

Algoritma *Artificial Neural Network* dalam Klasifikasi *Chest X-Rays* Pasien COVID-19

Muhamad Nurhikmat Zain, Dwi Agustin Nuraini Sirodj*

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

ARTICLE INFO

Article history :

Received : 19/8/2022

Revised : 24/11/2022

Published : 21/12/2022



Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Volume : 2

No. : 2

Halaman : 137-144

Terbitan : **Desember 2022**

ABSTRAK

Artificial neural network adalah jaringan yang dibentuk berdasarkan jaringan saraf manusia yang terdiri dari fungsi penjumlahan, fungsi aktivasi dan fungsi tujuan. Artificial neural network dapat digunakan untuk mengenali objek dari suatu citra. Output dari artificial neural network pada penelitian ini merupakan hasil klasifikasi. Penelitian ini akan melakukan klasifikasi chest x-rays pasien COVID-19. Proses klasifikasi dilakukan dengan menggunakan bantuan fungsi aktivasi sigmoid untuk kasus klasifikasi binary. Selama ini untuk mendiagnosa COVID-19 dilakukan dengan tes Reverse Transcription Polymerase Chain Reaction (RT-PCR). Namun ada beberapa masalah pada mekanisme pengujian tes (RT-PCR) diantaranya perlunya alat dan bahan khusus dan memakan waktu yang cukup lama. Solusi yang lebih cepat berbasis data sangat diperlukan dibandingkan melakukan tes (RT-PCR). Salah satu cara yang dapat membantu adalah dengan memanfaatkan citra digital berupa chest x-rays untuk mengambil informasi dan mengenali objek secara otomatis dengan menggunakan metode artificial neural network. Hasil klasifikasi chest x-rays dengan arsitektur artificial neural network yang telah dibangun mendapatkan nilai akurasi sebesar 88%, presisi sebesar 76% dan recall sebesar 76%.

Kata Kunci : Artificial Neural Network; Klasifikasi; COVID-19

ABSTRACT

Artificial neural network is a network formed based on a human neural network consisting of a summation function, activation function and goal function. Artificial neural networks can be used to recognize objects from an image. The output of the artificial neural network in this study is the result of classification. This study will classify chest x-rays of COVID-19 patients. The process of classification using sigmoid activation function for the case of binary classification. Currently, to diagnose COVID-19 using Reverse Transcription Polymerase Chain Reaction (RT-PCR). However, there are several problems with the test mechanism (RT-PCR), has to do it special tools, materials, and it takes quite a long time. Another solution that can help is, use digital images form of chest x-rays to retrieve information and recognize objects automatically using artificial neural network method. The results of chest x-rays classification using a artificial neural network method with sigmoid activation function, get an accuracy of 88%, precision of 76%, and recall of 76%.

Keywords : Artificial Neural Network; Classification; COVID-19

© 2022 Jurnal Riset Statistika Unisba Press. All rights reserved.

A. Pendahuluan

Machine learning adalah pemrograman komputer untuk mengoptimalkan performa menggunakan data sampel atau berdasarkan pengalaman masa lalu dengan menggunakan teori statistika untuk membangun model [1]. Metode *machine learning* yang cukup populer pada saat ini adalah Jaringan Saraf Tiruan atau *Artificial neural network* (ANN). Metode *Artificial neural network* dapat mempelajari hubungan yang tidak diketahui sebelumnya antara data *input* dan *output*. *Artificial neural network* merupakan metode *machine learning* yang memiliki tiga hingga empat *layer* [2]. Tiap *layer* memiliki banyak *neuron*. *Layer-layer* tersebut dinamakan *multilayer perceptron* (MLP) yang berfungsi menghubungkan tiap *neuron* dengan layer akhir berupa layer klasifikasi. *Neuron-neuron* tersebut tersusun dalam suatu lapisan (*layer*) yang terdiri dari *input layer*, *hidden layer* dan *output layer* [3]. Salah satu model NN yang biasa digunakan yaitu Probabilistic Neural Network (PNN). PNN adalah suatu algoritma yang menggunakan fungsi peluang [4]. Untuk kasus klasifikasi *binary* menggunakan fungsi *sigmoid*. Fungsi *sigmoid* akan memberikan nilai output dengan *range* 0-1 [5]. *Artificial neural network* dapat bekerja dengan data tabel, data citra maupun data teks.

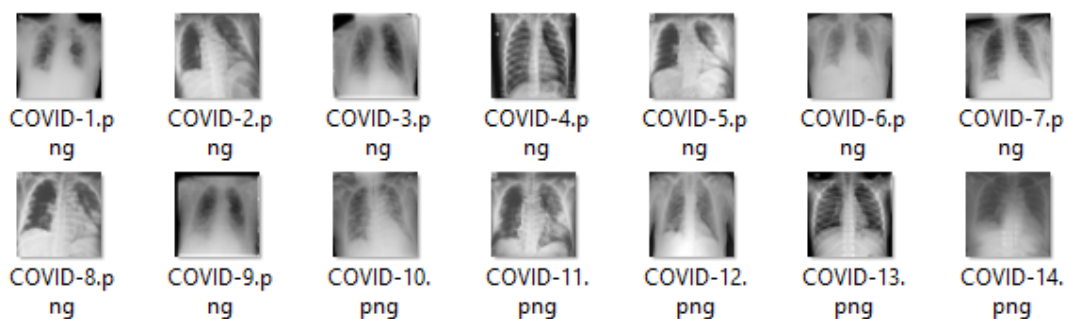
Selama ini untuk mendiagnosa COVID-19 yang dilakukan oleh dokter dan tenaga kesehatan dilakukan dengan tes *Reverse Transcription Polymerase Chain Reaction* (RT-PCR). Namun ada beberapa masalah pada mekanisme pengujian tes (RT-PCR) diantaranya perlunya alat dan bahan khusus, memakan waktu yang cukup lama, sensitivitas yang relatif rendah dan membuat pasien tidak nyaman [6]. Dengan semakin tingginya kasus COVID-19 dan terbatasnya sumber daya manusia, membuat dokter dan tenaga kesehatan kesulitan untuk mendiagnosa pasien COVID-19.

Solusi yang lebih cepat berbasis data sangat diperlukan dibandingkan melakukan tes (RT-PCR). Salah satu metode yang dapat membantu adalah dengan memanfaatkan citra digital *chest x-rays*. Pola radiografi pada *chest x-rays* terbukti menunjukkan sensitivitas dan spesifisitas yang lebih tinggi dibandingkan dengan tes (RT-PCR) [6]. Berdasarkan kelebihan dari metode *artificial neural network* yang dapat digunakan pada data citra sehingga penulis tertarik menggunakan metode *artificial neural network* untuk mengklasifikasikan citra *chest x-rays* pada pasien COVID-19 menggunakan fungsi aktivasi *sigmoid*.

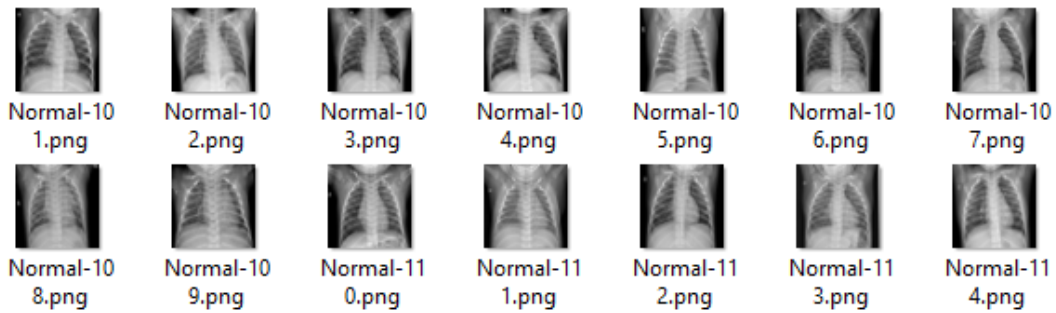
B. Metode Penelitian

Data dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari kaggle www.kaggle.com/ahmadalmahsiri/covid19-radiography-database. Variabel penelitian yang digunakan dalam penelitian ini adalah *chest x-rays* yang telah diberi label yaitu *citra chest x-rays* yang terkena penyakit COVID-19 dan *citra chest x-rays* yang tidak terkena penyakit COVID-19. Citra *chest x-rays* yang terkumpul selama periode Juni 2020 – Desember 2021 ada sebanyak 3616 citra *chest x-rays* yang terkena penyakit COVID-19 dan 10192 citra *chest x-rays* normal. Data citra *chest x-rays* yang diperoleh sebagai berikut:



Gambar 1. Citra *chest x-rays* COVID-19



Gambar 2. Citra chest x-rays normal

Tahapan Analisis

(1) Input data citra chest x-rays, dimisalkan variabel x sebagai citra x-rays dan variabel y sebagai label; (2) Melakukan preprocessing pada citra x-rays; (a) Menyeragamkan pixels pada semua citra chest x-rays yaitu 64x64 pixels; (b) Memberikan label data menjadi kategori, nilai 0 sebagai label COVID-19 dan nilai 1 sebagai label normal; (c) Membagi data citra menjadi data training dan data testing dengan proporsi 80:20; (d) Merubah nilai pada setiap pixels dengan melakukan operasi pembagian 1./225 sehingga mendapatkan nilai diantara 0-1.

(3) Membuat model arsitektur artificial neural network; (a) Merubah layer dua dimensi menjadi vektor satu dimensi yang akan digunakan sebagai input citra; (b) Membuat fully connected layer dengan vektor satu dimensi sebagai input citra yang terhubung ke dua hidden layer dan ke satu neuron pada output layer. Ukuran output pada flatten layer dan banyaknya parameter yang terbentuk pada fully connected layer dapat dihitung menggunakan persamaan: $Output = jumlah\ feature\ map\ sebelumnya \times ukuran\ output\ sebelumnya$
 $Parameter = (jumlah\ input + 1) \times jumlah\ neuron\ pada\ layer$; (d) Melakukan dropout pada fully connected layer untuk mengeluarkan neuron yang tidak akan digunakan pada proses pelatihan [2].

(4) Menetapkan epoch sebanyak 150 dan learning rate 0.1 untuk melakukan proses pengujian model dengan data training (Backpropagation) pada fully connected layer. Algoritma backpropagation terdiri dari dua tahap yaitu feedforward (feed forward) dan backpropagation (backward pass) [2]. Berikut algoritma backpropagation [7]: (a) Setiap unit hidden ($Z_j, j = 1, \dots, p$) menghitung nilai z_j dengan $z_j = v_{0j} + \sum_{i=1}^n x_i v_{ij}$; (b) Setiap unit output ($Y_k = k = 1, \dots, m$) menghitung nilai y_{in_k} dan y_k . $y_{in_k} = w_{0k} + \sum_{j=1}^p z_j w_{jk}$ $y_k = \frac{1}{1 + e^{-y_{in_k}}}$; (c) Setiap unit output ($Y_k = k = 1, \dots, m$) menghitung nilai loss pada output layer. Loss function merupakan suatu metode untuk mengevaluasi seberapa baik algoritma dalam memodelkan suatu data [8]. Loss function dikatakan baik ketika menghasilkan nilai loss yang rendah. Loss function yang digunakan yaitu binary cross entropy. $L = -y \log y_k - (1 - y) \log(1 - y_k)$; (d) Menghitung nilai gradient loss terhadap bobot dan bias pada output layer ke hidden layer. $\frac{\partial L}{\partial w_{jk}} = (y_k - y) z_j$ dan $\frac{\partial L}{\partial w_{0k}} = (y_k - y)$; (e) Update nilai bobot dan bias pada output layer ke hidden layer. $w_{jk}^{(l+1)} = w_{jk}^{(l)} - \alpha \frac{\partial L}{\partial w_{jk}}$ dan $w_{0k}^{(l+1)} = w_{0k}^{(l)} - \alpha \frac{\partial L}{\partial w_{0k}}$; (f) Menghitung nilai gradient loss terhadap bobot dan bias pada hidden layer ke input layer. $\frac{\partial L}{\partial v_{ij}} = (y_k - y) w_{jk}^{(l)} x_i$ dan $\frac{\partial L}{\partial v_{0j}} = (y_k - y) w_{jk}^{(l)}$; (g) Update nilai bobot dan bias pada hidden layer ke input layer. $v_{ij}^{(l+1)} = v_{ij}^{(l)} - \alpha \frac{\partial L}{\partial v_{ij}}$ dan $v_{0j}^{(l+1)} = v_{0j}^{(l)} - \alpha \frac{\partial L}{\partial v_{0j}}$

Uji kondisi berhenti. Kondisi berhenti ini terpenuhi jika nilai loss yang dihasilkan lebih kecil dari nilai target loss. Sebagai inisiasi awal ditetapkan target loss 0.16.

Keterangan:

- n = Banyaknya neuron pada input layer
- p = Banyaknya neuron pada hidden layer
- m = Banyaknya neuron pada output layer
- L = Nilai loss pada output layer

- x_i = Unit input ke-i
- z_j = Unit hidden ke-j
- y_k = Unit output ke-k
- α = Learning rate
- v_{0j} = Bias untuk unit hidden ke-j
- v_{ij} = Bobot dari unit input ke-i ke unit hidden ke-j
- w_{0k} = Bias untuk unit output ke-k
- w_{jk} = Bobot dari unit hidden ke-j ke unit output ke-k

(5) Melakukan pengujian model menggunakan data *testing*. Proses pengujian model menggunakan data *testing* sama seperti point nomor (iv) hanya saja pengujiannya hanya melakukan proses *feed forward*; (6) Membuat *confusion matrix* untuk mengevaluasi model dengan melihat tingkat akurasi, *presisi* dan *recall*. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi sebenarnya [9].

Tabel 1. Confusion Matrix

		Nilai Aktual	
		+	-
Nilai Prediksi	+	<i>True positive (TP)</i>	<i>False positive (FP)</i>
	-	<i>False negative (FN)</i>	<i>True negative (TN)</i>

Keterangan:

- TP = Jumlah data positif yang terklasifikasi dengan benar oleh sistem
- TN = Jumlah data negatif yang terklasifikasi dengan benar oleh sistem
- FP = Jumlah data positif yang terklasifikasi salah oleh sistem
- FN = Jumlah data negatif yang terklasifikasi salah oleh sistem

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} * 100\%$$

$$Presisi = \frac{TP}{TP+FP} * 100\%$$

$$Recall = \frac{TP}{TP+FN} * 100\%$$

Parameter hasil akurasi, *presisi*, dan *recall* dapat dikatakan menghasilkan hasil klasifikasi yang baik atau tidak menggunakan pedoman parameter hasil klasifikasi sebagai berikut [10]

Tabel 2. Pedoman parameter hasil klasifikasi

Rentang	Hasil Klasifikasi
90 – 100%	<i>Excellent classification</i>
80-90%	<i>Good classification</i>
70-80%	<i>Fair classification</i>
60-70%	<i>Poor classification</i>
50-60%	<i>Failure</i>

(7) Melakukan interpretasi dari hasil analisis yang telah dilakukan; (8) Menarik kesimpulan dan saran.

C. Hasil dan Pembahasan

Arsitektur Artificial Neural Network

Pada penelitian ini menggunakan input citra berukuran 64x64 pixels dengan dua hidden layer dan satu output layer. Pada output layer akan digunakan fungsi aktivasi sigmoid untuk mengubah nilai yang masuk dari hidden layer menjadi nilai yang berkisar antara 0 sampai 1. Model arsitektur artificial neural network yang terbentuk sebagai berikut:

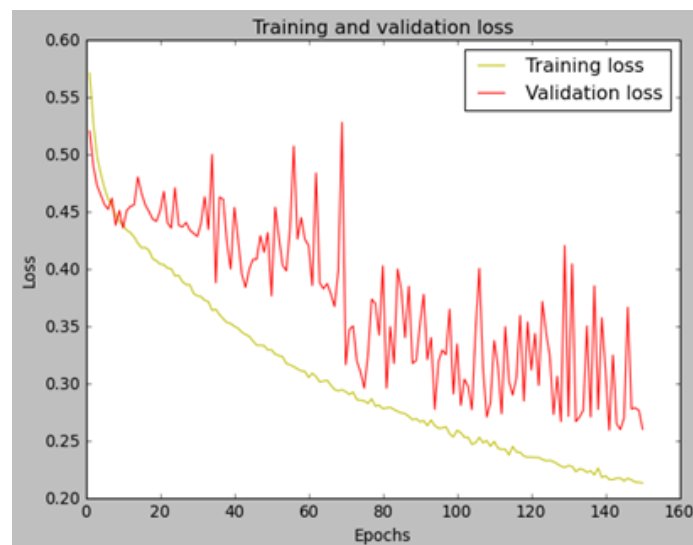
Tabel 3. Model Arsitektur Artificial Neural Network

Layer		Output Shape		Banyaknya Paramater
Klasifikasi	Flatten	Jumlah Input Citra	12288	0
	Dense	Jumlah neuron hidden layer	64	786496
	Dense	Jumlah neuron hidden layer	64	4160
	Dropout	Jumlah neuron hidden layer	64	0
	Dense	Jumlah neuron output layer	1	65
	Activation	Jumlah neuron output layer	1	0

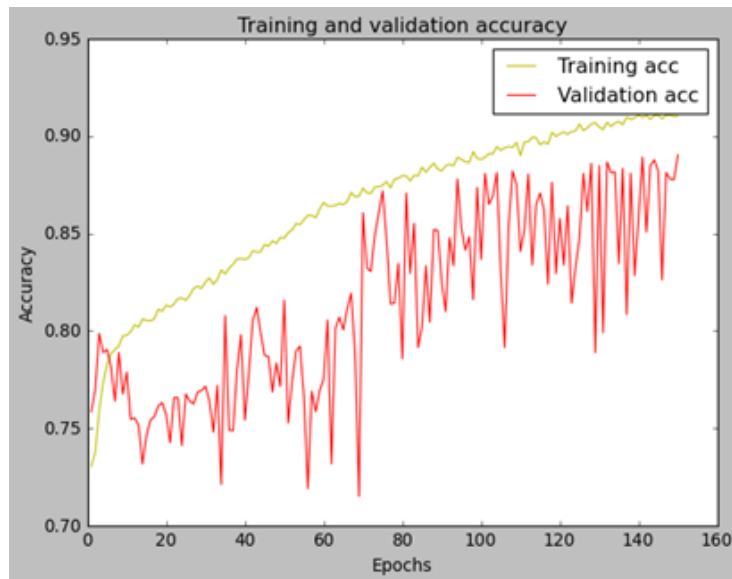
Tabel 3 merupakan model arsitektur artificial neural network yang telah dibangun untuk melakukan proses training. Pada flatten layer ada sebanyak 12288 yang akan akan dijadikan sebagai input citra. Pada hidden layer pertama ada sebanyak 64 input citra dengan 786496 banyaknya parameter yang terbentuk. Hidden layer kedua ada sebanyak 64 input citra dengan 4160 banyaknya parameter yang terbentuk dan satu buah output layer dengan 65 banyaknya parameter. Sehingga jumlah total parameter yang terbentuk sebanyak 790721 parameter.

Proses Training

Proses training terdiri dari tahapan feed forward dan backpropagation untuk mengupdate bobot, bias dan meminimalkan nilai loss. Hasil dari proses training ditunjukkan pada grafik berikut:



Gambar 3. Grafik dari proses training dengan nilai loss dari data training dan testing.

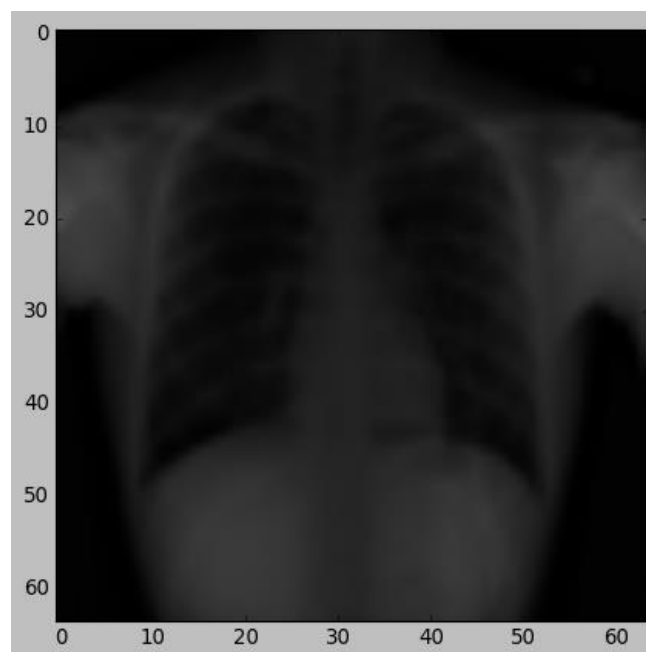


Gambar 4. Grafik dari proses training dengan nilai akurasi dari data training dan testing

Gambar 3 menunjukkan nilai *loss* dari data *training* dan *testing*. Gambar 4 menunjukkan nilai akurasi dari data *training* dan *testing*. Semakin banyak jumlah *epoch* semakin banyak bobot dan bias yang berubah sehingga semakin banyak jumlah waktu yang diperlukan untuk melakukan proses *training*. Semakin rendah nilai *loss* maka akurasi akan semakin meningkat. Berdasarkan Gambar nilai *loss* pada *epoch* ke-100 mencapai 0.16.

Proses Testing

Proses testing dilakukan dengan tahapan *feed forward* dengan menggunakan bobot dan bias dari hasil proses *training*. Output hasil klasifikasi berkisar diantara range 0-1. Ditetapkan *threshold* sebesar 0.5, jika nilai hasil klasifikasi dari fungsi sigmoid < *threshold* termasuk kedalam kategori COVID-19 dan jika nilai hasil klasifikasi dari fungsi sigmoid > *threshold* termasuk kategori normal.



Gambar 5. Data testing 64x64 pixels

Berdasarkan Gambar 4 setelah dilakukan proses *testing* dengan tahapan *feed forward* diperoleh nilai prediksi sebesar 0.99 sedangkan nilai aktual nya adalah 1 atau normal. Dengan nilai hasil klasifikasi > 0.5 maka data citra tersebut termasuk kedalam kategori normal dan termasuk sebagai representasi dari *true negative* (TN). Proses *testing* tersebut dilanjutkan sampai data *testing* terakhir dan disimpan kedalam *confusion matrix*. Berikut hasil *confusion matrix* menggunakan data *testing*:

Tabel 4. Confusion Matrix hasil klasifikasi menggunakan artificial neural network

		Nilai Aktual	
		COVID-19	Normal
Nilai Prediksi	COVID-19	504	163
	Normal	162	1933

Pada tabel 4 prediksi terhadap kategori COVID-19 yang diklasifikasikan dengan benar sebagai COVID-19 ada sebanyak 504. Prediksi terhadap kategori normal yang diklasifikasikan dengan benar sebagai normal ada sebanyak 1933. Prediksi terhadap kategori normal yang diklasifikasikan sebagai COVID-19 ada sebanyak 163. Dan prediksi terhadap kategori COVID-19 yang diklasifikasikan sebagai normal ada sebanyak 162. Dengan arsitektur *artificial neural network* yang telah dibangun, input citra 64x64 *pixels*, jumlah *epoch* sebanyak 150, nilai *loss* 0.16 dan nilai *learning rate* 0.1 diperoleh nilai akurasi dari data *testing* sebesar 88%, presisi 76% dan *recall* 76%. Nilai akurasi sebesar 88% menandakan model baik dalam melakukan klasifikasi (*good classification*). Nilai presisi sebesar 76% menandakan persentase pasien yang benar menderita COVID-19 dari keseluruhan pasien yang diprediksi menderita COVID-19. Nilai *recall* sebesar 76% menandakan persentase pasien yang diprediksi COVID-19 dibandingkan keseluruhan pasien yang sebenarnya COVID-19.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian yaitu berdasarkan hasil analisis pada arsitektur *artificial neural network* yang telah dibangun, model dapat memprediksi 504 citra *chest x-rays* yang termasuk kategori COVID-19 dari 666 citra *chest x-rays* COVID-19 dan 1933 citra *chest x-rays* yang termasuk kategori normal dari 2096 citra *chest x-rays* normal dengan nilai akurasi yang diperoleh sebesar 88%, presisi sebesar 76% dan *recall* sebesar 76%. Sehingga dapat disimpulkan bahwa dengan menggunakan algoritma *artificial neural network* dalam klasifikasi biner pasien COVID-19 menggunakan fungsi aktivasi *sigmoid* sudah baik dalam melakukan klasifikasi.

Daftar Pustaka

- [1] E. Alpaydm, *Introduction to Machine Learning, 4th ed.* London: MIT Press, 2020.
- [2] W. Setiawan, *Deep Learning Menggunakan Convolutional Neural Network.* Malang: MNC, 2020.
- [3] E. S. Noviando, E. ' . Ervianto, and I. ' . Yasri, "Studi Penerapan ANN (Artificial Neural Network) untuk Menghilangkan Harmonisa pada Gedung Pusat Komputer," *J. Online Mhs. Fak. Tek. Univ. Riau*, vol. 3, no. 2, pp. 1–6, 2016.
- [4] P. D. Inas Azizah, "Penerapan Probabilistic Neural Network pada Klasifikasi Berat Bayi Baru Lahir," *J. Ris. Stat.*, vol. 1, no. 2, pp. 152–159, Feb. 2022, doi: 10.29313/jrs.v1i2.524.
- [5] W. S. Eka Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.
- [6] S. Hassantabar, M. Ahmadi, and A. Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches.," *Chaos. Solitons. Fractals*, vol. 140, p. 110170, Nov. 2020, doi: 10.1016/j.chaos.2020.110170.

- [7] N. Srinivasan, V. Ravichandran, K. L. Chan, J. R. Vidhya, S. Ramakirishnan, and S. M. Krishnan, *Exponentiated backpropagation algorithm for multilayer feedforward neural networks*. 2002. doi: 10.1109/ICONIP.2002.1202187.
- [8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. 2016.
- [9] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009, doi: <https://doi.org/10.1016/j.ipm.2009.03.002>.
- [10] F. Gorunescu, "Data mining: Concepts, models and techniques," vol. 12, 2011, doi: 10.1007/978-3-642-19721-5.