

Model Prediksi Jumlah Kumulatif Kasus COVID-19 di Indonesia Menggunakan Metode *Neural Network*

Afifah Cahyaningsih ^{#1}, Novantri Prasetya Putra ^{*2}, Andre Pradika Ekoputro Pratama ^{#3}, Rafian Ramadhani ^{#4}

*Fakultas Informatika, Institut Teknologi Telkom Purwokerto
Kawasan Pendidikan Telkom, Jl. DI Panjaitan No 128 Purwokerto 53147 Indonesia*

¹ 19102123@ittelkom-pwt.ac.id

² 18102279@ittelkom-pwt.ac.id

³ 16102148@ittelkom-pwt.ac.id

⁴ 18102245@ittelkom-pwt.ac.id

Abstrak

Pandemi COVID-19 belum menunjukkan tanda-tanda akan berakhir. Hal tersebut ditunjukkan dengan jumlah kasus positif COVID-19 yang masih bertambah di dunia maupun di Indonesia. Jumlah kumulatif kasus positif COVID-19 berpengaruh dalam melakukan pertimbangan keputusan dan kebijakan terkait penanganan COVID-19 di Indonesia. Oleh karena itu, penelitian ini berfokus untuk membuat model prediksi jumlah kumulatif kasus positif COVID-19 di Indonesia. Data yang digunakan dalam penelitian ini diambil dari Satuan Tugas Penanganan COVID-19 Indonesia per 14 November 2020. Variabel yang digunakan pada penelitian ini adalah jumlah harian kumulatif kasus positif COVID-19 di Indonesia. *Neural network* digunakan untuk melakukan modeling data prediksi dengan tujuan untuk membuat model prediksi dari jumlah kumulatif kasus positif COVID-19 di Indonesia. Dari penelitian yang dilakukan, didapatkan hasil error *R-squared* sebesar 0.9793. Hasil tersebut, menjelaskan bahwa model yang digunakan baik. Semakin nilai error R-squared mendekati 1 maka model yang digunakan akan semakin baik. Hasil modeling diharapkan dapat digunakan pemerintah untuk dijadikan sebagai alternatif model yang digunakan untuk melakukan prediksi jumlah kumulatif kasus positif COVID-19 di Indonesia.

Keywords: covid-19, *machine learning*, *deep learning*, *neural network*, prediksi

I. PENDAHULUAN

Corona Virus Diseases 2019 (COVID-19) ditemukan pertama kali di Wuhan, Provinsi Huabei, China pada Desember 2019 [1]. Virus ini menyebar lewat percikan cairan dari mulut (*droplet*) yang menyebabkan infeksi pernafasan ringan seperti flu dan menyebabkan infeksi pernafasan berat seperti infeksi paru paru (*pneumonia*) [2]. Virus ini dapat menyerang siapa saja, bahkan pada kasus yang lebih parah virus ini dapat menyebabkan kematian. Sampai saat ini, belum ada obat khusus yang disarankan untuk mencegah atau mengobati penyakit yang disebabkan virus ini. Korban yang terinfeksi harus menerima perawatan yang tepat untuk meredakan dan mengobati gejala, dan pada kasus yang sakit serius harus dibawa ke rumah sakit [3]. Mudah-mudahan penyebaran virus ini mengakibatkan tingginya jumlah korban yang positif terjangkau, meskipun

berbagai negara telah mengupayakan untuk menekan jumlah korban. Oleh karena itu virus ini telah ditetapkan sebagai global pandemi oleh *World Health Organization* (WHO).

Penyebaran COVID-19 di Indonesia dimulai pada 2 Maret 2020 [4]. Saat ini jumlah kasus korban terjangkit virus ini telah mencapai 463.007 orang per 14 November 2020 [5]. Tingginya jumlah kasus ini telah menjadi ancaman nyata bagi keselamatan dan kesehatan masyarakat. Pemerintah telah mengeluarkan berbagai kebijakan yang bertujuan menekan jumlah kasus korban virus ini, salah satunya dengan melakukan pembelajaran secara daring, pengurangan jam kerja, PSBB (pembatasan sosial berskala besar), dan lain-lain [6]. Meningkatnya jumlah korban virus ini ditambah dengan munculnya kluster penyebaran baru, menjadi pengingat untuk melakukan berbagai upaya untuk menekan laju penyebaran virus ini. Salah satu upaya yang dapat dilakukan dengan melakukan pemodelan prediksi jumlah kumulatif kasus positif COVID-19. Model prediksi ini dapat digunakan sebagai proyeksi dan optimisasi bagi pemerintah untuk mengambil kebijakan dan mengevaluasi kebijakan yang ada [7]. Pada penelitian ini dilakukan pemodelan prediksi jumlah kumulatif kasus positif COVID-19 menggunakan algoritma *neural network*. Algoritma *neural network* dipilih karena telah banyak dipakai untuk pembuatan model prediksi seperti di sektor kesehatan, perindustrian, pertanian, dll [1][2][8][9]. Tidak hanya itu, model *neural network* memiliki keuntungan dalam hal pengolahan data, dikarenakan model *neural network* dapat mengolah data yang bersifat kompleks [10]. Dalam mengaplikasikan model dapat dilakukan pada suatu sistem yang bekerja secara *realtime*, sehingga proses pengambilan kebijakan dan evaluasi kebijakan pemerintah dalam penekanan laju penyebaran COVID-19 dapat dioptimasi.

II. TINJAUAN PUSTAKA

A. Studi Literatur

Penelitian ini menggunakan beberapa sumber referensi dari penelitian yang sudah diteliti sebelumnya, yang pertama ada penelitian dari Fra Siskus Dian Arianto, Noviyanti P (2020) [4] yang berjudul “Prediksi Kasus COVID-19 di Indonesia Menggunakan Metode Backpropagation dan Fuzzy Tsukamoto”. Pada penelitian ini dibuat model prediksi penambahan kasus COVID-19 di Indonesia menggunakan data *time series* dengan metode pembelajaran Fuzzy Tsukamoto dan menghasilkan MSE (*Mean Square Error*) pada data normalisasi sebesar 1,632337.

Penelitian selanjutnya dilakukan oleh Adhithio Satyo Bayangkari Karno, Widi Hastomo, Indra Sari Kusuma Wardhana (2020) [11] dengan judul “Prediksi Jangka Panjang COVID-19 Indonesia Menggunakan Deep Learning”. Penelitian ini lakukan berulang hingga mencapai prediksi selama 50 hari kedepan menggunakan metode *deep learning* berupa *Long Short Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) pada 4 hidden layer *machine learning* yang digunakan dan didapatkan hasil terbaik dengan perolehan nilai RSME (*Root Mean Square Error*) 206.632 untuk *epoch* sejumlah 5.

Dengan referensi jurnal lainnya oleh Prisca Deviani Pakan (2020) [12] berjudul “Peramalan Kasus Positif COVID-19 di Indonesia Menggunakan LSTM”. Dalam penelitian ini dikembangkan metode prediksi *time series* untuk memperkirakan jumlah kasus penyebaran COVID-19 dengan jaringan saraf tiruan utamanya *Long Short Term Memory* dan lebih baik kinerjanya dibandingkan dari algoritma ERNN dalam memprediksi jumlah kasus penyebaran pandemi ini. Didapatkan jumlah RMSE (*Root Mean Square Error*) sebesar 115,04 menggunakan LSTM (*Long Short Term Memory*) sedangkan untuk ERNN (*Elman Recurrent Neural Network*) diperoleh RMSE (*Root Mean Square Error*) sebesar 124,34.

Untuk penelitian lainnya oleh Wirawan Setialaksana, Dwi Rezky Anandari Sulaiman, Shabrina Syntha Dewi, Chairunnisa Ar Lamasitudju, Nini Rahayu Ashadi, Muhammad Asriadi (2020) [13] berjudul “Model Jaringan Syaraf Tiruan dalam Peramalan Kasus Positif Covid-19 di Indonesia”. Dalam penelitian ini dibandingkan prediksi total kasus pandemi COVID-19 dengan metode MLP (*Multi Layer Perceptron*) dan ELM (*Extreme Learning Machine*). Menggunakan tiga pengukuran akurasi prediksi yaitu MAE (*Mean*

Absolute Error), MAPE (*Mean Absolute Percentage Error*), dan RMSE (*Root Mean Square Error*) dengan perolehan nilai prediksi ketiganya lebih tinggi pada metode ELM dibanding metode MLP.

Perbandingan dari penelitian yang telah dilakukan sebelumnya adalah pada penelitian ini menggunakan model *neural network* untuk melakukan prediksi jumlah kumulatif kasus positif COVID-19 di Indonesia.

B. *Gathering Data*

Gathering data atau biasa disebut pengumpulan data merupakan proses paling awal dalam melakukan sebuah penelitian, dikarenakan penelitian tidak akan berjalan tanpa adanya data. Pengumpulan data dapat dilakukan secara langsung ataupun mengambil *dataset* dari *website* penyedia *dataset*. Pada penelitian ini, data yang digunakan merupakan data sekunder yang didapatkan dari *website* Satuan Tugas Penanganan COVID-19 Indonesia. *Website* tersebut merupakan *website* pemerintah untuk menampilkan data-data tentang COVID-19 secara *real time*. Data akan diakuisisi dengan menggunakan API. Data yang digunakan sebanyak 243 data dihitung dari awal kemunculan COVID-19 di Indonesia pada 2 Maret 2020. *Dataset* yang didapatkan memiliki beberapa variabel seperti jumlah meninggal, jumlah positif, jumlah sehat, jumlah dirawat dan masih ada beberapa variabel lain. Namun, pada penelitian ini variabel yang digunakan adalah jumlah kumulatif kasus positif.

C. *Data Preprocessing*

Kualitas sebuah model yang dibangun sangat bergantung kepada kualitas data yang digunakan. Kualitas data yang baik dapat didapatkan dari *data preprocessing*, proses *preprocessing* membuat data lebih mudah dipahami dan lebih terstruktur [14]. *Data processing* merupakan proses untuk menyiapkan data mentah untuk dapat digunakan lebih lanjut sesuai kebutuhan [15][16]. Proses pada *preprocessing* data meliputi pemilihan data, penghapusan data, pengubahan data, penambahan data, dan lain-lain.

D. *Windowing*

Windowing adalah teknik yang digunakan dalam menentukan data masukan dan data *output* dalam memprediksi data runtun waktu dengan menggunakan tipe *univariat*. Dengan teknik *windowing* maka data akan dibagi menjadi bagian masukan dan keluaran data [17]. Cara kerja teknik *windowing* sangat bergantung pada tipe *window* yang digunakan [18].

E. *Neural Network*

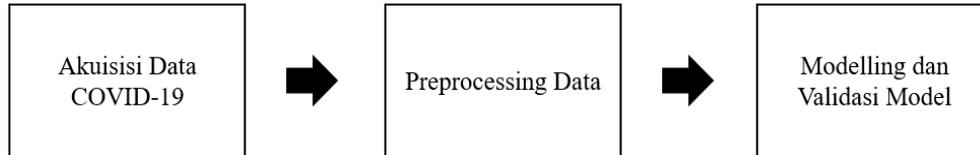
Neural Network atau jaringan saraf tiruan (*Artificial Neural Network*) adalah fungsi matematika dari penjabaran fungsi otak untuk menjalankan penghitungan secara paralel. Kegunaan dari *neural network* adalah untuk penyelesaian hitungan paralel serta hitungan rumit seperti pembuatan prediksi, pengenalan pola, dan permodelan. Proses yang terjadi pada *neural network* mirip seperti cara penyampaian perintah pada otak manusia yakni perintah berfungsi sebagai dendrit lalu hasil dari perintah berfungsi sebagai akson dengan fungsi aktivasi seperti fungsi dari sinapsis pada otak manusia. Penggunaan *neural network* pada komputer dapat membuat komputer bekerja layaknya jaringan syaraf pada manusia untuk dapat menyelesaikan proses perhitungan. Dengan cara meniru kerja pada sistem manusia dengan sistem pengenalan pola [19].

F. *R-square*

Merupakan sebuah proses untuk melakukan validasi terhadap suatu data setelah melalui proses modelling. *R-square* sering disebut dengan koefisien determinasi yang merupakan koefisien pengukur kebaikan model (*goodness of fit*) dari persamaan regresi, yaitu memberikan proporsi atau persentase variasi total dalam variabel terikat yang dijelaskan oleh variabel bebas. Untuk melakukan metode *R-Square* diperlukan beberapa tahap yaitu yang pertama dengan mencari nilai rata-rata data yang digunakan. Selanjutnya tahap kedua mencari *total sum of square* dan *residual sum of square*. Pada tahap terakhir akan dihitung *coefficient of determination* menggunakan *R-squared* [20][21].

III. METODE PENELITIAN

Secara umum, alur penelitian ditampilkan pada gambar 3.1.



Gambar 3.1 Metode Penelitian

1) Akuisisi Data COVID-19

Dataset COVID-19 didapatkan dari [3] yang berisi 243 baris dan 8 fitur data, pada *dataset* yang diperoleh akan dilakukan *filtering* untuk mendapatkan data fitur kolom jumlah kumulatif kasus positif COVID-19.

	tanggal	key	jumlah_item	jumlah_meninggal	jumlah_sembuh	jumlah_positif	jumlah_dirawat	jumlah_positif_kum	jumlah_sembuh_kum	jumlah_meni
0	2020-03-02	1583107200000	1	0	0	2	2	2	0	
1	2020-03-03	1583193600000	1	0	0	0	0	2	0	
2	2020-03-04	1583280000000	1	0	0	0	0	2	0	
3	2020-03-05	1583366400000	1	0	0	0	0	2	0	
4	2020-03-06	1583452800000	1	0	0	2	2	4	0	

Gambar 3.2 Data COVID-19

2) Data Preprocessing

Proses *data preprocessing* pada penelitian ini dilakukan dengan memilih data yang dibutuhkan untuk melakukan prediksi yaitu data fitur kolom jumlah kumulatif kasus positif COVID-19. Setelah proses *data preprocessing*, dilanjutkan dengan proses *windowing* yang memiliki tujuan agar data menjadi *multivariate*. *Windowing* ditentukan sebanyak 15 sehingga menghasilkan 15 fitur data hasil *windowing*.

	tanggal	jumlah_positif_kum
0	2020-03-02	2
1	2020-03-03	2
2	2020-03-04	2
3	2020-03-05	2
4	2020-03-06	4
5	2020-03-07	4
6	2020-03-08	6
7	2020-03-09	19
8	2020-03-10	27
9	2020-03-11	34

Gambar 3.3 Data Sebelum Dilakukan Windowing

	jumlah_positif_kum	jumlah_positif_kum_(t-1)	jumlah_positif_kum_(t-2)	jumlah_positif_kum_(t-3)	jumlah_positif_kum_(t-4)	jumlah_positif_kum_(t-5)	jumlah_positif_kum
0	172	134	117	96	69	34	
1	227	172	134	117	96	69	
2	309	227	172	134	117	96	
3	369	309	227	172	134	117	
4	450	369	309	227	172	134	

Gambar 3.4 Data Setelah Dilakukan Windowing

Setelah dilakukan proses *windowing*, kemudian dilakukan proses *split data* (pembagian data) menjadi data *training* sebanyak 174 baris data, data *validation* sebanyak 20 baris data, dan data *testing* sebanyak 49 baris data.

3) Modelling (Neural Network)

Setelah melalui proses *split data*, kemudian dilakukan modeling dengan menggunakan *neural network*. Metode *neural network* digunakan dalam membangun model prediksi dan *R-Square* digunakan untuk melakukan validasi model yang telah dibuat.

Layer (type)	Output Shape	Param #
input_5 (InputLayer)	[(None, 15)]	0
dense_16 (Dense)	(None, 32)	512
dense_17 (Dense)	(None, 64)	2112
dense_18 (Dense)	(None, 128)	8320
dropout_4 (Dropout)	(None, 128)	0
dense_19 (Dense)	(None, 1)	129
Total params: 11,073		
Trainable params: 11,073		
Non-trainable params: 0		

Gambar 3.5 Arsitektur Modelling

Model *neural network* yang dibuat terdiri dari 1 *input layer*, 4 *dense layer* dan 1 *dropud layer* dengan parameter *epoch* sebanyak 20. Waktu yang dibutuhkan untuk menjalankan training model dengan 20 epoch yaitu 0.87 detik. Sedangkan untuk menjalankan training model dengan menggunakan 10 epoch hanya membutuhkan 0.35 detik. Dapat disimpulkan bahwa semakin banyak epoch yang digunakan maka akan semakin lama waktu yang dibutuhkan dalam menjalankan training model.

IV. HASIL DAN PEMBAHASAN

Setelah melalui proses sesuai dengan metode penelitian, didapatkan *loss training* dan *loss validation* sebagai berikut :



Gambar 4.1 Data *Loss Training* dan *Validation Model*

Berdasarkan Gambar 4.1 nilai *loss training* dan *validation model* terkecil terdapat pada *epoch* ke-20. Hal tersebut menjelaskan bahwa estimasi output model prediksi mendekati output model sebenarnya. Didapatkan nilai error *R-squared* sebesar 0.9793 dan didapatkan model prediksi kumulatif kasus positif COVID-19 sebagai berikut.



Gambar 4.2 Model Prediksi COVID-19

Dilustrasikan pada Gambar 4.1. Data hasil prediksi model mendekati dengan data aktual. Dari hasil tersebut dapat disimpulkan bahwa model yang digunakan baik. Dikarenakan output data prediksi mendekati output data aktual.

V. KESIMPULAN

Berdasarkan hasil dari penelitian yang dilakukan, dapat disimpulkan bahwa *neural network* dapat digunakan untuk membangun alternatif model prediksi jumlah kasus COVID-19 dengan menghasilkan nilai error *R-squared* sebesar 0.9793. Model prediksi ini diharapkan dapat di buat *realtime* untuk dapat digunakan

sebagai alternatif model untuk melakukan prediksi jumlah kumulatif kasus positif COVID-19 di Indonesia, dan digunakan sebagai proyeksi dan optimasi bagi pemerintah dalam mengambil dan mengevaluasi kebijakan.

REFERENCES

- [1] B. Etikasari, T. D. Puspitasari, A. A. Kurniasari, and L. Perdanasari, "Sistem informasi deteksi dini Covid-19," *J. Tek. Elektro dan Komput.*, vol. 9, no. 2, pp. 101–108, 2020.
- [2] H. Wiguna, Y. Nugraha, F. R. R. A. Andika, and J. I. Kangrawan, "Kebijakan Berbasis Data : Analisis dan Prediksi Penyebaran COVID-19 di Jakarta dengan Metode Autoregressive Integrated Moving Average (ARIMA)," vol. 03, no. 02, pp. 74–83, 2020.
- [3] "Beranda | Satgas Penanganan COVID-19." <https://covid19.go.id/> (accessed Oct. 30, 2020).
- [4] F. S. D. Arianto and N. P., "Prediksi Kasus COVID-19 di Indonesia Menggunakan Metode Backpropagation dan Fuzzy Tsukamoto," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 120–127, 2020.
- [5] "Peta Sebaran | Satgas Penanganan COVID-19." <https://covid19.go.id/peta-sebaran> (accessed Oct. 30, 2020).
- [6] M. N. Kholis, Fratemesi, and L. O. Wahidin, "Prediksi Dampak Covid-19 Terhadap Pendapatan Nelayan Jaring Insang Di Kota Bengkulu," *ALBACORE J. Penelit. Perikan. Laut*, vol. 4, no. 1, pp. 001–011, 2020, doi: 10.29244/core.4.1.001-011.
- [7] F. R. Pratiko, "Prediksi Akhir Pandemi COVID-19 di Indonesia dengan Simulasi Berbasis Model Pertumbuhan Parametrik," *J. Rekayasa Sist. Ind.*, vol. 9, no. 2, pp. 63–68, 2020, doi: 10.26593/jrsi.v9i2.4018.63-68.
- [8] F. Ilmu and K. Universitas, "PERBANDINGAN RADIAL BASIS FUNCTION DAN RECURRENT NEURAL NETWORK PADA PREDIKSI CURAH Scanned by TapScanner," 2019.
- [9] A. T. B. Parlina, "InfoTekJar : Jurnal Nasional Informatika dan Artificial Neural Network Pada Industri Non Migas Sebagai Langkah Menuju," *Artif. Neural Netw. Pada Ind. Non Migas Sebagai Langkah Menuju Revolusi Ind. 4.0 Iin*, vol. 1, pp. 6–11, 2019, [Online]. Available: <https://core.ac.uk/download/pdf/276551346.pdf>.
- [10] A. Ramadhan, S. Trinitasjati, C. Limoa, and M. K. Mubarrok, "ANALISIS KUALITAS WEBSITE TANGGAP COVID-19 JAWA TIMUR MENGGUNAKAN END USER COMPUTING SATISFACTION DENGAN NEURAL NETWORK QUALITY ANALYSIS OF EAST JAVA TANGGAP COVID-19 WEBSITE USING END-USER COMPUTING SATISFACTION WITH NEURAL NETWORKS," pp. 571–579, 2020.
- [11] A. S. B. Karno, W. Hastomo, and I. S. K. Wardhana, "Prediksi Jangka Panjang COVID-19 Indonesia Menggunakan Deep," pp. 483–490, 2020.
- [12] P. D. Pakan, "Peramalan Kasus Positif di Indonesia Menggunakan LSTM 13," vol. 6, no. 1, pp. 12–15, 2020.
- [13] W. Setialaksana *et al.*, "Model Jaringan Syaraf Tiruan dalam Peramalan Kasus Positif Covid-19 di Indonesia," *J. Media Pendidik. Tek. Inform. dan Komput.*, vol. 2, no. 2, pp. 53–56, 2020.
- [14] J. Speed and S. P. Engineering, "speed.web.id," vol. 12, no. 4, pp. 1–7, 2020.
- [15] D. Sutarya, "Tungku Sinter Degussa," pp. 1–12, 1979.
- [16] Z. Indra and L. Trisnawati, "Zul Indra, 2) Liza Trisnawati," vol. 3, no. 1, pp. 47–57, 2018.
- [17] I. Suryani and R. S. Wahono, "Penerapan Exponential Smoothing untuk Transformasi Data dalam Meningkatkan Akurasi Neural Network pada Prediksi Harga Emas," *J. Intell. Syst.*, vol. 1, no. 2, pp. 67–75, 2015.
- [18] L. LIDYAWATI, P. RAHMIATI, and Y. SUNARTI, "Implementasi Filter Finite Impulse Response (FIR) Window Hamming dan Blackman menggunakan DSK TMS320C6713," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 4, no. 1, p. 16, 2018, doi: 10.26760/elkomika.v4i1.16.
- [19] N. F. D, R. G. H, S. K. S, and T. Salsabila, "Perbandingan metode double exponential smoothing dan artificial neural network untuk meramalkan perkembangan covid-19 di Indonesia," pp. 312–318, 2020.
- [20] U. Azmi, Z. N. Hadi, and S. Soraya, "ARDL METHOD: Forecasting Data Curah Hujan Harian NTB," *J. Varian*, vol. 3, no. 2, pp. 73–82, 2020, doi: 10.30812/varian.v3i2.627.
- [21] D. Pratidana, "Hak cipta dan penggunaan kembali : Lisensi ini mengizinkan setiap orang untuk menggubah , memperbaiki , dan membuat ciptaan turunan bukan untuk kepentingan komersial , selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat ya," *J. Exp. Psychol. Gen.*, vol. 136, no. 1, pp. 23–42, 2017, [Online]. Available: [http://kc.umn.ac.id/5548/1/BAB II.pdf](http://kc.umn.ac.id/5548/1/BAB%20II.pdf).