Variance-Skewness-Kurtosis Method in Indonesia Stocks. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *23*(2), 88–97. https://doi.org/10.24036/eksakta/vol23-iss02/316

Ardesfira, G., Zedha, H. F., Fazana, I., Rahmadhiyanti, J., Rahima, S., & Anwar, S. (2022). Peramalan Nilai Tukar Rupiah terhadap Dollar Amerika dengan Menggunakan Metode Autoregressive Integrated Moving Average (Arima) [Number: 2]. *Jambura Journal of Probability and Statistics*, *3*(2), 71–84. https://doi.org/10.34312/jjps.v3i2.15469

Arisena, A., Noviyanti, L., Soleh, A. Z., & Indrayatna, F. (2023). Return Portofolio Optimal dengan Pendekatan Single Index Model, Treynor Black Model, Dan Black-Litterman Model [Number: 2]. *VARIANCE: Journal of Statistics and Its Applications*, *5*(2), 117–130. https://doi.org/10.30598/variancevol5iss2page117-130

Atta Mills, E. F. E., Yan, D., Yu, B., & Wei, X. (2016). Research on regularized mean–variance portfolio selection strategy with modified Roy safety-first principle. *SpringerPlus*, *5*(1), 919. https://doi.org/10.1186/s40064-016-2621-7

Bastian, M. E., Rahayudi, B., & Ratnawati, D. E. (2021). Prediksi Trend Harga Saham Jangka Pendek berdasarkan Fitur Technical Analysis dengan menggunakan Algoritma Random Forest [Number: 10]. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *5*(10), 4536–4542. https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/9945

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (Fifth edition). Wiley.

Cholissodin, I., & Soebroto, A. (2019, July). *Buku Ajar AI, Machine Learning & Deep Learning*. Fakultas Ilmu Komputer (FILKOM), Universitas Brawijaya (UB).

Desiyanti, R. (2017). *Teori Investasi Dan Portofolio*. Bung Hatta University Press.

Fadhila, S. N., & Zuliana, S. U. (2023). Optimasi Portofolio Saham Menggunakan Model Markowitz Berdasarkan Prediksi Harga Saham. *Kaunia: Integration and Interconnection Islam and Science Journal*, *18*(1), 35–40. https://doi.org/10.14421/kaunia.3948

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Second edition). O'Reilly.

Gunadi, N. L. D. S., & Widyatama, J. (2021). Perhitungan Sebagai Seorang Investor Saham Atas Besaran Pajak Yang Harus Dibayarkan Kepada Negara. *Jurnal Locus Delicti*, *2*(1), 13–23. https://doi.org/10.23887/jld.v2i1.455

Hu, Z., Zhao, Y., & Khushi, M. (2021). A Survey of Forex and Stock Price Prediction Using Deep Learning. *Applied System Innovation*, *4*(1), 9. https://doi.org/10.3390/asi4010009

Ji, X., Wang, J., & Yan, Z. (2021). A stock price prediction method based on deep learning technology. *International Journal of Crowd Science*, *5*(1), 55–72. https://doi.org/10.1108/IJCS-05-2020-0012

Jogunuri, S., F.T, J., Stonier, A. A., Peter, G., Jayaraj, J., S, J., J, J. J., & Ganji, V. (2024). Random forest machine learning algorithm based seasonal multi-step ahead short-term solar photovoltaic power output forecasting. *IET Renewable Power Generation*, rpg2.12921. https://doi.org/10.1049/rpg2.12921

Logubayom, A. I., & Victor, T. A. (2019). Portfolio Optimization of Some Stocks on the Ghana Stock Exchange Using the Markowitz Mean-Variance Approach. *Journal of Financial Risk Management*, *08*(01), 29–41. https://doi.org/10.4236/jfrm.2019.81003

Mercurio, P. J., Wu, Y., & Xie, H. (2020). An Entropy-Based Approach to Portfolio Optimization. *Entropy*, *22*(3), 332. https://doi.org/10.3390/e22030332

Pinelis, M., & Ruppert, D. (2022). Machine learning portfolio allocation. *The Journal of Finance and Data Science*, *8*, 35–54. https://doi.org/10.1016/j.jfds.2021.12.001

Putra, J. W. G. (2020). *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning* (Edisi 1.4).

Rathi, V., Kshirsagar, M., & Ryan, C. (2024). Enhancing Portfolio Performance: A Random Forest Approach to Volatility Prediction and Optimization. *Proceedings of the 16th International Conference on Agents and Artificial Intelligence*, 1278–1285. https://doi.org/10.5220/0012464600003636

Rivalno, F., Murni, D., & Sriningsih, R. (2019). Portofolio Mean Variance Efficient Dua Konstrain Pada 15 Saham Indeks LQ45. *Journal of Mathematics UNP*, *4*(1).

Tyralis, H., & Papacharalampous, G. (2017). Variable Selection in Time Series Forecasting Using Random Forests. *Algorithms*, *10*(4), 114. https://doi.org/10.3390/a10040114

Wang, L., Ahmad, F., Luo, G.-l., Umar, M., & Kirikkaleli, D. (2022). Portfolio optimization of financial commodities with energy futures. *Annals of Operations Research*, *313*(1), 401–439. https://doi.org/10.1007/s10479-021-04283-x

Yin, L., Li, B., Li, P., & Zhang, R. (2023). Research on stock trend prediction method based on optimized random forest. *CAAI Transactions on Intelligence Technology*, *8*(1), 274–284. https://doi.org/10.1049/cit2.12067

[This page intentionally left blank.]