

Prediksi Penyebaran COVID-19 Di Indonesia Menggunakan Long Short Term Memory

Haviluddin^{a,1}, Novianti Puspitasari^{a,2}, Raynold Raynaldo^{a,3}

^a Departemen Informatika, Fakultas Teknik, Universitas Mulawarman, Jalan Sambaliung, Samarinda, 75119, Indonesia

¹ haviluddin@unmul.ac.id; ² miechan.novianti@gmail.com; ³ raynoldraynaldo@gmail.com

ABSTRAK

ARTICLE INFO

Penelitian ini bertujuan untuk memprediksi arah tren penyebaran COVID-19 di Indonesia, serta memperkirakan jumlah kumulatif kasus terkonfirmasi dan kematian dalam 30 hari ke depan. Model *Long Short-Term Memory* (LSTM) digunakan karena kemampuannya dalam mengolah data *time-series*, sementara *Mean Squared Error* (MSE) digunakan sebagai fungsi kerugian untuk mengukur akurasi prediksi. Sebanyak sembilan skenario pengujian dilakukan untuk masing-masing kategori kasus terkonfirmasi dan kematian, guna menentukan model terbaik. Berdasarkan hasil pengujian, model menunjukkan arah tren yang meningkat (uptrend) untuk kasus kumulatif terkonfirmasi maupun kematian. Prediksi menunjukkan bahwa pada tanggal 19 Juli 2021, jumlah kasus terkonfirmasi di Indonesia akan mencapai 2.917.557 kasus. Sementara itu, prediksi untuk jumlah kasus kematian hingga 9 Agustus 2021 diperkirakan mencapai 73.265 kasus. Temuan ini diharapkan dapat menjadi acuan bagi pemangku kebijakan dalam menyusun strategi penanganan pandemi secara lebih tepat dan terukur.



Article history:

Received: 27 Desember 2024

Revised: 20 Januari 2025

Accepted: 12 Februari 2025

Keywords:

COVID-19
Indonesia
LSTM
MSE
Prediction

This is an open-access article under the [CC-BY-SA](#) license.

I. Pendahuluan

The formatter will need to create these components, incorporating the applicable criteria that follow. This template is designed to assist you in preparing your manuscript; it is an exact representation of the format expected by the editor. To use this template, please just Save As to your document, then copy and paste your document here. The work should not have been published or submitted for publication elsewhere.

Seluruh dunia menderita pandemi yang sedang berlangsung karena Penyakit Coronavirus yang disebabkan oleh sindrom pernapasan akut coronavirus 2 (SARS-CoV-2) [1]. Virus ini diidentifikasi pada 7 Januari dan diamati menyebar melalui penularan dari manusia ke manusia melalui tetesan atau kontak langsung. Infeksi diperkirakan pada masa inkubasi rata-rata 6,4 hari dan angka reproduksi dasar. Sejak diidentifikasi, penyakit ini telah menyebar dengan cepat ke seluruh dunia, oleh karena itu Organisasi Kesehatan Dunia (WHO) telah menyatakan COVID-19 sebagai pandemi global pada 11 Maret 2020 [2].

Situasi di Indonesia (15 Agustus 2020), kasus positif baru bertambah per hari ada 2.345 kasus. Jumlah kasus aktif COVID-19 sebanyak 40.076 kasus. Total kumulatif hingga saat ini telah mencapai 137.468 kasus. Penambahan kasus baru terbanyak berada di DKI Jakarta dengan 583 kasus dan secara kumulatif mencapai 28.882 kasus. Disusul Jawa Timur dengan penambahan 436 kasus baru dan secara kumulatif mencapai 27.415 kasus. Kemudian Jawa Barat dengan jumlah harian tertinggi ketiga menambah kasus baru sebanyak 237 kasus dengan jumlah kumulatif 8.512 kasus. Kasus kematian hari ini juga bertambah lagi sebanyak 50 kasus. Total kematian pasien mencapai 6.071 kasus. Kematian harian tertinggi terjadi di Jawa Timur dengan 21 kasus.

Jumlah kematian pasien di provinsi ini mencapai 1988 kasus. Selain itu, sebaran wilayah masih terjadi di 34 provinsi. Namun ada penambahan 1 wilayah di tingkat Kabupaten/Kota, kini menjadi 482 wilayah (Tim Komunikasi Panitia Penanganan Corona Virus Disease 2019 (Covid-19) dan Pemulihan Ekonomi Nasional, 2020). Hal ini menunjukkan urgensi situasi dengan penyebaran virus COVID-19 di Indonesia pada situasi saat ini dan perlunya prediksi sebagai dasar acuan dalam pengambilan kebijakan yang tepat. Dalam kondisi pandemi, prediksi berperan penting dalam hal pengambilan kebijakan, baik medis maupun pemerintahan [3]. Prediksi berperan dalam menggambarkan pola penyebaran pandemi di masa depan yang terbentuk berdasarkan data historis dari pola penyebaran pada waktu sebelumnya [4].

Setiap wabah penyakit menular menunjukkan pola dan pola ini perlu diidentifikasi berdasarkan dinamika penularan wabah tersebut [5]. Tindakan intervensi untuk memberantas penyakit menular bergantung pada metode yang digunakan untuk mengevaluasi wabah saat terjadi [6]. Wabah apa pun di suatu negara atau provinsi biasanya terjadi pada tingkat besaran yang berbeda terhadap waktu, yaitu, perubahan musim, adaptasi virus dari waktu ke waktu [7]. Pola wabah tertentu memungkinkan kita untuk memprediksi pola penyebaran di masa depan dengan harapan dapat memitigasi peningkatan penyebaran COVID-19 di Indonesia.

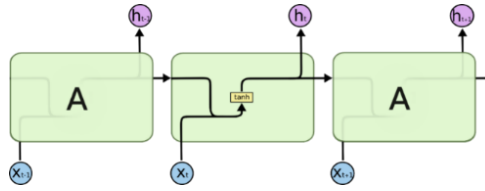
Perkembangan teknologi semakin pesat dan terdapat berbagai macam disiplin ilmu yang ada, manusia dapat mengatasi berbagai permasalahan yang ada di dunia ini, dan salah satu disiplin ilmu yang dapat digunakan dalam kasus Prediksi Penyebaran COVID-19 di Indonesia adalah Deep Learning Algorithm (DL) [8]. Saat ini, algoritma DL merupakan algoritma baru dan terus digunakan untuk menganalisis data yang besar dan kompleks, memberikan kinerja tinggi dalam mengidentifikasi pasien COVID-19 yang berguna untuk mengendalikan wabah SARS-CoV-2 [9]. Ada banyak varian algoritma DL mulai dari Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Hybrids, Long Short-Term Memory (LSTM) dan masih banyak lagi [10].

LSTM dianggap sebagai salah satu solusi yang paling layak, karena mengantisipasi prakiraan masa depan bergantung pada berbagai fitur yang disorot yang ada dalam kumpulan data [11]. Berdasarkan prinsip LSTM, data bergerak melalui komponen yang disebut status sel dan dapat secara akurat mengingat atau mengabaikan sesuatu. Informasi yang dikumpulkan selama kerangka waktu progresif digambarkan sebagai data deret waktu dan untuk menghasilkan prakiraan dengan nilai data ini. LSTM umumnya diusulkan menjadi metodologi yang stabil [12]. Hal inilah yang menjadi dasar pemilihan algoritma LSTM sebagai algoritma yang digunakan dalam penelitian ini.

Artikel ini menggunakan metode Long Short-Term Memory (LSTM) untuk menganalisa dan memprediksi penyebaran COVID-19 di Indonesia. Tujuan penelitian ini adalah untuk membantu pihak-pihak terkait dalam pengambilan keputusan terkait penanganan penyebaran COVID-19. Artikel ini terdiri dari motivasi penulisan artikel pada bagian pertama. Kedua, menjelaskan model kerja algoritma LSTM. Ketiga, analisa hasil percobaan. Kesimpulan penelitian pada bagian akhir.

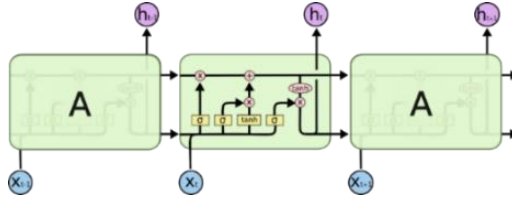
II. Methods and Materials

Penelitian ini berfokus pada pencarian model terbaik terlebih dahulu menggunakan Jaringan LSTM khususnya untuk kasus kumulatif terkonfirmasi dan kasus kumulatif kematian di Indonesia. Jaringan Long Short-Term Memory disingkat "LSTM" adalah jenis khusus dari RNN, yang mampu mempelajari dependensi jangka panjang. Mereka diperkenalkan oleh Hochreiter dan Schmidhuber [13], dan disempurnakan dan dipopulerkan oleh banyak orang dalam karya berikut. Mereka bekerja sangat baik pada berbagai macam masalah, dan sekarang banyak digunakan. LSTM secara eksplisit dirancang untuk menghindari masalah ketergantungan jangka panjang [14]. Mengingat informasi untuk jangka waktu yang lama pada dasarnya adalah perilaku bawaan mereka, bukan sesuatu yang sulit untuk mereka pelajari. RNN memiliki bentuk rangkaian modul jaringan syaraf tiruan yang berulang[15]. Dalam RNN standar, modul iteratif ini akan memiliki struktur yang sangat sederhana[16], seperti lapisan tanh tunggal, Gambar 1.



Gambar 1. Modul Iteratif dalam RNN Standar Berisi Satu Lapisan

LSTM juga memiliki struktur seperti rantai ini, tetapi modul berulang memiliki struktur yang berbeda. Alih-alih memiliki satu lapisan jaringan saraf, ada empat lapisan, yang berinteraksi dengan cara yang sangat khusus, Gambar 2.



Gambar 2. Modul Iteratif dalam LSTM Berisi Empat Lapisan

Gambar 2 menjelaskan bagaimana alur kerja sel memori pada setiap neuron LSTM bekerja. Terdapat empat proses fungsi aktivasi pada setiap masukan ke neuron, yang selanjutnya disebut sebagai unit gerbang. Unit gerbang tersebut adalah forget-gates, input-gates, cell-gates, and output-gates.

Pada forget-gates, informasi pada setiap data input akan diproses dan data mana yang akan disimpan atau dibuang di sel memori [17]. Fungsi aktivasi yang digunakan pada forget-gates ini adalah fungsi aktivasi sigmoid. Dimana outputnya antara 0 dan 1. Jika outputnya 1 maka semua data akan disimpan dan sebaliknya jika outputnya 0 maka semua data akan dibuang menggunakan Persamaan 1.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Pada input-gates ada dua gerbang yang akan diimplementasikan, pertama akan ditentukan nilai mana yang akan diupdate menggunakan fungsi aktivasi sigmoid. Kemudian fungsi aktivasi tanh akan membuat vektor nilai baru yang akan disimpan di sel memori menggunakan Persamaan 2 dan 3.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

Cell-gates akan menggantikan nilai pada sel memori sebelumnya dengan nilai sel memori yang baru. Dimana nilai ini didapat dari penggabungan nilai-nilai yang terdapat pada forget-gates dan input-gates menggunakan Persamaan 4.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

Pada gerbang keluaran terdapat dua gerbang yang akan diimplementasikan, pertama akan ditentukan bagian mana dari nilai sel memori yang akan dikeluarkan menggunakan fungsi aktivasi sigmoid. Selanjutnya, nilai tersebut akan ditempatkan di sel memori menggunakan fungsi aktivasi tanh. Terakhir, kedua gerbang tersebut dikalikan sehingga menghasilkan nilai yang akan dikeluarkan menggunakan Persamaan 5 dan 6

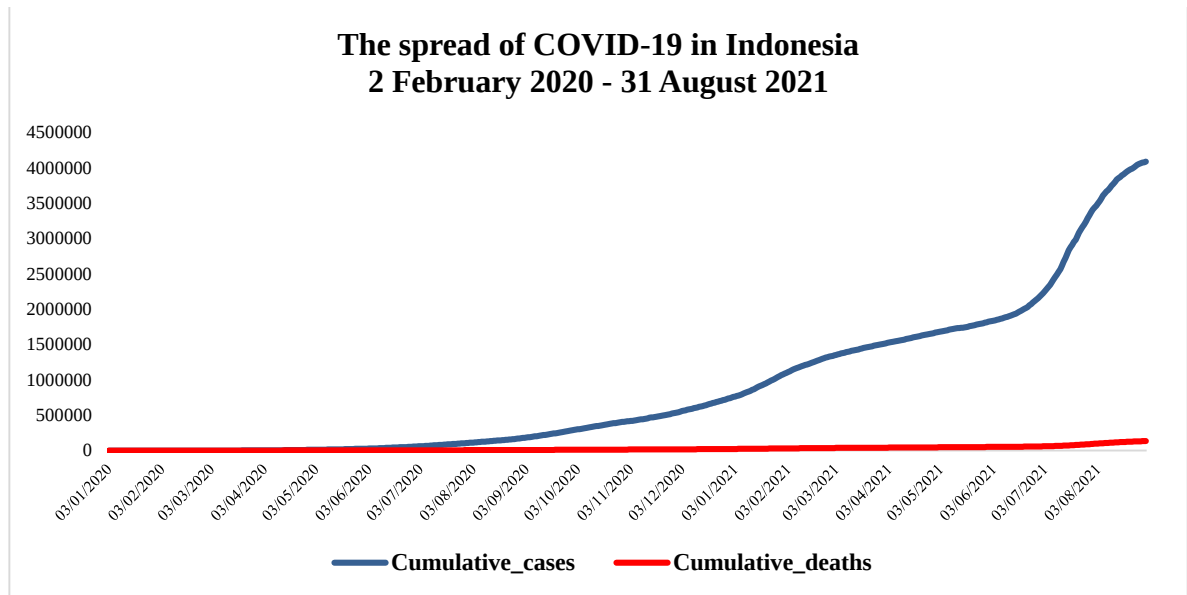
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

Dalam persamaan di atas, tanh digunakan untuk menskalakan nilai ke dalam rentang 1 hingga 1, adalah fungsi aktivasi yang diambil sebagai sigmoid dan W adalah matriks bobot yang sesuai.

Studi ini berfokus pada kasus dan kematian terkonfirmasi positif harian untuk Indonesia untuk memprediksi kasus 30 hari ke depan. Dataset ini kami peroleh dari WHO. Data set kasus terkonfirmasi positif dikumpulkan dari awal penyakit ini muncul di Indonesia yakni tanggal 03 Januari 2020 sampai tanggal 20 Juni 2021 dan data set kasus kematian dikumpulkan dari 03 Januari

2020 sampai tanggal 11 Juli 2021. Pada bagian ini, pertama-tama kita membahas secara singkat kumpulan data ini, Semua kumpulan data dan kode yang akan digunakan di bagian ini tersedia untuk umum (<https://covid19.who.int/WHO-COVID-19-global-data.csv>).



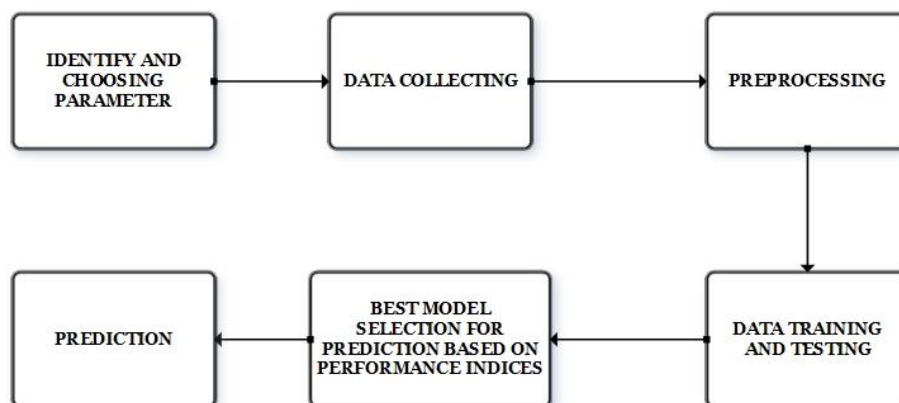
Gambar 3. Sebaran COVID-19 di Indonesia (World Health Organization, 2020b)

Hasil utama adalah prediksi arah penyebaran dan jumlah kasus 30 hari ke depan untuk kasus kumulatif terkonfirmasi dan kasus kumulatif kematian berdasarkan data historikal yang dihasilkan untuk waktu tertentu.

Kinerja sistem dilakukan dengan menggunakan perhitungan Mean Squared Error (MSE). Model terbaik adalah model dengan nilai MSE terkecil dibawah 0,5. Sedangkan MSE menggunakan Persamaan 7.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (7)$$

Dimana, N adalah Jumlah data; $\hat{Y}_i$ adalah nilai yang didapat dari model atau hasil prediksi; Y_i adalah nilai data sebenarnya. Pemilihan model akurasi MSE dilakukan karena MSE sudah banyak digunakan dalam prediksi time-series [19].



Gambar 4. Tahapan Penelitian

III. Hasil dan Pembahasan

Pada bagian ini, kami akan menguraikan pekerjaan kami berdasarkan Gambar 3. Pertama, kami akan mengidentifikasi parameter yang akan diuji untuk fase pembelajaran. Kedua, kami mengumpulkan kumpulan data. Ketiga, dilakukan preprocessing atau normalisasi data menggunakan min-max scaler dari python library, sehingga data akan dikompresi antara 0 sampai 1. Keempat, menjalankan training dan testing data menggunakan LSTM Network. Kelima, memilih model terbaik untuk prediksi. Keenam, memprediksi jumlah kasus kumulatif terkonfirmasi, kumulatif kasus kematian, dan arah tren penyebaran COVID-19 di Indonesia.

Tabel 1. Skema yang diusulkan dengan parameter dan nilainya untuk model prediksi

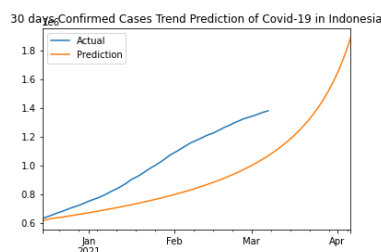
Parameters	Values
Ratio Data (Train-Test)	50:50; 70:30; 80:20
Memory Block	150, 350, 500
Epoch	100
Optimizer	Adam

Parameter yang kami gunakan dalam percobaan ini didasarkan pada Tabel 1. Ini adalah parameter yang akan diuji untuk memilih model terbaik untuk prediksi. Kumpulan data akan dikompresi terlebih dahulu, sebelum dilatih dan diuji (Disajikan pada Tabel 2).

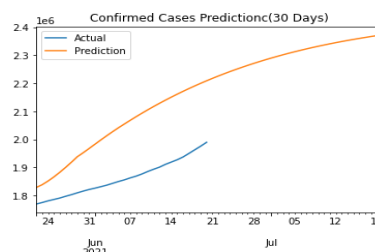
Tabel 2. Kumpulan data yang telah diproses sebelumnya

Date	Country	Cumulative Cases [0,1]	Cumulative Deaths [0,1]
1/3/2020	Indonesia	0	0
1/4/2020	Indonesia	0	0
1/5/2020	Indonesia	0	0
1/6/2020	Indonesia	0	0
1/7/2020	Indonesia	0	0
1/8/2020	Indonesia	0	0
1/9/2020	Indonesia	0	0
1/10/2020	Indonesia	0	0
1/11/2020	Indonesia	0	0
1/12/2020	Indonesia	0	0
...	...	...	...
6/17/2021	Indonesia	0.998176508	0.966767204
6/18/2021	Indonesia	0.998632381	0.975075403
6/19/2021	Indonesia	0.999088254	0.983383602
6/20/2021	Indonesia	0.999544127	0.991691801
6/21/2021	Indonesia	1	1

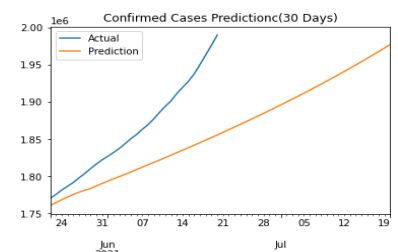
Setelah data dinormalisasi atau diproses sebelumnya (Disajikan pada Tabel 2). Kami melakukan percobaan menggunakan parameter yang diusulkan sebelumnya. Ini menghasilkan 9 model prediksi masing-masing untuk kasus kumulatif yang dikonfirmasi (lihat Gambar 4) dan kasus kumulatif kematian (lihat Gambar 5) yang akan dipilih untuk prediksi berdasarkan indeks kinerjanya. Indeks kinerja yang kami gunakan adalah MSE, semakin rendah skor semakin baik (Disajikan pada Tabel 3 dan Tabel 4).



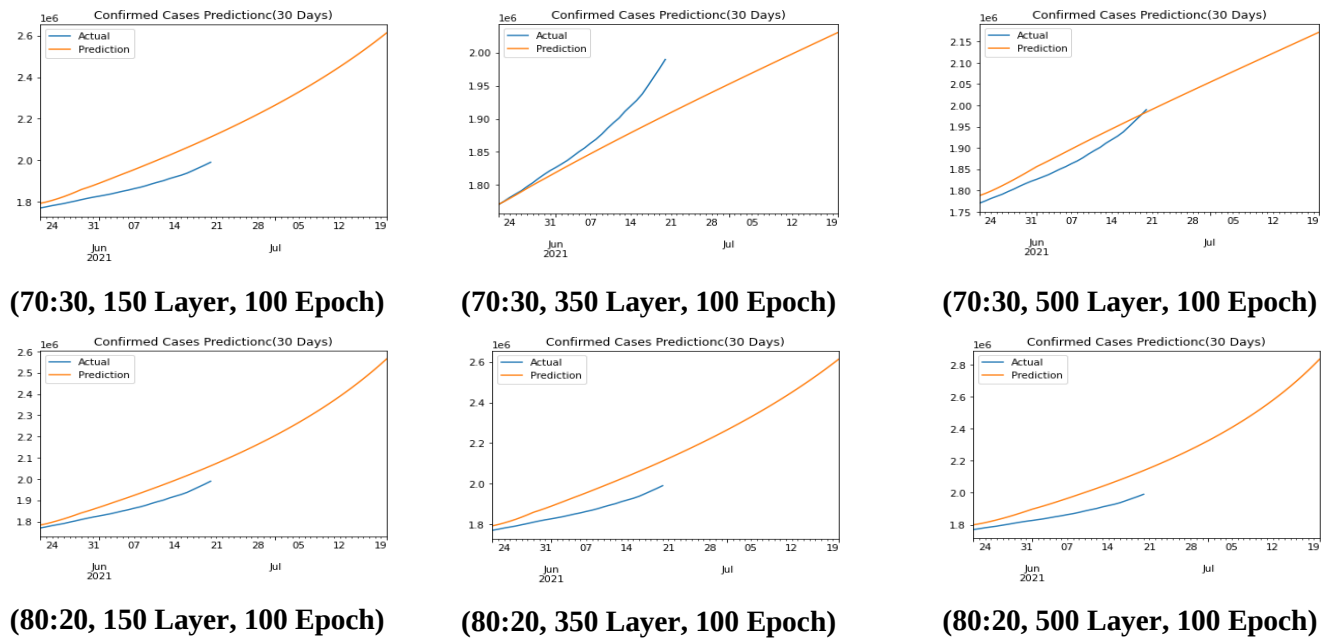
(50:50, 150 Layer, 100 Epoch)



(50:50, 350 Layer, 100 Epoch)



(50:50, 500 Layer, 100 Epoch)



Gambar 5. Prediksi *confirmed_cases* dengan 9 skenario

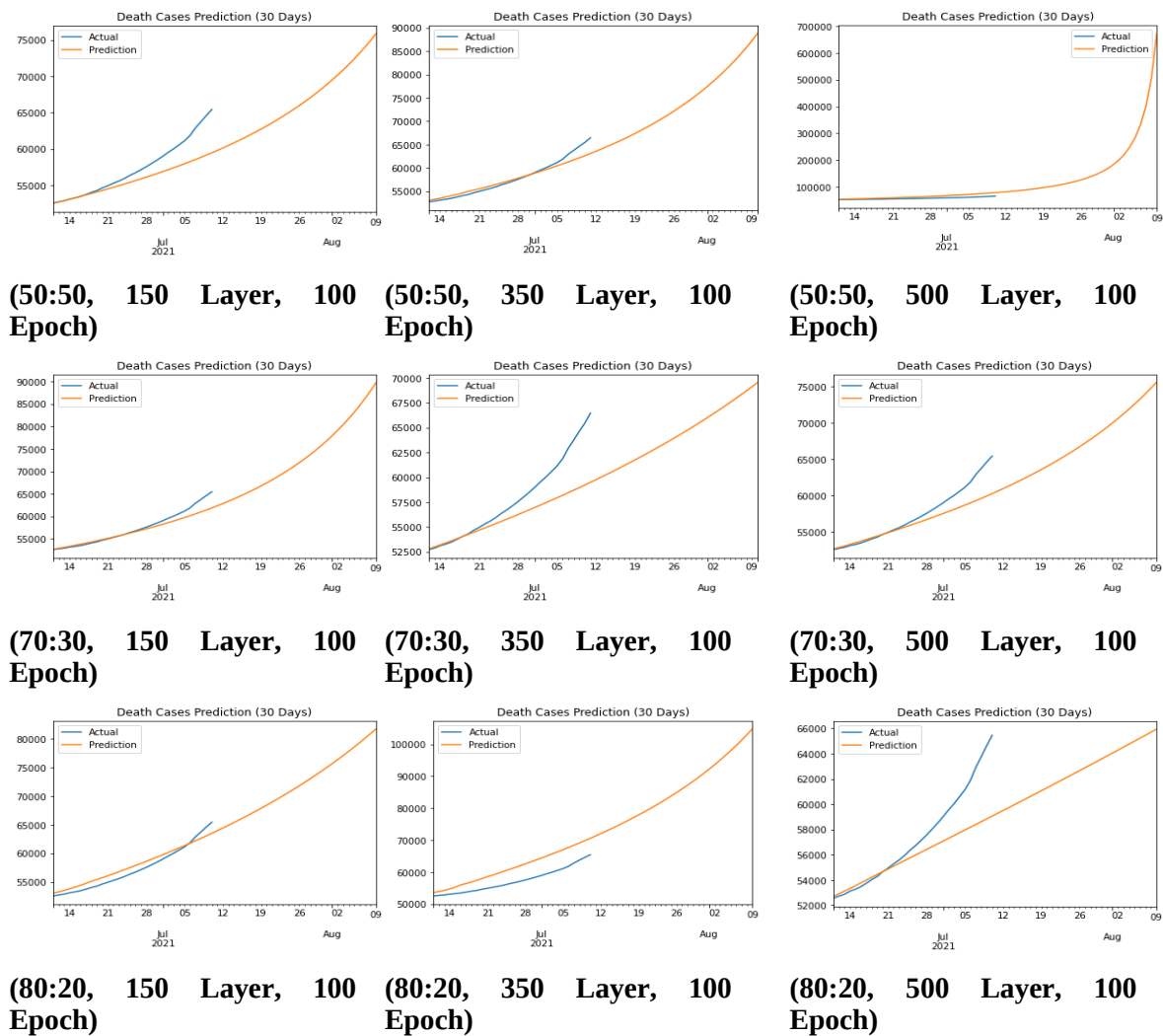
Tabel 3. Pemilihan Model Prediksi untuk Kasus Kumulatif yang Dikonfirmasi Berdasarkan Indeks

No	Data Ratio (Train-Test)	Memory Block	Epoch	MSE Value
1	50:50	150	100	0.0039
2	50:50	350	100	0.0031
3	50:50	500	100	0.0032
4	70:30	150	100	0.0032
5	70:30	350	100	0.0049
6	70:30	500	100	0.0035
7	80:20	150	100	0.0027
8	80:20	350	100	0.0034
9	80:20	500	100	0.0035

Dilakukan sebanyak sembilan skenario pengujian untuk menguji arah *tren* dari **kasus terkonfirmasi positif**. Didapatkan pada hasil pengujian, model data dengan perbandingan 80:20 (*Train:Test*), 150 *Layers*, dan 100 *epochs* adalah model dengan nilai MSE terkecil dan menjadi model pengujian terbaik dalam menguji arah *tren* terkonfirmasi positif dengan jumlah data 516.

Dilakukan sebanyak sembilan skenario pengujian untuk menguji arah *tren* dari kasus kematian. Didapatkan pada hasil pengujian, model data dengan perbandingan 70:30 (*train:test*), 150 *layers*, dan 100 *epochs* adalah model dengan nilai MSE terkecil dan menjadi model pengujian terbaik dalam menguji arah *tren* terkonfirmasi positif dengan jumlah data 537.

Berdasarkan Tabel 3 dan Tabel 4, masing-masing sembilan skenario pengujian telah dilakukan untuk kasus kumulatif terkonfirmasi dan kasus kumulatif kematian untuk menguji model prediksi terbaik berdasarkan skor nilai MSE-nya. Model prediksi dengan rasio data 80:20 (*Train-Test*), 150 *Layer*, dan 100 *epoch* merupakan model prediksi terbaik untuk data kasus kumulatif terkonfirmasi. Sedangkan untuk kasus kumulatif kematian, model prediksi dengan rasio data 70:30 (*train-test*), 150 *layer*, dan 100 *epoch* merupakan model terbaik karena keduanya menghasilkan nilai MSE terkecil. Model prediksi ini akan digunakan di bagian prediksi.

Gambar 6. Prediksi *death_cases* dengan 9 skenario

Tabel 4. Pemilihan Model Prediksi untuk Kasus Kumulatif Kematian Berdasarkan Indeks Kinerja

No	Data Ratio (Train-Test)	Memory Block	Epoch	MSE Value
1	50:50	150	100	0.0027
2	50:50	350	100	0.0025
3	50:50	500	100	0.0038
4	70:30	150	100	0.0021
5	70:30	350	100	0.0033
6	70:30	500	100	0.0037
7	80:20	150	100	0.0030
8	80:20	350	100	0.0081
9	80:20	500	100	0.0033

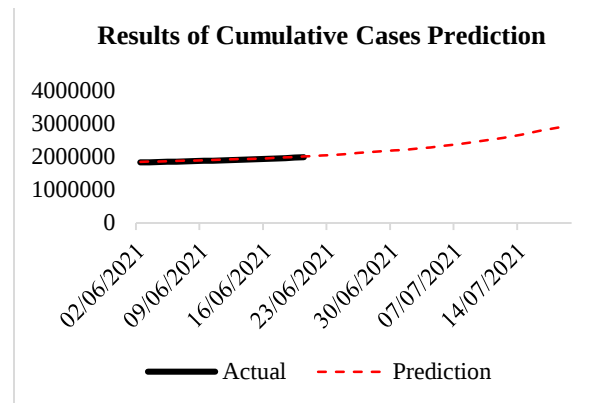
Perbedaan model prediksi rasio data, blok memori, dan epoch untuk kasus kumulatif yang dikonfirmasi dan kasus kumulatif kematian adalah karena dataset dihasilkan dalam periode waktu yang berbeda yang mengarah ke jumlah dataset yang berbeda. Kumpulan data yang digunakan untuk pengujian kasus kumulatif yang dikonfirmasi lebih rendah daripada kumpulan data untuk pengujian kasus kumulatif kematian. Jumlah set data untuk pengujian kasus kumulatif terkonfirmasi adalah 516 data, sedangkan jumlah set data untuk pengujian kasus kumulatif kematian adalah 537.

Tabel 5. Prediction of Confirmed Cases (Rasio Data 80:20, 150 Layers, 100 Epochs) and Death Cases (Rasio Data 70:30, 150 Layers, 100 Epochs)

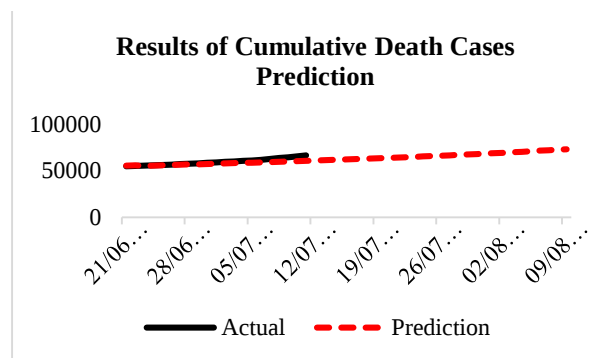
Date	Confirmed Cumulative Cases		Date	Death Cases	
	Actual	Prediction		Actual	Prediction
6/2/2021	1831773	1850090.73	6/21/2021	54956	55505.56
6/3/2021	1837126	1855497.26	6/22/2021	55291	55843.91
6/4/2021	1843612	1862048.12	6/23/2021	55594	55344.15144
6/5/2021	1850206	1868708.06	6/24/2021	55949	55638.05968
6/6/2021	1856038	1874598.38	6/25/2021	56371	55925.5349
6/7/2021	1863031	1881661.31	6/26/2021	56729	56209.26633
6/8/2021	1869325	1888018.25	6/27/2021	57138	56495.02125
6/9/2021	1877050	1895820.50	6/28/2021	57561	56765.08911
6/10/2021	1885942	1904801.42	6/29/2021	58024	57032.90278
6/11/2021	1894025	1912965.25	6/30/2021	58491	57292.18085
6/12/2021	1901490	1920504.90	7/1/2021	58995	57587.36119
6/13/2021	1911358	1930471.58	7/2/2021	59534	57880.34006
6/14/2021	1919547	1938742.47	7/3/2021	60027	58173.38485
6/15/2021	1927708	1946985.08	7/4/2021	60582	58467.66879
6/16/2021	1937652	1957028.52	7/5/2021	61140	58763.59393
6/17/2021	1950276	1969778.76	7/6/2021	61868	59064.33724
6/18/2021	1963266	1982898.66	7/7/2021	62908	59370.74899
6/19/2021	1976172	1995933.72	7/8/2021	63760	59684.78016
6/20/2021	1989909	2009808.09	7/9/2021	64631	60001.3094
6/21/2021		2024489.45	7/10/2021	65457	60321.59562
6/22/2021		2038294.13	7/11/2021	66464	60646.42318
6/23/2021		2053755.21	7/12/2021		60976.3721
6/24/2021		2074534.95	7/13/2021		61311.98286
6/25/2021		2093595.67	7/14/2021		61653.27524
6/26/2021		2114901.62	7/15/2021		62000.22946
6/27/2021		2136457.04	7/16/2021		62352.54892
6/28/2021		2157357.98	7/17/2021		62710.7675
6/29/2021		2178029.65	7/18/2021		63075.25432
6/30/2021		2200054.72	7/19/2021		63446.31256
7/1/2021		2207514.216	7/20/2021		63824.10042
7/2/2021		2233395.876	7/21/2021		64208.84858
7/3/2021		2261364.702	7/22/2021		64600.78775
7/4/2021		2288652.168	7/23/2021		65000.1552
7/5/2021		2318456.658	7/24/2021		65407.3464
7/6/2021		2349708.036	7/25/2021		65822.7107
7/7/2021		2384155.794	7/26/2021		66246.41945
7/8/2021		2422623.576	7/27/2021		66678.67699
7/9/2021		2460823.824	7/28/2021		67119.8722
7/10/2021		2495988.012	7/29/2021		67570.49942
7/11/2021		2532257.406	7/30/2021		68030.94094
7/12/2021		2572765.26	7/31/2021		68501.47358
7/13/2021		2620760.058	8/1/2021		68982.22918
7/14/2021		2675386.092	8/2/2021		69473.84048
7/15/2021		2732256.606	8/3/2021		69976.92707
7/16/2021		2786364.606	8/4/2021		70491.91736
7/17/2021		2838420.51	8/5/2021		71019.41117
7/18/2021		2883230.952	8/6/2021		71559.8369
7/19/2021		2917556.466	8/7/2021		72113.80097
			8/8/2021		72681.81088
			8/9/2021		73264.11052

Berdasarkan hasil prediksi arah tren cumulative cases dan death cases pada tabel 5 untuk bulan Juli-Agustus 2021 menghasilkan adanya peningkatan yang sangat signifikan pada bulan-bulan Juli dan Agustus di Indonesia. Prediksi jumlah kasus terkonfirmasi positif atau confirmed cumulative cases pada tanggal 19 Juli 2021 dengan mempelajari historikal data sebelumnya adalah sebanyak 2.917.557 (Pembulatan) kasus. Sedangkan, prediksi jumlah kasus kematian atau death cases pada tanggal 9 Agustus 2021 dengan mempelajari historikal data sebelumnya adalah sebanyak 73.265 (Pembulatan) kasus akan terjadi di Indonesia. Data kasus kumulatif terkonfirmasi positif dan kumulatif kematian yang digunakan adalah data yang diperoleh dari WHO. Data uji dan latih pada kasus kumulatif terkonfirmasi positif yang digunakan untuk prediksi adalah data dari tanggal 3 Januari 2020 sampai 20 Juni 2021. Sedangkan, data uji dan latih kasus kumulatif kematian diperoleh dari tanggal 3 Januari 2020 sampai 11 Juli 2021. Data ini merupakan gambaran pola di masa lampau yang menjadi dasar penambahan kasus pada hasil prediksi yang ada pada Tabel 5.

Hasil prediksi tren kasus kumulatif terkonfirmasi dan kasus kumulatif kematian pada Tabel 5 Juli-Agustus 2021 menghasilkan peningkatan yang sangat signifikan pada Juli & Agustus di Indonesia. Prediksi jumlah kumulatif kasus terkonfirmasi pada 19 Juli 2021 sebanyak 2.917.557 kasus (dibulatkan). Sementara itu, prediksi jumlah kumulatif kasus kematian pada 9 Agustus 2021 sebanyak 73.265 (dibulatkan) kasus yang akan terjadi di Indonesia dengan mempelajari data historis sebelumnya sebagai dasar pola penambahan kasus.



Gambar 7. Prediksi Cumulative Cases



Gambar 8. Prediksi Death Cases

Berdasarkan data uji dan latih yang berperan sebagai data historikal yang dipelajari dengan LSTM dan hasil prediksi yang telah diperoleh pada tabel 5, didapatkan arah tren COVID-19 di Indonesia untuk kasus kumulatif terkonfirmasi positif yakni arah tren naik/meningkat dan arah tren kasus kumulatif kematian yakni arah tren naik/meningkat (Dapat dilihat pada Gambar 6 dan 7).

IV. Kesimpulan

Metode LSTM untuk memprediksi kasus COVID-19 di Indonesia telah diterapkan dengan memperhatikan 3 (tiga) parameter seperti: data rasio, layer dari memory block, dan juga nilai epoch. Berdasarkan hasil uji coba, didapatkan bahwa prediksi dengan data latih dan uji sebesar 80:20, layer memory block sebanyak 150, dan epoch sebesar 100 menghasilkan MSE yang paling kecil dengan nilai MSE sebesar 0.0027 atau dibawah error rate 0.05. Sedangkan, perbandingan data latih dan uji sebesar 70:30, layer memory block sebanyak 150, dan epoch sebesar 100 menghasilkan MSE yang paling kecil dengan nilai MSE sebesar 0.0021 atau dibawah error rate 0.05. Hasil prediksi jumlah kasus terkonfirmasi positif atau confirmed cumulative cases pada tanggal 19 Juli 2021 adalah sebanyak 2.917.557 (Pembulatan) kasus. Sedangkan, prediksi jumlah kasus kematian atau cumulative death cases pada tanggal 9 Agustus 2021 adalah sebanyak 73.265 (Pembulatan) kasus akan terjadi di Indonesia. Hal ini berarti bahwa semakin banyak data latih, semakin akurat hasil prediksi yang dihasilkan.

Daftar Pustaka

- [1] A. Synowiec, A. Szczepański, E. Barreto-Duran, L. K. Lie, and K. Pyrc, "Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): a Systemic Infection," *Clin. Microbiol. Rev.*, vol. 34, no. 2, Mar. 2021, doi: 10.1128/CMR.00133-20.
- [2] A. AminJafari and S. Ghasemi, "The possible of immunotherapy for COVID-19: A systematic review," *Int. Immunopharmacol.*, vol. 83, p. 106455, Jun. 2020, doi: 10.1016/j.intimp.2020.106455.
- [3] R. Grieve *et al.*, "The importance of investing in data, models, experiments, team science, and public trust to help policymakers prepare for the next pandemic," *PLOS Glob. Public Heal.*, vol. 3, no. 11, p. e0002601, Nov. 2023, doi: 10.1371/journal.pgph.0002601.
- [4] S. Shah, A. Mulahuwaish, K. Z. Ghafoor, and H. S. Maghdid, "Prediction of global spread of COVID-19 pandemic: a review and research challenges," *Artif. Intell. Rev.*, vol. 55, no. 3, pp. 1607–1628, Mar. 2022, doi: 10.1007/s10462-021-09988-w.
- [5] S. Ryu *et al.*, "Epidemiology and Transmission Dynamics of Infectious Diseases and Control Measures," *Viruses*, vol. 14, no. 11, p. 2510, Nov. 2022, doi: 10.3390/v14112510.
- [6] J. H. Ellwanger, A. B. G. da Veiga, V. de L. Kaminski, J. M. Valverde-Villegas, A. W. Q. de Freitas, and J. A. B. Chies, "Control and prevention of infectious diseases from a One Health perspective," *Genet. Mol. Biol.*, vol. 44, no. 1 suppl 1, 2021, doi: 10.1590/1678-4685-gmb-2020-0256.
- [7] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons & Fractals*, vol. 135, p. 109864, Jun. 2020, doi: 10.1016/j.chaos.2020.109864.
- [8] T. Alafif, A. M. Tehame, S. Bajaba, A. Barnawi, and S. Zia, "Machine and Deep Learning towards COVID-19 Diagnosis and Treatment: Survey, Challenges, and Future Directions," *Int. J. Environ. Res. Public Health*, vol. 18, no. 3, p. 1117, Jan. 2021, doi: 10.3390/ijerph18031117.
- [9] S. Zheng *et al.*, "Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study," *BMJ*, p. m1443, Apr. 2020, doi: 10.1136/bmj.m1443.
- [10] J. Naskath, G. Sivakamasundari, and A. A. S. Begum, "A Study on Different Deep Learning Algorithms Used in Deep Neural Nets: MLP SOM and DBN," *Wirel. Pers. Commun.*, vol. 128, no. 4, pp. 2913–2936, Feb. 2023, doi: 10.1007/s11277-022-10079-4.
- [11] K. Ijaz, Z. Hussain, J. Ahmad, S. F. Ali, M. Adnan, and I. Khosa, "A Novel Temporal Feature Selection Based LSTM Model for Electrical Short-Term Load Forecasting," *IEEE Access*, vol. 10, pp. 82596–82613, 2022, doi: 10.1109/ACCESS.2022.3196476.
- [12] S. Arora, K. D. Bhaukhandi, and P. K. Mishra, "Coronavirus lockdown helped the environment to bounce back," *Sci. Total Environ.*, vol. 742, p. 140573, Nov. 2020, doi: 10.1016/j.scitotenv.2020.140573.
- [13] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [14] J. Shen, W. Wu, and Q. Xu, "Accurate Prediction of Temperature Indicators in Eastern China Using a Multi-Scale CNN-LSTM-Attention model," Dec. 2024, [Online]. Available: <http://arxiv.org/abs/2412.07997>
- [15] F. M. Salem, "Recurrent Neural Networks (RNN)," in *Recurrent Neural Networks*, Cham: Springer International Publishing, 2022, pp. 43–67. doi: 10.1007/978-3-030-89929-5_3.
- [16] Y. Zhang, "Research on Text Classification Method Based on LSTM Neural Network Model," in *2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, IEEE, Apr. 2021, pp. 1019–1022. doi: 10.1109/IPEC51340.2021.9421225.
- [17] D. Innovation, "Machine Learning in Business Analytics: Advancing Statistical Methods for Data-Driven Innovation," no. August, pp. 104–111, 2023, doi: 10.32996/jcsts.
- [18] World Health Organization, "Situation by Country, Territory & Area."
- [19] W. Ahmed and M. Bahador, "The accuracy of the LSTM model for predicting the S&P 500 index and the difference between prediction and backtesting," *KTH R. Inst. Technol.*, p. 35, 2018.