

Prediksi *Return* Saham Perbankan dengan Metode LSTM dan Estimasi *Value at Risk* dengan Copula Ali-Mikhail-Haq Menggunakan Korelasi Kendall's Tau

Diva Ryan Mahendra dan Ulil Azmi
Departemen Aktuaria, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: ulil_azmi@its.ac.id

Abstrak—Prediksi *return* saham dan estimasi *Value at Risk* (VaR) adalah hal yang penting dalam pengelolaan portofolio investasi. Metode LSTM telah menunjukkan potensi untuk prediksi harga saham yang akurat, sementara Copula Ali-Mikhail-Haq dengan korelasi Kendall's Tau digunakan untuk mengatasi distribusi non-normal dalam estimasi VaR. Penelitian ini bertujuan untuk menerapkan metode LSTM dalam prediksi *return* saham periode 2021-2023 dan estimasi VaR menggunakan Copula Ali-Mikhail-Haq dengan korelasi Kendall's Tau pada saham perbankan. Berdasarkan hasil penelitian, metode LSTM menunjukkan model terbaik untuk saham BMRI dengan konfigurasi 100 epoch, 64 unit, dan dropout sebesar 0,1 serta menghasilkan MAE sebesar 0,0174. Sedangkan untuk saham BBRI, model terbaik memiliki konfigurasi yang sama tetapi dengan dropout sebesar 0,2 menghasilkan MAE sebesar 0,0222. Secara keseluruhan, model LSTM menunjukkan kemampuan yang baik dalam memprediksi harga dan *return* saham. Dalam estimasi VaR, diperoleh koefisien Kendall's Tau sebesar 0,0676 dan estimasi parameter Copula Ali-Mikhail-Haq sebesar 0,281. Estimasi VaR portofolio saham pada tingkat kepercayaan 99%, 95%, dan 90% berturut-turut adalah -0,03001060; -0,01875786; dan -0,01350327.. Hasil analisis ini memberikan pemahaman yang lebih baik tentang prediksi *return* saham menggunakan LSTM dan estimasi VaR menggunakan Copula Ali-Mikhail-Haq.

Kata Kunci—Asuransi Umum, Bootstrapping, Cadangan IBNR, Mack Chain Ladder.

I. PENDAHULUAN

PASAR modal Indonesia, terutama sektor perbankan, telah menjadi perhatian utama bagi investor dan pelaku pasar dalam beberapa tahun terakhir. Saham-saham perbankan kerap dipilih karena kinerjanya yang relatif stabil dan potensi keuntungan yang menarik. Namun, ketidakstabilan harga saham seringkali menjadi penghalang bagi investor yang menginginkan *return* tinggi dengan risiko rendah [1]. Oleh karena itu, prediksi yang akurat mengenai pergerakan harga saham serta pengukuran risiko investasi menjadi sangat penting bagi pengambilan keputusan investasi yang lebih baik.

Dalam upaya mengatasi ketidakpastian ini, metode deep learning, seperti Long Short Term Memory (LSTM), telah mendapatkan perhatian khusus. LSTM, yang diperkenalkan oleh Hochreiter & Schmidhuber pada tahun 1997, mampu mempelajari dan mengingat hubungan jangka panjang dalam data time series, seperti data historis pasar saham, sehingga memungkinkan prediksi yang lebih akurat [2]. Beberapa penelitian menemukan bahwa LSTM dapat memberikan prediksi yang lebih akurat dibandingkan metode konvensional.

Selain prediksi, pengelolaan risiko investasi juga menjadi fokus penting. *Value at Risk* (VaR) merupakan metode yang

umum digunakan untuk mengukur potensi kerugian maksimum dalam kondisi pasar normal [3]. Namun, fluktuasi harga saham perbankan yang seringkali tidak mengikuti distribusi normal menuntut pendekatan alternatif seperti copula. Copula, khususnya Copula Archimedes seperti Ali-Mikhail-Haq, mampu menggambarkan hubungan struktural antara saham dengan fleksibilitas yang lebih tinggi [4].

Penelitian ini bertujuan untuk menerapkan metode LSTM dalam prediksi *return* saham dan estimasi VaR dengan Copula Ali-Mikhail-Haq pada saham perbankan Indonesia periode 2021-2023. Diharapkan penelitian ini dapat memberikan kontribusi signifikan dalam pengembangan metode prediksi harga saham dan pengukuran risiko investasi, serta memberikan panduan praktis bagi para investor dalam mengelola portofolio investasi di sektor perbankan.

II. KAJIAN PUSTAKA

A. Investasi

Investasi merupakan alokasi sumber daya saat ini untuk memperoleh aset atau barang modal yang diharapkan menghasilkan pendapatan di masa depan, dengan prinsip bahwa risiko saat ini dapat mendatangkan keuntungan lebih besar di masa mendatang. Tujuan utama investasi adalah meningkatkan kesejahteraan investor serta membangun kekayaan jangka Panjang. Namun, setiap keputusan investasi melibatkan risiko yang harus dikelola dengan baik untuk menghindari kerugian finansial [5].

B. Saham Perbankan

Saham perbankan adalah instrumen investasi yang sangat diminati di pasar modal Indonesia, sering kali mengungguli pertumbuhan IHSG. Analisis fundamental, yang menilai tingkat kesehatan bank melalui rasio keuangan, pertumbuhan pendapatan, manajemen risiko, dan kualitas aset, menjadi kunci dalam menilai saham perbankan. Investor mencari bank dengan kinerja kuat dan manajemen baik karena kesehatan bank berdampak signifikan pada perubahan harga saham [6].

C. Return Saham

Metode *Return* saham mengukur perubahan nilai harga saham dari periode t ke periode $t-1$, mencerminkan kinerja historis saham serta menjadi dasar bagi evaluasi portofolio investasi. *Return* saham dibagi menjadi *return* realisasi, yang mencerminkan perubahan aktual nilai saham, dan *return* ekspektasi, yang merupakan proyeksi nilai saham di masa mendatang. Persamaan matematis untuk menghitung *return* saham ditunjukkan oleh Persamaan (1).

Tabel 1.
Struktur Data LSTM Untuk Satu Variabel Saham

t	y_t	y_{t-k_1}	y_{t-k_2}	...	y_{t-k_p}
$k_p + 1$	y_{k_p+1}	$y_{k_p+1-k_1}$	$y_{k_p+1-k_2}$...	y_1
$k_p + 2$	y_{k_p+2}	$y_{k_p+2-k_1}$	$y_{k_p+2-k_2}$...	y_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
n	y_n	y_{n-k_1}	y_{n-k_2}	...	y_{n-k_p}

Tabel 2.
Struktur Copula Ali-Mikhail-Haq

t	u_t	v_t
1	u_1	v_1
2	u_2	v_2
\vdots	\vdots	\vdots
n	u_n	v_n

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right). \quad (1)$$

D. Portofolio Saham

Portofolio saham adalah kumpulan aset yang dimiliki investor, dengan teori portofolio menekankan pentingnya mempertimbangkan risiko dan pengembalian dalam pembentukannya. *Mean Variance Efficient Portfolio* (MVEP) adalah salah satu metode untuk mencari vektor pembobotan agar portofolio memiliki varian minimum dengan persamaan [8]:

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}_N}{\mathbf{1}_N^T \Sigma^{-1} \mathbf{1}_N}. \quad (2)$$

E. Preprocessing Data

Preprocessing data merupakan tahap penting dalam analisis data untuk membersihkan, mempersiapkan, dan mengorganisir data sebelum analisis lebih lanjut. Penanganan *missing value*, seperti menggunakan metode *Moving Average*, mengisi data yang hilang dengan rata-rata dari sejumlah data sebelumnya atau sesudahnya dalam rentang waktu tertentu untuk menjaga stabilitas pola data. *Central Moving Average* dihitung dengan rumus berikut [9]:

$$N_{mt} = \frac{1}{n} \left(\sum_{i=t-(n-1)/2}^{t+(n-1)/2} N_{Qi} \right). \quad (3)$$

Di samping penanganan *missing value*, normalisasi data juga penting untuk menghindari perbedaan rentang nilai antar fitur yang dapat mempengaruhi performa algoritma atau model. Salah satu teknik normalisasi yang umum adalah *Min-Max Normalization*, dengan rumus berikut [10]:

$$x'_i = New_{min} + \left(\frac{x_i - x_{min}}{x_{max} - x_{min}} \right) \times (New_{max} - New_{min}) \quad (4)$$

F. Uji Stasioneritas

Stasioneritas menunjukkan bahwa deret waktu tidak mengalami perubahan signifikan dalam nilai selama periode observasi, dengan fluktuasi data yang stabil. Uji *Augmented Dickey Fuller* (ADF) adalah metode umum untuk menguji stasioneritas, menggunakan Persamaan (5).

$$\nabla Z_t = \theta_0 + \phi Z_{t-1} + \sum_{i=1}^k \delta_i \nabla Z_{t-i} + a_t, \quad (5)$$

dengan menggunakan statistik uji seperti yang ditunjukkan oleh Persamaan (6).

Tabel 3.
Statistika Deskriptif Data

Statistika Deskriptif	Closing Price	
	BMRIJK	BBRIJK
Rata-rata	Rp4.305,77	Rata-rata
Standar Deviasi	1001,45	Standar Deviasi
Nilai Minimum	Rp2.825,00	Nilai Minimum
Q1	Rp3.484,37	Q1
Median	Rp4.100,00	Median
Q3	Rp5.115,62	Q3
Nilai Maksimum	Rp6.125,00	Nilai Maksimum

Tabel 4.
Hasil Uji ADF Sebelum Differencing

Saham	P-Value
BMRIJK	0,7765456895735311
BBRIJK	0,9356197563055578

Tabel 5.
Hasil Uji ADF Setelah Differencing

Saham	P-Value
BMRIJK	0
BBRIJK	0

$$T_{hit} = \frac{\hat{\phi} - 1}{se(\hat{\phi})}. \quad (6)$$

Hipotesis nol ditolak jika $T_{hit} >$ titik kritis pada tabel Dickey Fuller atau jika nilai *p-value* $< \alpha$, menunjukkan bahwa data stasioner dalam rata-rata.

G. Autocorrelation Function (ACF)

Autocorrelation Function (ACF) mengukur korelasi antara data pada waktu ke- t dengan data sebelumnya, menggambarkan kovarians dan korelasi antar pengamatan. Nilai ACF dihitung dengan persamaan berikut [11]:

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}}. \quad (7)$$

Jika plot ACF menunjukkan nilai lag turun cepat mendekati nol, ini menandakan data stasioner. Jika data tidak stasioner, diferensiasi digunakan untuk menghitung selisih antara nilai observasi pada periode waktu tertentu. Persamaan diferensiasi adalah [12]:

$$Z_t^{(k)} = \nabla^k Z_t = (1 - B)^k Z_t = Z_t - Z_{t-k}. \quad (8)$$

H. Partial Autocorrelation Function (PACF)

Partial Autocorrelation Function (PACF) mengukur korelasi antara pengamatan pada waktu tertentu setelah menghilangkan pengaruh dari pengamatan sebelumnya, penting untuk memahami hubungan langsung antar pengamatan pada interval waktu tertentu. Koefisien PACF dihitung dengan persamaan berikut [11]:

$$P_k = \frac{Cov(Z_t - \hat{Z}_t, Z_{t+k} - \hat{Z}_{t+k})}{\sqrt{Var(Z_t - \hat{Z}_t)}\sqrt{Var(Z_{t+k} - \hat{Z}_{t+k})}}. \quad (9)$$

Lag PACF yang melewati batas confident interval menunjukkan korelasi signifikan, digunakan dalam olah data.

I. Artificial Neural Network (ANN)

Artificial Neural Network (ANN) adalah model komputasi yang terinspirasi dari cara kerja jaringan saraf biologis manusia, digunakan untuk aplikasi seperti pengenalan pola dan prediksi. ANN terdiri dari *input layer*, *hidden layer*, dan *output layer*, di mana data diproses secara bertahap melalui

Tabel 6.
Input Data Saham BMRI.JK

Saham	Lag Signifikan	Input
BMRI	1	Z_{t-1}, Z_{t-2}
	2	$Z_{t-1}, Z_{t-2}, Z_{t-3}$
	32	$Z_{t-1}, Z_{t-32}, Z_{t-33}$
	57	$Z_{t-1}, Z_{t-57}, Z_{t-58}$
	1,2	$Z_{t-1}, Z_{t-2}, Z_{t-3}$
	1,32	$Z_{t-1}, Z_{t-2}, Z_{t-32}, Z_{t-33}$
	1,57	$Z_{t-1}, Z_{t-2}, Z_{t-57}, Z_{t-58}$
	2,32	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-32}, Z_{t-33}$
	2,57	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-57}, Z_{t-58}$
	32,57	$Z_{t-1}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$
	1,2,32	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-32}, Z_{t-33}$
	1,2,57	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-57}, Z_{t-58}$
	1,32,57	$Z_{t-1}, Z_{t-2}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$
	2,32,57	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$
	1,2,32,57	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$
BBRI	1	Z_{t-1}, Z_{t-2}
	2	$Z_{t-1}, Z_{t-4}, Z_{t-5}$
	32	$Z_{t-1}, Z_{t-11}, Z_{t-12}$
	57	$Z_{t-1}, Z_{t-29}, Z_{t-30}$
	1,2	$Z_{t-1}, Z_{t-2}, Z_{t-4}, Z_{t-5}$
	1,32	$Z_{t-1}, Z_{t-2}, Z_{t-11}, Z_{t-12}$
	1,57	$Z_{t-1}, Z_{t-2}, Z_{t-29}, Z_{t-30}$
	2,32	$Z_{t-1}, Z_{t-4}, Z_{t-5}, Z_{t-11}, Z_{t-12}$
	2,57	$Z_{t-1}, Z_{t-4}, Z_{t-5}, Z_{t-29}, Z_{t-30}$
	32,57	$Z_{t-1}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$
	1,2,32	$Z_{t-1}, Z_{t-2}, Z_{t-4}, Z_{t-5}, Z_{t-11}, Z_{t-12}$
	1,2,57	$Z_{t-1}, Z_{t-2}, Z_{t-4}, Z_{t-5}, Z_{t-29}, Z_{t-30}$
	1,32,57	$Z_{t-1}, Z_{t-2}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$
	2,32,57	$Z_{t-1}, Z_{t-4}, Z_{t-5}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$
	1,2,32,57	$Z_{t-1}, Z_{t-2}, Z_{t-4}, Z_{t-5}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$

tiap lapisan dengan bobot yang dihitung secara matematis. ANN seperti *Multi-Layer Perceptron (MLP)*, *Convolutional Neural Network (CNN)*, dan *Recurrent Neural Network (RNN)* dapat digunakan untuk berbagai tugas prediksi berdasarkan pola data historis [13].

J. *Recurrent Neural Network (RNN)*

Recurrent Neural Network (RNN) dirancang khusus untuk memproses data berurutan, seperti data deret waktu. Dengan kemampuannya dalam mempertahankan dan menggunakan informasi dari masa lalu untuk memprediksi informasi di masa depan, RNN sangat berguna dalam aplikasi forecasting dan pengolahan data time series. Varian RNN termasuk *Simple RNN*, *Long Short Term Memory (LSTM)*, dan *Gated Recurrent Unit (GRU)*, masing-masing dengan keunggulan dan kelemahan dalam pemrosesan data berurutan. Persamaan matematis yang digunakan dalam RNN adalah sebagai berikut [14]:

$$h_t = \tanh(W_h h_{t-1} + W_x x_t + b_x + b_h). \quad (10)$$

Meski RNN berguna untuk data berurutan, ia memiliki kelemahan dalam memproses data jangka panjang, seperti masalah *vanishing* dan *exploding gradient*.

K. *Long Short Term Memory (LSTM)*

Long Short Term Memory (LSTM) adalah jenis RNN yang dikembangkan khusus untuk mengatasi kendala utama dalam proses pembelajaran mesin, terutama dalam memproses data sekuensial. LSTM muncul sebagai turunan dari RNN dengan tujuan untuk mengatasi kendala utama RNN seperti *vanishing gradient*, di mana gradien-nilai yang mendekati nol dapat menyebabkan masalah dalam proses pembelajaran. Pada bagian *hidden layer*, LSTM memiliki *recurrent hidden layer* yang mengandung *memory cell* serta tiga unit *gate*, yaitu *input gate*, *forget gate*, dan *output gate*. Ketiga *gate* ini

berperan dalam membaca, menyimpan, dan memperbarui informasi yang dilewatkan dalam proses pembelajaran [15]. Berikut adalah rumus-rumus yang diperkenalkan oleh [16]:

1) Rumus untuk *Forget Gate* f_t :

$$f_t = \sigma(W_f \cdot x_t + W_h \cdot h_{t-1} + b_f) \quad (11)$$

2) Rumus untuk *Input Gate* i_t :

$$i_t = \sigma(W_i \cdot x_t + W_h \cdot h_{t-1} + b_i) \quad (12)$$

3) Rumus untuk *Candidate State* \tilde{C} :

$$\tilde{C} = \tanh(W_C \cdot x_t + W_h \cdot h_{t-1} + b_C) \quad (13)$$

4) Rumus untuk *Update Cell State* C_t :

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C} \quad (14)$$

5) Rumus untuk *Output Gate* O_t :

$$O_t = \sigma(W_o \cdot x_t + W_h \cdot h_{t-1} + b_o) \quad (15)$$

6) Rumus untuk *Hidden State* h_t :

$$h_t = O_t \times \tanh(C_t) \quad (16)$$

dengan struktur yang lebih kompleks, LSTM dapat mempertahankan informasi jangka panjang dengan lebih baik dibandingkan RNN konvensional.

L. *Hyperparameter*

Hyperparameter adalah parameter yang ditentukan pada awal pembentukan model *Neural Network* dan tetap tidak berubah selama pelatihan. Optimasi *hyperparameter* bertujuan untuk menemukan kombinasi nilai yang tepat guna mencapai kinerja maksimum. Beberapa hyperparameter penting termasuk *epoch*, *iterasi*, *batch size*, *hidden layer*, dan *dropout*. *Tuning hyperparameter* memegang peranan penting dalam meningkatkan akurasi model dan mencegah *overfitting* [17].

Tabel 7.
Contoh Normalisasi Data

BMRI		BBRI	
Sebelum Normalisasi	Setelah Normalisasi	Sebelum Normalisasi	Setelah Normalisasi
Rp3.187,50	0,109848	Rp4.245,38	0,370999
Rp3.162,50	0,102273	Rp4.209,01	0,355541
Rp3.075,00	0,075758	Rp4.245,38	0,370999

Tabel 8.
Hasil Pemodelan BMRI

Epoch	Unit	Dropout	Input	MAE
100	64	0,1	$Z_{t-1}, Z_{t-2}, Z_{t-57}, Z_{t-58}$	0,0174
50	16	0,2	Z_{t-1}, Z_{t-2}	0,017
⋮	⋮	⋮	⋮	⋮
100	8	0,05	$Z_{t-1}, Z_{t-2}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$	0,159
50	8	0,1	$Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-32}, Z_{t-33}, Z_{t-57}, Z_{t-58}$	0,170

M. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) digunakan sebagai metrik evaluasi untuk mengukur seberapa baik model *Neural Network* dalam memprediksi nilai yang sebenarnya. MAE merupakan hasil rata-rata dari selisih absolut antara nilai yang diprediksi oleh model dan nilai yang sebenarnya. Semakin kecil nilai MAE, semakin akurat model dalam memprediksi nilai yang sebenarnya. Persamaan untuk MAE adalah [18]:

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \tag{17}$$

N. Copula

Copula adalah fungsi distribusi bersama yang diperoleh dari penggabungan marginal yang berdistribusi uniform [0,1]. Salah satu keunggulannya adalah kemampuannya menunjukkan pola sebaran data pada tail distribusi tanpa asumsi distribusi normal. Copula juga fleksibel dalam menggambarkan hubungan struktural antara saham, yang cocok untuk data *return* keuangan yang tidak berdistribusi normal [4].

Copula Archimedes adalah kelompok copula yang sering digunakan, dan salah satunya adalah copula Ali-Mikhail-Haq. Secara umum, Copula Archimedes dinyatakan dengan rumus:

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v)). \tag{18}$$

Dalam konteks pengukuran dependensi, korelasi Kendall's Tau menjadi pilihan yang populer. Korelasi Kendall's Tau dapat dihitung menggunakan rumus berikut data [19]:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}. \tag{19}$$

Estimasi parameter copula Archimedes melalui tau Kendall dapat dihitung langsung dari fungsi generator copula, seperti pada persamaan [20]:

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt. \tag{20}$$

O. Copula Ali-Mikhail-Haq (AMH)

Copula Ali-Mikhail-Haq (AMH) adalah varian dari Copula Archimedes yang memungkinkan penggambaran dependensi negatif dan positif dengan parameter θ yang bervariasi dari -1 hingga 1 [21]. AMH copula cocok untuk hubungan korelasi



Gambar 1. Grafik perbandingan data aktual dan prediksi harga saham BMRI.

Tabel 9.
Hasil Pemodelan BBRI

Epoch	Unit	Dropout	Input	MAE
100	64	0,2	Z_{t-1}, Z_{t-2}	0,0222
75	64	0,1	Z_{t-1}, Z_{t-2}	0,0225
⋮	⋮	⋮	⋮	⋮
75	8	0,05	$Z_{t-1}, Z_{t-2}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$	0,1140
100	16	0,05	$Z_{t-1}, Z_{t-11}, Z_{t-12}, Z_{t-29}, Z_{t-30}$	0,1201

rendah dan dapat menangkap dependensi kompleks. Bentuk matematis Copula AMH:

$$C_{(u,v)} = \frac{uv}{1 - \theta(1-u)(1-v)}. \tag{21}$$

Fungsi generator Copula AMH adalah:

$$\phi_{\theta}(t) = \ln \left(\frac{(1 - \theta(1-t))}{t} \right). \tag{22}$$

P. Monte Carlo

Metode Monte Carlo adalah teknik yang digunakan untuk menghasilkan data acak guna memperkirakan distribusi parameter risiko keuangan. Metode ini sangat fleksibel dan dapat diterapkan pada berbagai distribusi multivariat, memungkinkan penggambaran ketergantungan antar variabel dengan lebih akurat [1].

Q. Value at Risk (VaR)

Value at Risk (VaR) adalah ukuran statistik yang mengindikasikan kerugian maksimum yang mungkin dialami oleh suatu portofolio investasi pada tingkat kepercayaan tertentu selama periode waktu tertentu [1]. Dihitung dengan:

$$P(r \leq VaR) = 1 - \alpha. \tag{23}$$

Persamaan (23) menunjukkan bahwa VaR adalah nilai dimana probabilitas *return* r kurang dari atau sama dengan VaR adalah $1 - \alpha$. Rumus umum untuk menghitung VaR adalah:

$$VaR = W_0 \cdot R^*, \tag{24}$$

dimana W_0 adalah nilai investasi awal dan R^* adalah kuantil distribusi *return* pada tingkat kepercayaan $1 - \alpha$. Sedangkan rumus untuk menghitung VaR portofolio Monte Carlo adalah:

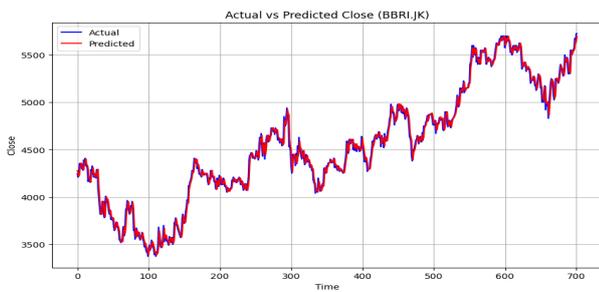
$$VaR(c, T) = E(F_T) - Q((F_T, c), \tag{25}$$

dimana c adalah tingkat kepercayaan dan F_T adalah portofolio hasil simulasi.

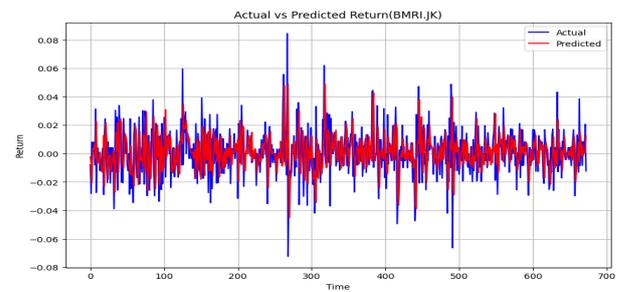
III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapatkan melalui *finance.yahoo.com* berupa data harga penutupan saham (*closing price*) harian dalam



Gambar 2. Grafik perbandingan data aktual dan prediksi harga saham BBRI.



Gambar 3. Grafik perbandingan data aktual dan prediksi return BMRI.

satuan rupiah dari Januari 2021 sampai dengan Desember 2023. Saham 1 adalah BMRI.JK dan saham 2 adalah BBRI.JK.

B. Variabel

Variabel-variabel yang digunakan dalam penelitian adalah data *closing price* saham dan input *lag* saham disajikan dalam Tabel 1. Variabel y_{t-k_p} adalah nilai lag yang didapatkan berdasarkan analisis grafik PACF dengan k_i sebagai lag yang signifikan dengan i adalah 1, 2, 3, hingga p dan $k_1 < k_2 < \dots < k_p$. Variabel n adalah jumlah pengamatan dalam dataset. Dalam model LSTM, y_t adalah variabel respons, sedangkan y_{t-k_1} , y_{t-k_2} , hingga y_{t-k_p} adalah variabel prediktor. Sedangkan, struktur data yang digunakan dalam pemodelan metode Copula Ali-Mikhail-Haq dimuat dalam Tabel 2.

Variabel u mewakili residual dari saham 1, sementara v merupakan residual dari saham 2 yang diperoleh melalui proses prediksi menggunakan metode Long Short-Term Memory (LSTM).

C. Tahapan Penelitian

Langkah-langkah yang dilakukan adalah sebagai berikut:

1) *Studi Literatur*

Teori, konsep, dan penelitian terdahulu tentang prediksi harga saham, estimasi VaR, metode LSTM, dan metode Copula Ali-Mikhail-Haq dikumpulkan untuk jadi referensi.

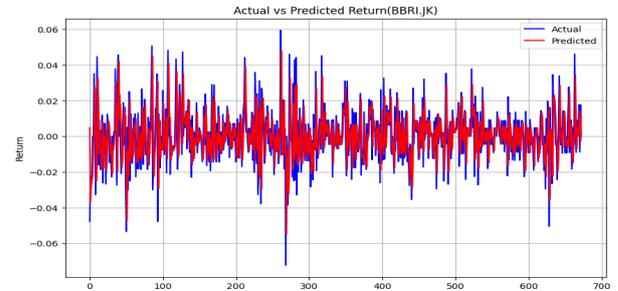
2) *Pengumpulan Data*

Data historis harga saham perbankan dikumpulkan dari sumber terpercaya seperti Yahoo Finance.

3) *Pra-pemrosesan Data*

Memproses terlebih dahulu data agar menjadi data yang siap diproses dalam model:

1. Analisis statistika deskriptif dilakukan untuk memahami karakteristik data.
2. Nilai yang hilang diisi menggunakan metode *Central Moving Average* dengan $n=5$.
3. Uji stasioneritas dan differencing menggunakan metode Augmented Dickey-Fuller (ADF) atau melalui analisis plot *Autocorrelation Function* (ACF) untuk memastikan stasioneritas data. Jika diperlukan, dilakukan proses *differencing*.
4. Identifikasi pola signifikan lag PACF dengan analisis Partial Autocorrelation Function (PACF) untuk dijadikan input penelitian.
5. Normalisasi data menggunakan teknik Min-Max Normalization untuk menstandarisasi rentang nilai data ke dalam interval [0,1].



Gambar 4. Grafik perbandingan data aktual dan prediksi return BBRI.

6. Pembagian data dengan memisahkan data menjadi dua bagian, yaitu *data training* (80%) dan *data testing* (20%), sesuai dengan praktek umum dalam pemodelan prediktif.

4) *Pemodelan LSTM*

Model dibentuk dengan *input lag return* saham 1 dengan *output closing price* saham 1 dan *input lag closing price* saham 2 dengan *output closing price* saham 2:

1. Menentukan hyperparameter yang optimal untuk model LSTM, seperti jumlah neuron, jumlah layer, dan fungsi aktivasi.
2. Membangun arsitektur model LSTM menggunakan data training untuk menghasilkan prediksi harga saham berdasarkan lag-lag *return* saham sebelumnya.

5) *Evaluasi Model*

Data testing digunakan untuk mengevaluasi kinerja model LSTM. Salah satu metrik evaluasi yang umum digunakan adalah *Mean Absolute Error* (MAE), yang mengukur rata-rata kesalahan prediksi model terhadap nilai aktual.

6) *Prediksi Return dan Residual*

Prediksi *return* saham dibuat dari data prediksi harga saham yang dihasilkan oleh model LSTM. Kemudian, residual didapatkan dari selisih antara nilai prediksi dan nilai aktual.

7) *Pembobotan dengan MVEP*

Melakukan pembobotan portfolio menggunakan metode *Mean Variance Efficient Portfolio* (MVEP) untuk mencari kombinasi bobot saham yang menghasilkan portofolio dengan risiko minimum pada tingkat *return* yang diharapkan.

8) *Pembangunan Model Copula AMH*

1. Uji distribusi normalitas terhadap residual saham menggunakan metode Kolmogorov-Smirnov.
2. Korelasi atau dependensi antar-variabel didapatkan dengan menggunakan koefisien Kendall's Tau.
3. Estimasi parameter Copula Ali-Mikhail-Haq didapatkan berdasarkan koefisien Kendall's Tau.

Tabel 10.
Hasil Prediksi Return

Tanggal	Prediksi Return BMRI	Prediksi Return BBRI
2021-03-30	-0,00229	0,00476
2021-03-31	-0,00726	-0,03714
⋮	⋮	⋮
2023-12-22	0,00074	-0,00280
2023-12-27	0,01246	0,01134

Tabel 11.
Hasil Residual Return

Tanggal	Residual BMRI	Residual BBRI
2021-03-30	0,00558	0,05249
2021-03-31	0,02079	0,01467
⋮	⋮	⋮
2023-12-22	0,01987	0,02043
2023-12-27	0,02478	0,01134

9) *Estimasi Value at Risk*

1. Distribusi *return* saham disimulasikan menggunakan metode Monte Carlo berdasarkan parameter dari model Copula Ali-Mikhail-Haq.
2. Menghitung VaR pada tingkat kepercayaan yang ditentukan (misalnya 90%, 95%, dan 99%) untuk mengevaluasi potensi kerugian maksimum yang mungkin terjadi pada portofolio investasi.

10) *Penarikan Kesimpulan*

Hasil analisis akan dibuat kesimpulan dan saran.

IV. ANALISIS DAN PEMBAHASAN

A. *Karakteristik Data*

Data yang digunakan adalah data harga penutupan BMRI.JK dan BBRI.JK, dari Januari 2021 hingga Desember 2023. Harga penutupan saham BMRI menunjukkan tren yang dinamis selama periode tersebut. Harga saham mengalami penurunan dari awal tahun 2021 hingga pertengahan 2022, kemudian mengalami pemulihan yang kuat hingga akhir 2023. Sedangkan pada harga penutupan saham BBRI juga menunjukkan tren yang serupa dengan BMRI. Harga saham BBRI mengalami penurunan pada awal periode, namun kemudian mengalami pemulihan dan menunjukkan tren kenaikan yang stabil hingga akhir 2023.

B. *Preprocessing Data*

Data yang digunakan telah melewati tahap pra-pemrosesan untuk mempersiapkannya sebelum dimasukkan ke dalam model.

1) *Eksplorasi Data*

Data akan terlebih dahulu di cari statistika deskriptifnya dan didapatkan hasil seperti pada Tabel 3. Statistik deskriptif dari kedua saham menunjukkan variasi harga yang signifikan selama periode observasi. BMRI memiliki volatilitas yang lebih besar dibandingkan BBRI. Distribusi harga penutupan saham BMRI lebih melebar, sedangkan BBRI cenderung memiliki distribusi harga yang lebih terpusat di sekitar nilai median.

2) *Uji Stasioneritas*

Selanjutnya dilakukan uji stasioneritas dengan Uji ADF dan didapatkan hasil yang dimuat dalam Tabel 4. Dari Tabel 4, terlihat bahwa p-value untuk BMRI adalah 0,9356 dan untuk BBRI adalah 0,7765. Kedua nilai ini lebih besar dari taraf signifikansi 0,05, yang menunjukkan bahwa data tidak stasioner. Setelah differencing 1 kali, didapatkan hasil Uji ADF yang ditunjukkan dalam Tabel 5. P-value sudah 0 berarti data sudah stasioner.

3) *Pemilihan Input Model*

Input model LSTM berasal dari lag-lag yang signifikan dari data yang telah mengalami beberapa pengolahan

sebelumnya. Oleh karena itu, diperlukan identifikasi lag signifikan dari PACF data tersebut. Pada saham BMRI.JK lag yang signifikan adalah lag 1, 2, 32 dan 57. Sedangkan pada saham BBRI.JK lag yang signifikan adalah lag 1, 4, 11, dan 29. Untuk mendapatkan input, lag tersebut dimasukkan ke fungsi autoregression :

$$a_t = \phi_1(B)\nabla kZ_t \tag{26}$$

Setelah itu, didapatkan input BMRI sebagai yang disajikan oleh Tabel 6. Dengan demikian, input-input yang di dapat bisa digunakan ke dalam pembangunan model yang akurat.

4) *Normalisasi Data*

Setelah pemilihan input dilakukan, langkah selanjutnya dalam persiapan data adalah normalisasi. Normalisasi data *closing price* saham dilakukan menggunakan metode min-max normalization dengan rentang antara 0 dan 1. Contoh data hasil normalisasi ditunjukkan oleh Tabel 7.

5) *Pembagian Data Train dan Test*

Pembagian data menjadi data training dan data testing. Pembagian ini dilakukan dengan proporsi 80:20 atau 80% *training* 20% *testing*. Dengan demikian, data *training* BMRI sejumlah 539 sedangkan BBRI sejumlah 561. Sedangkan *data test* BMRI adalah 135 dan BBRI adalah 141.

C. *Pembuatan Model LSTM*

Setelah *preprocessing* data selesai, data yang telah dipersiapkan kemudian dapat diproses untuk pembuatan model LSTM.

1) *Hyperparameter*

Dalam pembuatan model menggunakan LSTM, pengaturan *hyperparameter* penting untuk membentuk model. Pada penelitian ini digunakan *hyperparameter* berupa epoch sejumlah 50, 75, 100; unit sejumlah 8, 16, 32, 64; *batch size* sejumlah 32; dan *dropout* sejumlah 5%, 10%, 20%.

2) *Hasil Pemodelan*

Dengan input yang telah ditentukan sebelumnya dan kombinasi hyperparameter yang dipilih, dilakukan pemodelan LSTM untuk memprediksi harga saham. Hasil dari pemodelan tersebut disajikan dalam Tabel 8.

Model LSTM terbaik untuk BMRI memiliki 100 epoch, 64 unit, dan dropout sebesar 0,1 dengan variabel input $Z_{t-1}, Z_{t-2}, Z_{t-57}, Z_{t-58}$ yang menghasilkan Mean Absolute Error (MAE) sebesar 0,0174. Selain itu, hyperparameter seperti Unit dan dropout juga berpengaruh terhadap hasil model. Semakin banyak unit akan semakin kecil MAE yang dihasilkan. Semakin besar dropout maka akan semakin kecil MAE yang dihasilkan. Sedangkan hasil keakuratan prediksi dengan model terbaik pada *closing price* saham BMRI dapat dilihat pada grafik Gambar 1.

Secara umum, garis prediksi (merah) sangat dekat dengan garis aktual (biru), menunjukkan bahwa model LSTM memiliki kinerja yang baik dalam memprediksi harga penutupan saham BMRI. Ini berarti model mampu

Tabel 12.
Hasil Uji Kolmogorov-Smirnov

Saham	P-Value	Dhitung
BMRI	$< 2,2 \times 10^{-16}$	0,5
BBRI	$< 2,2 \times 10^{-16}$	0,5

Tabel 13.
Hasil Value at Risk

Tingkat kepercayaan	Value at Risk
99%	-0,03001060
95%	-0,01875786
90%	-0,01350327

menangkap pola tren dan fluktuasi harga dengan akurasi yang cukup tinggi. Selanjutnya untuk hasil pemodelan saham BBRI dimuat dalam Tabel 9. Pemodelan LSTM untuk saham BBRI menunjukkan hasil yang sedikit lebih variatif dibandingkan dengan BMRI. Model terbaik memiliki 100 epoch, 64 unit, dan dropout sebesar 0,2 dengan variabel input Z_{t-1} , Z_{t-2} yang menghasilkan MAE sebesar 0,0222. Meskipun demikian, variasi dalam MAE tidak terlalu signifikan di antara beberapa konfigurasi model. Sedangkan keakuratannya ditunjukkan oleh Gambar 2.

Secara keseluruhan, garis prediksi sangat dekat dengan garis aktual, menandakan bahwa model LSTM mampu memprediksi harga penutupan saham BBRI dan mampu menangkap pola tren dan fluktuasi dengan akurasi tinggi.

D. Prediksi Return dan Residual

Return saham dihitung dari prediksi harga saham menggunakan persamaan yang telah ada sebelumnya, sedangkan residual didapatkan dari selisih antara nilai aktual dan nilai prediksi.

1) Prediksi Return

Hasil prediksi return saham BMRI dan BBRI dari harga penutupan yang diprediksi sebelumnya dengan model LSTM dapat dilihat dalam Tabel 10. Hasil prediksi return saham BMRI menunjukkan variasi yang cukup signifikan antara periode-periode yang berbeda. Seperti pada prediksi return BMRI, hasil prediksi return BBRI juga menunjukkan variasi nilai antara periode-periode yang berbeda. Sedangkan keakuratannya dapat dilihat pada plot oleh Gambar 3 dan Gambar 4.

Terlihat bahwa keakuratan model dalam prediksi return dapat dikatakan akurat karena garis merah (return prediksi) bergerak tidak jauh dari garis biru (return aktual). Selanjutnya data return akan dicari residualnya untuk kebutuhan pengolahan data berikutnya, yaitu pemodelan Copula Ali-Mikhail-Haq.

2) Residual Return

Tabel 11 memuat hasil residual return untuk kedua saham. Hasil residual yang lebih rendah menunjukkan fluktuasi yang lebih kecil.

E. Uji Kolmogorov Smirnov Residual

Dilakukan uji normalitas pada residual return menggunakan uji Kolmogorov-Smirnov. Hasilnya disajikan oleh Tabel 12. Hasil p-value dari kedua residual $< 0,05$ dan

Dhitung $> 0,052$. Oleh karena itu, keputusannya adalah tolak H_0 artinya residual return LSTM tidak berdistribusi normal.

F. Uji Kendall's Tau

Sebelum melakukan pemodelan Copula, dilakukan uji korelasi menggunakan Kendall's Tau. Hasilnya adalah 0,0676. Hasil itu lah yang akan digunakan dalam Copula.

G. Estimasi Parameter Copula Ali-Mikhail-Haq

Setelah koefisien Kendall's Tau dimasukkan ke dalam rumus, didapatkan estimasi parameternya adalah 0,281. Dari parameter tersebut didapatkan persamaan berikut :

$$C_{0,281}^{AMH}(u, v) = \frac{uv}{1 - 0,281(1-u)(1-v)} \quad (27)$$

Selanjutnya, persamaan tersebut akan digunakan dalam pencarian estimasi VaR dari portofolio kedua saham.

H. Mean Variance Efficient Portfolio

Metode Mean Variance Efficient Portfolio digunakan dalam pembobotan saham dalam pembentukan portofolio. Hasil pembobotannya didapatkan bobot saham BMRI sebesar 0,46 dan BBRI sebesar 0,54.

I. Estimasi Value at Risk

Selanjutnya akan dicari estimasi Value at Risk menggunakan simulasi Monte Carlo. Simulasi Monte Carlo dilakukan sebanyak 1000 kali untuk pada tiga tingkat kepercayaan, yaitu 1%, 5%, dan 10%. Hasilnya dimuat dalam Tabel 13.

Tabel 13 menunjukkan bahwa nilai VaR portofolio saham pada tingkat kepercayaan 99%, 95%, dan 90% berturut-turut sebesar -0,03001060; -0,01875786; dan -0,01350327. Jika investasi awal sebesar Rp100.000.000,00, nilai VaR pada tingkat kepercayaan 99%, 95%, dan 90% adalah Rp3.001.060,00; Rp1.875.786,00; dan Rp1.350.327,00 untuk 1 hari kedepan.

V. KESIMPULAN

A. Kesimpulan

Berdasarkan hasil pembahasan, dapat disimpulkan: (1) Model terbaik memiliki konfigurasi 100 epoch, 64 unit, dan dropout sebesar 0,1 dengan variabel input ['X1', 'X2', 'X6', 'X7'], menghasilkan Mean Absolute Error (MAE) sebesar 0,0174. Sedangkan pada BBRI model terbaik memiliki konfigurasi 100 epoch, 64 unit, dan dropout sebesar 0,2 dengan variabel input ['X1', 'X2'], menghasilkan MAE sebesar 0,0222. (2) Korelasi Kendall's Tau yang diperoleh sebesar 0,0676 dan estimasi parameter model Copula Ali-Mikhail-Haq menghasilkan estimasi parameter sebesar 0,281 dengan persamaan $C_{0,281}^{AMH}(u_1, u_2) = \frac{u_1 u_2}{1 - 0,281(1-u_1)(1-u_2)}$. (2) Nilai VaR portofolio saham pada tingkat kepercayaan 99%, 95%, dan 90% berturut-turut sebesar -0,03001060; -0,01875786; dan -0,01350327.

B. Saran

Saran untuk penelitian selanjutnya adalah sebagai berikut: (1) Mencoba konfigurasi parameter dan kombinasi variabel input lainnya. (2) Prediksi harga dan return saham dapat dilakukan dengan berbagai metode lain. (3) Penelitian lebih lanjut dapat dilakukan dengan

menggunakan data yang lebih banyak. (4) Menggunakan metode Copula jenis lainnya.

DAFTAR PUSTAKA

- [1] B. S. D., *Pasar Modal: Mekanisme Perdagangan Saham, Analisis Sekuritas, dan Strategi Investasi di Bursa Efek Indonesia (BEI)*. Yogyakarta: UPP STIM YKPN, ISBN: 978-979-3532-38-7, 2009.
- [2] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Phys. A Stat. Mech. its Appl.*, vol. 519, no. C, pp. 127–139, Apr. 2019, doi: 10.1016/j.physa.2018.11.061.
- [3] P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk 3rd Edition*. USA: McGraw Hill, ISBN: 978-0071464956, 2006.
- [4] L. G. Udayani, I. W. Sumarjaya, and M. Susilawati, "Analisis hubungan produksi padi dan indikator Enso di Kabupaten Tabanan dengan pendekatan Copula," *E-Jurnal Mat.*, vol. 5, no. 4, p. 164, Nov. 2016, doi: 10.24843/MTK.2016.v05.i04.p136.
- [5] Afriyeni and D. Marlius, "Analisis Faktor-Faktor yang Berpengaruh Terhadap Ketepatan Waktu Penyampaian Laporan Keuangan Pada Perusahaan yang Listing di Bursa Efek Indonesia," Padang, OSF Preprints, 2019.
- [6] A. W. P. Setyawan, "Pengaruh Komponen Risk Based Bank Rating Terhadap Harga Saham Perusahaan Perbankan yang Go Public di Bursa Efek Indonesia (BEI) Tahun 2008-2011," Universitas Diponegoro, Semarang, 2012.
- [7] M. M. Hanafi and A. Halim, *Analisis Laporan Keuangan*, 4th ed. Yogyakarta: UPP STIM YKPN, ISBN: 978-8170-43-1, 2012.
- [8] D. A. I. Maruddani and A. Purbowati, "Pengukuran Value at Risk pada aset tunggal dan portofolio dengan simulasi Monte Carlo," *MEDIA Stat.*, vol. 2, no. 2, pp. 93–104, Mar. 2012, doi: 10.14710/medstat.2.2.93-104.
- [9] R. Alhabib and A. A. Salama, "Using moving averages to pave the neutrosophic time series," *Int. J. Neutrosophic Sci.*, vol. 3, no. 1, pp. 14–20, 2020.
- [10] W. li and Z. Liu, "A method of SVM with normalization in intrusion detection," *Procedia Environ. Sci.*, vol. 11, pp. 256–262, 2011, doi: 10.1016/j.proenv.2011.12.040.
- [11] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed. USA: Pearson Addison Wesley, ISBN: 0-321-32216-9, 2006.
- [12] F. C. Garini and W. Anbiya, "Application of GARCH forecasting method in predicting the number of rail passengers (Thousands of people) in Jabodetabek region," *J. Mat. Stat. dan Komputasi*, vol. 18, no. 2, pp. 198–223, Jan. 2022, doi: 10.20956/j.v18i2.18382.
- [13] J. Singh and R. Banerjee, "A Study on Single and Multi-layer Perceptron Neural Network," in *3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Mar. 2019, pp. 35–40, doi: 10.1109/ICCMC.2019.8819775.
- [14] Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, "Recurrent neural networks for multivariate time series with missing values," *Sci. Rep.*, vol. 8, no. 1, p. 6085, Apr. 2018, doi: 10.1038/s41598-018-24271-9.
- [15] Rowan, L. Muflikhah, and I. Cholissodin, "Peramalan kasus positif COVID-19 di Jawa Timur menggunakan metode hybrid ARIMA-LSTM," *J. Pengemb. Teknol. Inf dan Ilmu Komput.*, vol. 6, no. 9, pp. 4146–4153, 2022.
- [16] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. C. Khatri, "Predicting stock market index using LSTM," *Mach. Learn. with Appl.*, vol. 9, no. 2, p. 100320, Sep. 2022, doi: 10.1016/j.mlwa.2022.100320.
- [17] A. Khumaidi and I. A. Nirmala, *Algoritma Long Short Term Memory dengan Hyperparameter Tuning: Prediksi Penjualan Produk*. Yogyakarta: Deepublish, ISBN: 978-623-02-4913-6, 2022.
- [18] N. Ahmad, Y. Ghadi, M. Adnan, and M. Ali, "Load forecasting techniques for fower system: Research challenges and survey," *IEEE Access*, vol. 10, pp. 71054–71090, 2022, doi: 10.1109/ACCESS.2022.3187839.
- [19] R. B. Nelsen, *An Introduction to Copulas*, 2nd ed. New York: Springer, ISBN: 978-0387-28659-4, 2006.
- [20] E. Luxviantono, A. Setiawan, and L. R. Sasongko, "Analisis hubungan IHK (Indeks harga konsumen) dan kurs beli IDR-USD melalui pendekatan Copula," *d'CARTESIAN J. Mat. dan Apl.*, vol. 7, no. 2, pp. 52–58, 2018.
- [21] P. Kumar, "Probability distributions and estimation of Ali-Mikhail-Haq copula," *Appl. Math. Sci.*, vol. 4, no. 14, pp. 657–666, 2010.