

Prediksi Pergerakan Harga Saham Harian Menggunakan Model *Differencing Vector Autoregressive* dan *Vector Autoregressive Moving Average* (Studi Kasus Saham PT Bank Neo Commerce Tbk)

Hadi Sabililhaq¹, Indwiarti², Aniq Atiqi Rohmawati³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung
¹hadisabililhaq@students.telkomuniversity.ac.id, ²indwiarti@telkomuniversity.ac.id,
³aniqatiqi@telkomuniversity.ac.id

Abstrak

Perkembangan teknologi informasi telah membuka peluang baru di sektor investasi. Pembukaan rekening saham yang mudah membantu meningkatkan minat investasi di bursa saham Indonesia. Perbankan online menawarkan prospek investasi yang menarik, salah satunya saham PT Bank Neo Commerce Tbk (BBYB.JK). Peramalan dilakukan untuk mengetahui prediksi harga saham tersebut. Penggunaan analisis *time series* bisa membantu dalam memprediksi peramalan harga saham mendatang. Model *time series* yang digunakan dalam penelitian ini adalah model *Vector Autoregressive* (VAR) dan *Vector Autoregressive Moving Average* (VARMA), dengan melibatkan dua variabel data. Pemilihan kedua model tersebut didasarkan pada penggunaan dua variabel, yakni harga saham dan volume saham sebagai variabel kedua. Volume perdagangan saham berfungsi sebagai indikator aktivitas saham dan memberikan wawasan tentang permintaan beli dan jual saham. Metode *differencing* digunakan untuk menangani data yang tidak memiliki kestasioneran. Model dievaluasi menggunakan nilai *Mean Absolute Percentage Error* (MAPE) dan *Mean Absolute Error* (MAE). Model VAR (1) dengan rasio data *training* 90% dan *testing* 10% memberikan hasil paling akurat dengan nilai MAPE 2.08% dan nilai MAE 11.04. Semakin rendah nilai MAPE dan MAE mengindikasikan hasil peramalan yang lebih akurat.

Kata kunci : Prediksi, Saham, *Time series*, *Differencing*, VAR, VARMA

Abstract

The development of information technology has opened new opportunities in the investment sector. The ease of opening securities accounts helps boost investment interest in the Indonesian stock exchange. Online banking offers attractive investment prospects, one of which is the stock of PT Bank Neo Commerce Tbk (BBYB.JK). Forecasting is conducted to predict the stock price. Forecasting can be analyzed using time series analysis to make predictions of future stock prices. In this study, the Vector Autoregressive (VAR) and Vector Autoregressive Moving Average (VARMA) models were employed, integrating two variables. The selection of both models was based on the inclusion of two variables, namely stock price and stock volume as the second variable. Stock trading volume serves as an indicator of stock activity and provides insights into buying and selling demand. Differencing methods were used to handle non-stationary data. The models were evaluated using Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). The VAR (1) model trained with 90% training data and 10% testing data, yielding the most accurate results with a MAPE value of 2.08% and MAE value of 11.04. The greater the accuracy of the forecasting results is suggested by lower values of MAPE and MAE.

Keywords: Prediction, Stocks, Time series, Differencing, VAR, VARMA

1. Pendahuluan

Latar Belakang

Pada zaman yang modern ini, dengan kemajuan teknologi yang begitu pesat, menciptakan peluang baru bagi masyarakat Indonesia yang ingin masuk ke dunia investasi pasar modal. Simplifikasi pembukaan rekening investasi pasar modal berperan dalam peningkatan investor pasar modal di Indonesia [1]. Penambahan jumlah investor pasar modal di Indonesia saat pandemi terus bertambah, mencapai 5.088 juta pada bulan April tahun 2021, yang didominasi oleh investor pasar modal yang berumur 30 tahun ke bawah, dengan demografi mencapai 57,40% dari total investor pasar modal. Oleh karena itu, pentingnya pemahaman literasi investasi agar para investor memilih investasi pasar modal yang tepat [2]. Investasi yang menarik di pasar modal salah satunya adalah perusahaan bank digital karena perkembangan bank digital di Indonesia berpotensi mendisrupsi sektor perbankan tradisional, serta memperluas akses layanan keuangan bagi masyarakat [3]. Sehingga, potensi pertumbuhan industri perbankan digital di Indonesia sangat menjanjikan [3]. Artikel berita sebelumnya menyampaikan bahwa saham PT Bank Neo Commerce Tbk (BBYB.JK) merupakan salah satu saham perbankan digital yang direkomendasikan untuk di investasi karena mempunyai ekosistem Akulaku yang telah didukung oleh Alipay.

Oleh karena itu, saham BBYB.JK menjadi rekomendasi pilihan untuk investasi jangka menengah [4]. Dalam hal ini, diperlukan peramalan untuk mengetahui prediksi harga saham BBYB.JK.

Peramalan tersebut dapat dianalisis menggunakan analisis *time series*. Tujuan dari pemodelan *time series* ini untuk mengoleksi dan menganalisa data dari deret waktu tertentu untuk membuat ramalan pada periode mendatang [5]. Pada penelitian yang berkaitan dengan analisis *time series* sebelumnya tahun 2021 menggunakan model LSTM dan menggunakan data harga saham BBCA.JK, menyimpulkan model LSTM dapat digunakan sebagai prediktor jangka pendek dengan nilai akurasi sebesar 94,57% [6]. Nilai akurasi didapatkan dengan RMSE diperoleh dari penggunaan data *training* satu tahun dengan jumlah *epoch* yang tinggi pada fase penelitian [6]. Kemudian, penelitian yang berkaitan dengan analisis *time series* sebelumnya tahun 2022 menggunakan model ARIMA(1,1,0)-GARCH(1,1) dan menggunakan data harga minyak internasional menghasilkan nilai MAPE 0.045% dan RMSE 0.071 [7]. Selanjutnya, penelitian yang berkaitan dengan analisis *time series* sebelumnya tahun 2023 menggunakan model ARIMA, SARIMA, dan Holt-Winter *Multiplicative Approach* menggunakan data kasus malaria pakistan menghasilkan nilai RMSE 560.92 menggunakan model ARIMA (1,1,0), dan model SARIMA (2,1,2) (1,1,0) menghasilkan nilai RMSE 476.92, dan Holt-Winter *multiplicative* menghasilkan nilai RMSE 417.42 [8]. Sebelum melakukan pemodelan peramalan, data yang diuji diubah menjadi stasioner dengan metode *differencing* untuk mengatasi data yang tidak stasioner [9].

Pada penelitian ini, model yang digunakan adalah VAR dan VARMA. Pemilihan model VAR dan VARMA dalam penelitian ini, karena model VAR dan VARMA dapat menangani dua variabel *endogenous* yang saling mempengaruhi karena merupakan model untuk multivariat sedangkan AR dan ARMA merupakan model untuk univariat, yang hanya mempertimbangkan satu variabel *endogenous* dan tidak memperhitungkan hubungan variabel [9], [10]. Pada penelitian sebelumnya, dengan menggunakan model VAR untuk peramalan harga saham ASII.JK menggunakan data harga penutupan dan harga terendah saham menunjukkan nilai pengujian Confusion Matrix dengan akurasi 50.98% pada model VAR [10]. Pada penelitian lainnya, dengan menggunakan model VARMA(1,1) menggunakan dua data dari harga penutupan saham KAEF.JK dan INAF.JK menghasilkan nilai MAPE di bawah 10% [9]. Pada penelitian lainnya dengan menggunakan model VAR menghasilkan nilai MAPE 0.35%, 2.03%, dan 3.75% untuk memprediksi kasus baru Covid-19 di tiga negara [11]. Pada penelitian lainnya, dengan menggunakan model VARMA(1,1) untuk menganalisa prediksi inflasi dan suku bunga Indonesia menghasilkan nilai MAPE total 0.048754 dan nilai MSE total 0.054876 [12]. Berdasarkan penelitian sebelumnya [9][10][11][12], menggunakan model VAR dan VARMA menghasilkan nilai akurasi yang tinggi untuk pemodelan *time series*. Oleh karena itu penelitian ini menggunakan model VAR dan VARMA.

Penelitian ini mengkaji perbandingan prediksi harga saham BBYB.JK menggunakan model VAR dan VARMA dengan melibatkan dua variabel dari data harga saham dan volume saham BBYB.JK. Pemilihan volume saham, karena volume perdagangan saham merupakan suatu indikator aktivitas saham dan memberikan gambaran tentang tingkat permintaan jual beli saham [13]. Untuk mendapatkan evaluasi terbaik dari kedua model tersebut digunakan analisa *Mean Absolute Percentage Error* (MAPE) dan *Mean Absolute Error* (MAE).

Topik dan Batasannya

Dalam penelitian ini, topik dan batasan masalah adalah mengetahui apakah pemodelan VAR dan VARMA dapat meramalkan harga saham perusahaan PT Bank Neo Commerce Tbk (BBYB.JK) yang akurat dengan membandingkan performansi model VAR dan VARMA berdasarkan nilai Mean Absolute Percentage Error (MAPE) dan nilai *Mean Absolute Error* (MAE)?.

Tujuan

Tujuan dari penelitian ini adalah mendapatkan pemodelan terbaik dalam meramalkan harga saham perusahaan PT Bank Neo Commerce Tbk (BBYB.JK) dengan menggunakan model VAR dan VARMA, berdasarkan nilai *Mean Absolute Percentage Error* (MAPE) dan nilai *Mean Absolute Error* (MAE).

2. Studi Terkait

2.1 Prediksi

Prediksi atau peramalan merupakan bidang yang berkaitan dengan pembuatan pernyataan tentang hal-hal yang tidak diketahui pada masa kini [14]. Sebuah prediksi atau peramalan diperlukan jika ada ketidakpastian [14]. Oleh karena itu, prediksi atau peramalan membantu untuk mengurangi ketidakpastian dalam membuat suatu keputusan [14].

2.2 Saham

Saham adalah instrumen keuangan yang menunjukkan bukti kepemilikan pihak yang mempunyai surat berharga terhadap suatu perusahaan. [15]. pihak yang mempunyai surat berharga

tersebut mempunyai hak memperoleh bagian dari dividen perusahaan yang menerbitkan surat berharga tersebut [15].

2.3 Time Series

Time series adalah kumpulan data yang dicatat secara berurutan sesuai dengan urutannya [5]. *Time series* yang hanya memakai satu variabel disebut sebagai univariat. Namun, jika memakai lebih dari satu variabel yang dipertimbangkan, hal tersebut disebut sebagai multivariat. [5]. tujuan utama dari *time series modelling* adalah mengoleksi dan menganalisa data dari deret waktu tertentu untuk membuat ramalan pada periode mendatang [5].

2.4 Uji Stasioneritas Time Series

Pengujian akar unit merupakan metode untuk menguji stasioneritas data. Pengujian *Augmented Dickey Fuller (ADF)* merupakan teknik pengujian akar yang diterapkan, berikut merupakan persamaan dari pengujian *Augmented Dickey Fuller (ADF)* [16]:

$$\Delta(y_t) = \theta_1 + \theta_2 t + \beta y_{t-1} + \sum_{t=1}^p \alpha_p \Delta y_{t-p} + \varepsilon_t \quad (1)$$

Keterangan: y_t adalah variabel data pada waktu t , Δ adalah operator *first difference*, y_{t-1} adalah variabel data pada waktu $t - 1$, β_3 adalah koefisien regresi y_t dan y_{t-1} dan $\rho - 1$, θ adalah koefisien pada tren waktu, ε_t adalah komponen residual, α adalah koefisien signifikansi [16].

Dengan asumsi,

H_0 : Data yang tidak memiliki kestasioneran, $p - value > \alpha$ (mengindikasikan adanya akar unit)

H_1 : Data yang memiliki kestasioneran, $p - value \leq \alpha$ (mengindikasikan tidak adanya akar unit akar)

2.5 Pearson Correlation

Pearson Correlation adalah metode penghitungan korelasi antara dua variabel penelitian. Penghitungan ini menguji apakah dua variabel penelitian saling memiliki korelasi positif atau negatif, nilai koefisien dari -1 sampai 1, nilai koefisien tertinggi menandakan korelasi tinggi. Berikut adalah persamaan korelasi pearson [17]:

$$r = \frac{\frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\left[\left(\frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})^2 \right) \left(\frac{1}{n} \sum_{i=1}^n (b_i - \bar{b})^2 \right) \right]^{1/2}} \quad (2)$$

Keterangan: r merupakan korelasi *pearson*, n adalah jumlah data, a merupakan nilai dari variabel bebas a dalam sampel, b merupakan nilai dari variabel bebas b dalam sampel, \bar{a} adalah mean variabel a dari sampel data, \bar{b} adalah mean variabel b dalam sampel [17].

Berikut adalah skala koefisien korelasi Pearson: $0 < r \leq 0.19$ menunjukkan Korelasi yang sangat lemah, $0.2 \leq r \leq 0.39$ menunjukkan Korelasi yang lemah, $0.4 \leq r \leq 0.59$ menunjukkan Korelasi yang moderat, $0.6 \leq r \leq 0.79$ menunjukkan korelasi yang kuat, $0.8 \leq r \leq 1.0$ menunjukkan korelasi yang sangat kuat [18].

2.6 Differencing

Differencing data digunakan untuk mengatasi data yang tidak stasioner [9]. Metode differencing merubah data yang tidak stasioner menjadi data stasioner [9]. Berikut adalah persamaan dari perhitungan differencing [19]:

$$\Delta(y_t) = y_t - y_{t-1} \quad (3)$$

Keterangan: $\Delta(y_t)$ adalah variabel *differencing* pada waktu t , y_t adalah variabel asli pada waktu t , dengan t , sama dengan $1, 2, 3, \dots, n$ dan y_{t-1} adalah variabel asli tepat pada $t - 1$ [19].

2.7 Akaike's Information Criterion (AIC)

AIC merupakan metode menentukan model terunggul dengan membandingkan model statistik yang berbeda [20]. Nilai AIC yang paling kecil merupakan nilai yang terbaik, karena model yang terkecil berhasil memprediksi data dengan baik tanpa menggunakan banyak parameter [20]. Berikut adalah persamaan untuk mendapatkan nilai AIC [20]:

$$AIC = -2L + 2r \quad (4)$$

Keterangan: L adalah Likelihood, dan r adalah jumlah parameter dalam pemodelan dari rumus AIC, Likelihood menggambarkan sejauh mana model tersebut sesuai dengan data yang diberikan, dan $2r$ memastikan pemilihan model yang lebih sederhana kompleksitasnya dengan memberikan penalti berdasarkan jumlah parameter r [20].

2.8 Vector Autoregressive (VAR)

Model *Vector Autoregressive* (VAR) merupakan peramalan deret waktu dengan data *multivariate* yang dapat menangani lebih dari satu variabel *endogenous* yang saling mempengaruhi [10]. Model VAR merupakan model *multivariate* dari model *Autoregressive* (AR) [10]. Berikut adalah persamaan *Vector Autoregressive* (VAR) dengan ordo p dan dengan variabel k [21]:

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t \quad (5)$$

Keterangan: v adalah koefisien intersep ($k \times 1$) mempertimbangkan $E(y_t)$, A_i adalah estimasi ($k \times k$) dengan parameter matrix ($i = 1, 2, \dots, p$), ε_t adalah dimensi k yang tidak dapat diamati *white noise* [21].

2.9 Vector Autoregressive Moving Average (VARMA)

Model *Vector Autoregressive Moving Average* (VARMA) merupakan model peramalan *time series* dengan data *multivariate* yang dapat menangani lebih dari satu variabel *endogenous* yang saling mempengaruhi. Model VARMA merupakan model *multivariate* dari model *Autoregressive Moving Average* (ARMA) [9]. Model VARMA adalah kombinasi antara model *Vector Autoregressive* (VAR) orde p dengan *Vector Moving Average* (VMA) orde q [9]. Berikut adalah persamaan dari model VARMA (p, q) dengan variabel n [21]:

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t + M_1 \varepsilon_{t-1} + \dots + M_q \varepsilon_{t-q} \quad (6)$$

Keterangan: y_t adalah data *time series* pada waktu t berukuran ($n \times 1$), v adalah koefisien intersep ($n \times 1$), A_i adalah estimasi ($n \times n$) dengan parameter matriks VAR ($i = 1, 2, \dots, p$), M_j merupakan parameter matriks VMA ($j = 1, 2, \dots, q$), ε_t adalah dimensi n yang tidak dapat diamati *white noise* [21].

2.10 Mean Absolute Percentage Error (MAPE)

Penggunaan *Mean Absolute Percentage Error* (MAPE) bertujuan untuk menilai akurasi prediksi model. MAPE dihitung dengan cara mengambil *absolute error* pada tiap periode dan kemudian membaginya dengan nilai yang diamati pada periode tersebut. Setelah itu, diambil rata-rata dari persentase-persentase kesalahan absolut tersebut [22], [23]. Berikut adalah persamaan dari perhitungan MAPE [22], [23]:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (7)$$

Keterangan: y_t adalah nilai aktual pada waktu t , \hat{y}_t adalah nilai prediksi pada waktu t , n adalah jumlah data dalam peramalan [22], [23].

2.11 Mean Absolute Error (MAE)

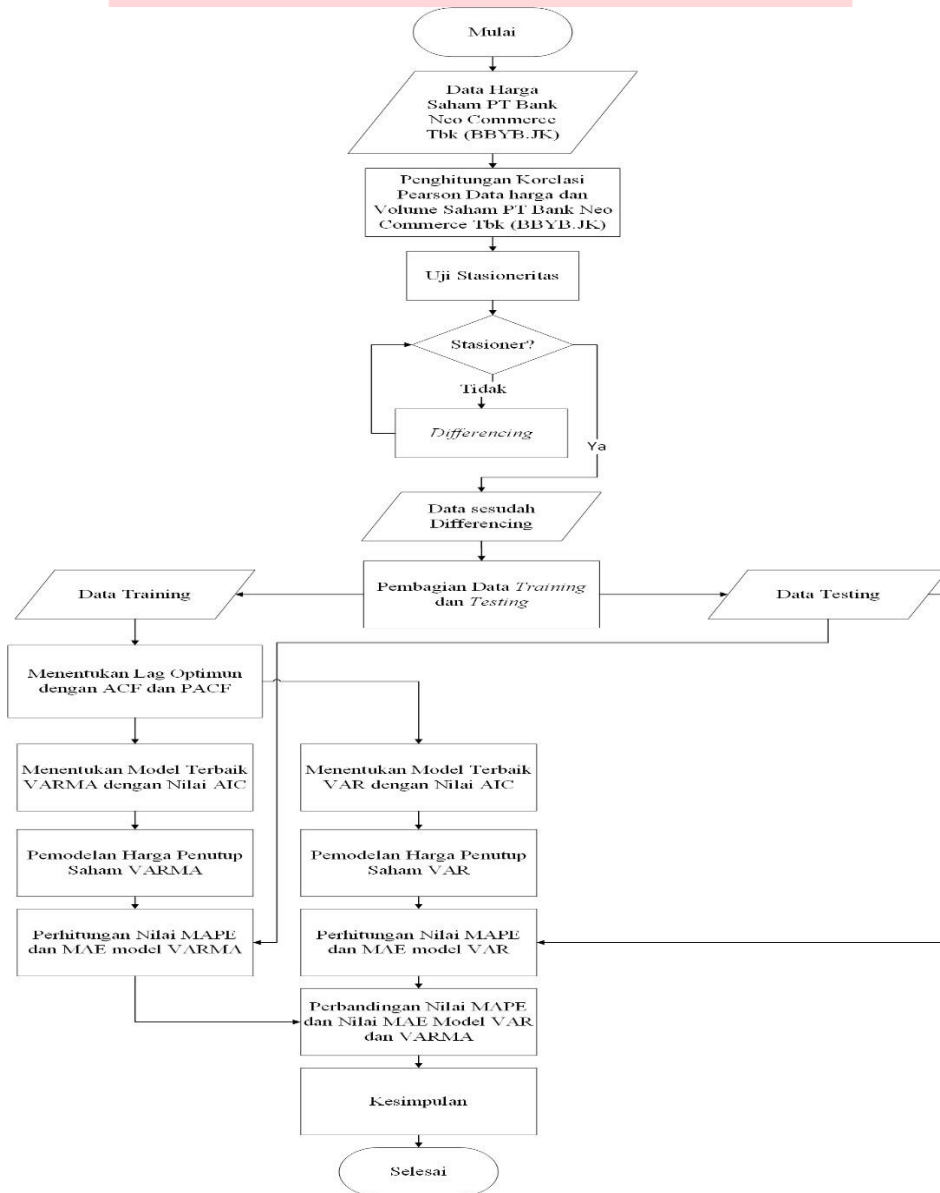
Mean Absolute Error (MAE) digunakan untuk mengukur akurasi prediksi model. MAE dihitung dengan cara mengambil *absolute error* pada setiap periode, yaitu selisih antara nilai aktual dan nilai prediksi. Setelah itu, diambil rata-rata dari *absolute error* tersebut [24]. Berikut adalah persamaan dari perhitungan MAE [24]:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (8)$$

Keterangan: y_t merupakan nilai aktual pada waktu t , \hat{y}_t merupakan nilai prediksi pada waktu t , n adalah jumlah data dalam peramalan [24].

3. Sistem yang Dibangun

Sistem pemodelan menggunakan model *Vector Autoregressive* dan *Vector Autoregressive Moving Average* dengan menggunakan data historis saham PT Bank Neo Commerce Tbk (BBYB.JK). Berikut adalah alur pemodelan sistem:



Gambar 1. Flowchart Alur Pemodelan

Keterangan:

1. Penginputan dataset harian saham perusahaan PT Bank Neo Commerce Tbk (BBYB.JK) yang diambil dari website yahoo *finance* yang memiliki rentang waktu periode 3 September 2018 sampai 31 Agustus 2023. Pada tahap ini juga dilakukan analisa data.
2. Penghitungan korelasi pearson untuk menguji korelasi tertinggi dari dataset harga saham dengan volume saham.
3. Uji Stasioneritas dengan melakukan Pengujian *Augmented Dickey Fuller (ADF)*.
4. Jika data tidak stasioner dilakukan *Differencing* data agar data menjadi stasioner.
5. Pembagian data *training* dan *testing* dengan rasio sebesar 70:30, 80:20, dan 90:10.
6. Menentukan lag optimum menggunakan ACF dan PACF
7. Menentukan pemodelan VAR dan VARMA terbaik menggunakan AIC
8. Pemodelan model *Vector Autoregressive (VAR)* dan *Vector Autoregressive Moving Average (VARMA)*.
9. *Re-differencing data*
10. Menilai pemodelan dengan melihat nilai yang dihasilkan oleh *Mean Absolute Percentage Error (MAPE)* buat mengukur akurasi prediksi dari pemodelan *Vector Autoregressive (VAR)* dan *Vector Autoregressive Moving Average (VARMA)*.
11. Pengevaluasian penelitian dan membuat kesimpulan dari hasil penelitian berdasarkan nilai MAPE dan nilai MAE pada model VAR dan VARMA.

3.1 Dataset

Pada penelitian ini, data yang dipakai adalah data harian harga saham dari perusahaan PT Bank Neo Commerce Tbk (BBYB.JK) yang diambil dari website *finance.yahoo.com*. Data yang diuji memiliki rentang waktu periode 3 September 2018 sampai 31 Agustus 2023. Data yang diuji terdiri dari 1234 baris.

Tabel 1. Sampel Data Saham (3 September 2018 Sampai 31 Agustus 2023)

Date	Open	High	Low	Close	Adj Close	Volume
2018-09-03	306.0	306.0	304.0	306.0	305.634766	1179000.0
2018-09-04	306.0	306.0	300.0	306.0	305.634766	1433200.0
2018-09-05	304.0	304.0	298.0	304.0	299.641937	757600.0
2018-09-06	300.0	304.0	298.0	300.0	297.644318	1190700.0
2018-09-07	300.0	300.0	298.0	300.0	299.641937	1569300.0
...
2023-08-25	364.0	370.0	352.0	352.0	352.0	19212000.0
2023-08-28	352.0	358.0	348.0	354.0	354.0	12952200.0
2023-08-29	354.0	364.0	352.0	356.0	356.0	10835500.0
2023-08-30	356.0	362.0	332.0	332.0	332.0	28539200.0
2023-08-31	334.0	352.0	332.0	344.0	344.0	43599400.0

Dari dataset tabel 1, data yang diuji dalam penelitian ini adalah data harga tertinggi saham, data pembukaan saham, harga penutup saham, harga terendah saham, dan *adjusted close* dengan volume saham sebagai variabel pendukung. Dari kelima variabel pada tabel 1, akan dipilih salah satu variabel yang memiliki korelasi hubungan tertinggi dengan data volume saham.

4. Evaluasi

4.1 Hasil Korelasi Pearson

Penghitungan korelasi Pearson pada data harian saham PT Bank Neo Commerce Tbk (BBYB.JK) untuk menentukan korelasi hubungan tertinggi antara variabel data harga saham dengan volume saham pada dataset tabel 1, berikut adalah hasil korelasi Pearson harga saham dengan volume saham:

Tabel 2. Hasil Korelasi Pearson Harga Saham dengan Volume Saham.

Variabel Data	Koefisien korelasi pearson
Harga penutup saham dengan volume	0.48
Harga tertinggi saham dengan volume	0.49
Harga terendah saham dengan volume	0.45
Harga pembukaan saham dengan volume	0.46
<i>adjusted close</i> dengan volume	0.48

Dapat diambil dari kesimpulan dari tabel 2, menunjukkan korelasi positif antara semua variabel yang dianalisis. Nilai korelasi pearson tertinggi adalah korelasi harga tertinggi saham dengan volume dengan menunjukkan nilai pearson 0.49. Korelasi antara harga tertinggi saham dan volume menunjukkan korelasi moderat karena skala korelasi koefisien pearson diantara 0.4 sampai 0.59 [18]. Oleh karena itu, variabel data harga tertinggi saham dengan volume dipilih pada pemodelan VAR dan VARMA.

4.2 Hasil Uji Stasioneritas *Augmented Dickey Fuller*

Pengujian uji stasioneritas data ini menggunakan *Augmented Dickey Fuller*, berikut adalah pengujian stasioneritas data harga tertinggi saham dan volume saham PT Bank Neo Commerce Tbk (BBYB.JK) sebelum *differencing*:

Tabel 3. Hasil Uji Stasioneritas ADF Sebelum Differencing

Variabel Data	$p - value$	koefisien signifikan (α)	Keterangan
Harga tertinggi saham PT Bank Neo Commerce Tbk (BBYB.JK)	0.4029	0,05	Data Tidak Stasioner
Volume saham PT Bank Neo Commerce Tbk (BBYB.JK)	0.0042	0,05	Data Stasioner

Dengan hipotesis,

H_0 : Data tidak stasioner, $p - value > \alpha$ (m mengandung akar unit)

H_1 : Data Stasioner, $p - value \leq \alpha$ (tidak mengandung akar unit)

Dapat diambil kesimpulan dari tabel 3 bahwa hipotesis H_0 diterima pada data harga tertinggi saham dan hipotesis H_1 diterima pada data volume saham PT Bank Neo Commerce Tbk (BBYB.JK). Agar data menjadi stasioner dilakukan metode *differencing*. *Differencing* data digunakan untuk mengatasi data yang tidak stasioner. Berikut adalah hasil pengujian uji stasioneritas ini menggunakan *Augmented Dickey Fuller* dengan data yang sudah di *differencing*.

Tabel 4. Hasil Uji Stasioneritas ADF Sesudah Differencing

Variabel Data	$p - value$	koefisien signifikan (α)	Keterangan
Harga tertinggi saham PT Bank Neo Commerce Tbk (BBYB.JK)	$8.5347e^{-11}$	0,05	Data Stasioner
Volume saham PT Bank Neo Commerce Tbk (BBYB.JK)	$9.0490e^{-14}$	0,05	Data Stasioner

Dengan hipotesis,

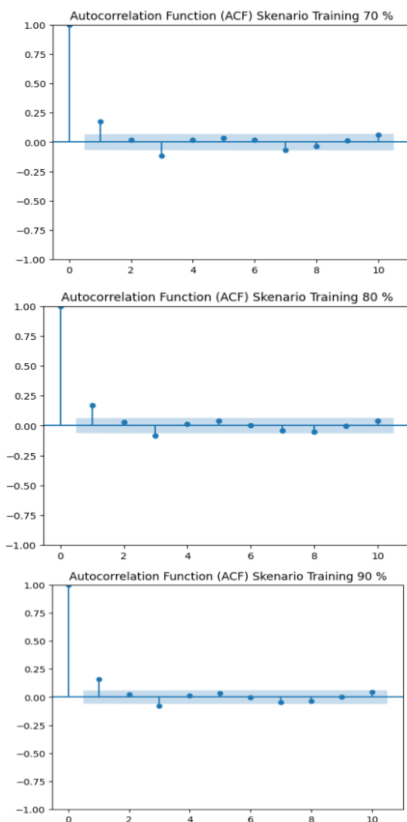
H_0 : Data tidak stasioner, $p - value > \alpha$ (mengandung akar unit)

H_1 : Data Stasioner, $p - value \leq \alpha$ (tidak mengandung akar unit)

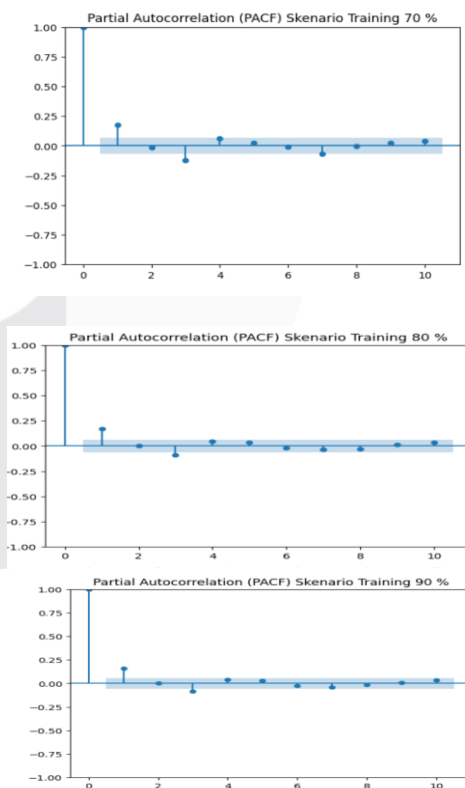
Dapat diambil kesimpulan dari tabel 4 bahwa hipotesis H_1 diterima pada data harga tertinggi saham dan volume saham PT Bank Neo Commerce Tbk (BBYB.JK) maka kedua data harga tertinggi saham dan volume saham sudah stasioner setelah *differencing*.

4.3 Hasil ACF dan PACF

Menentukan orde pada model VAR dan VARMA menggunakan ACF untuk menentukan orde optimal *Moving Average* (MA) dan PACF untuk menentukan orde optimal *Autoregressive* (AR), berikut adalah hasil dari ACF dan PACF:



Gambar 2. Plot ACF pada Tiga Skenario Rasio Pembagian Data



Gambar 3. Plot PACF pada Tiga Skenario Rasio Pembagian Data

Dapat diambil kesimpulan dari gambar 2 ACF untuk tiga skenario pembagian data bahwa menentukan orde optimal *Moving Average* (MA), dilihat dari cut-off pada orde lag plot ACF [10]. Lag yang diambil untuk tiga skenario pembagian data adalah lag-1 untuk orde *Moving Average* (MA) karena terjadi *cut-off* setelah lag-1. Kemudian pada gambar 3 PACF, dapat disimpulkan bahwa dari tiga skenario pembagian data untuk menentukan orde optimal *Autoregressive* (AR), dilihat dari *cut-off* orde lag pada plot PACF. Lag yang diambil untuk tiga skenario pembagian data adalah lag-1 untuk orde *Autoregressive* (AR) karena terjadi *cut off* setelah lag-1. Jadi, pada pemodelan VAR orde AR(p) yang memiliki korelasi signifikan adalah orde AR(1), dan pada pemodelan VARMA kombinasi orde AR(p) dan MA(q) yang memiliki korelasi signifikan adalah AR(1) dan MA(1). Oleh karena itu, penghitungan AIC tidak dilakukan karena dari hasil ACF dan PACF hanya model VAR(1) dan VARMA(1,1) yang nilai korelasi yang signifikan.

4.4 Pemodelan VAR (1)

Setelah mendapatkan orde AR(p) dengan ACF dan PACF, menunjukkan bahwa model VAR (1) merupakan model terbaik untuk digunakan pada penelitian pembagian data *training* dan *testing* dengan rasio sebesar 70:30, 80:20, dan 90:10, berikut adalah model VAR (1) memprediksi harga tertinggi saham dengan volume saham sebagai variabel kedua:

$$\Delta(y_{1t}) = v_1 + A_1 \times y_{1t-1} + A_2 \times y_{2t-1} + \epsilon_{1t} \quad (9)$$

Dari persamaan diatas, terdapat y_t adalah data *time series* pada waktu t berukuran (2×1) dengan 2 variabel harga tertinggi saham dan volume saham, Δ adalah operator *first difference*, v adalah koefisien intersep (2×1) , A_1 parameter matriks tertinggi saham, A_2 adalah parameter matriks volume saham VAR(1), ϵ_t adalah *random vector* pada waktu t dengan ukuran (2×1) yang mempunyai variansi 1 dan mean 0 yang merupakan bilangan acak harga tertinggi saham [21]. Berikut adalah hasil koefisien harga tertinggi saham model VAR (1) skenario rasio pembagian data 70:30, 80:20, dan 90:10:

Tabel 5. Hasil koefisien Harga Tertinggi Saham pada Model VAR (1) dengan Skenario Rasio Pembagian Data 70:30, 80:20, dan 90:10.

Rasio Skenario Pembagian Data	Parameter dan Koefisien		
	Intercept (v_1)	LI.high (A_1)	LI. Volume (A_2)
70:30	1.7555	0.2006	$-5.975e^{-08}$
80:20	0.7744	0.1968	$-6.47e^{-08}$
90:10	0.3006	0.1855	$-7.17e^{-08}$

Dapat dilihat pada tabel 5 terdapat koefisien parameter pada model VAR (1) harga tertinggi saham dengan volume saham sebagai variabel kedua pada skenario rasio pembagian data 70:30, 80:20, dan 90:10. Koefisien dari parameter tersebut diambil dari library python statsmodels.tsa.statespace.varmax. Dari koefisien yang dihasilkan pada tabel 5, persamaan model VAR (1) yang dihasilkan adalah:

$$\Delta(y_{1t}) = 1.7555 + ((0.2006 \times y_{1t-1}) + (-0.00000005975 \times y_{2t-1})) + \epsilon_{1t} \quad (10)$$

$$\Delta(y_{1t}) = 0.7744 + ((0.1968 \times y_{1t-1}) + (-0.0000000647 \times y_{2t-1})) + \epsilon_{1t} \quad (11)$$

$$\Delta(y_{1t}) = 0.3006 + ((0.1855 \times y_{1t-1}) + (-0.0000000717 \times y_{2t-1})) + \epsilon_{1t} \quad (12)$$

Persamaan 10,11,12 diatas adalah persamaan skenario dari rasio pembagian data 70:30, 80:20, dan 90:10 model VAR(1) yang sudah dimasukan nilai koefisien dari parameter v_1 , A_1 , dan A_2 ke dalam persamaan 9.

4.5 Pemodelan VARMA (1,1)

Setelah mendapatkan orde terbaik AR(p) dan MA(q) dengan ACF dan PACF untuk menentukan orde terbaik model VARMA (p,q). Dari ACF dan PACF menunjukan bahwa model VARMA (1,1) merupakan model terbaik untuk digunakan pada penelitian dengan rasio pembagian data *training* dan *testing* 70:30, 80:20, dan 90:10, berikut adalah model VARMA (1,1) memprediksi harga tertinggi saham dengan volume saham sebagai variabel kedua:

$$\Delta(y_{1t}) = v_1 + A_1 \times y_{1t-1} + A_2 \times y_{2t-1} + \epsilon_{1t} + M_1 \times \epsilon_{1t-1} + M_2 \times \epsilon_{2t-1} \quad (13)$$

Dari persamaan diatas, terdapat y_t adalah data *time series* pada waktu t berukuran (2×1) dengan 2 variabel harga tertinggi saham dan volume saham, Δ adalah operator *first difference*, v adalah koefisien intersep (2×1) , A_1 adalah parameter koefisien harga tertinggi dan A_2 adalah parameter koefisien volume saham pada model VAR pada model VARMA (1,1), M_1 adalah parameter koefisien harga tertinggi dan M_2 adalah parameter koefisien volume saham pada model VMA pada model VARMA (1,1), ϵ_t adalah *random vector* pada waktu t dengan ukuran (2×1) yang mempunyai variansi 1 dan mean 0 yang merupakan bilangan acak harga tertinggi saham [21]. Berikut adalah hasil koefisien harga tertinggi saham model VARMA (1,1) skenario rasio pembagian data 70:30, 80:20, dan 90:10:

Tabel 6. Hasil Koefisien Harga Tertinggi Saham pada Model VARMA (1,1) dengan Skenario Rasio Pembagian Data 70:30, 80:20, dan 90:10.

Rasio Skenario Pembagian Data	Parameter dan Koefisien				
	Intercept (v_1)	L1.high (A_1)	L1.volume (A_2)	L1.e(high) (M_1)	L1.e(volume) (M_2)
70:30	1.7527	0.2794	-3.738e ⁻⁰⁷	-0.1231	-3.973e ⁻⁰⁷
80:20	0.7740	0.2611	-3.32e ⁻⁰⁷	-0.0991	3.365e ⁻⁰⁷
90:10	-0.0233	1.0063	1.472e ⁻⁰⁷	-0.9980	-2.696e ⁻⁰⁸

Dapat dilihat dari tabel 6 library Python statsmodels.tsa.statespace.varmax digunakan untuk mendapatkan koefisien dari parameter pada model VARMA (1,1) harga tertinggi saham dengan volume saham sebagai variabel kedua pada skenario rasio pembagian data 70:30, 80:20, dan 90:10, berikut adalah persamaan model pada tiga skenario pembagian data setelah koefisien dimasukkan:

$$\Delta(y_{1t}) = 1.7527 + ((0.2794 \times y_{1t-1}) + -0.0000003738 \times y_{2t-1})) + \epsilon_{1t} + ((-0.1231 \times \epsilon_{t-1}) + (0.0000003973 \times \epsilon_{t-1})) \quad (14)$$

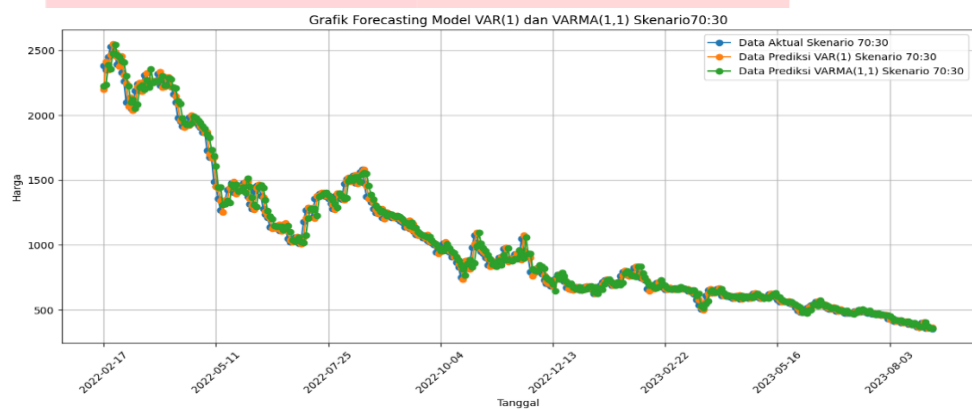
$$\Delta(y_{1t}) = 0.7740 + ((0.2611 \times y_{1t-1}) + -0.000000332 \times y_{2t-1})) + \epsilon_{1t} + ((-0.0991 \times \epsilon_{t-1}) + (0.0000003365 \times \epsilon_{t-1})) \quad (15)$$

$$\Delta(y_{1t}) = -0.0233 + ((1.0063 \times y_{1t-1}) + 0.0000001472 \times y_{2t-1})) + \epsilon_{1t} + ((-0.9980 \times \epsilon_{t-1}) + (-0.00000002696 \times \epsilon_{t-1})) \quad (16)$$

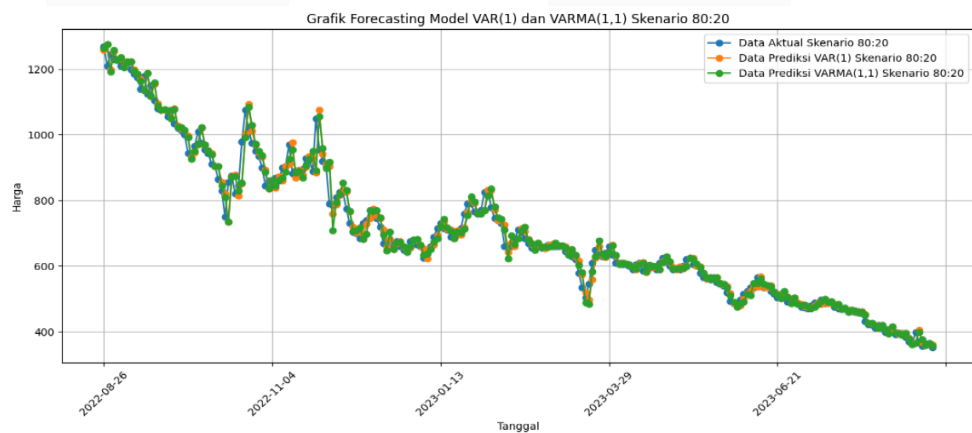
Persamaan 14, 15, dan 16 diatas adalah persamaan skenario rasio pembagian data 70:30, 80:20, dan 90:10 model VARMA(1,1) yang telah dimasukan koefisien dari parameter v_1 , A_1 , A_2 , M_1 , dan M_2 ke dalam persamaan 13.

4.6 Perbandingan Model VAR(1) dengan VARMA(1)

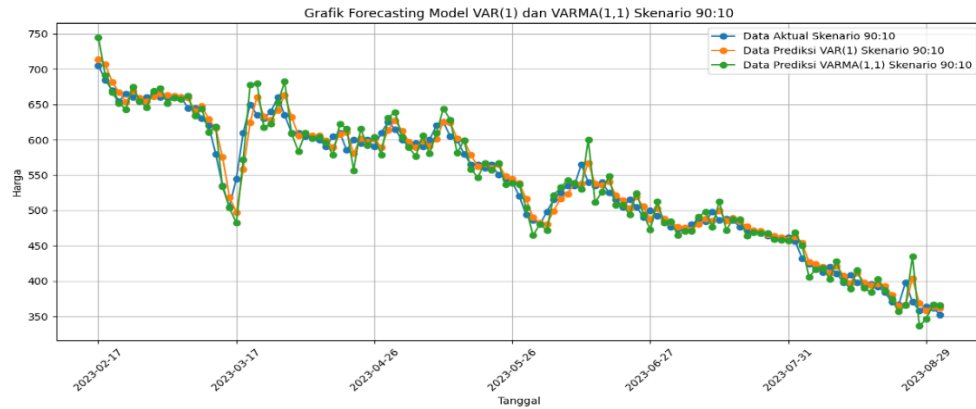
Berdasarkan pemodelan VAR(1) dan VARMA(1,1) dihasilkan grafik perbandingan data aktual dan data prediksi pada tiga skenario pembagian data, berikut adalah grafik perbandingan pemodelan VAR(1) dan VARMA(1,1):



Gambar 4. Grafik Forecasting Model VAR(1) dan VARMA(1,1) pada Skenario Rasio pembagian data 70:30



Gambar 5. Grafik Forecasting Model VAR(1) dan VARMA(1,1) pada Skenario Rasio pembagian data 80:20



Gambar 6. Grafik Forecasting Model VAR(1) dan VARMA(1,1) pada Skenario Rasio pembagian data 90:10

Pada gambar 4,5, dan 6 terdapat grafik *forecasting* model VAR(1) dan VARMA(1) pada skenario rasio pembagian data 70:30, 80:20, dan 90:10. Pada gambar visualisasi tersebut, nilai aktual dan nilai prediksi grafiknya hampir identik. Observasi ini menunjukkan bahwa kedua model tersebut mampu memberikan prediksi dengan tingkat akurasi yang tinggi. Berikut adalah perbandingan nilai MAPE dan nilai MAE dengan tiga skenario pembagian data model VAR(1) dan VARMA(1,1):

Tabel 7. Hasil Nilai MAPE dan MAE Model VAR (1) dan VARMA (1,1)

Rasio Skenario Pembagian data	Nilai MAPE		Nilai MAE	
	Model VAR (1)	Model VARMA (1,1)	Model VAR (1)	Model VARMA (1,1)
70:30	2.81%	3.95%	28.98	40.98
80:20	2.71%	2.77%	19.96	20.38
90:10	2.08%	2.78%	11.04	14.61

Dapat diambil kesimpulan dari tabel 7 bahwa model VAR(1) lebih baik dari model VARMA(1,1), hal ini dilihat dari perbandingan nilai MAPE dan nilai MAE model VAR(1) dengan nilai MAPE dan MAE model VARMA(1,1). Dari seluruh skenario, model VAR(1) menghasilkan nilai MAPE dan nilai MAE lebih rendah dibandingkan model VARMA(1,1), hal ini mengindikasikan nilai prediksi VAR(1) lebih akurat dibandingkan model VARMA(1,1).

5. Kesimpulan

Pada penelitian ini, hubungan korelasi Pearson tertinggi pada dataset adalah harga tertinggi saham dengan volume saham PT Bank Neo Commerce Tbk (BBYB.JK). Penentuan orde optimum ditentukan dengan melihat *cut off* pada gambar plot ACF dan plot PACF. Orde model terbaik yang dihasilkan adalah VAR (1) dan VARMA (1,1), karena terjadi *cut off* pada gambar plot ACF dan plot PACF setelah lag-1 pada tiga skenario pembagian data. Skenario pembagian data *training* dan *testing* dengan rasio sebesar 90:10 menggunakan model VAR(1) memberikan hasil paling akurat, hal ini dilihat dari nilai MAPE model VAR (1) pada skenario rasio pembagian data 90:10 sebesar 2.08%, dan nilai MAE 11.04. Artinya, jarak antara nilai aktual dan nilai prediksi sangat kecil, sehingga model ini cukup bagus untuk meramalkan harga saham perusahaan PT Bank Neo Commerce Tbk (BBYB.JK).

Saran penelitian selanjutnya dapat mempertimbangkan penambahan variabel lainnya selain volume saham dalam model VAR atau VARMA. Selain itu, pertimbangkan juga penambahan variabel *exogenous* menggunakan penerapan model VAR-X dan VARMA-X.

Referensi

- [1] Y. Sukmana, "Naik 89,58 Persen, Jumlah Investor Pasar Modal Indonesia Capai 7,3 Juta," *Money Kompas*, 26 December 2021. [Online]. Available: <https://money.kompas.com/read/2021/12/26/074549626/naik-8958-persen-jumlah-investor-pasar-modal-indonesia-capai-73-juta?page=all#:~:text=Whats%20New-,Naik%2089%2C58%20Persen%2C%20Jumlah%20Investor%20Pasar%20Modal,Indonesia%20Capai%207%2C3%20Juta&text=JAKA>. [Accessed 11 March 2022].
- [2] R. H. Yustrianthe and R. Tjandra, "Determinants Of Investment Interest From Young Accountants," *Jurnal Akutansi*, vol. 27, no. 2, pp. 242-260, May 2023.
- [3] S. Nurhaliza, "Bank Digital Bisa Dorong Minat Investasi di Pasar Modal," *Bisnis.com*, 28 October 2021. [Online]. Available: <https://www.idxchannel.com/banking/bank-digital-bisa-dorong-minat-investasi-di-pasar-modal>. [Accessed 11 March 2022].
- [4] D. M. Hutauruk and A. Wikanto, "Harga saham bank digital AGRO, ARTO, BBYB, BBHI, BABP melesat, ini rekomendasi analis," *Investasi kontan*, 14 september 2021. [Online]. Available: <https://investasi.kontan.co.id/news/harga-saham-bank-digital-agro-arto-bbyb-bbhi-babp-melesat-ini-rekomendasi-analis>. [Accessed 11 March 2022].
- [5] R. Adhikari and R. K. Agrawal, "An Introductory Study on Time Series Modeling and Forecasting,," 2013. [Online]. Available: <https://doi.org/10.48550/arXiv.1302.6613>.
- [6] W. Budiharto, "Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM)," *J Big Data* 8, 47, 2021. [Online]. Available: <https://doi.org/10.1186/s40537-021-00430-0>.
- [7] Y. Xiang, "Using ARIMA-GARCH Model to Analyze Fluctuation Law of International Oil Price," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3936414, 7 pages, 2022. [Online]. Available: <https://doi.org/10.1155/2022/3936414>.
- [8] M. Riaz, M. H. Sial, S. Sharif and Q. Mehmood, "Epidemiological Forecasting Models Using ARIMA, SARIMA, and Holt–Winter Multiplicative Approach for Pakistan," *Journal of Environmental and Public Health*, vol. 2023, Article ID 8907610, 8 pages, 2023. [Online]. Available: <https://doi.org/10.1155/2023/8907610>.
- [9] A. Ulya, "Peramalan harga saham penutupan menggunakan model vector autoregressive moving average (VARMA),," *Universitas Islam Negeri Maulana Malik Ibrahim*, 2019.
- [10] A. Wulandari and A. A. Rohmawati, "Prediksi Pergerakan Harga Saham PT. Astra Internasional tbk Menggunakan Vector Autoregressive (VAR) dan Logistic Regression," in *e-Proceeding of Engineering: Vol 7, No 2, 2614.*, 2020.
- [11] K. Rajab, F. Kamalov and A. K. Cherukuri, "Forecasting COVID-19: Vector Autoregression-Based Model," *Arab J Sci Eng*, vol. 47(6), no. Epub 2022 Jan 3. PMID: 35004125; PMCID: PMC8722659., pp. 6851-6860, 2022.
- [12] Hermansah, D. Rosadi, Abdurakhman, H. Utami and G. Darmawan, "Multivariate Time Series Data Forecasting Using Multi-Output NARNN Model," *Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences (ICRIEMS 2020)*, 2021. [Online]. Available: <https://doi.org/10.2991/assehr.k.210305.041>.
- [13] L. F. Triani, "FAKTOR-FAKTOR YANG MEMPENGARUHI PERUBAHAN INDEKS HARGA SAHAM DI JAKARTA ISLAMIC INDEX SELAMA TAHUN 2011," *Jurnal Organisasi Dan Manajemen*, 9(2), 162–178., 2013. [Online]. Available: <https://doi.org/10.33830/jom.v9i2.45.2013..>
- [14] A. Graefe, K. Green and J. S. Armstrong, *Forecasting*. 10.1007/978-1-4419-1153-7_357., *Encyclopedia of Operations Research and Management Science* (pp.539-604)Edition: 3Publisher: SpringerEditors: Saul I. Gass, Michael C. Fu.
- [15] A. Maskuri, Hartono and A. Pranaditya, "PENGARUH PROFITABILITAS TERHADAP HARGA SAHAM PERUSAHAAN RETAIL YANG TERDAFTAR DI BURSA EFEK INDONESIA," *Jurnal Ilmiah Mahasiswa SI Akuntansi Universitas Pandanaran*, Vols. 3, No 3., no. ISSN : 2502-7697, 2017.
- [16] A. Rahmawati, D. A. I. Maruddani and A. Hoyyi, "STRUCTURAL VECTOR AUTOREGRESSIVE UNTUK ANALISIS DAMPAK SHOCK NILAI TUKAR RUPIAH TERHADAP DOLAR AMERIKA SERIKAT PADA INDEKS HARGA SAHAM GABUNGAN," *Jurnal Gaussian*, vol. 6(3), pp. 291-302., 2017.

- [17] Y. E. Windarto, "ANALISIS PENYAKIT KARDIOVASKULAR MENGGUNAKAN METODE KORELASI PEARSON, SPEARMAN DAN KENDALL," *J. Saintekom*, Vols. 10,no. 2, no. doi: 10.33020/saintekom.v10i2.149., p. 119–127, 2020.
- [18] M. Selvanathan, N. Jayabalan, G. K. Saini, M. Supramaniam and N. Hussain, "Employee Productivity in Malaysian Private Higher Educational Institutions.," *PalArch s Journal of Archaeology of Egypt*, vol. Egyptology 17(8), pp. 66-79, 2020.
- [19] D. N. Gujarati, *Basic Econometrics*, Fourth Edition, New York: The McGraw-Hill, 2004.
- [20] A. H. Al-Nasser and L. T. Abdullah, "The Estimators of Vector Autoregressive Moving Average Model VARMA of Lower Order: VARMA (0,1), ARMA (1,0), VARMA (1,1), VARMA (1,2), VARMA (2,1), VARMA (2,2) with Forecasting," *J. Phys.: Conf. Ser. 1818 012145*, vol. 1818, no. DOI 10.1088/1742-6596/1818/1/012145, 2021.
- [21] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, New York: Springer, 2005.
- [22] H. K. Sungil Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting*, vol. 32(3), no. <https://doi.org/10.1016/j.ijforecast.2015.12.003>, pp. 669-679, 2016.
- [23] U. Khair, H. Fahmi, S. A. Hakim and R. Rahim, "Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error," *Journal of Physics Conference 930(1):012002*, 2017.
- [24] A. A. Suryanto and A. Muqtadir, "PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI," *SAINTEKBU*, vol. 11, no. 1, pp. 78-83, 2019.