



Bagian 2: Model Arsitektur *Neural Network* dengan Kombinasi *K-Medoids* dan *Backpropagation* pada kasus Pandemi COVID-19 di Indonesia

Agus Perdana Windarto^{1,*}, Jufriadif Na'am², Yuhandri², Anjar Wanto³, Mesran⁴

¹ Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

² Program Studi Teknik Informatika, Universitas Putra Indonesia YPTK Padang, Indonesia

³ Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

⁴ Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email Penulis Korespondensi: ^{1,*}agus.perdana@amiktunasbangsa.ac.id

Abstrak— Tujuan dari penelitian untuk membuat model prediksi pada arsitektur neural network terbaik dengan mengkombinasi metode k-medoids dan backpropagation pada kasus pandemi COVID-19 di Indonesia. Data diperoleh dari Kementerian Kesehatan, dicuplik dan diolah dari covid19.go.id dan bnpb.go.id. Kasus yang diangkat adalah jumlah persebaran pandemi COVID-19 di Indonesia per tanggal 7 Juli 2020 yang terdiri dari 34 record. Variabel yang digunakan pada penelitian adalah jumlah kasus positif (x_1), jumlah kasus sembuh (x_2) dan jumlah kasus meninggal dunia (x_3) berdasarkan provinsi. Proses analisis data menggunakan bantuan software RapidMiner. Solusi yang diberikan adalah dengan mengkombinasi metode k-medoids dan backpropagation. Dimana metode k-medoids melakukan pemetaan berupa kluster yang ditentukan. Label kluster yang digunakan adalah kluster tinggi (C1= zona merah), kluster waspada (C2= zona kuning), kluster rendah (C3= zona hijau). Hasil dari pemetaan kluster dilanjutkan ke metode backpropagation untuk memprediksi hasil akurasi dari kluster yang ada. Dengan menggunakan model arsitektur terbaik 3-2-1 diperoleh nilai akurasi 94,17% dengan $learning_rate = 0.696$. Hasil pemetaan kluster diperoleh 9 provinsi berada di kluster tinggi (C1= zona merah), 3 provinsi berada di kluster waspada (C2= zona kuning) dan 22 provinsi berada di kluster rendah (C3= zona hijau). Diharapkan hasil dari penelitian dapat memberikan informasi kepada pemerintah berupa pemetaan kluster terhadap wilayah di Indonesia.

Kata Kunci: pandemi COVID-19, k-medoids, backpropagation, prediksi, model arsitektur, neural network

Abstract—The aim of the research is to create a prediction model on the best neural network architecture by combining the k-medoids and backpropagation methods in the case of the COVID-19 pandemic in Indonesia. Data obtained from the Ministry of Health is sampled and processed from covid19.go.id and bnpb.go.id. The case raised was the number of the spread of the COVID-19 pandemic in Indonesia as of July 7, 2020, with 34 records. The variables used in this study are the number of positive cases (x_1), the number of cases cured (x_2), and the number of deaths (x_3) by province. The process of data analysis uses the help of RapidMiner software. The solution provided is to combine the k-medoids and backpropagation methods. Where the k-medoids method is mapping the specified cluster. The cluster labels used are high cluster (C1 = red zone), alert cluster (C2 = yellow zone), low cluster (C3 = green zone). The results of cluster mapping are continued to the backpropagation method to predict the accuracy of the existing cluster results. By using the best architectural model 3-2-1, the accuracy value is 94.17% with $learning_rate = 0.696$. Cluster mapping results obtained nine provinces are in the high cluster (C1 = red zone), three provinces are in the alert cluster (C2 = yellow zone), and 22 provinces are in the low cluster (C3 = green zone). It is expected that the results of the research can provide information to the government in the form of cluster mapping of regions in Indonesia.

Keywords: COVID-19 Pandemic, K-Medoids, Backpropagation, Prediction, Architectural Models, Neural Network

1. PENDAHULUAN

Pada awal tahun 2020 ini, dunia dikejutkan dengan wabah Coronavirus 2019 (COVID-19) yang menginfeksi hampir seluruh negara di dunia [1]. WHO semenjak Januari 2020 telah menyatakan dunia masuk ke dalam darurat global terkait virus ini [2]. Penyakit COVID-19 telah menjadi pandemi dan merupakan ancaman besar bagi kesehatan masyarakat di Indonesia [3]. Dengan adanya COVID-19 pemerintah Indonesia telah membuat beberapa kebijakan untuk menghentikan penyebaran wabah ini, seperti melakukan *lockdown* di daerah yang sudah termasuk ke dalam zona merah penyebaran virus, lalu *physical quarantine* untuk menghindari penyebaran virus secara kontak fisik [4].

Banyak penelitian terkait COVID-19 yang dilakukan untuk melakukan pencegahan terhadap virus COVID-19. Seperti yang dilakukan Nuraini et al, 2020 [5] tentang model sederhana untuk memprediksi endemik di Indonesia yang didasarkan pada Kurva Richard yang mewakili persamaan logistik yang dimodifikasi. Hasil penelitian tersebut menunjukkan bahwa endemik akan berakhir pada April 2020 dengan jumlah total kasus lebih dari 8000. Berikutnya penelitian yang dilakukan Swastika, 2020 [6] tentang deteksi awal COVID-19 menggunakan citra CT berbasis *deep learning*. Hasil penelitian menyebutkan bahwa akurasi yang diperoleh sebesar 92,86% untuk mengklasifikasikan infeksi COVID-19 dan normal. Nilai spesifisitas dan sensitivitas sebesar 100% dan 85,71% untuk pelatihan menggunakan optimizer SGD.

Dalam hal ini penelitian yang dilakukan membuat model prediksi pada arsitektur *neural network* dengan mengkombinasi metode *k-medoids* dan *backpropagation* pada kasus pandemi COVID-19 di Indonesia. Ada banyak alat analisis berupa *prediction model* (model prediksi) yang memungkinkan membuat prediksi secara cepat untuk masalah yang bersifat kompleks. Salah satu model tersebut adalah metode *backpropagation* [7] dan



k-medoids [8]. Kedua metode tersebut dikombinasikan untuk memprediksi kasus pandemi COVID-19 di Indonesia. Setiap metode memiliki fungsi dan tugas masing-masing. Metode *k-medoids* melakukan pemetaan berupa kluster yang telah ditentukan.

Terdapat banyak kelebihan yang dimiliki *k-medoids*. Salah satunya metode *k-medoids* muncul sebagai penanggulangan kelemahan metode *k-means* yang sensitif terhadap *outlier* [9]. Hasil dari kluster kemudian diprediksi dengan metode *backpropagation* untuk melihat hasil akurasi dari kluster yang terbentuk. Selain itu metode *backpropagation* sangat baik digunakan pada deret waktu yang bersifat *nonlinear* [10] sehingga metode *backpropagation* sangat baik dalam meramalkan. Berdasarkan hal tersebut diharapkan hasil penelitian dapat memberikan informasi kepada pemerintah berupa pemetaan kluster wilayah pada kasus pandemi COVID-19 di Indonesia.

2. METODOLOGI PENELITIAN

Data yang digunakan adalah data hasil penelitian yang dilakukan Windarto, et al (2020) [11] tentang kombinasi metode klustering dan klasifikasi dengan kasus Pandemi COVID-19 di Indonesia. Data yang digunakan data dari Kementerian Kesehatan dan diolah dari COVID19.go.id dan bnpb.go.id. terhadap jumlah persebaran pandemi COVID-19 di Indonesia pada tanggal 7 juli 2020 yang terdiri dari 34 *record*. Berikut data yang digunakan seperti yang ditunjukkan pada Tabel 1 Berikut:

Tabel 1. Jumlah persebaran COVID-19 di Indonesia (7 juli 2020)

No	Provinsi	Jumlah Kasus Positif	Jumlah Kasus Sembuh	Jumlah Kasus Meninggal
1	Aceh	88	42	3
2	Bali	1.94	1.034	25
3	Banten	1.531	952	80
4	Bangka Belitung	171	136	2
5	Bengkulu	144	96	13
6	DI Yogyakarta	346	282	8
7	DKI Jakarta	12.86	8.277	661
8	Jambi	121	81	1
9	Jawa Barat	3.779	1.763	180
10	Jawa Tengah	4.878	1.617	220
11	Jawa Timur	14.6	5.114	1.079
12	Kalimantan Barat	344	305	4
13	Kalimantan Timur	603	452	11
14	Kalimantan Tengah	1.058	464	62
15	Kalimantan Selatan	3.695	1.064	205
16	Kalimantan Utara	206	176	2
17	Kepulauan Riau	313	263	16
18	Nusa Tenggara Barat	1.392	889	65
19	Sumatera Selatan	2.356	1.21	112
20	Sumatera Barat	780	637	31
21	Sulawesi Utara	1.252	317	92
22	Sumatera Utara	1.821	494	108
23	Sulawesi Tenggara	487	285	8
24	Sulawesi Selatan	6.192	2.242	202
25	Sulawesi Tengah	191	164	6
26	Lampung	201	159	12
27	Riau	236	214	11
28	Maluku Utara	967	124	32
29	Maluku	830	414	17
30	Papua Barat	266	185	4
31	Papua	2.057	963	19
32	Sulawesi Barat	127	89	2
33	Nusa Tenggara Timur	118	54	1
34	Gorontalo	276	227	15

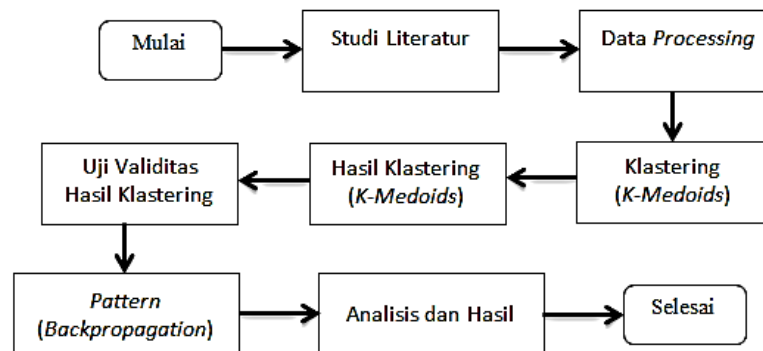
Sumber: data publikasi (Windarto et al., 2020) pada Jurnal Media Informatika Budidarma (MIB)

Volume 4, Nomor 3, Juli 2020, Page 855-862

Pada tahapan ini, penelitian mengkombinasikan kedua metode yakni *k-medoids* dan *backpropagation*. Metode *k-medoids* melakukan *mapping* berupa kluster terhadap wilayah. Hasil dari *mapping* dilanjutkan untuk dilakukan prediksi menggunakan *backpropagation*. Hal ini untuk melihat sejauh mana nilai akurasi yang diperoleh dari hasil *mapping* yang terbentuk. Penelitian ini berbeda dengan penelitian yang dilakukan Windarto, et al (2020) tentang kombinasi klustering dan klasifikasi. Dimana hasil dari pemetaan berupa klustering diteruskan dengan menggunakan metode klasifikasi untuk melihat aturan berupa pohon keputusan. Proses analisis menggunakan



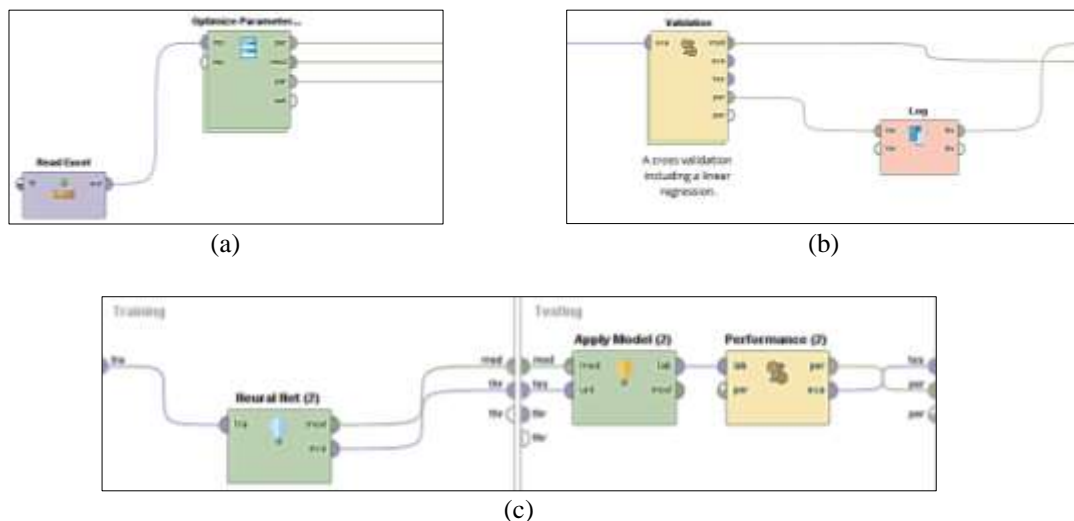
bantuan *software RapidMiner*. Berikut gambar metodologi penelitian seperti yang ditunjukkan pada Gambar 1 berikut:



Gambar 1. Metodologi Penelitian

3. HASIL DAN PEMBAHASAN

Pada tahap hasil dan pembahasan mengarah pada kerangka penelitian dimana proses analisis menggunakan bantuan *software RapiMiner*. Berikut beberapa rancangan yang digunakan dengan memanfaatkan *software RapidMiner* seperti yang ditunjukkan Gambar 2 berikut:



Gambar 2. Rancangan *software RapidMiner*

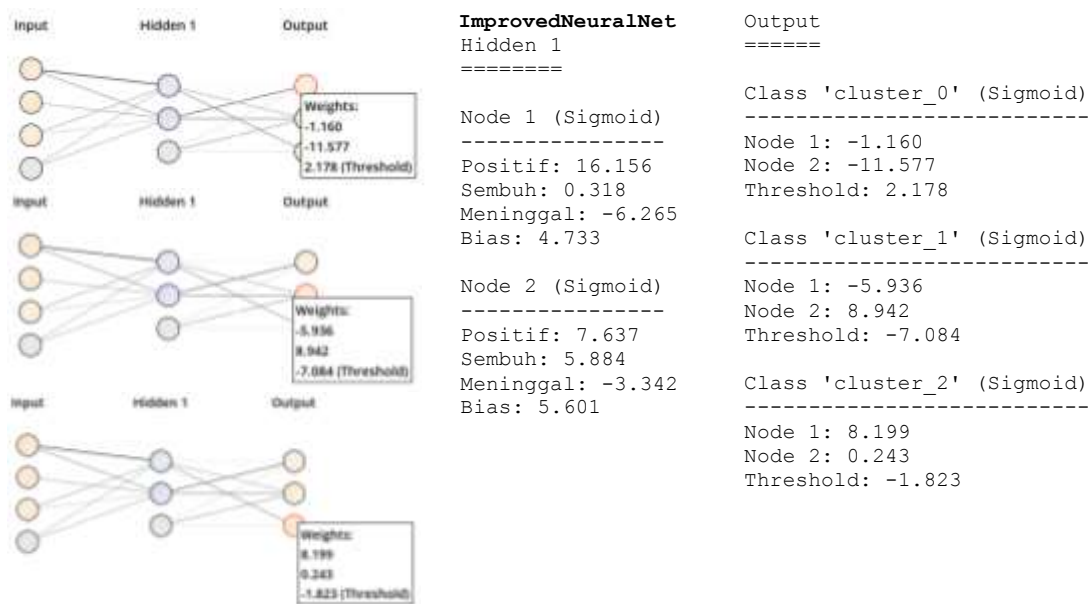
Pada Gambar 2 dijelaskan tahapan proses *input* menggunakan *tool read excel* dengan memanfaatkan data *excel*. Data *excel* yang digunakan merupakan hasil *mapping* terhadap kluster wilayah dengan menggunakan metode *k-medoids*. Dimana hasil kluster yang diperoleh memiliki 3 label yakni kluster tinggi (C1= zona merah), kluster waspada (C2= zona kuning), kluster rendah (C3= zona hijau). Berdasarkan hasil *mapping* berdasarkan wilayah bahwa label kluster tinggi adalah *cluster_0*; kluster waspada adalah *cluster_1*; dan kluster rendah adalah *cluster_2* yang proses penentuan kluster dilihat dari hasil nilai akhir *centroid*. Selanjutnya operator *read excel* dihubungkan ke operator *Optimize Parameters (Grid)* yang berfungsi untuk menemukan nilai optimal dari parameter yang dipilih untuk operator dalam subprocessnya (a). Dalam operator *Optimize Parameters (Grid)* terdapat proses *validation* yang berfungsi melakukan validasi silang untuk memperkirakan kinerja statistik model pembelajaran (b). Dalam proses *validation* terdapat model *neural network (backpropagation)* yang melakukan prediksi terhadap hasil kluster (c). Dimana pada model *neural network (backpropagation)* menggunakan *training* dan *testing* melalui *apply model* untuk melihat hasil akurasi berupa *learning rate*, *momentum*, dan *performance* yang disimpan menggunakan operator *log*. Selain itu operator *log* dapat diplot oleh GUI ketika eksekusi proses selesai melalui Informasi yang disimpan.

Pada proses prediksi menggunakan *backpropagation* model arsitektur yang diajukan adalah 3-2-1; 3-5-1; 3-10-1; 3-5-10-1 dan 3-10-5-1 dengan model arsitektur terbaik 3-2-1. Adapun parameter yang digunakan adalah



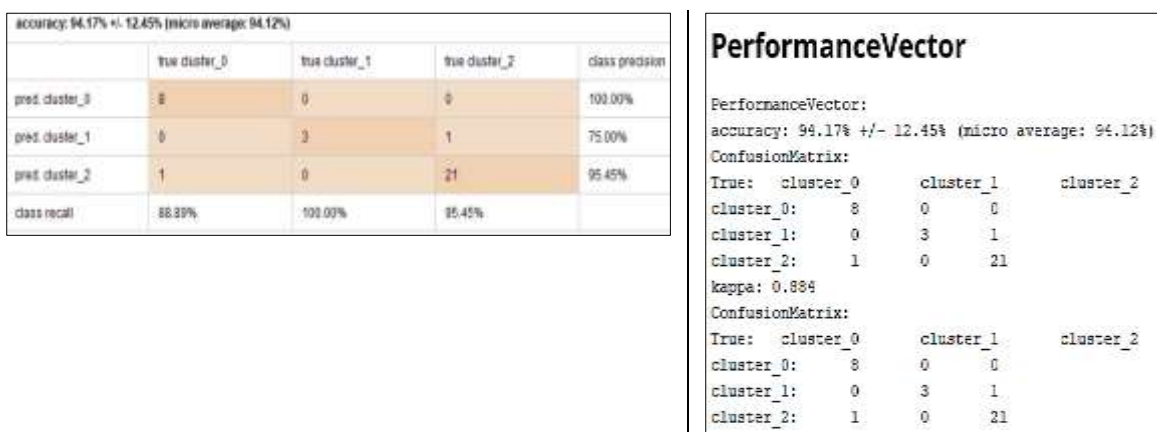
Gambar 3. Parameter *backpropagation*

Pada Gambar 3 dijelaskan bahwa proses training dan testing pada model arsitektur *backpropagation* menggunakan *epoch max*= 200; *learning rate*= 0.01; *momentum*= 0.9. Berikut hasil dari model arsitektur terbaik 3-2-1 seperti yang ditunjukkan pada gambar berikut:



Gambar 4. Hasil perhitungan *backpropagation*

Pada gambar 4 dapat dijelaskan bahwa hasil prediksi menggunakan model arsitektur terbaik 3-2-1 memiliki *output* yang berbeda sesuai dengan hasil klastering yang dilakukan sebelumnya (klaster tinggi (C1= zona merah), klaster waspada (C2= zona kuning), klaster rendah (C3= zona hijau)). Setiap *output* memiliki nilai bobot (*weight*) dan *threshold* yang berbeda berdasarkan setiap klaster. Dimana fungsi aktivasi yang digunakan adalah fungsi *sigmoid*. Berikut hasil lengkap prediksi *backpropagation* yang dapat dilihat nilai akurasi, *learning rate* dan *momentum* pada model arsitektur terbaik 3-2-1 seperti yang ditunjukkan pada gambar berikut:



Gambar 5. Hasil prediksi *backpropagation*



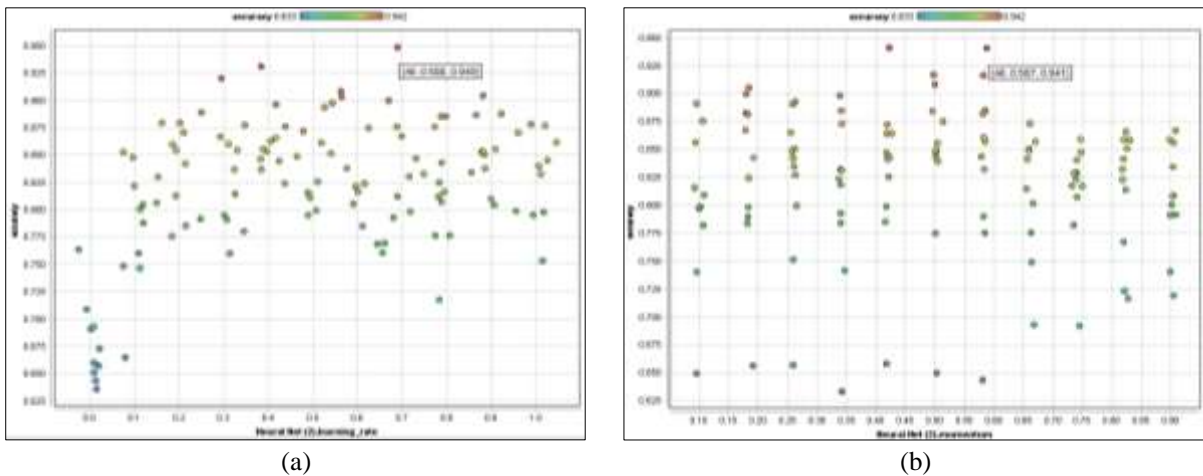
```

ParameterSet
Parameter set:

Performance:
PerformanceVector [
----accuracy: 94.17% +/- 12.45% (micro average: 94.12%)
ConfusionMatrix:
True:   cluster_0 cluster_1 cluster_2
cluster_0:   8         0         0
cluster_1:   0         3         1
cluster_2:   1         0        21
----kappa: 0.884
ConfusionMatrix:
True:   cluster_0 cluster_1 cluster_2
cluster_0:   8         0         0
cluster_1:   0         3         1
cluster_2:   1         0        21
]
Neural Net (2).learning_rate = 0.696
Neural Net (2).momentum     = 0.58
    
```

Gambar 5. Hasil prediksi *backpropagation* (lanjutan)

Hasil yang ditunjukkan pada Gambar 5, bahwa *backpropagation* dapat melakukan prediksi terhadap hasil klastering yang terbentuk dengan menggunakan metode *k-medoids*. Nilai prediksi sebesar 94.17% dengan *Neural Net (2).learning_rate* = 0.696 dan *Neural Net (2).momentum* = 0.58. Pada gambar prediksi yang dilakukan *backpropagation* untuk *cluster_0* adalah 8 (*true*) dan (*false*=1) untuk total klaster adalah 9. Sementara untuk *cluster_1* adalah 3 (*true*) dan (*false*=0) untuk total klaster adalah 3 dan *cluster_2* adalah 21 (*true*) dan (*false*=1) untuk total klaster adalah 22. Berikut adalah grafik akurasi terbaik berdasarkan *learning rate* dan *momentum* pada *backpropagation*:



Gambar 6. Grafik nilai akurasi berdasarkan *learning rate* dan *momentum*

Pada Gambar 6, grafik berupa scatter plot dalam menentukan nilai akurasi terbaik dengan menggunakan nilai maksimal *learning rate* dan *momentum* pada proses *backpropagation*. Pada grafik *scatter learning rate* (a) diperoleh nilai berkisar 0.688 untuk nilai akurasi mencapai 94% sedangkan grafik *scatter momentum* (b) diperoleh nilai berkisar 0.587 untuk nilai akurasi mencapai 94%. Hal ini sesuai dengan hasil prediksi yang terdapat pada penjelasan Gambar 5.

4. KESIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa kombinasi *k-medoids* dan *backpropagation* dapat diterapkan pada kasus persebaran pandemi COVID-19 di Indonesia dengan memanfaatkan *software* bantu RapidMiner. Dengan menggunakan kombinasi tersebut untuk hasil pemetaan wilayah berupa klaster diperoleh 9 provinsi berada di klaster tinggi (C1= zona merah), 3 provinsi berada di klaster waspada (C2= zona kuning) dan 22 provinsi berada di klaster rendah (C3= zona hijau) diperoleh nilai akurasi 94,17% dengan *learning_rate*= 0.696 dan *momentum*= 0.58 pada menggunakan model arsitektur terbaik 3-2-1.

REFERENCES

[1] D. R. Buana, “Analisis Perilaku Masyarakat Indonesia dalam Menghadapi Pandemi Covid-19 dan Kiat Menjaga



- Kesejahteraan Jiwa,” *SALAM; J. Sos. Budaya Syar-i*, vol. 7, no. 3, pp. 217–226, 2020, doi: 10.15408/sjsbs.v7i3.15082.
- [2] M. Pradana, S. Syahputra, A. Wardhana, B. R. Kartawinata, and C. Wijayangka, “The Effects of Incriminating COVID-19 News on the Returning Indonesians’ Anxiety,” *J. Loss Trauma*, vol. 0, no. 0, pp. 1–6, 2020, doi: 10.1080/15325024.2020.1771825.
- [3] I. M. A. Wirawan and P. P. Januraga, “Forecasting COVID-19 Transmission and Healthcare Capacity in Bali, Indonesia,” *J. Prev. Med. Public Health*, vol. 53, no. 3, pp. 158–163, 2020, doi: 10.3961/jpmph.20.152.
- [4] Nurkholis, “Dampak Pandemi Novel-Corona Virus Disiase (Covid-19) Terhadap Psikologi Dan Pendidikan Serta Kebijakan Pemerintah,” *J. PGSD*, vol. 6, no. 1, pp. 39–49, 2020, doi: 10.32534/jps.v6i1.1035.
- [5] N. Nuraini, K. Khairudin, and M. Apri, “Modeling Simulation of COVID-19 in Indonesia based on Early Endemic Data,” *Commun. Biomath. Sci.*, vol. 3, no. 1, pp. 1–8, 2020, doi: 10.5614/cbms.2020.3.1.1.
- [6] W. Swastika, “Studi Awal Deteksi Covid-19 Menggunakan Citra Ct Berbasis Deep,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 629–634, 2020, doi: 10.25126/jtiik.202073399.
- [7] A. Pujianto, K. Kusriani, and A. Sunyoto, “Perancangan Sistem Pendukung Keputusan Untuk Prediksi Penerima Beasiswa Menggunakan Metode Neural Network Backpropagation,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 157, 2018, doi: 10.25126/jtiik.201852631.
- [8] F. Rahman, I. I. Ridho, M. Muflih, S. Pratama, M. R. Raharjo, and A. P. Windarto, “Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination C,” 2020, doi: 10.1088/1757-899X/835/1/012058.
- [9] I. Kamila, U. Khairunnisa, and Mustakim, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, pp. 119–125, 2019.
- [10] H. Havaluddin, Z. Arifin, A. H. Kridalaksana, and D. Cahyadi, “Prediksi Kedatangan Turis Asing ke Indonesia Menggunakan Backpropagation Neural Networks,” *J. Teknol. dan Sist. Komput.*, vol. 4, no. 4, p. 485, 2016, doi: 10.14710/jtsiskom.4.4.2016.485-490.
- [11] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, “Bagian 1: Kombinasi Metode Klastering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia),” *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 855, 2020, doi: 10.30865/mib.v4i3.2312.