

Paper

Implementasi Algoritma Gated Recurrent Unit Dalam Melakukan Prediksi Harga Kelapa Sawit Dengan Memanfaatkan Model Recurrent Neural Network (RNN)

Author: Alwi Aulia, Andi Marwan Elhanafi, Haida Dafitri

Implementasi Algoritma Gated Recurrent Unit Dalam Melakukan Prediksi Harga Kelapa Sawit Dengan Memanfaatkan Model Recurrent Neural Network (RNN)

Alwi Aulia¹, Andi Marwan Elhanafi², Haida Dafitri³

^{1,2,3}Universitas Harapan, Medan, Indonesia

¹alwiaulia92@gmail.com, ²elh645@gmail.com, ³aida.stth@gmail.com

Abstrak- Kelapa sawit atau Palm Oil (PO) merupakan salah satu eksportir unggulan Indonesia, salah satu produk perkebunan. Harga kelapa sawit sering mengalami kenaikan harga dan penurunan harga yang fluktuatif sehingga menyebabkan berkembangnya penelitian tentang prediksi harga kelapa sawit. Penelitian menggunakan algoritma Deep Learning GRU(Gated Recurrent Unit) untuk Time Series Forecasting atau Peramalan Seri Waktu dari perubahan harga kelapa sawit. Algoritma GRU diterapkan pada aplikasi prediksi ini untuk menghasilkan nilai akurasi yang dapat digunakan sebagai acuan dalam melakukan penjualan kelapa sawit. Hasil dari penelitian ini menunjukkan data aktual dan data prediksi tidak jauh berbeda. Pada algoritma GRU menggunakan learning rate 0,01, batch size, 100, epoch 15000, hidden state 512 dan windows size 30.

Kata Kunci: *Kelapa Sawit, GRU, Prediksi*

Abstract- Palm oil (PO) is one of Indonesia's leading exporters, one of the plantation products. Oil prices often experience price increases and fluctuating price decreases, causing the development of research on palm oil price predictions so that research uses the deeplearning GRU (Gated Recurrent Unit) for timeseriesforecasting of changes in palm oil prices. The GRU algorithm is applied to this prediction application to produce an accuracy value that can be used as a reference in selling palm oil. The results of this study show that the actual data and predictive data are not much different. The GRU algorithm uses a learning rate of 0.01, batch size, 100, epoch 15000, hidden state 512 and windows size 30.

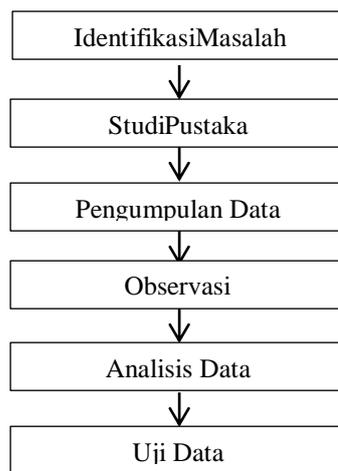
Keywords: *Palm Oil, GRU, Prediction*

1. PENDAHULUAN

Kelapa sawit atau Palm Oil (PO) merupakan salah satu eksportir unggulan Indonesia, salah satu produk perkebunan[1]. Harga kelapa sawit sering mengalami kenaikan harga dan penurunan harga yang fluktuatif sehingga menyebabkan berkembangnya penelitian tentang prediksi harga kelapa sawit ini telah dilakukan oleh banyak ahli, diantaranya yaitu, menggunakan metode gabungan antara *Artificial Neural Network* dan *Genetic Algorithm* [2]. Penelitian lainnya dengan menggunakan metode gabungan antara *Multilayer Perception* dan *Elman Recurrent Neural Network (ERNN)* menghasilkan nilai prediksi dari model yang diusulkan pada harga kelapa sawit yang sesuai dengan nilai sebenarnya. Prediksi dengan model *ERNN* diusulkan paling mendekati data aktual, sehingga harga kelapa sawit dapat berubah setiap hari.

2. METODE PENELITIAN

Pada penelitian ini memuat beberapa tahapan-tahapan yang akan dilakukan, adapun tahapan-tahapannya dapat dilihat pada gambar 1.



Gambar 1. Metode Penelitian

Pada tahap-tahap yang terdapat pada gambar 1 akan dijelaskan sebagai berikut:

1. Identifikasi Masalah
Tahapan ini merupakan tahapan dimana akan dilakukan identifikasi masalah yang akan dilakukan
2. Studi Pustaka
Pada tahap ini, menganalisa serta mengevaluasi hasil penelitian teori-teori.
3. Pengumpulan Data
Selanjutnya ialah pengumpulan data yang dibutuhkan, melakukan perancangan konsep dan ide.
4. Observasi
Pada tahap ini, melakukan pengamatan langsung terhadap sistem manual dan pencatatan secara cermat dan sistematis untuk mengumpulkan data-data agar diperoleh informasi yang dibutuhkan.
5. Analisis Data
Pada tahap ini, melakukan analisis data yang akan digunakan pada sebuah data besar.
6. Uji Data
Langkah terakhir dari tahap penelitian ini ialah pengujian data yang sudah dibangun dan dilakukan analisa.

2.1 Jaringan Saraf Tiruan

Sebuah jaringan saraf adalah sebuah sistem yang dibentuk dari sejumlah elemen pemroses sederhana yang bekerja secara paralel dimana fungsinya ditentukan oleh struktur jaringan, kekuatan hubungan, dan pengolahan dilakukan pada komputasi elemen atau node. *Neural Network (NN)* adalah suatu struktur pemroses informasi yang terdistribusi dan bekerja secara paralel, yang terdiri atas elemen pemroses (yang memiliki memori lokal dan beroperasi dengan informasi lokal) yang diinterkoneksi bersamadengan alur sinyal searah yang disebut koneksi[3].

2.2 Recurrent Neural Network

Recurrent Neural Network atau biasa disingkat RNN adalah jenis jaringan saraf tiruan untuk memproses data sekuensial seperti pengenalan ucapan, pemodelan Bahasa dan lain lain. RNN bekerja dimana pemrosesannya dilakukan secara berulang[4].

2.3 Gate Recurrent Unit

Gated Recurrent Unit (GRU) adalah salah satu variasi dari *Recurrent Neural Network* yang dibuat untuk menghindari masalah ketergantungan jangka panjang pada *Recurrent Neural Network (RNN)*. *GRU* dapat mengingat informasi jangka panjang sama seperti *Recurrent Neural Network*, *GRU* juga terdiri dari modul pemrosesan berulang[5].

2.4 Peramalan

Gated Recurrent Unit (GRU) adalah salah satu variasi dari *Recurrent Neural Network* yang dibuat untuk menghindari masalah ketergantungan jangka panjang pada *Recurrent Neural Network (RNN)*. *GRU* dapat mengingat informasi jangka panjang sama seperti *Recurrent Neural Network*, *GRU* juga terdiri dari modul pemrosesan berulang[6].

2.5 Kelapa Sawit

Kelapa sawit adalah tumbuhan industri/perkebunan yang berguna sebagai penghasil minyak masak, minyak industri, maupun bahan bakar. Pohon kelapa sawit terdiri dari dua *spesies* yaitu *elaeis guineensis* dan *elaeis oleifera* yang digunakan untuk pertanian komersil dalam pengeluaran minyak kelapa sawit. Pohon KelapaSawit *elaeis guineensis*, berasal dari Afrika barat diantara Angola dan Gambia, pohon kelapa sawit *elaeis oleifera*, berasal dari Amerika tengah dan Amerika selatan[7].

2.6 Time Series

Time series atau deret waktu merupakan pengamatan satu atau beberapa *variabel* yang diambil secara beruntun terhadap *interval* waktu yang tetap. Pada tahun 1976 George Box dan Gwilyn Jenkins memperkenalkan analisis *time series* untuk pertama kalinya. Analisis *time series* adalah salah satu prosedur statistika yang digunakan pada peramalan kejadian di masa depan. Analisis *time series* menggunakan data yang terpaut oleh waktu, sehingga korelasi antara kejadian saat ini dengan periode waktu sebelumnya akan terjadi. Selain berhubungan antara waktu *time series* juga terdapat kemungkinan adanya hubungan antara dimensi lain seperti wilayah ataupun dimensi lain yang saling berkaitan[8].

2.7 Python

Python adalah bahasa pemrograman interpretative multiguna. Tidak seperti bahasa lain yang susah untuk dibaca dan dipahami, python lebih menekankan pada keterbacaan kode agar lebih mudah untuk memahami sintaks. Hal ini membuat Python sangat mudah dipelajari baik untuk pemula maupun untuk yang sudah menguasai bahasa pemrograman lain. Bahasa ini muncul pertama kali pada tahun 1991, dirancang oleh seorang bernama *Guido van Rossum*. Sampai saat ini python masih dikembangkan oleh *python software foundation*. Bahasa python mendukung hamper semua system operasi, bahkan untuk system operasi linux, hamper semua distronya sudah menyertakan Python di dalamnya. Dengan kode yang simpel dan mudah diimplementasikan, seorang *programmer* dapat lebih mengutamakan pengembangan aplikasi yang dibuat, bukan malah sibuk mencari *syntax error*[9].

2.8 Flowchart

Flowchart adalah bagan – bagan yang mempunyai arus yang menggambarkan langkah –langkah penyelesaian suatu masalah. *Flowchart* merupakan cara penyajian dari suatu algoritma. *Flowchart* adalah untuk menggambarkan sebuah algoritma yang terstruktur dan mudah dipahami oleh orang lain khususnya *programmer* yang bertugas mengimplementasikan program, maka dibutuhkan alat bantu yang berbentuk diagram alur[10].

3. HASIL DAN PEMBAHASAN

3.1 Analisis Kebutuhan Sistem

Analisis system menjelaskan tentang berbagai analisis yang berhubungan dengan sistem yang akan di buat. Tahap analisis ini bertujuan untuk memperoleh informasi-informasi yang dibutuhkan dan memperoleh konsep peramalan harga kelapa sawit yang akan dibuat. Ditahap ini, penelitian di awali dengan melakukan observasi terhadap informasi-informasi yang dibutuhkan di dalam pembuatan system peramalan harga kelapa sawit. Informasi tersebut antara lain data harga kelapa sawit dan kebutuhan pembuatan sistem. Dalam menerapkan system peramalan pada penelitian ini akan menggunakan metode *gated recurrent unit* yang mencakup proses *preprocessing*, *split data*, *inisialisasi hyper parameter*, *learning rate*, *hidden layer*. Hal di dalam penelitian ini mencakup tentang analisis sistem yang akan di buat, analisis kebutuhan fungsional dan analisis kebutuhan non-fungsional.

3.2 Sumber Data

Tabel 1. Data Harga KelapaSawit

Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%
02/08/2021	4.445,00	4.490,00	4.490,00	4.430,00	0,38K	-4,53%
30/07/2021	4.656,00	4.695,00	4.740,00	4.651,00	0,74K	-0,36%
29/07/2021	4.673,00	4.640,00	4.722,00	4.640,00	0,19K	1,81%
28/07/2021	4.590,00	4.674,00	4.699,00	4.562,00	0,35K	-2,11%
27/07/2021	4.689,00	4.665,00	4.720,00	4.630,00	0,97K	1,27%
26/07/2021	4.630,00	4.523,00	4.658,00	4.489,00	0,46K	1,94%
23/07/2021	4.542,00	4.440,00	4.586,00	4.440,00	0,96K	2,48%
22/07/2021	4.432,00	4.359,00	4.450,00	4.351,00	1,30K	0,84%
21/07/2021	4.395,00	4.370,00	4.400,00	4.300,00	2,60K	1,29%
19/07/2021	4.339,00	4.343,00	4.343,00	4.301,00	1,95K	0,91%
16/07/2021	4.300,00	4.227,00	4.317,00	4.205,00	2,70K	3,64%
15/07/2021	4.149,00	4.150,00	4.150,00	4.149,00	0,11K	1,34%
14/07/2021	4.094,00	4.100,00	4.130,00	4.042,00	0,08K	1,14%
13/07/2021	4.048,00	4.014,00	4.050,00	4.014,00	0,03K	1,48%
12/07/2021	3.989,00	3.985,00	4.010,00	3.985,00	0,08K	-0,77%
09/07/2021	4.020,00	4.000,00	4.020,00	4.000,00	0,13K	2,55%
08/07/2021	3.920,00	3.910,00	3.950,00	3.895,00	0,10K	-0,41%
07/07/2021	3.936,00	3.913,00	3.943,00	3.870,00	0,31K	-0,28%
06/07/2021	3.947,00	4.020,00	4.040,00	3.946,00	0,62K	-0,63%
05/07/2021	3.972,00	3.887,00	3.980,00	3.880,00	0,95K	1,90%

3.3 Rancangan Sistem Prediksi

Sub bab ini membahas tentang tahapan penelitian yang dilakukan mulai dari penyiapan data, pelaksanaan dan evaluasi model serta interpretasi hasil penelitian untuk mendapatkan kesimpulan yang baik. Peneliti menggunakan bahasa pemrograman python dengan bantuan beberapa modul diantaranya *pytorch*, *pandas*, *numpy*, *sklearn* dan *matplotlib*.

3.3.1 Preprocessing

Preprocessing diperlukan sebelum melakukan prediksi pada data yang akan digunakan. Pada penelitian ini terdapat dua tahapan untuk *preprocessing* data yaitu data *splitting* dan *normalisasi* data. Berikut ini *coding* dalam melakukan *preprocessing* yang hanya melakukan *cleaning* terhadap data:

```
# Fix datatype onto datetime
df['Date'] = pd.to_datetime(df['Date'])
# Remove comma (,)
df['Price'] = df['Price'].str.replace(',', '').astype(float)
df['Open'] = df['Open'].str.replace(',', '').astype(float)
df['High'] = df['High'].str.replace(',', '').astype(float)
df['Low'] = df['Low'].str.replace(',', '').astype(float)
```

3.3.2 Split Data

Pada tahap *splitting* data, data akan dibagi menjadi data *training*, *validasi*, dan *testing*. 200 data *training* digunakan untuk pemodelan menggunakan GRU. Setelah mendapatkan model terbaik dari data *training*, selanjutnya model tersebut diuji cobakan pada total 69 data *testing* untuk melihat apakah model yang dihasilkan akan memiliki akurasi yang tinggi atau tidak berikut ini *script* dari *split* data.

```
train_start = str(dt.date(2019,12,2))
train_end = str(dt.date(2020,7,31))
train_data = series.loc[train_start:train_end] # Output Train Data
# Validation Data
```

```

val_start = str(dt.date(2020,8,1))
val_end = str(dt.date(2020,10,31))
val_data = series.loc[val_start:val_end] # Output Validation Data

# Test Data
test_start = str(dt.date(2020,11,1))
test_end = str(dt.date(2020,12,31))
test_data = series.loc[test_start:test_end] # Output Test Data

```

3.3.3 Normalisasi Data

Normalisasi adalah proses pemberian skala nilai atribut pada data sehingga data tersebut berada dalam rentang tertentu. Pada penelitian ini normalisasi data dilakukan dengan cara membagi data dengan nilai data terbesar sehingga data berada pada interval 0 sampai 1. Langkah ini sangat penting karena dapat meminimalisir kesalahan. Tujuan dari normalisasi data adalah untuk menghindari fitur dengan nilai yang lebih besar mendominasi fitur dengan nilai yang lebih kecil.

3.3.4 Inisialisasi Hyper parameter model GRU

Implementasi dilakukan sesudah perancangan selesai dilakukan dan selanjutnya akan diimplementasikan di bahasa pemrograman yang akan dipergunakan pada penelitian ini menggunakan Bahasa pemrograman *python*. Tujuan dari Implementasi ini adalah untuk mengkonfirmasikan modul program perancangan dan melakukan prediksi harga emas antam per-gram.

3.3.5 Learning Rate

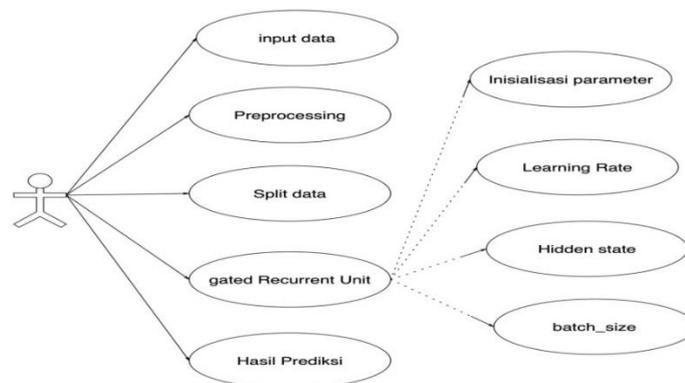
Pemilihan nilai *learning rate* yang tepat dapat membantu mendapatkan hasil *forecast* yang optimal. *Learning rate* merupakan salah satu parameter pelatihan untuk menghitung nilai koreksi bobot selama proses pelatihan. Nilai dari berada dalam rentang nol (0) hingga (1). Semakin tinggi *learning rate* maka semakin cepat proses pelatihan akan berjalan. Namun, jika kecepatan pembelajaran relative terlalu besar, proses pelatihan dapat melebihi keadaan optimal ketika nilai kesalahan minimum tercapai. Dengan kata lain, kecepatan belajar mempengaruhi akurasi jaringan suatu sistem. Semakin besar *learning rate* maka akurasi jaringan akan semakin menurun, tetapi sebaliknya jika *learning rate* semakin kecil maka akurasi jaringan akan semakin besar atau meningkat dengan konsekuensi proses pelatihan akan memakan waktu yang lebih lama. Berikut asumsi *learning rate* dalam penelitian ini, $lr = 0,001$.

3.3.6 Hidden State

Hidden state yang optimal adalah barisan *state* yang mempunyai peluang tertinggi dalam menghasilkan barisan observasi yang telah diketahui sebelumnya. Misalkan didalam h_t dan h_{t-1} dan z_t artinya h_t adalah fungsi dari h_{t-1} dan z_t atau ditulis $h_t = f(h_{t-1}, z_t)$. Namun z_t juga fungsi dari x_t atau dapat di tulis $z_t = g(x_t)$, akibatnya, juga fungsi dari x_t dapat di tulis $h_t = f(h_{t-1}, g(x_t))$.

3.3.7 Perancangan Sistem

Dalam melakukan perancangan sistem mengacu pada pendekatan berorientasi objek sehingga digunakanlah UML (*Unified Modelling Language*). Pada penelitian membangun sistem prediksi harga kelapa sawit penulis menggunakan diagram *usecase*, *activity* dan *sequence* diagram.



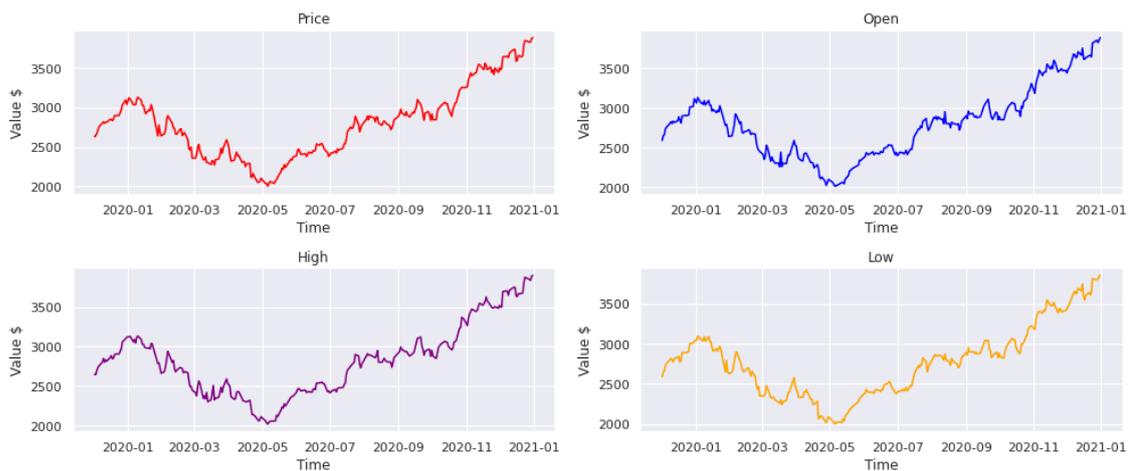
Gambar 2. Use case Algoritma GRU

3.4 Implementasi

Implementasi dilakukan setelah perancangan selesai dilakukan dan selanjutnya akan diimplementasikan pada bahasa pemrograman yang akan digunakan pada penelitian ini menggunakan Bahasa pemrograman *python*. Tujuan implementasi adalah untuk mengkonfirmasi modul program perancangan dan melakukan prediksi harga kelapa sawit.

3.4.1 Visualisasi OHLC Pada Data

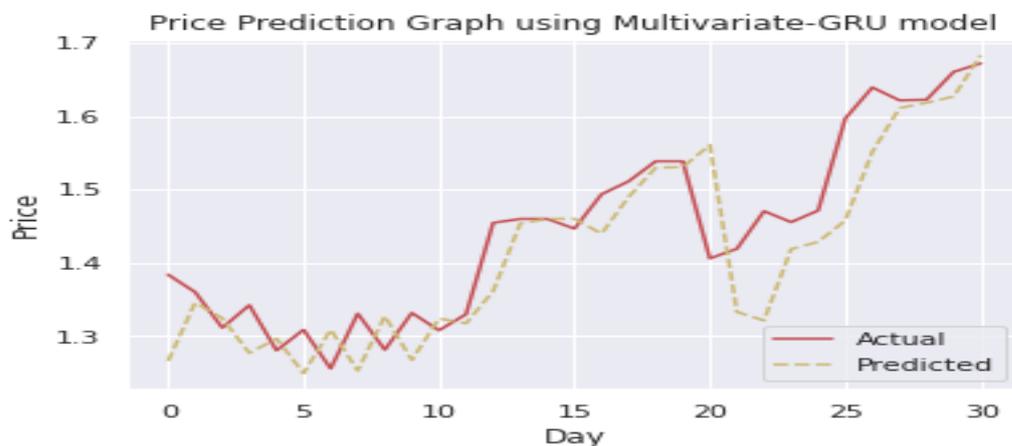
Visualisasi *OHLC (Open High Low Close)* merupakan representasi grafis dari data harga kelapa sawit dan informasi. Dengan menggunakan elemen-elemen *visual*, visualisasi data juga mempermudah untuk menarik kesimpulan dengan melihat tren atau pola yang ada.



Gambar 3. Visualisasi OHLC Pada Data

Visualisasi data OHLC (*Open High Low Close*) masing-masing secara *individual* untuk melihat pola dari data secara *intuitif* apakah terdapat *tren* dan *seasonal* atau tidak.

3.4.2 Visualisasi Perbandingan antara prediksi data dan actual data uji



Gambar 4. Visualisasi Perbandingan antara Prediksi Data dan Actual Data Uji

Visualisasi perbandingan antara data *actual* dengan data prediksi, dimana dalam hal ini data yang diprediksi adalah *price*, sedangkan data yang digunakan sebagai fitur adalah *Open High Low Close (OHLC) Volume Change* dan *Price* itu sendiri.

4. KESIMPULAN

Dalam uraian rangkaian mulai dari proses prediksi harga kelapa sawit. Secara umum langkah-langkah dalam membangun sistem yang dibuat terdiri dari *ipreprocessing* data, *inisialisasi hyperparameter*, *training* GRU dan melakukan test pada data *testing.*, dapat ditarik beberapa kesimpulan penting antara lain:

- a. Dengan melakukan percobaan terasi 125 maka didapat jumlah MSE terkecil pada iterasi 125 dengan kombinasi model search menggunakan learning rate 0,01, batch size 100, hidden state epoch 512 dan windows size 30.
- b. Ketelitian dan ketelitian dalam prediksi atau peramalan pada model ini dengan nilai mape sebesar 4,84% yang menunjukkan bahwa kemampuan prediksinya baik.

DAFTAR PUSTAKA

- [1] Radifan, F. (2014). Faktor-Faktor Yang Mempengaruhi Ekspor Crude Palm Oil Indonesia Dalam Perdagangan Internasional. *Economics Development Analysis Journal*, 3(2), 259–267.
- [2] Mahdiani, M. R., & Khomehchi, E. (2016). A Modified Neural Network Model For Predicting The Crude Oil Price. *Intellectual Economics*, 10(2), 71–77. <https://doi.org/10.1016/j.intele.2017.02.001>
- [3] Hrp, A. I., Lubis, J. R., Ferianto, I., Djasmayena, S., Dewi, S., & Aritonang, R. (2021). Analisis Penerapan Jaringan Saraf Tiruan Untuk Mendeteksi Gangguan Psikologi Pada Manusia Menggunakan Metode Backpropagation. *Jurnal Susi*, 1(02), 1–4.
- [4] Ibrahim, M. (2018). Sistem Pendukung Keputusan Pengajuan Dan Penilaian Beasiswa Berbasis Website Menggunakan Metode Topsis. *Doubleclick: Journal Of Computer And Information Technology*, 2(1), 22. <https://doi.org/10.25273/Doubleclick.V2i1.3214>
- [5] Zaman, L., Sumpeno, S., & Hariadi, M. (2019). Analisis Kinerja Lstm Dan Gru Sebagai Model Generatif Untuk Tari Remo. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (Jnteti)*, 8(2), 142–150.
- [6] Hernadewita, Hadi, Y. K., Syaputra, M. J., & Setiawan, D. (2020). Peramalan Penjualan Obat Generik Melalui Time Series Forecasting Model Pada Perusahaan Farmasi Di Tangerang: Studi Kasus. *Journal Industrial Engineering & Management Research (Jiemar)*, 1(2), 35–49.
- [7] Firdaus, D. (2017). Penggunaan Data Mining Dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer. *Jurnal Format*, 6(2), 91–97.
- [8] Ainun Faulina, N. (2020). *Perbandingan Metode Arimax Dan Varimax Untuk Peramalan Jumlah Penumpang Kereta Api Menurut Wilayah*. Muhammadiyah University, Semarang.
- [9] Xia, J., Wan, W., Liu, R., Chen, G., & Feng, Q. (2015). Distributed Web Crawling: A Framework For Crawling Of Micro-Blog Data. *Iet Conference Publications, 2015(Cp672)*, 62–68. <https://doi.org/10.1049/Cp.2015.0255>.
- [10] Green, M. S. (2018). The Erie Doctrine: A Flowchart. *Akron L. Rev.*, 52, 215.