

OPTIMALISASI DATA TERBATAS PREDIKSI JANGKA PANJANG COVID-19 DENGAN KOMBINASI LSTM DAN GRU

Adhitho Satyo Bayangkari Karno¹, Widi Hastomo², dan Sutarno²

⁽¹⁾Universitas Gunadarma

Jl. Margonda Raya No. 100, Depok, Jawa Barat 16424

⁽²⁾STMIK Jakarta STI&K

Jl. BRI No.17, Radio Dalam, Kebayoran Baru, Jakarta Selatan 12140

{adh1t10.2, widie.has}@gmail.com

ABSTRAK

Jumlah data yang terbatas menjadi masalah utama Deep Learning (DL) untuk dapat menghasilkan prediksi data time series COVID-19 yang baik. Dalam penelitian ini dipergunakan 2 metode untuk dapat melakukan optimalisasi, pertama adalah memperoleh hasil prediksi dengan nilai error yang minimum, yaitu dengan mengkombinasikan hidden layer dengan 2 metode LSTM (Long Short Term Memory) dan GRU (Gated Recurrent Unit) dan yang kedua adalah memperkecil nilai time step agar dapat melipat gandakan jumlah data menjadi lebih banyak. Algoritma Deep Learning umumnya hanya dirancang untuk dapat melakukan prediksi jangka pendek (satu langkah kedepan setelah prediksi data testing). Dalam memperoleh prediksi jangka panjang yaitu dengan menjadikan hasil prediksi satu langkah kedepan menjadi satu data baru. Hal ini dilakukan secara berulang hingga mencapai prediksi jangka panjang yang diharapkan. Hasil dari penelitian ini adalah grafik prediksi dengan nilai RMSE (Root Mean Square Error) minimum untuk 100 hari kedepan. Pada fitur new_cases COVID-19 Indonesia diperoleh dengan nilai RMSE = 57.497234, epoch = 25, dengan susunan hidden layer pertama sampai keempat adalah LSTM, LSTM, LSTM, LSTM. Sedangkan untuk fitur new_deaths diperoleh nilai RMSE = 3.2645181, epoch = 55, dengan urutan hidden layer adalah GRU, GRU, GRU, GRU.

Kata Kunci : LSTM, GRU, Deep Learning, COVID-19

PENDAHULUAN

Berdasarkan sebaran data dari Gugus Tugas Percepatan Penanganan COVID-19 di Indonesia, penambahan kasus (new cases) dan kematian (new deaths) masih menunjukkan trend yang terus bertambah, kasus baru hingga tanggal 9 Agustus 2020 mencapai 1.893 kasus, hal ini seperti terlihat pada gambar 1.



Gambar 1. Data pergerakan kasus Covid-19[1]

Berbagai upaya telah dijalankan dengan maksimum oleh pemerintah Indonesia, namun hasil yang diperoleh masih jauh dari harapan. Sementara dampak pandemi ini telah menimbulkan macetnya roda perekonomian[2],[3], pengurangan

bertambah[4]–[6], serta kemungkinan besar akan timbulnya dampak buruk lainnya jika pandemi ini tidak cepat berakhir, yaitu resesi ekonomi[7], [8].

Harapan masyarakat untuk mengetahui kapan berakhirnya masa pandemi ini, merupakan salah satu pendorong dari penelitian ini untuk dapat mengetahui prediksi jangka panjang kedepan berdasarkan data yang ada. Dengan metode Deep Learning (DL) yang dipergunakan dalam penelitian ini terdapat beberapa tantangan yang dihadapi, jumlah data yang minim menjadi kendala untuk menghasilkan prediksi yang baik dan metode algoritma yang ada dalam DL hanya menghasilkan prediksi satu langkah kedepan menjadikan tantangan untuk dapat menjadikan DL mampu melakukan dan menghasilkan prediksi jangka panjang yang diharapkan.

Penelitian terkait prediksi penyebaran covid-19 telah banyak dilakukan menggunakan model statistik[9],[10]–[13] matematika[7],[14] dan mesin

belajar[15],[16],[17]. Namun prediksi jangka panjang menggunakan mesin belajar masih terbatas. Hasil dari penelitian [18] prediksi covid-19 dengan menggunakan metode LSTM eror mortality rate 3,2% pada hari ke 14, recovery rate mencapai 2,5%. Penelitian yang dilakukan [19] tentang penyebaran covid-19 dengan metode Support Vector Regression and stacking-ensemble akurasi error hingga 0.87%-3.51% pada hari pertama, sedangkan untuk tiga hari kedepan 1.02%–5.63% dan 0.95% -6.90% pada enam hari kedepan, jumlah objek 40.581. [20] berpendapat hasil prediksi image x-ray covid-19 dengan kombinasi CNN dengan LSTM menghasilkan akurasi mencapai 97%, spesifisitas 91%, dan sensitivitas 100% namun kendala pada keterbatasan jumlah data image.

Kebaruan pada penelitian ini yaitu mampu memprediksi penyebaran covid-19 di Indonesia dalam 100 hari kedepan.

METODE PENELITIAN

Dengan menggunakan paket *library seaborn*[21]–[23] di python, hasil penghitungan korelasi dapat ditunjukkan di matrik heatmap dalam bentuk numerik dan tingkatan warna.

$$r_{xi,y} = \frac{n(\sum X_i Y) - (\sum X_i)(\sum Y)}{\sqrt{\{n(\sum X_i^2) - (\sum X_i)^2\} \{n(\sum Y^2) - (\sum Y)^2\}}}$$

Dimana:

$\sum X_i$ = Jumlah data X_i

$\sum Y$ = Jumlah dari Y

$\sum X_i Y$ = Jumlah dari $X_i \cdot Y$

$\sum X_i^2$ = Jumlah dari X_i^2

Rumus korelasi Pearson, [24]

Hasil numerik korelasi berupa bilangan dengan rentang -1 dan 1, dimana nilai ± 1 menunjukkan hubungan yang kuat, dan nilai 0 menunjukkan hubungan yang rendah antara dua data.

Tingkat akurasi ditunjukkan selain secara visual dalam bentuk grafik juga dalam bentuk kuantitatif dengan mengukur nilai RMSE (*Root Mean Square Error*) dan MSE (*Mean Square Error*). Secara garis besar nilai RMSE dipahami untuk mengetahui seberapa jauh nilai kesalahan

dari suatu model terhadap garis regresi linier [25]. RMSE berhubungan dengan variasi sebaran frekuensi (*frequency distribution*) dari besar kesalahan yang diperoleh, tapi tidak dengan variasi kesalahan. Sedangkan MSE merupakan kuadrat nilai dari RMSE. Hasil dari RMSE mampu meredam perubahan kesalahan yang besar, sebaliknya MSE mampu melihat perubahan kesalahan yang kecil.

Dalam mesin learning biasanya MSE dipergunakan untuk mengukur kesalahan untuk setiap 1 epoch, dan RMSE mengukur kesalahan dalam sejumlah epoch tertentu.

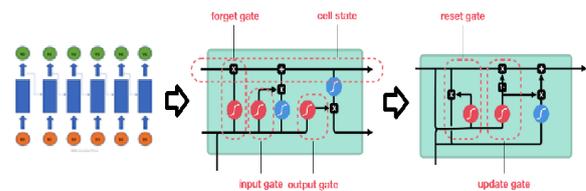
Rumus RMSE dan MSE:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_1 - \hat{y}_1)^2}{n}} \quad \text{dan}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_1 - \hat{y}_1)^2$$

$\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ = nilai yang diprediksi
 y_1, y_2, \dots, y_n = nilai yang diamati
 n = jumlah data yang diamati

Dengan terus bertambahnya jumlah layer dan node yang ada dalam suatu jaringan neural network, membuat suatu klasifikasi tersendiri dari mesin learning yang dinamakan Deep Learning (DL), seperti halnya 2 metode ML yang dipergunakan dalam penelitian ini (LSTM dan GRU).



Gambar 2. RNN, LSTM, GRU [26]

Formula GRU

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

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

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

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

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = O_t * \tanh(C_t)$$

Formula LSTM

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

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

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

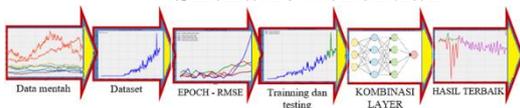
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = O_t * \tanh(C_t)$$

Model mesin learning LSTM dirancang untuk mengatasi kekurangan yang ada di model sebelumnya yaitu RNN (Recurrent Neural Network). Masalah *vanishing gradient* yang ada di RNN disebabkan karena ketidak mampuan untuk mengatasi informasi yang panjang [27]. Berbeda dengan RNN, LSTM berbentuk sel dengan rangkaian *gate* (4 *gate*) yang cukup rumit. Dengan jumlah *gate* yang banyak dan kerumitan jaringan di LSTM tentunya juga akan berpengaruh terhadap durasi proses pengolahan data. Berdasarkan hal tersebut maka dikembangkan lagi suatu model yaitu GRU. GRU berupa sel yang berisi hanya 2 *gate* dengan rangkaian yang lebih sederhana dibandingkan dengan LSTM. Kemampuan GRU lebih baik dari LSTM terutama untuk dataset yang jumlahnya sedikit [28].

Skenario Penelitian



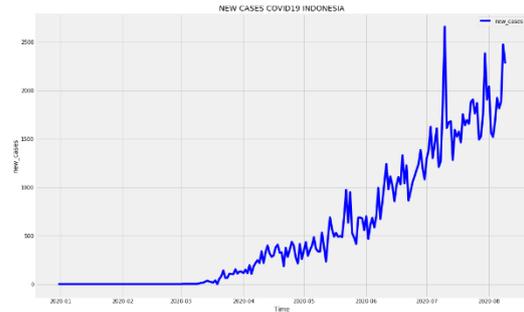
Gambar 3. Skenario Penelitian Prediksi COVID-19 Indonesia

Sumber data yang dipergunakan dalam penelitian ini diperoleh dari <https://ourworldindata.org>, berisidata COVID-19 seluruh negara di dunia. Dari data mentah dilakukan pemilihan data, dengan hanya mengambil data untuk negara Indonesia berisidat 35 fitur (kolom), 216 baris (31-12-2019 s/d 9-8-2020). Total dari 35 fitur, dilakukan pemilihan lagi untuk dipergunakan sebagai dataset dalam penelitian ini yaitu data penambahan kasus baru (*new_cases*) dan data kematian baru (*new_deaths*).

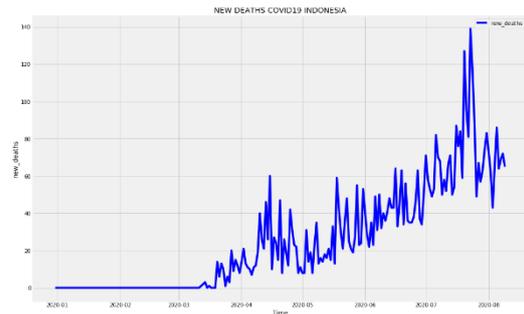
date	iso_code	continent	location	total_cases	new_cases	total_deaths	new_deaths	total_cases_per_million	new_cases_per_million	total
2019-12-31	IDN	Asia	Indonesia	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2020-01-01	IDN	Asia	Indonesia	0.0	0.0	0.0	0.0	0.0	0.0	0.0

date	iso_code	continent	location	total_cases	new_cases	total_deaths	new_deaths	total_cases_per_million	new_cases_per_million	total
2020-08-08	IDN	Asia	Indonesia	121228.0	2473.0	5583.0	72.0	443.201	0.041	
2020-08-09	IDN	Asia	Indonesia	123803.0	2277.0	5588.0	65.0	451.528	8.325	

Gambar 4. Data COVID-19 Indonesia dengan 35 fitur

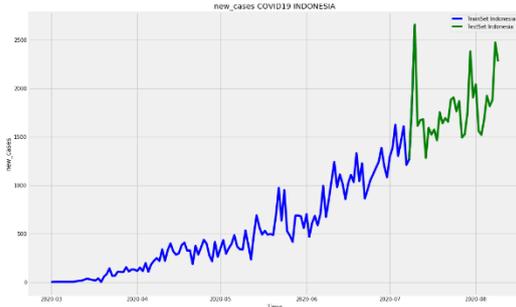


Grafik1-a. Sebaran *new_cases* COVID-19 Indonesia

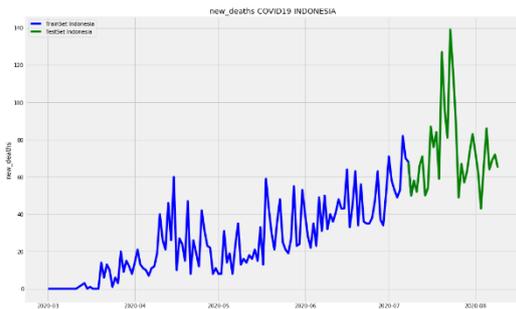


Grafik1-b. Grafik sebaran *new_deaths* COVID-19 Indonesia

Sebelum mesin melakukan prediksi, terlebih dahulu mesin akan melakukan proses belajar. Hasil belajar yang telah dilakukan, selanjutnya mesin akan diuji untuk mengetahui seberapa jauh hasil dari proses belajar yang telah dilakukan. Agar mesin dapat melakukan proses belajar dan pengujian, maka data harus dipisahkan untuk kedua proses tersebut dalam cara membagi data yang ada menjadi dua bagian yaitu data training (70%-80%) dan data testing (20%-30%).



Grafik2-a.Sebaran data training dan data testing new_cases COVID-19 Indonesia



Grafik 2-b.Sebaran data training dan data testing new_deaths COVID-19 Indonesia

Semakin banyak jumlah dataset yang dipergunakan maka mesin akan lebih mampu menghasilkan data prediksi dengan tingkat presisi yang baik. Jumlah data COVID-19 sangat terbatas (jumlah data sedikit) untuk kebutuhan yang diperlukan dalam proses DL. Agar mesin mampu menghasilkan prediksi data sebaik mungkin dengan menggunakan data yang terbatas ini, maka dilakukan upaya menggandakan data yaitu dengan memperkecil parameter time step dari 7 (7 hari dalam 1 minggu) menjadi 3. Parameter time step dengan 3 data akan ini dijadikan sebagai data input dalam proses DL nantinya. Semula data yang berupa barisan akan dirubah menjadi bentuk matrik dengan ukuran baris = 3 dan kolom = N/3 (N = jumlah data)

Target	Predict	DATA SET												
		X1	X2	X3	X4	X5	XN-2	XN-1	XN			
X4	P4	X1	X2	X3										
X5	P5		X2	X3	X4									
...							
X10	PN								XN-3	XN-2	XN-1			
	FN+1								XN-2	XN-1	XN			

Gambar 5.Matrik input dengan timestep=3

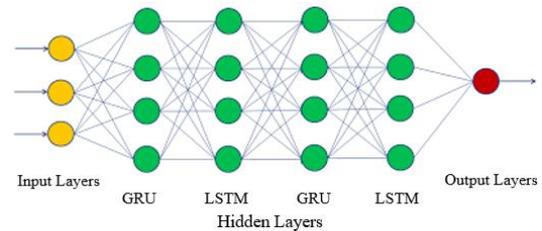
$$\begin{bmatrix} x_1 & x_2 & \dots & x_{N-3} & x_{N-2} \\ x_2 & x_3 & \dots & x_{N-2} & x_{N-1} \\ x_3 & x_4 & \dots & x_{N-1} & x_N \end{bmatrix}$$

Kombinasi layer

Arsitektur DL yang dipergunakan dalam penelitian ini melibatkan 4 hidden layer. Agar memperoleh hasil yang terbaik dilakukan 7 tipe kombinasi layer dengan menggunakan 2 model yaitu LSTM dan GRU.

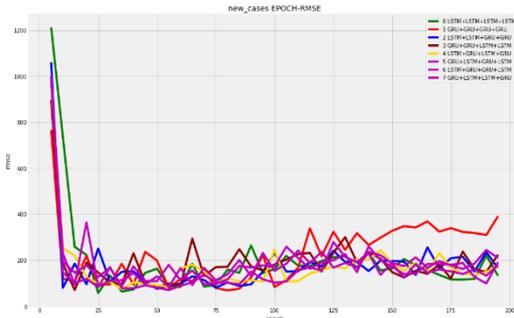
Tabel 1.7 tipe kombinasi hidden layer

Tipe	Layer-1	Layer-2	Layer-3	Layer-4
0	LSTM	LSTM	LSTM	LSTM
1	GRU	GRU	GRU	GRU
2	LSTM	LSTM	GRU	GRU
3	GRU	GRU	LSTM	LSTM
4	GRU	LSTM	GRU	LSTM
5	LSTM	GRU	GRU	LSTM
6	GRU	LSTM	LSTM	GRU

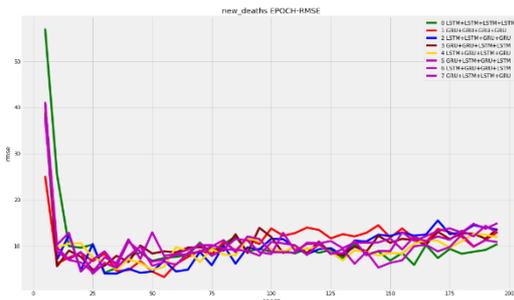


Gambar 6.Kombinasi hidden layer tipe 4

Salah satu kesulitan lain untuk menghasilkan prediksi yang akurat, adalah penetapan nilai epoch dalam proses DL. Tindakan yang umum dilakukan adalah dengan cara coba salah (try-error), yang memakan waktu lama, dan hasil yang kurang baik. Untuk itu sebelum menggunakan parameter epoch dalam proses DL, akan lebih baik diketahui dahulu karakteristik epoch dan RMSE terhadap data yang digunakan. Dalam penelitian ini dibuat program untuk menghasilkan grafik dan numerik yang menunjukkan karakteristik epoch dan RMSE dari setiap data dengan tipe hidden layer yang berbeda. Dengan diketahuinya grafik dan numerik karakteristik epoch dan RMSE, dapat dengan mudah menentukan nilai epoch yang akan dipergunakan dalam prosen DL.



Grafik 3-a. Epoch-RMSE new_cases COVID-19 untuk 7 tipe kombinasi hidden layer.



Grafik 3-b. Epoch-RMSE new_deaths COVID-19 untuk 7 tipe kombinasi hidden layer.

Algoritma DL untuk prediksi data time series [29],[30] hanya mampu melakukan satu langkah prediksi kedepan. Kemudian hasil prediksi akan dihitung nilai errornya dengan data real. Kemudian nilai error ini dipergunakan untuk update nilai bobot untuk iterasi selanjutnya sehingga dihasilkan prediksi satu langkah kedepan lagi (prediksi langkah kedua). Hal ini dilakukan terus menerus hingga mencapai data real akhir (akhir data real). Setelah itu DL tidak dapat melakukan prediksi untuk langkah selanjutnya karena tidak ada lagi data real.

Agar dapat memperoleh prediksi jangka panjang, hasil prediksi terakhir dijadikan sebagai data baru dengan melakukan sedikit kombinasi dengan nilai RMSE. Dengan adanya data baru ini maka DL akan mampu melakukan satu langkah prediksi lebih kedepan lagi. Hal ini dilakukan berulang hingga mencapai jumlah langkah kedepan yang diinginkan (dalam penelitian ini dilakukan sampai 100 langkah (hari) kedepan).

ITER = 1										
Target	Predict	DATA SET								
		X1	X2	X3	X4	X5	XN-2	XN-1
X4	P4	X1	X2	X3						
X5	P5		X2	X3	X4					
...			
X10	PN							XN-3	XN-2	XN-1
	FN+1								XN-2	XN-1

Gambar 7-a. Gerak time step (input) setelah mencapai ujung data real, menghasilkan satu langkah prediksi (F_{N+1}).

ITER = 2										
Target	Predict	DATA SET								
		X1	X2	X3	X4	X5	XN-2	XN-1
X4	P4	X1	X2	X3						
X5	P5		X2	X3	X4					
...			
X10	P10							XN-3	XN-2	XN-1
XN+1	FN+1								XN-2	XN-1
	FN+2								XN-1	XN

Gambar 7-b. Menjadikan F_{N+1} sebagai data baru (x_{N+1}), sehingga menghasilkan prediksi kedua (F_{N+2}).

ITER = 3										
Target	Predict	DATA SET								
		X1	X2	X3	X4	X5	XN-2	XN-1
X4	P4	X1	X2	X3						
X5	P5		X2	X3	X4					
...			
X10	P10							XN-3	XN-2	XN-1
XN+1	FN+1								XN-2	XN-1
XN+2	FN+2								XN-1	XN
	FN+3								XN	XN+1

Gambar 7-c. Menjadikan F_{N+2} sebagai data baru (x_{N+2}), sehingga menghasilkan prediksi ketiga (F_{N+3}).

ITER = 4										
Target	Predict	DATA SET								
		X1	X2	X3	X4	X5	XN-2	XN-1
X4	P4	X1	X2	X3						
X5	P5		X2	X3	X4					
...			
X10	P10							XN-3	XN-2	XN-1
XN+1	FN+1								XN-2	XN-1
XN+2	FN+2								XN-1	XN
XN+3	FN+3								XN	XN+1
	FN+4								XN+1	XN+2

Gambar 7-d. Menjadikan F_{N+3} sebagai data baru (x_{N+3}), sehingga menghasilkan prediksi keempat (F_{N+4}), dan seterusnya.

HASIL DAN PEMBAHASAN

Terlihat pada gambar 10-a dan 10-b, dengan mengambil nilai epoch dengan RMSE minimum untuk tiap tipe kombinasi layer dapat dinyatakan secara numerik dalam tabel berikut ini: (Tabel 2).

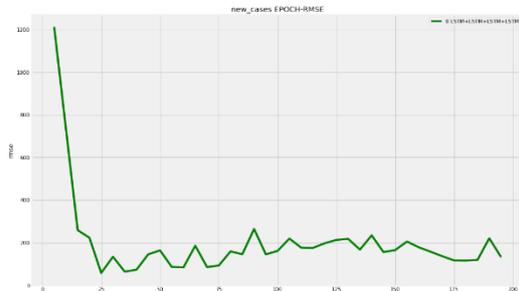
Tabel 2. Tabel epoch RMSE minimum untuk 7 tipe layer data new_cases dan new_deaths COVID-19 Indonesia.

Tipe Layer	new_cases		new_deaths	
	epoch	RMSE	epoch	RMSE
0	25	57.497234	30	4.2025036
1	80	68.973643	55	3.2645181
2	50	75.423218	35	4.0128964
3	15	70.581916	25	4.8586305

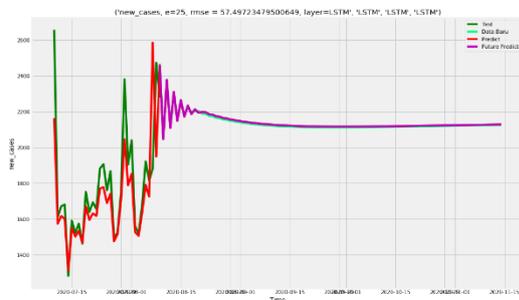
4	35	74.295289	50	4.8382693
5	55	83.452558	25	4.3652808
6	35	86.514773	20	4.4383484

Pada tabel 2, terlihat bahwa untuk data dengan fitur new_cases hasil prediksi terbaik yang diperoleh ada di arsitektur hidden layer tipe 0 (LSTM,LSTM,LSTM,LSTM), dan untuk fitur new_deaths ada di tipe 1 (GRU,GRU,GRU,GRU).

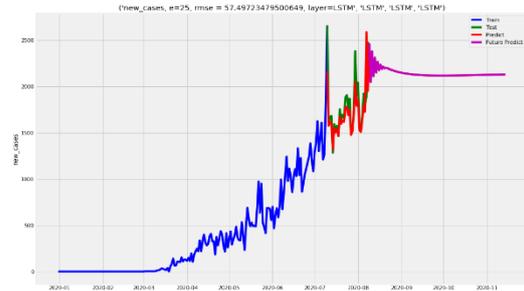
Hasil tampilan grafik dari tabel 2 secara lengkap terlampir di akhir tulisan ini. Sedangkan hasil grafik dengan prediksi terbaik sesuai dengan tabel 2 (berwarna merah) untuk fitur new_cases terlihat pada gambar di bawah ini: (grafik 4).



Grafik 4-a. Karakteristik epoch-RMSE untuk tipe 0, new_cases COVID-19 Indonesia

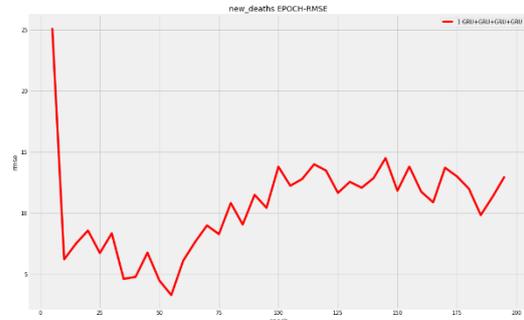


Grafik 4-b. Sebaran dataset testing, prediksi dan future prediksi untuk tipe 0, new_cases COVID-19 Indonesia

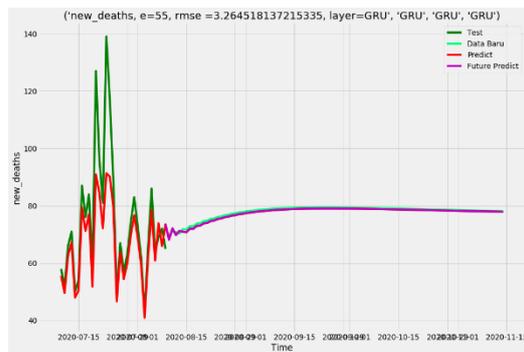


Grafik 4-c. Sebaran dataset training, testing, prediksi dan future prediksi untuk tipe 0, new_cases COVID-19 Indonesia

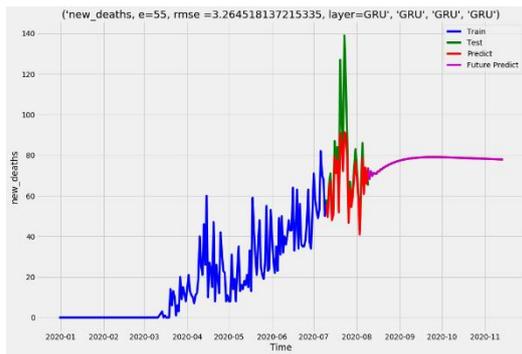
Hasil grafik dengan prediksi terbaik sesuai dengan tabel 2 (berwarna merah) untuk fitur new_deaths terlihat pada grafik di bawah ini: (grafik 5).



Grafik 5-a. Karakteristik epoch-RMSE untuk tipe 1, new_deaths COVID-19 Indonesia



Grafik 5-b. Sebaran dataset testing, prediksi dan future prediksi untuk tipe 1, new_deaths COVID-19 Indonesia



Grafik5-c. Sebaran dataset training, testing, prediksi dan future prediksi untuk tipe 1, new_deaths COVID-19 Indonesia

PENUTUP

- Hasil dari penelitian ini memberikan prediksi yang sangat baik, terlihat secara grafik sebaran data prediksi dapat mengikuti sebaran data testing. Prediksi terbaik untuk data new_cases dengan nilai RMSE = 57.497234 dan untuk data new_death dengan nilai RMSE = 3.2645181.
- Hasil dari tabel dan grafik terlihat bahwa optimalisasi dengan mengkombinasikan 4 hidden layer dengan model LSTM dan GRU, mampu memberikan pilihan tipe layer dengan hasil prediksi yang lebih baik.
- Dengan melakukan penggandaan data (augmented data), keterbatasan data COVID-19 dapat teratasi sehingga proses DL dapat terlaksana dengan hasil yang sangat baik.
- Diketahuinya prediksi jangka panjang dengan menggunakan DL dapat dilakukan dengan menjadikan data prediksi akhir menjadi data baru. Sehingga trend untuk 100 hari kedepan setelah data real dapat terlihat dalam bentuk grafik.
- Interpretasi dari sebaran grafik yang diperoleh dalam penelitian ini dapat dikatakan bahwa untuk kasus baru (new_cases) akan sedikit mengalami penurunan dan untuk data kematian (new_deaths) akan mengalami kenaikan.
- Masih perlu dilakukan penelitian dengan menggunakan kombinasi hidden layer lain atau menambah jumlah hidden layer sehingga variasi kombinasi dapat

lebih banyak lagi untuk menghasilkan prediksi yang lebih baik.

- Masih diperlukan cara yang lebih baik untuk dapat menghasilkan fluktuasi prediksi jangka panjang lebih dari sekedar penelitian ini yang hanya mampu menunjukkan trend jangka panjang.

DAFTAR PUSTAKA

- [1] Antonius Budi Nurcahyo, "Data Pergerakan Kasus Positif Harian Covid-19 sampai 9 Agustus 2020," 2020. [Online]. Available: <https://www.beritasatu.com/berita-grafik/663651-data-pergerakan-kasus-positif-harian-covid19-sampai-9-agustus-2020>.
- [2] M. Olimov and I. Khotamov, "Negative Impacts of Coronavirus Pandemic on Industrial Production in Uzbekistan and Measures To Overcome," *Архив Научных Исследований*, vol. 1, no. 4, pp. 266–271, 2020.
- [3] S. R. Gern, Klaus-Jürgen; Möhle, "The impact of the COVID-19 pandemic on the global economy: Survey-based evidence from free zones," 2020.
- [4] E. Kong and D. Prinz, "The Impact of Non-Pharmaceutical Interventions on Unemployment During a Pandemic," *SSRN Electron. J.*, 2020.
- [5] G. Acs and M. Karpman, "Employment , Income , and Unemployment Insurance during the COVID-19 Pandemic," pp. 1–11, 2020.
- [6] L. C. Pullen, "Pandemic-related unemployment impacts transplant in the US," *Am. J. Transplant.*, vol. 20, no. 8, pp. 1957–1958, 2020.
- [7] L. Elliott, "Spillover of COVID-19: Impact on the Global Economy," *SSRN Electron. J.*, no. 99850, 2020.
- [8] C. Hevia and A. Neumeyer, "A Conceptual Framework for Analyzing the Economic Impact of COVID-19 and its Policy Implications," *COVID19 Policy Doc. Ser.*, no. 1, pp. 1–18, 2020.
- [9] N. Chintalapudi, G. Battineni, and F.

- Amenta, "COVID-19 virus outbreak forecasting of registered and recovered cases after sixty day lockdown in Italy: A data driven model approach," *J. Microbiol. Immunol. Infect.*, vol. 53, no. 3, pp. 396–403, 2020.
- [10] M. Henrique, D. Molin, R. Gomes, V. Cocco, and S. Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil," no. January, 2020.
- [11] R. K. Singh *et al.*, "Prediction of the COVID-19 Pandemic for the Top 15 Affected Countries: Advanced Autoregressive Integrated Moving Average (ARIMA) Model," *JMIR Public Heal. Surveill.*, vol. 6, no. 2, p. e19115, 2020.
- [12] H. Tandon, P. Ranjan, T. Chakraborty, and V. Suhag, "Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future," no. January 2020, pp. 1–11, 2020.
- [13] T. Dehesh, H. A. Mardani-Fard, and P. Dehesh, "Forecasting of COVID-19 Confirmed Cases in Different Countries with ARIMA Models," pp. 1–12, 2020.
- [14] I. Ciufolini and A. Paolozzi, "Mathematical prediction of the time evolution of the COVID-19 pandemic in Italy by a Gauss error function and Monte Carlo simulations," *Eur. Phys. J. Plus*, vol. 135, no. 4, pp. 1–8, 2020.
- [15] F. Rustam *et al.*, "COVID-19 Future Forecasting Using Supervised Machine Learning Models," *IEEE Access*, vol. 8, pp. 101489–101499, 2020.
- [16] S. K. Bandyopadhyay and S. Dutta, "Machine Learning Approach for Confirmation of COVID-19 Cases: Positive, Negative, Death and Release," *medRxiv*, vol. 00, no. Cdc, p. 2020.03.25.20043505, 2020.
- [17] G. Pinter, I. Felde, A. Mosavi, P. Ghamisi, and R. Gloaguen, "COVID-19 Pandemic Prediction for Hungary; A Hybrid Machine Learning Approach," *SSRN Electron. J.*, 2020.
- [18] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons and Fractals*, vol. 135, 2020.
- [19] Z. Car, S. B. Š, N. An, I. Lorencin, and V. Mrzljak, "Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron," vol. 2020, 2020.
- [20] M. Z. Islam, M. M. Islam, and A. Asraf, "A Combined Deep CNN-LSTM Network for the Detection of Novel Coronavirus (COVID-19) Using X-ray Images," no. June, pp. 1–20, 2020.
- [21] S. Mohanty, A. Mishra, and A. Saxena, "EasyChair Preprint Medical Data Analysis Using Machine Learning with KNN MEDICAL DATA ANALYSIS USING MACHINE," 2020.
- [22] Blessing Nwamaka Iduh, "Python programming — plotting," *Int. J. Latest Technol. Eng. Manag. Appl. Sci.*, vol. 9, no. 2, p. 26, 2020.
- [23] and H. F. Flores, Andrew, "Sentiment Analysis on Airline Tweets Using Naïve Bayes Classifier," *athena.ecs.csus.edu*, 2020.
- [24] K. Pearson, "Notes on Regression and Inheritance in the Case of Two Parents Proceedings of the Royal Society of London," vol. 58, pp. 240–242, 1895.
- [25] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: an analysis and review," *Int. J. Forecast. Elsevier*, vol. 16, no. 4, pp. 437–450, 2000.
- [26] Anonim, "Illustrated guide to lstms and gru step by step explanation," *Medium*, 2019. [Online]. Available: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb8>. [Accessed: 19-Aug-2020].
- [27] S. Hochreiter and J. Uergen Schmidhuber, "Long Short Term Memory," *Neural Comput.*, vol. 9, no. 8, p. 17351780, 1997.

- [28] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," pp. 1–9, 2014.
- [29] W. Hastomo and A. Satyo, "Long Short Term Memory Machine Learning Untuk Memprediksi Akurasi Nilai Tukar IDR Terhadap USD," vol. 3, 2019.
- [30] Adhitio Satyo Bayangkari Karno, "Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long ShortTerm Memory)," vol. 1, no. 1, pp. 1–8, 2020.

LAMPIRAN
Data New_Cases

Tipe kombinasi layer	Grafik data testing, prediksi, future prediksi	Grafik data training, testing, prediksi, future prediksi	Epoch-RMSE tiap tipe	Epoch-RMSE gabung
0				
1				
2				
3				
4				
5				
6				

Data New_Deaths

Tipe kombinasi layer	Grafik data testing, prediksi, future prediksi	Grafik data training, testing, prediksi, future prediksi	Epoch-RMSE tiap tipe	Epoch-RMSE gabung
0				
1				

