

Model Jaringan Syaraf Tiruan dalam Peramalan Kasus Positif Covid-19 di Indonesia

Wirawan Setialaksana¹, Dwi Rezky Anandari Sulaiman², Shabrina Syntha Dewi³, Chairunnisa Ar Lamasitudju⁴, Nini Rahayu Ashadi⁵, Muhammad Asriadi⁶

Universitas Negeri Makassar

¹wirawans@unm.ac.id, ²dwirezky@unm.ac.id, ³shabrinasd@unm.ac.id, ⁴nisalamasitudju@unm.ac.id, ⁵nini.rahayu.ashadi@unm.ac.id, ⁶muhammadasriadi@unm.ac.id

Abstract — Mitigation steps to control Covid-19 outbreak in Indonesia need to take. One of those step is forecasting the spread of the disease. This study compare two artificial neural network models in catching the pattern of Covid-19 positive total cases in Indonesia. Data Training used in this study is Indonesian total positive cases of Covid-19 from March 2 until May 26. The next 10 days data become data testing to show the model accuracy in predicting Covid-19 total cases. MLP shows a better prediction comparing to ELM. Three different prediction accuracy measurement is used – MAE, MAPE, and RMSE. All of them shows less value in MLP than in ELM.

Keywords: Covid-19, Forecasting, MLP, ELM

I. PENDAHULUAN

Sars-Cov-2 merupakan salah satu jenis virus corona yang menyebabkan penyakit pada saluran nafas yang disebut Covid-19[1]. World Health Organization (WHO) mengkonfirmasi bahwa penyakit ini berasal dari China dan sangat infeksius[2]. Hal ini menyebabkan Sars-Cov-2 mudah untuk berpindah dan menyebar.

Penyebaran Covid-19 telah terjadi di seluruh dunia. Covid-19 telah masuk ke Indonesia sejak 2 Maret 2019. Hingga 5 Juni 2020 telah menjangkiti total 29.521 warga Indonesia[3], [4]. Untuk mengendalikan penyakit ini, langkah-langkah mitigasi harus dilakukan.

Langkah mitigasi yang dapat dilakukan untuk mengontrol penyebaran dan penularan covid-19 adalah melakukan peramalan kasus covid-19, baik kasus positif, sembuh, maupun meninggal. Untuk melakukan peramalan, model statistik dan model matematika dapat menjadi dasar peramalan. Salah satu metode statistik yang digunakan dalam melakukan peramalan dampak Covid-19 berupa model runtun waktu[5]–[7]. Metode peramalan dengan model runtun waktu dapat digunakan untuk meramalkan data jangka pendek[8]. Salah satu model klasik runtun waktu yang digunakan dalam meramalkan penyebaran penyakit adalah ARIMA[9]–[11]. Model ini juga digunakan dalam peramalan penyebaran Covid-19[12].

Dalam perkembangan model-model peramalan statistik, model jaringan syaraf tiruan merupakan model yang dianggap baik dalam melakukan peramalan. JST merupakan model yang dapat menjelaskan masalah-masalah peramalan yang kompleks dengan hubungan non linear[13]. Model JST menunjukkan performa yang lebih baik dibandingkan model runtun waktu klasik

seperti ARIMA[14].

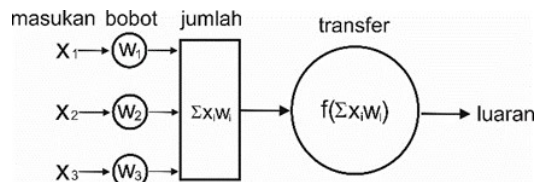
Salah satu model JST yang digunakan dalam peramalan adalah model Multi Layer Perceptron (MLP). Multi Layer Perceptron (dikenal juga sebagai multilayer feed-forward networks) merupakan model JST yang paling populer dan paling banyak digunakan[13]. MLP menunjukkan performa yang baik dibandingkan beberapa model lainnya baik model regresi linear ganda maupun model JST lainnya[15]. Dalam peramalan, MLP juga mampu menunjukkan performa yang lebih baik dibanding model klasik ARIMA[16].

Dalam studi ini, model MLP akan digunakan untuk melakukan peramalan kasus positif Covid-19 di Indonesia. Sebagai pembandingan, model ELM yang juga merupakan model JST akan digunakan.

II. LANDASAN TEORI

A. Multi Layer Perceptron

JST merupakan model statistik yang struktur dan fungsinya mirip dengan otak manusia. Neuron tiruan pada JST memiliki fungsi yang mirip dengan neuron pada sistem syaraf manusia. Gambar 1 menunjukkan model neuron tiruan. JST terdiri atas neuron tiruan yang terdiri atas lapisan-lapisan (layers) yang digabungkan dalam satu model[13].

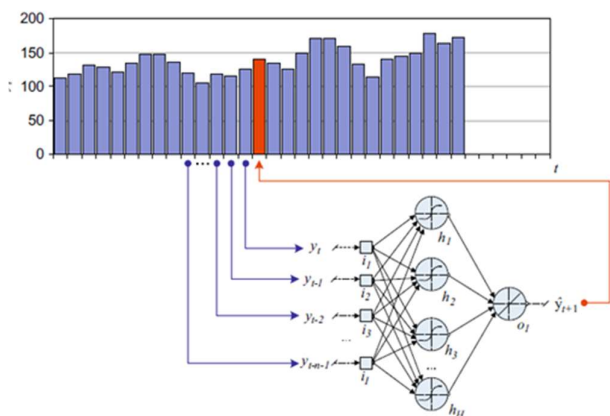


Gambar 1 Model Neuron Tiruan

Multi Layer Perceptron terdiri atas lapisan masukan dan lapisan luaran serta lapisan terselubung (hidden layer). Namun jumlah lapisan terselubung pada MPL bisa bermacam-macam. Jumlah lapisan masukan sama dengan jumlah variabel bebas dalam model. Sedangkan jumlah neuron pada variabel luarannya sama dengan variabel terikat pada model[13]. Gambar 2 menunjukkan model Multi Layer Perceptron (MLP)[17].

Pada model runtun waktu menggunakan MLP, sebuah titik di waktu t dengan peramalan h -langkah ke depan dihitung dengan menggunakan $n+1$ lag dari observasi yaitu yang merupakan n titik ke belakang dari waktu dengan menyatakan jumlah masukan pada neuron MLP. Bentuk fungsional dari MLP dengan lapisan tunggal dengan node tunggal adalah sebagai berikut:

Dengan merupakan vektor yang berisi observasi berbentuk data runtun waktu dengan diketahui jumlah masukan pada MLP. Parameter jaringan dituliskan sebagai bobot dan dari luaran dan lapisan tersembunyi dari MLP. dan merupakan bias dari setiap neuron. I dan H menyatakan banyaknya masukan dan neuron pada lapisan tersembunyi pada jaringan MLP. Fungsi g merupakan fungsi transfer non linear, yang biasanya menggunakan logistik sigmoid atau fungsi tangen hiperbolik.[17].



B. Extreme Learning Machine

Extreme Learning Machine adalah single hidden layer feed forward neural networks dengan bobot-bobot masukan dan biasnya dibangkitkan secara acak sedangkan luarannya dihitung secara analitik. Mekanisme dari ELM akan dijelaskan dalam langkah-langkah berikut.

Misal ELM dengan lapisan tersembunyi sebanyak k anda fungsi aktivasi untuk memodelkan sampel data \mathbf{X} , secara matematis model ELM dapat dituliskan sebagai Parameter merupakan bobot dari hubungan antara masukan dan neuron pada

lapisan tersembunyi. Sedangkan parameter merupakan bobot penghubung lapisan tersembunyi dan luaran. Parameter merupakan bias dari node ke- j di lapisan tersembunyi. Persamaan (1) di atas dapat dituliskan dalam bentuk sebagai berikut:

Matriks H merupakan matriks luaran dari lapisan tersembunyi. Himpunan parameter diperoleh dengan menyelesaikan persamaan berikut:

Proses estimasi di atas ekuivalen dengan meminimalkan fungsi berikut:

Dengan diketahui nilai dari \mathbf{H} dan \mathbf{T} , maka persamaan (2) menjadi sistem linear, dengan bobot luarannya dapat diestimasi dengan persamaan:

Matriks H merupakan matriks invers tergeneralisasi Moore-Penrose dari matriks H [18]. Salah satu cara yang dapat digunakan untuk menentukan nilai dari matriks H adalah dengan menggunakan singular value decomposition (SVD) karena keumuman dari metode ini[19].

C. Data

Data yang digunakan di penelitian ini adalah data total positif Covid-19 Indonesia yang diperoleh di <https://www.worldometers.info/coronavirus/country/indonesia/>. Data total positif yang digunakan adalah data dari 2 Maret 2020 hingga 26 Mei 2020. Data ini digunakan sebagai data training. Sedangkan untuk menguji akurasi prediksi data, data mulai dari 27 Mei 2020 hingga 5 Juni 2020 digunakan sebagai data testing.

D. Ukuran Keakuratan Prediksi

Ukuran yang digunakan untuk mengevaluasi hasil prediksi dari kedua model adalah ukuran Mean Absolute Percent Error (MAPE), Root Mean Square Error (RMSE), dan Mean Absolute Error (MAE). Semakin kecil ukuran akurasi ini, maka semakin baik hasil prediksi dari model yang digunakan.

1. MAPE

MAPE merupakan ukuran akurasi hasil prediksi yang mengukur jarak antara data real dan data hasil prediksi yang kemudian dibuat dalam bentuk proporsi terhadap data real. Formula dari MAPE adalah:

2. RMSE

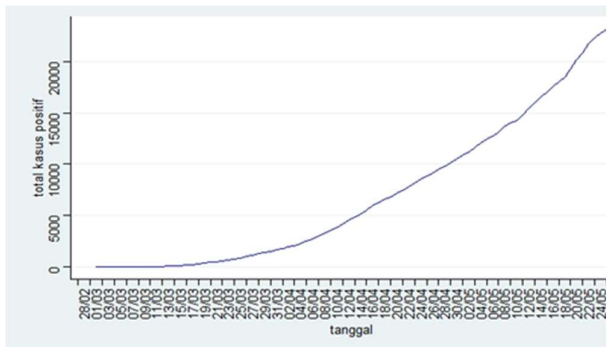
Mean Square Error (MSE) diperoleh dari memangkatkan hasil pengurangan data real dan data hasil prediksi kemudian dibagi dengan banyaknya data. RMSE diperoleh dari mengakarkan MSE.

3. MAE

MAE merupakan ukuran yang mirip dengan MAPE. Namun MAE tidak menggambarkan proporsi error terhadap data real.

III. HASIL DAN PEMBAHASAN

Gambar 4 menunjukkan pergerakan data total positif covid-19 di Indonesia. Total positif covid-19 di Indonesia masih menunjukkan tren peningkatan. Covid-19 diumumkan secara resmi oleh pemerintah Indonesia pada 2 Maret 2020. Hingga 26 Mei 2020, total kasus positif di Indonesia mencapai 23165 kasus positif.

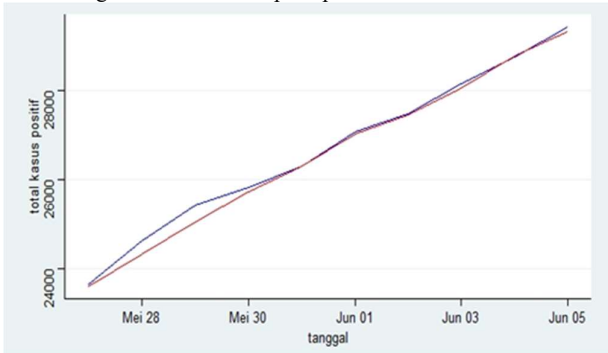


Gambar 3. Total Kasus Positif Covid-19 Indonesia hingga 26 Mei 2020.

Data pada gambar 4 dijadikan data training untuk membentuk model JST.

A. Hasil Prediksi

Gambar 5 menunjukkan hasil prediksi dari model MLP dan model ELM. Hasil prediksi dari kedua model JST yang digunakan tidak berbeda signifikan. Terutama pada prediksi setelah 30 Mei 2020.



Gambar 4. Prediksi Kasus Positif Covid-19 Indonesia hingga 26 Mei 2020. Garis merah merupakan hasil prediksi ELM, sedangkan garis biru menunjukkan prediksi dari model MLP

B. Akurasi Hasil Prediksi

Evaluasi terhadap hasil prediksi dapat dilihat di Tabel 1. Nilai mutlak residu diperoleh dengan mengurangkan hasil prediksi dengan data yang sesungguhnya di tanggal yang sama. Hasil pengurangan kemudian dimutlakan.

| Hari ke- | Prediksi | | Nilai Mutlak Residu | |
|----------|----------|-------|---------------------|-----|
| | ELM | MLP | ELM | MLP |
| 87 | 23607 | 23641 | 244 | 210 |
| 88 | 24326 | 24624 | 212 | 86 |
| 89 | 25036 | 25415 | 180 | 199 |
| 90 | 25713 | 25821 | 60 | 48 |

| | | | | |
|-------|-------|-------|------|------|
| 91 | 26283 | 26284 | 190 | 189 |
| 92 | 27015 | 27076 | 75 | 136 |
| 93 | 27440 | 27460 | 109 | 89 |
| 94 | 28040 | 28149 | 193 | 84 |
| 95 | 28780 | 28742 | 38 | 76 |
| 96 | 29322 | 29411 | 199 | 110 |
| Total | | | 1500 | 1277 |

Kedua model JST mampu memprediksi total kasus positif covid-19 Indonesia dengan baik. Semua hasil prediksi harian menunjukkan residu yang kurang dari 1% kecuali prediksi hari pertama dari ELM yang sedikit berada di atas 1%. Persentase mutlak dari residu ELM merentang dari 0,132% hingga 1,023%. Sedangkan untuk MLP, merentang dari 0,186% hingga 0,880%. Perbandingan rentang persentase mutlak menunjukkan bahwa model MLP lebih presisi dibanding model ELM. Hal ini disebabkan rentang persentase mutlak residu dari MLP lebih kecil dibanding ELM.

Nilai MAPE dari masing-masing model adalah 0,571% untuk ELM dan 0,468% untuk MLP. Dengan nilai , dapat disimpulkan bahwa MLP lebih akurat dibandingkan ELM.

Ukuran akurasi terakhir untuk mengevaluasi 2 model JST yang digunakan adalah RMSE. Tabel 3 menunjukkan data yang akan digunakan untuk menghitung RMSE dari MLP dan ELM.

Tabel 3. Hasil Prediksi dengan Residu Kuadrat

| Hari ke- | Nilai Residu | | Residu Kuadrat | |
|----------|--------------|------|----------------|--------|
| | ELM | MLP | ELM | MLP |
| 87 | -244 | -210 | 59536 | 44100 |
| 88 | -212 | 86 | 44944 | 7396 |
| 89 | -180 | 199 | 32400 | 39601 |
| 90 | -60 | 48 | 3600 | 2304 |
| 91 | -190 | -189 | 36100 | 35721 |
| 92 | 75 | 136 | 5625 | 18496 |
| 93 | -109 | -89 | 11881 | 7921 |
| 94 | -193 | -84 | 37249 | 7056 |
| 95 | -38 | -76 | 1444 | 5776 |
| 96 | -199 | -110 | 39601 | 12100 |
| Total | | | 272380 | 180471 |

Untuk menghitung RMSE, terlebih dahulu dihitung nilai dari MSE.

Selanjutnya, nilai MSE yang diperoleh merupakan hasil kuadrat dari residu. Sehingga untuk memperoleh ukuran yang tepat dalam menggambarkan akurasi prediksi, nilai MSE diakar untuk memperoleh RMSE.

Nilai lebih besar dari . Sehingga, model MLP menunjukkan performa yang lebih baik dalam melakukan prediksi dibandingkan model ELM.

Tabel 4. Ringkasan Ukuran Akurasi

| Model | Ukuran Akurasi | | |
|----------------------|----------------|--------|--------|
| | MAE | MAPE | RMSE |
| MLP | 150 | 0,468% | 134,34 |
| ELM | 127,7 | 0,571% | 165,04 |
| Model Terbaik | MLP | MLP | MLP |

Tabel 4 menunjukkan ringkasan dari 3 ukuran akurasi prediksi yang digunakan. MLP mengungguli ELM dalam prediksi total pasien positif Covid-19 di Indonesia berdasarkan 3 ukuran tersebut.

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

Kedua model JST yang digunakan menunjukkan hasil yang baik. Dengan MAPE yang berada di bawah 1% kecuali pada ELM. Kedua model dapat digunakan untuk memprediksi kasus positif covid-19 di Indonesia. Namun model MLP menunjukkan hasil yang lebih baik dengan MAPE, RMSE, dan MAE yang lebih kecil dibanding ELM.

MAPE pada MLP juga menunjukkan bahwa MLP memiliki akurasi dalam prediksi jangka pendek/harian. MAPE pada MLP selalu berada dibawah 1% dari 10 titik waktu yang digunakan untuk menguji model.

B. Saran

Penelitian ini masih memiliki banyak sekali batasan. Model-model JST lainnya dapat menjadi pembanding bagi MLP. Pada penelitian selanjutnya, akurasi MLP dalam prediksi jangka panjang juga perlu diukur.

DAFTAR PUSTAKA

[1] A. E. Gorbalenya et al., "The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2," *Nat. Microbiol.*, vol. 5, no. 4, pp. 536–544, 2020.

[2] WHO, "Novel Coronavirus (2019-nCoV): Situation Report-3," 2020.

[3] E. Pranita, "Diumumkan Awal Maret, Ahli: Virus Corona Masuk Indonesia dari Januari." [Online]. Available: <https://www.kompas.com/sains/read/2020/05/11/130600623/diumumkan-awal-maret-ahli--virus-corona-masuk-indonesia-dari-januari>. [Accessed: 06-Jun-2020].

[4] T. Maharani, "UPDATE 5 Juni: Bertambah 703, Kasus Covid-19 di Indonesia Capai 29.521." [Online]. Available: <https://nasional.kompas.com/read/2020/06/05/16061321/update-5-juni-bertambah-703-kasus-covid-19-di-indonesia-capai-29521>. [Accessed: 06-Jun-2020].

[5] 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.

[6] S. Singh, K. Singh, J. Kumar, S. Jitendra, and S. Makkhan, "Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19," *Chaos, Solitons Fractals*, vol. 135, pp. 1–8, 2020.

[7] X. Duan and X. Zhang, "ARIMA modelling and forecasting of irregularly patterned COVID-19 outbreaks using Japanese and South Korean data," *Data Br.*, p. 105779, 2020.

[8] M. Henrique, D. Molin, and R. Gomes, "Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil," *Chaos, Solitons and Fractals*, p. 109853, 2020.

[9] Z. He and H. Tao, "Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study," *Int. J. Infect. Dis.*, vol. 74, pp. 61–70, 2018.

[10] S. Paul, O. Mgbere, R. Arafat, B. Yang, and E. Santos, "Modeling and Forecasting Influenza-like Illness (ILI) in Houston, Texas Using Three Surveillance Data Capture Mechanisms," *Online J. Public Health Inform.*, vol. 9, no. 2, 2017.

[11] M. J. Kane, N. Price, M. Scotch, and P. Rabinowitz, "Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks," *BMC Bioinformatics*, vol. 15, no. 1, 2014.

[12] D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti, and M. Ciccozzi, "Data in brief Application of the ARIMA model on the COVID- 2019 epidemic dataset," *Data Br.*, vol. 29, p. 105340, 2020.

[13] J. Szoplik, "Forecasting of natural gas consumption with artificial neural networks," *Energy*, vol. 85, pp. 208–220, 2015.

[14] P. J. García Nieto, F. Sánchez Lasheras, E. García-Gonzalo, and F. J. de Cos Juez, "PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study," *Sci. Total Environ.*, vol. 621, pp. 753–761, 2018.

[15] B. Choubin, S. Khalighi-Sigaroodi, A. Malekian, and Ö. Kişi, "Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals," *Hydrol. Sci. J.*, vol. 61, no. 6, pp. 1001–1009, 2016.

[16] A. Pal, J. P. Singh, and P. Dutta, "Path length prediction in MANET under AODV routing: Comparative analysis of ARIMA and MLP model," *Egypt. Informatics J.*, vol. 16, no. 1, pp. 103–111, 2015.

[17] S. F. Crone and N. K. Å, "Neurocomputing Feature selection for time series prediction – A combined filter and wrapper approach for neural networks," *Neurocomputing*, vol. 73, no. 10–12, pp. 1923–1936, 2010.

[18] X. Chen et al., "Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping," pp. 1–8, 2012.

[19] G. Huang, L. Chen, and C. Siew, "Universal Approximation Using Incremental Constructive Feedforward Networks With Random Hidden Nodes," vol. 17, no. 4, pp. 879–892, 2006.