

**PREDIKSI COVID-19 DI PROVINSI LAMPUNG
MENGUNAKAN METODE *ARTIFICIAL NEURAL NETWORK*
DENGAN *K-FOLD CROSS VALIDATION***

(Skripsi)

Oleh

**NUR ALIFIAH
1857031002**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

ABSTRACT

PREDICTION OF COVID-19 IN LAMPUNG PROVINCE USING THE ARTIFICIAL NEURAL NETWORK METHOD WITH K-FOLD CROSS VALIDATION

By

Nur Alifiah

This study predicts the number of new cases, recovered cases, and Covid-19 deaths in Lampung Province using the Artificial Neural Network (ANN) method with k-fold cross-validation. The variables used in this study are three variables, namely new cases, recovered cases, and death cases in the period from March 19, 2020, to October 26, 2021. Furthermore, data on new cases recovered cases, and cases of death from October 27 to January 24 2022, will be compared with the forecasting data obtained. This method is tested with k-fold cross-validation, and the best parameters are obtained by hyper-tuning. The optimal parameters of k-fold cross validation ($k = 10$) for the new case are one hidden layer with 30 nodes, 0.2 dropouts, 16 batch sizes, 100 epochs, and a 0.01 learning rate. Furthermore, the cases recovered: 1 hidden layer of 30 nodes, dropout of 0.2, batch size of 32, an epoch of 50, and learning rate of 0.01. While the death cases, namely: 3 layers hidden layer, with the first layer consisting of 30 nodes, the second layer with 20 nodes, and the third layer with 13 nodes, dropout of 0.2, batch size of 32, an epoch of 100, and learning rate of 0.01. The three variables use the ReLU activation function. The prediction results of MAPE values obtained in new cases and recovered cases. Cases of death were 1.78%, 1.92% and 0.95%, so it can be interpreted that these results have excellent predictive abilities with model accuracy values of 98.22% for new cases, 98.08% for recovered, and 99.05% for death cases. Therefore, this method is very well used to predict Covid-19 cases in the future period.

Keywords: COVID-19, Data Mining, Machine Learning, Artificial Neural Network, K-Fold Cross Validation

ABSTRAK

PREDIKSI COVID-19 DI PROVINSI LAMPUNG MENGUNAKAN METODE *ARTIFICIAL NEURAL NETWORK* DENGAN *K-FOLD CROSS VALIDATION*

Oleh

Nur Alifiah

Penelitian ini memprediksi banyaknya penambahan kasus baru, kasus sembuh, dan kasus kematian Covid-19 di Provinsi Lampung menggunakan metode *Artificial Neural Network* (ANN) dengan *k-fold cross validation*. Variabel yang digunakan dalam penelitian ada 3 variabel, yaitu kasus baru, kasus sembuh, dan kasus kematian pada periode tanggal 19 Maret 2020 sampai dengan 26 Oktober 2021. Selanjutnya, data kasus baru, kasus sembuh, dan kasus kematian pada periode tanggal 27 Oktober sampai 24 Januari 2022 akan dijadikan perbandingan dengan data hasil peramalan yang didapatkan. Metode ini dilakukan pengujian dengan *k-fold cross validation* dan parameter terbaik diperoleh dengan *hypertuning*. Parameter optimal yang diperoleh dengan *k-fold cross validation* ($k = 10$) untuk kasus baru, yaitu: *hidden layer* sebanyak 1 *layer* dengan 30 *node*, *dropout* sebesar 0.2, *batch size* sebesar 16, *epoch* sebesar 100, dan *learning rate* sebesar 0.01. Selanjutnya, kasus sembuh, yaitu: *hidden layer* sebanyak 1 *layer* 30 *node*, *dropout* sebesar 0.2, *batch size* sebesar 32, *epoch* sebesar 50, dan *learning rate* sebesar 0.01. Sedangkan kasus kematian, yaitu: *hidden layer* sebanyak 3 *layer*, dengan *layer* pertama terdiri dari 30 *node*, dan *layer* kedua dengan 20 *node*, dan *layer* ketiga dengan 13 *node*, *dropout* sebesar 0.2, *batch size* sebesar 32, *epoch* sebesar 100, dan *learning rate* sebesar 0.01. Ketiga variabel tersebut menggunakan fungsi aktivasi ReLU. Hasil prediksi nilai MAPE yang diperoleh pada kasus baru, kasus sembuh dan kasus kematian sebesar 1.78%, 1.92% dan 0.95%, sehingga dapat diartikan bahwa hasil tersebut memiliki kemampuan prediksi sangat baik dengan nilai akurasi model sebesar 98.22% untuk kasus baru, 98.08% untuk kasus sembuh, dan 99.05% untuk kasus kematian. Oleh karena itu, metode ini sangat baik digunakan untuk meramal kasus Covid-19 pada periode ke depan.

Kata kunci: Kasus COVID-19, Data Mining, *Machine Learning*, *Artificial Neural Network*, *K-Fold Cross Validation*.

**PREDIKSI COVID-19 DI PROVINSI LAMPUNG
MENGUNAKAN METODE *ARTIFICIAL NEURAL NETWORK*
DENGAN *K-FOLD CROSS VALIDATION***

Oleh

Nur Alifiah

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

Judul Skripsi : **PREDIKSI COVID-19 DI PROVINSI LAMPUNG
MENGUNAKAN METODE ARTIFICIAL NEURAL
NETWORK DENGAN K-FOLD CROSS VALIDATION**

Nama Mahasiswa : **Nur Afifah**

Nomor Pokok Mahasiswa : **1857031002**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. **Komisi Pembimbing**

A handwritten signature in black ink, appearing to be 'Dian Kurniasari', written over a white background.

Dian Kurniasari, S.Si., M.Sc.
NIP 19690305 199603 2 001

A handwritten signature in black ink, appearing to be 'Amanto', written over a white background.

Amanto, S.Si., M.Si.
NIP 19730314 200012 1 002

2. **Ketua Jurusan Matematika**

A handwritten signature in blue ink, appearing to be 'Aang Nuryaman', written over a white background.

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : Dian Kurniasari, S.Si., M.Sc.



Sekretaris : Amanto, S.Si., M.Si.



**Penguji
Bukan Pembimbing : Ir. Warsono, M.S., Ph.D.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Sripto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001



Tanggal Lulus Ujian Skripsi : 20 Juli 2022

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : **Nur Alifiah**

Nomor Pokok Mahasiswa : **1857031002**

Jurusan : **Matematika**

Judul Skripsi : **PREDIKSI COVID-19 DI PROVINSI
LAMPUNG MENGGUNAKAN METODE
ARTIFICIAL NEURAL NETWORK DENGAN
K-FOLD CROSS VALIDATION**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 28 Juli 2022

Penulis



Nur Alifiah

RIWAYAT HIDUP

Penulis bernama Nur Alifiah dilahirkan di Kota Bekasi, Provinsi Jawa Barat pada 28 Mei 2000. Penulis merupakan anak pertama dari dua bersaudara pasangan Bapak Arfa Ageng Sendang dan Ibu Ika Chasanah.

Penulis pertama kali menempuh pendidikannya di Taman Kanak-Kanak Islam Al-Qudus Bekasi pada tahun 2005-2006 dan melanjutkan pendidikannya ke Sekolah Dasar di SDN Sumber Jaya 04 Tambun Selatan pada tahun 2006-2012. Kemudian, penulis melanjutkan jenjang pendidikannya di SMP Negeri 3 Kota Serang pada tahun 2012-2015. Jenjang pendidikan selanjutnya di SMA Negeri 1 Kota Serang pada tahun 2015-2018.

Pada tahun 2018 penulis diterima sebagai mahasiswa S1 Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur Seleksi Mandiri Masuk Perguruan Tinggi Negeri (SMMPTN). Selama menjadi mahasiswa penulis juga aktif dalam berorganisasi Himpunan Mahasiswa Matematika (HIMATIKA) FMIPA UNILA dan Himpunan Mahasiswa Banten (HMB). Pada tahun 2021 penulis melakukan Kuliah Kerja Nyata (KKN) di Desa Cikoneng, Kecamatan Anyar, Serang, Banten, Kerja Praktik (KP) di BPS Provinsi Banten, dan Kampus Mengajar Di SDIT Mega Utama Tunas Bangsa. Pada tahun 2022 penulis mengikuti kegiatan Studi Independen di Zenius dengan program *Accelerated Machine Learning*.

KATA INSPIRASI

*Maka Sesungguhnya bersama kesulitan itu ada kemudahan
(Q.S. Al-Insyirah: 5)*

*Jangan terlalu keras pada dirimu sendiri, karena hasil akhir dari semua urusan
di dunia ini sudah ditetapkan oleh Allah.
(Umar bin Khattab)*

*Terus berjalan jangan memikirkan apa yang kamu lakukan dan jangan sedih
karena kamu tidak merasa menjadi yang terbaik, suatu saat kamu akan
meningkat menjadi lebih baik sebelum kamu menyadarinya
(Penulis)*

*Jika seseorang bepergian dengan tujuan mencari ilmu, maka Allah akan
menjadikan perjalanannya seperti perjalanan menuju surge
(HR. Bukhari)*

PERSEMBAHAN

Dengan mengucapkan puji dan syukur atas kehadiran Allah SWT, atas segala rahmat dan hidayah-Nya skripsi ini dapat terselesaikan dengan baik, serta solawat serta salam kepada junjungan kita Nabi Muahammad SAW. Dengan penuh ketulusan saya persembahkan karya kecil ini untuk:

Bapak Arfa Ageng Sendang dan Ibu Ika Chasanah

Terima kasih kepada kedua orang tua saya yang selalu memberikan dukungan dan saran dalam setiap keputusan, kasih sayang serta doa yang tak pernah putus dalam setiap langkah yang saya tempuh

Adik Alfenia Rahma

Terima kasih telah memberikan doa, semangat, serta dukungan selama ini

Dosen Pembimbing dan Pembahas

Terima kasih kepada bapak dan ibu dosen yang sangat berjasa, membantu, memberikan arahan, serta masukan dan ilmu yang bermanfaat

Teman-teman yang telah membantu, menemani, serta mendukung setiap langkahnya dari awal, hingga saat ini, dan seterusnya;

Almamater Tercinta, Universitas Lampung

SANWACANA

Puji dan syukur penulis panjatkan kehadirat Allah SWT karena berkat segala rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi dengan judul “Prediksi Covid-19 di Provinsi Lampung Menggunakan Metode *Artificial Neural Network* dengan *K-Fold Cross Validation*”.

Dalam menyusun laporan ini penulis banyak mendapatkan bantuan. Oleh karena itu, penulis mengucapkan terimakasih sebesar-besarnya kepada:

1. Ibu Dian Kurniasari, S.Si., M.Sc., selaku Dosen Pembimbing I yang telah bersedia membimbing, memberi saran, bantuan, motivasi, dan arahan dalam menyelesaikan skripsi ini.
2. Bapak Amanto, S.Si., M.Si., selaku Pembimbing II yang telah memberikan saran serta masukan kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Bapak Ir. Warsono, M.S., Ph.D., selaku dosen penguji yang telah memberikan kritik dan saran selama proses penyusunan skripsi.
4. Ibu Dr. Khoirin Nisa, M.Si., selaku dosen Pembimbing Akademik.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika.
6. Bapak Dr. Eng. Suropto Dwi Yuwono, S.Si., M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen, staff, karyawan Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Orang tua, feni, dan keluarga besar yang selalu mendoakan dan memberikan dukungan.
9. Mega, Martha, Dalfa, Udil, Vinny, Nanda, Bunga, Ajeng, Lisur, Tyas, Resita yang memberi dukungan selama ini.

10. Ferzy, Zaenal, Farel, Joshua, Putsal, Maydia, Sulis, Shafira, Luthfia, Febi, Oktin, Virda yang telah memberikan dukungan, bantuan, serta menemani penulis selama menyusun skripsi ini.
11. Semua pihak yang telah membantu dalam menyelesaikan laporan kerja praktik ini yang tidak bisa penulis sebutkan satu persatu.

Penulis menyadari masih banyak kekurangan dalam laporan ini. Oleh karena itu, kritik dan saran sangat diharapkan agar dapat menjadi pelajaran dan perbaikan untuk kedepannya. Semoga skripsi ini dapat memberikan manfaat baik bagi penulis maupun bagi pihak yang membutuhkan.

Bandar Lampung, Juli 2022
Penulis,

Nur Alifiah

DAFTAR ISI

	Halaman
DAFTAR ISI	ii
DAFTAR TABEL	iv
DAFTAR GAMBAR	v
I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	4
II. TINJAUAN PUSTAKA	5
2.1 Data Mining	5
2.2 <i>Machine Learning</i>	7
2.3 <i>Artificial Neural Network (ANN)</i>	9
2.4 <i>Arsitektur Artificial Neural Network</i>	9
2.5 <i>Struktur Artificial Neural Network</i>	11
2.6 Fungsi Aktivasi	11
2.7 <i>Algoritma Backpropagation</i>	12
2.8 <i>Hyperparameter Tuning</i>	13
2.8.1. <i>Nodes Pada Hidden Layer</i>	14
2.8.2. <i>Dropout</i>	14
2.8.3. <i>Epochs</i>	15
2.8.4. <i>Batch Size</i>	15
2.8.5. <i>Learning Rate</i>	15
2.9 <i>K-Fold Cross Validation</i>	16
2.10 Covid-19	16
III. METODOLOGI PENELITIAN	18
3.1 Waktu dan Tempat Penelitian.....	19
3.2 Data Penelitian	19
3.3 Metode Penelitian	20

IV. HASIL DAN PEMBAHASAN	22
4.1 Proses Eksplorasi Data	22
4.1.1 <i>Input Data</i>	23
4.1.2 Visualisasi Data	24
4.1.3 <i>Preprocessing Data</i>	24
4.1.4 Transformasi Data	25
4.2 Membangun Model <i>Artificial Neural Network</i> (ANN).....	26
4.3 Pengujian Model	28
4.4 Hasil Prediksi Kasus Covid-19 di Provinsi Lampung.....	34
4.5 Evaluasi Model <i>Artificial Neural Network</i> (ANN)	36
4.6 Peramalan Kasus Covid-19 di Provinsi Lampung	37
V. KESIMPULAN	42
DAFTAR PUSTAKA	43
LAMPIRAN	

DAFTAR GAMBAR

Gambar	Halaman
1. Proses <i>Knowledge Discovery in Database</i>	6
2. Struktur <i>Single Layer Network</i>	10
3. Struktur <i>Multilayer Network</i>	10
4. Struktur <i>Artificial Neural Network</i>	11
5. Fungsi Aktivasi <i>Rectified Linear Unit</i> (ReLU)	12
6. Contoh <i>K-Fold Cross Validation</i> dengan $k=5$	16
7. <i>Flowchart</i> proses metode ANN	20
8. Plot Data Covid-19 di Provinsi Lampung.....	24
9. Struktur <i>Artificial Neural Network</i> (ANN) Pada Kasus Baru.....	27
10. Struktur <i>Artificial Neural Network</i> (ANN) Pada Kasus Sembuh	27
11. Struktur <i>Artificial Neural Network</i> (ANN) Pada Kasus Kematian.....	28
12. Grafik <i>Loss</i> Untuk <i>Fold 1</i> Pada Data Kasus Baru	29
13. Grafik <i>Loss</i> Untuk <i>Fold 2</i> Pada Data Kasus Baru	29
14. Grafik <i>Loss</i> Untuk <i>Fold 3</i> Pada Data Kasus Baru	29
15. Grafik <i>Loss</i> Untuk <i>Fold 4</i> Pada Data Kasus Baru	29
16. Grafik <i>Loss</i> Untuk <i>Fold 5</i> Pada Data Kasus Baru	30
17. Grafik <i>Loss</i> Untuk <i>Fold 6</i> Pada Data Kasus Baru	30
18. Grafik <i>Loss</i> Untuk <i>Fold 7</i> Pada Data Kasus Baru	30

19. Grafik <i>Loss</i> Untuk <i>Fold</i> 8 Pada Data Kasus Baru	30
20. Grafik <i>Loss</i> Untuk <i>Fold</i> 9 Pada Data Kasus Baru	30
21. Grafik <i>Loss</i> Untuk <i>Fold</i> 10 Pada Data Kasus Baru	30
22. Grafik <i>Loss</i> Untuk <i>Fold</i> 1 Pada Data Kasus Sembuh	31
23. Grafik <i>Loss</i> Untuk <i>Fold</i> 2 Pada Data Kasus Sembuh	31
24. Grafik <i>Loss</i> Untuk <i>Fold</i> 3 Pada Data Kasus Sembuh	31
25. Grafik <i>Loss</i> Untuk <i>Fold</i> 4 Pada Data Kasus Sembuh	31
26. Grafik <i>Loss</i> Untuk <i>Fold</i> 5 Pada Data Kasus Sembuh	31
27. Grafik <i>Loss</i> Untuk <i>Fold</i> 6 Pada Data Kasus Sembuh	31
28. Grafik <i>Loss</i> Untuk <i>Fold</i> 7 Pada Data Kasus Sembuh	31
29. Grafik <i>Loss</i> Untuk <i>Fold</i> 8 Pada Data Kasus Sembuh	31
30. Grafik <i>Loss</i> Untuk <i>Fold</i> 9 Pada Data Kasus Sembuh	32
31. Grafik <i>Loss</i> Untuk <i>Fold</i> 10 Pada Data Kasus Sembuh	32
32. Grafik <i>Loss</i> Untuk <i>Fold</i> 1 Pada Data Kasus Kematian.....	32
33. Grafik <i>Loss</i> Untuk <i>Fold</i> 2 Pada Data Kasus Kematian.....	32
34. Grafik <i>Loss</i> Untuk <i>Fold</i> 3 Pada Data Kasus Kematian.....	33
35. Grafik <i>Loss</i> Untuk <i>Fold</i> 4 Pada Data Kasus Kematian.....	33
36. Grafik <i>Loss</i> Untuk <i>Fold</i> 5 Pada Data Kasus Kematian.....	33
37. Grafik <i>Loss</i> Untuk <i>Fold</i> 6 Pada Data Kasus Kematian.....	33
38. Grafik <i>Loss</i> Untuk <i>Fold</i> 7 Pada Data Kasus Kematian.....	33
39. Grafik <i>Loss</i> Untuk <i>Fold</i> 8 Pada Data Kasus Kematian.....	33
40. Grafik <i>Loss</i> Untuk <i>Fold</i> 9 Pada Data Kasus Kematian.....	33
41. Grafik <i>Loss</i> Untuk <i>Fold</i> 10 Pada Data Kasus Kematian.....	33
42. Plot Perbandingan Data Aktual dengan Data Prediksi Kasus Baru	34

43. Plot Perbandingan Data Aktual dengan Data Prediksi Kasus Sembuh.....	35
44. Plot Perbandingan Data Aktual dengan Data Prediksi Kasus Kematian	36
45. Plot Hasil Peramalan dengan Data Aktual Pada Kasus Baru.....	39
46. Plot Hasil Peramalan dengan Data Aktual Pada Kasus Sembuh	40
47. Plot Hasil Peramalan dengan Data Aktual Pada Kasus Kematian.....	41

DAFTAR TABEL

Tabel	Halaman
1. Data Harian Kasus Covid-19 di Provinsi Lampung	19
2. Data Kasus Covid-19 di Provinsi Lampung	22
3. Data <i>Input</i> Kasus Covid-19 di Provinsi Lampung	23
4. <i>Missing Value</i> Pada Data Covid-19	25
5. Hasil Data <i>Scaling</i>	25
6. Parameter optimal dari proses <i>hypertuning</i>	26
7. Nilai <i>accuracy</i> hasil pengujian model <i>artificial neural network</i> dengan <i>k-fold cross validation</i>	29
8. Evaluasi Model Kasus Covid-19 di Provinsi Lampung.....	37
9. Hasil Peramalan Covid-19 di Provinsi Lampung	37

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

World Health Organization (WHO) secara resmi mengemukakan bahwa Covid-19 yang telah menyebar ke berbagai negara di dunia ditetapkan sebagai pandemi pada tanggal 11 Maret 2020 (Virtyani, dkk., 2021). Covid-19 pertama di Indonesia dilaporkan masuk pada tanggal 2 Maret 2020 dengan jumlah terkonfirmasi positif sebanyak 1 kasus. Covid-19 merupakan penyakit yang disebabkan oleh SARS-CoV-2 yang menyerang sistem pernapasan. Kasus penyebaran Covid-19 mengalami peningkatan setiap harinya.

Pemerintah Indonesia merencanakan pemberlakuan vaksinasi di saat kondisi yang sedang darurat ini. Pemerintah berharap masyarakat mau menerima vaksinasi dikarenakan vaksin berfungsi untuk meningkatkan kekebalan atau memperlambat laju penularan. Akan tetapi, vaksin bukan sepenuhnya menangkal Covid-19, kemungkinan besar masyarakat masih terkena Covid-19 setelah vaksinasi. Saat ini masyarakat masih diminta untuk selalu mewaspadaai prediksi gelombang ketiga Covid-19. Penambahan kasus Covid-19 setiap harinya belum bisa diprediksi menjadikan penanganan, seperti penyediaan layanan, fasilitas, dan tenaga medis belum bisa diperkirakan. Angka tersebut akan selalu bertambah apabila masyarakat tidak memiliki kewaspadaan dan berperan aktif mengurangi laju penambahan kasus baru (Kemenkes, 2021). Oleh karena itu, kesadaran dan kewaspadaan publik perlu ditingkatkan dengan menyajikan informasi prediksi kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 agar dapat dijadikan acuan pemerintah dalam

mengambil dan menetapkan suatu kebijakan untuk mengatasi penyebaran Covid-19.

Prediksi adalah memperkirakan segala sesuatu yang akan terjadi. Prediksi dilakukan guna memperoleh informasi, kajian, serta wawasan berdasarkan fakta dan faktor-faktor yang mempengaruhi. Prediksi dilakukan berdasarkan pengumpulan data dengan tenggat waktu tertentu dan dianalisis dengan metode statistik (Sulastri dan Nugroho, 2017). Kemajuan teknologi saat ini dapat dimanfaatkan untuk melakukan prediksi, yaitu dengan *artificial intelligence*. *Artificial intelligence* merupakan sebuah mesin berbasis teknologi dengan kecerdasan seperti manusia berpikir. *Artificial intelligence* sangat efektif digunakan dikarenakan dapat meminimalisir *human error*. Sub bidang dalam *artificial intelligence* diantaranya *machine learning*. *Machine learning* merupakan inovasi dari mesin yang dikembangkan untuk belajar menghasilkan suatu model dari sekumpulan data. Proses belajar pada *machine learning* menggunakan algoritma khusus yang lebih sering disebut dengan *machine learning algorithms* (Santoni, dkk., 2020).

Algoritma *machine learning* yang paling sering digunakan adalah *artificial neural networks* (ANN). ANN dapat digunakan untuk menyelesaikan permasalahan yang rumit, seperti prediksi dan pemodelan. Secara garis besar proses *artificial neural networks* memasukkan *input* yang berfungsi sebagai *dendrite* pada otak manusia, kemudian *output* seperti *akson* dengan fungsi aktivasi sebagai sinapsis. Selanjutnya, metode ini menggunakan *K-Fold Cross Validation*. *K-Fold* metode yang populer dengan cara melipat data sebanyak k bagian, dimana data *testing* menggunakan satu bagian dan $k-1$ bagian sebagai data *training*, kemudian tiap k akan dicari tingkat akurasi, sehingga memperoleh rata-rata akurasi (Ramadhona, dkk., 2018).

Berdasarkan beberapa penelitian terkait penerapan algoritma *artificial neural network* telah banyak digunakan sebelumnya. Penelitian Santoni, dkk. (2020), mengenai prediksi hipertensi menggunakan *decision tree*, *naïve bayes*, dan *artificial neural network* pada *software* KNIME menghasilkan kesimpulan bahwa algoritma *artificial neural network* lebih efektif dibandingkan algoritma *decision tree*, dan *naïve bayes* dengan nilai akurasi sebesar 94,7% pada metode *artificial neural network*, 83% pada metode *decision tree*, dan 74,5% pada metode *naïve bayes*. Pada penelitian Ramadhona, dkk. (2018), mengenai prediksi produktivitas padi menggunakan jaringan syaraf tiruan *backpropagation* menghasilkan kesimpulan bahwa nilai *Root Mean Square Error* (RMSE) yang didapatkan sebesar 8.6918, sedangkan pada pengujian *k-fold cross validation* menghasilkan nilai rata-rata RMSE sebesar 8.2126.

Oleh karena itu, peneliti tertarik melakukan riset dengan menggunakan algoritma *machine learning*, yaitu *artificial neural network* dengan *k-fold cross validation* untuk memprediksi kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung. Penelitian ini diharapkan mampu mendapatkan tingkat akurasi yang baik.

1.2 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Menentukan struktur jaringan terbaik untuk memprediksi kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung menggunakan metode *Artificial Neural Network* (ANN) dengan *k-fold cross validation*.
2. Melakukan prediksi mengenai kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung menggunakan metode *Artificial Neural Network* (ANN) dengan *k-fold cross validation*.

3. Melakukan peramalan kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung menggunakan metode *Artificial Neural Network* (ANN) dengan *k-fold cross validation*.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

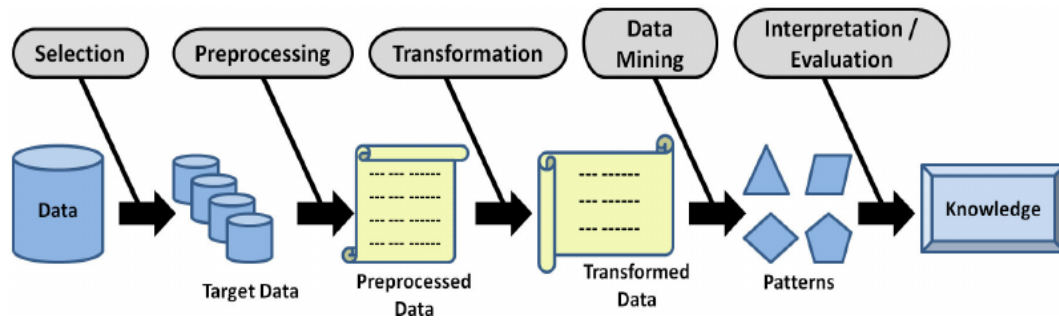
1. Memberikan pengetahuan mengenai metode *Artificial Neural Network* dengan *k-fold cross validation*.
2. Memberikan gambaran terkait kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung sehingga dapat membantu pemerintah dalam menentukan kebijakan, serta perkiraan penyediaan layanan, fasilitas, dan tenaga medis.

II. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah proses menemukan korelasi, pola, dan trend yang berguna untuk data yang besar yang disimpan dalam repository, menggunakan pola pengenalan teknologi serta teknik statistik dan matematika. Beberapa tugas data mining yang paling umum, yaitu estimasi, prediksi/peramalan, klasifikasi, *clustering* (Larose dan Larose, 2014). Sumber data yang tersimpan pada sistem mencakup *database*, *data warehouse*, *web*, maupun *repository*. Oleh karena itu, data mining memiliki sumber ilmu sebagai referensi, seperti *artificial intelligence*, *machine learning*, statistik, dan *database* (Han dan Kamber, 2006). Data mining sebagian besar berkaitan dengan membangun model. Model hanyalah sebuah algoritma atau seperangkat aturan yang menghubungkan sekumpulan *input* ke hasil tertentu (Berry dan Linoff, 2004).

Menurut Hermawati (2013), data mining merupakan proses mengerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan memperoleh wawasan secara otomatis dari suatu kumpulan data. Proses data mining merupakan salah satu bagian dari *Knowledge Discovery in Database* (KDD) yang kerap kali dimanfaatkan untuk memperoleh informasi dari sekumpulan data.



Gambar 1. Proses *Knowledge Discovery in Database*. (Sumber: Gullo, 2015)

Secara umum, proses *Knowledge Discovery in Database* (KDD) adalah sebagai berikut:

1. *Data Selection* merupakan proses seleksi data dari keseluruhan data operasional. *Data Selection* dilakukan sebelum proses penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining akan disimpan dalam suatu berkas terpisah dari basis data operasional.
2. *Pre-processing/cleaning* merupakan tahap yang dilakukan sebelum proses data mining. Tujuan dari tahap ini guna menghapus duplikasi data, dan memperbaiki kesalahan pada data. Pada proses ini juga dilakukan proses *enrichment* yaitu memperbanyak data yang relevan dari data sebelumnya, seperti pengetahuan eksternal.
3. *Transformation* merupakan proses untuk merubah skala pengukuran data dari bentuk asli kedalam bentuk lain sehingga dapat digunakan untuk analisis dan asumsi-asumsi tertentu. Transformasi memiliki tujuan untuk menyelaraskan data yang diolah berdasarkan pada algoritma serta *software* yang digunakan. Transformasi data dapat dilakukan dengan menggunakan *standard scaler*. *Standard Scaler* merupakan proses penskalaan ulang nilai distribusi sehingga rata-rata nilai yang diamati adalah 0 dan standar deviasinya adalah 1. Persamaan pada *standard scaler*, yaitu (Halim, dkk., 2020):

$$x^* = \frac{x - \mu_x}{\sigma_x} \quad (2.1)$$

dimana:

x^* = nilai x baru

μ_x = rata-rata nilai x

σ_x = standar deviasi nilai x

4. Data Mining merupakan tahap untuk mencari pola atau informasi dalam data yang terpilih dengan menggunakan metode tertentu. Teknik, metode, maupun algoritma dalam data mining sangat beragam.
5. *Interpretation/Evaluation* merupakan tahap memeriksa pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak. Hasil yang diperoleh perlu disajikan dalam bentuk yang mudah dimengerti.

2.2. *Machine Learning*

Menurut Putra (2020), *machine learning* merupakan cara yang dilakukan untuk memperoleh kesimpulan dari data yang ada melalui pendekatan matematis. Secara terperinci, *machine learning* adalah untuk membentuk model matematis yang merefleksikan pola-pola data. Tujuan dari *machine learning* pada umumnya ada dua, yaitu memprediksi masa yang akan datang dan wawasan. Tujuan tersebut dapat tercapai melalui *dataset* atau kumpulan data/sampel dalam statistik, kemudian membuat model untuk menggeneralisasi pola data sehingga dapat digunakan untuk mendapatkan informasi/membuat keputusan. Dalam pembangunan model *machine learning* terdapat dua istilah penting, yaitu *training* dan *testing*. *Training* merupakan tahap membentuk model dan *testing* merupakan tahap pengujian performa model. Pada umumnya, rasio pembagian *dataset* menjadi *training* dan *testing* (90% : 10%), (80% : 20%), (70% : 30%), atau (50% : 50%).

Raw data atau data mentah merupakan beberapa fakta yang kemungkinan besar tidak memberikan informasi apapun. *Raw data* memiliki sifat tidak rapi, misalnya

mengandung data yang hilang. *Raw data* dapat dianalisis dengan menggunakan teknik *machine learning* yang dikenal dengan istilah *pre-processing*. Hal pertama yang dapat dilakukan data dirapikan sesuai dengan format yang diinginkan, kemudian mencari fitur yang dapat mempresentasikan data. Setelah *pre-processing*, kita menggunakan teknik yang ada untuk menemukan pola pada data (Putra, 2020).

Menurut Sathya dan Abraham (2013), *machine learning* secara garis besar terbagi menjadi 2, yaitu:

1. *Supervised Learning*

Supervised learning merupakan pembelajaran menyelesaikan problematika melalui latihan atau *training* di dalam sistem syaraf buatan melalui data yang sudah diberikan dan dikategorikan atau *labeling*. *Supervised learning* didasarkan pada pelatihan sampel data dari sumber data dengan klasifikasi yang sudah ditugaskan. Teknik tersebut digunakan dalam model *feedforward* atau *MultiLayer Perceptron* (MLP). Beberapa masalah *supervised learning*, diantaranya:

a. Klasifikasi

Klasifikasi merupakan suatu objek perlu ditugaskan ke dalam kelompok atau kelas yang telah ditentukan berdasarkan sejumlah pengamatan atribut yang terkait dengan objek tersebut.

b. Prediksi

Prediksi merupakan dugaan nilai sebuah variabel berdasarkan nilai dari variabel tersebut. Prediksi juga merupakan proses menentukan data di masa mendatang dengan menganalisis data masa lalu hingga saat ini.

2. *Unsupervised Learning*

Unsupervised Learning mengacu pada kemampuan untuk mempelajari dan mengatur informasi tanpa memberikan *error* untuk mengevaluasi solusi. Metode ini digunakan untuk mengidentifikasi pola tersembunyi dalam data input yang tidak berlabel.

2.3. *Artificial Neural Network (ANN)*

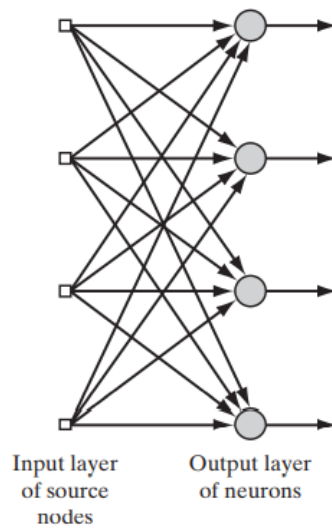
Menurut Santoso, dkk., (2021), *artificial neural network* merupakan sistem proses informasi yang memiliki karakteristik performa seperti jaringan *neural* secara biologi. *Artificial Neural Network (ANN)* atau jaringan syaraf tiruan merupakan penjabaran fungsi otak manusia dalam bentuk fungsi matematika yang dapat digunakan untuk menyelesaikan tugas-tugas yang rumit, seperti prediksi dan pemodelan. Secara garis besar proses dari *Artificial Neural Network (ANN)* memasukkan *input* yang berfungsi sebagai *dendrite* pada otak manusia lalu *output* seperti akson dengan fungsi aktivasi sebagai sinapsis. *Neural network* dibangun dari banyak *node/unit* yang dihubungkan oleh *link* secara langsung. *Link* dari *unit* yang satu ke *unit* yang lainnya digunakan untuk melakukan propagasi aktivasi dari *unit* pertama ke *unit* selanjutnya. Setiap *link* memiliki bobot numerik. Bobot ini menentukan kekuatan serta penanda dari sebuah konektivitas (Dewi, dkk, 2020).

2.4. *Arsitektur Artificial Neural Network*

Menurut Haykin (2009), arsitektur *artificial neural network* terbagi menjadi beberapa jenis, yaitu:

a. *Single layer network*

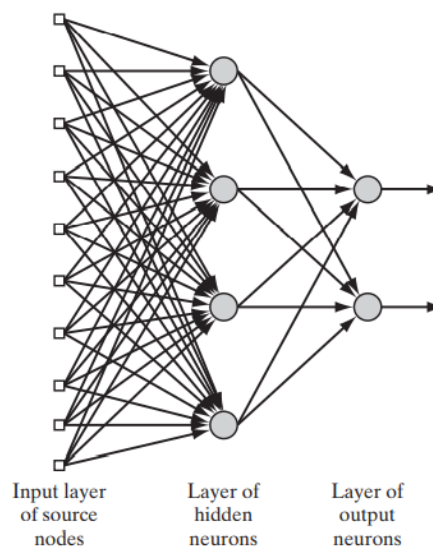
Single layer network merupakan jaringan yang paling sederhana yang terdiri dari lapisan *input* dan *output*. Pada jaringan ini lapisan *input* memproyeksikan langsung ke lapisan *output*. Semua *neuron* pada lapisan satu dan dua di hubungan dengan bobot yang bersesuaian.



Gambar 2. Struktur *Single Layer Network*. (Sumber: Haykin, 2009)

b. *Multilayer Network*

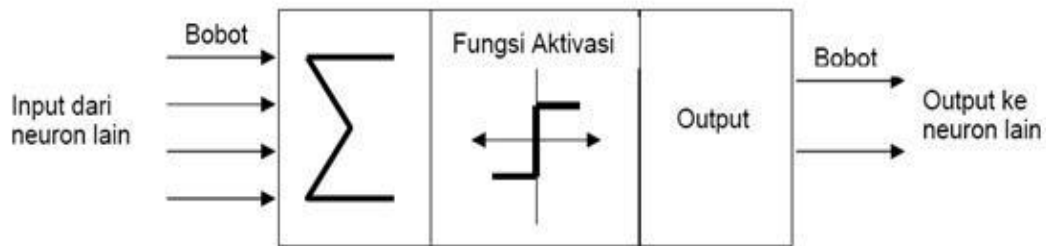
Multilayer Network merupakan jaringan yang mempunyai lapisan lebih dari 1 atau bisa disebut *hidden layer*. Lapisan ini dinamakan *hidden* karena diantara lapisan *input* dan *output*. Lapisan *input* diproyeksikan ke dalam lapisan *hidden* setelah itu diproyeksikan ke dalam lapisan *output*. Jaringan ini menyelesaikan masalah lebih rumit dari pada *single*.



Gambar 3. Struktur *Multilayer Network*. (Sumber: Haykin, 2009)

2.5. Struktur *Artificial Neural Network*

Menurut Aina (2018), struktur neuron pada *Artificial Neural Network* atau jaringan syaraf tiruan terdiri dari beberapa *neuron* dan terdapat penghubung antara *neuron-neuron* yang dikenal dengan bobot seperti halnya otak manusia.



Gambar 4. Struktur *Artificial Neural Network*. (Sumber: Aina, 2018)

Berdasarkan gambar tersebut, komponen-komponen *artificial neural network* sebagai berikut:

- a. *Input* adalah sebagai penerima informasi masukan dari *neuron* lain.
- b. *Neuron* adalah sebagai komponen yang bertugas untuk memproses informasi.
- c. Bobot adalah untuk menunjukkan kekuatan hubungan antara *neuron* satu dengan yang lainnya.
- d. Fungsi aktivasi adalah suatu nilai tertentu yang memetakan fungsi hasil penjumlahan yang diterima oleh semua *input* dari suatu *neuron*.
- e. *Output* adalah proses pembelajaran atau proses perhitungan suatu fungsi aktivasi yang akan menghasilkan suatu *output* dari jaringan yang telah di *input* atau bahkan akan menjadi *input* bagi *neuron* yang lain.

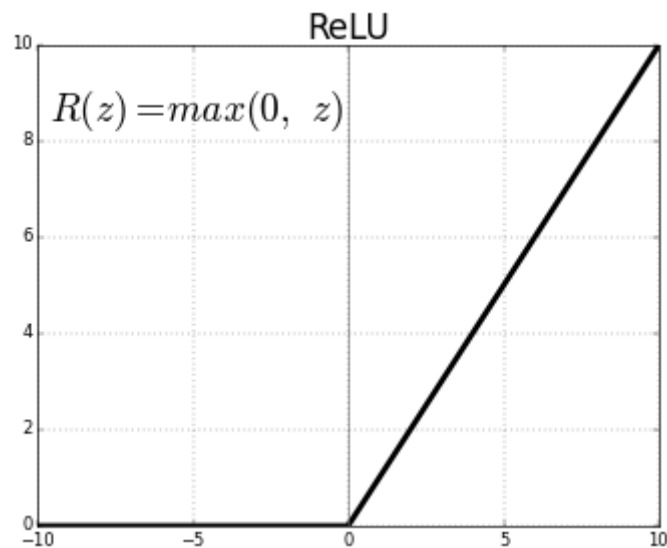
2.6. Fungsi Aktivasi

Menurut Sitepu dan Sigiro (2021), fungsi aktivasi adalah fungsi yang mentransformasi suatu *input* menjadi suatu *output*. Fungsi aktivasi memiliki peran penting dalam suatu jaringan syaraf tiruan dimana penggunaannya tergantung

sesuai kebutuhan dan target yang diinginkan serta fungsi aktivasi ini yang akan menentukan besarnya bobot. Fungsi aktivasi yang umum digunakan, yaitu *Rectified Linear Unit* (ReLU). ReLU adalah fungsi aktivasi yang digunakan untuk menormalisasikan nilai yang dihasilkan *layer*. Persamaan pada fungsi aktivasi ini, yaitu:

$$f(x) = \max(0, x) \quad (2.2)$$

ReLU hanya membuat pembatas pada bilangan nol, artinya ketika $x \leq 0$ maka nilai $x = 0$, dan ketika $x > 0$, maka $x = x$. Setiap *input* negatif yang diberikan akan selalu dipetakan menjadi 0 dan setiap *input* positif akan dipertahankan nilainya, sehingga tidak terdapat hasil yang bernilai negatif.



Gambar 5. Fungsi Aktivasi *Rectified Linear Unit* (ReLU). (Sumber: Sarkar, 2018)

2.7. Algoritma *Backpropagation*

Menurut Jumarwanto, dkk., (2009), algoritma *backpropagation* melatih jaringan mendapatkan keseimbangan antara kemampuan jaringan untuk mengenali pola yang digunakan selama pelatihan serta kemampuan untuk memberi respon yang

benar terhadap pola masukan yang serupa dengan pola yang dipakai selama pelatihan. Pelatihan *backpropagation* meliputi tiga fase yaitu:

1. Propagasi maju

Pada proses ini, sinyal *input* diproyeksikan ke *hidden layer* menggunakan fungsi aktivasi yang telah ditentukan. Setiap unit *hidden layer* memiliki keluaran, selanjutnya dilanjutkan ke *hidden layer* lainnya menggunakan fungsi aktivasi yang telah ditentukan sampai menghasilkan nilai *output* pada *output layer*. Berikutnya, nilai *output* dibandingkan dengan target yang harus dicapai. Cara membandingkannya dengan mencari nilai *error*. Jika *error* yang dihasilkan lebih kecil dari batas toleransi yang ditentukan, maka iterasi dihentikan. Apabila *error* masih lebih besar dari batas toleransinya, maka bobot setiap garis dalam jaringan akan dimodifikasikan untuk mengurangi kesalahan yang terjadi.

2. Propagasi mundur

Proses ini dilakukan apabila nilai yang dihasilkan pada lapisan *output* memiliki tingkat *error* lebih besar dari data yang ditentukan. Berdasarkan nilai *error*, dihitung *error gradient* yang digunakan untuk mengubah bobot garis yang menghubungkan langsung dengan *output layer*. Dengan menggunakan cara yang sama, nilai bobot diubah pada dari setiap unit di *hidden layer* ke lapisan di bawahnya.

3. Perubahan bobot

Setelah semua faktor dihitung, bobot semua garis dimodifikasi bersamaan. Perubahan bobot suatu garis didasarkan atas faktor *neuron* di lapis atasnya. Sebagai contoh, perubahan bobot garis yang menuju ke lapis keluaran didasarkan atas dasar yang ada di unit keluaran.

2.8. *Hyperparameter Tuning*

Menurut Michelucci (2018), *hyperparameter* merupakan parameter yang terdapat dalam suatu model dan telah ditentukan sebelum proses *training* dilakukan. Kombinasi parameter yang optimal akan menghasilkan model yang baik. Kombinasi parameter yang optimal dapat dicari dengan menggunakan *hypertuning*,

akan tetapi *hypertuning* membutuhkan waktu yang lama. Oleh karena itu, proses *hypertuning* ini akan menggunakan *early stopping* agar proses training yang dilakukan pada model berhenti ketika kondisi telah terpenuhi. *Hyperparameter* yang umum digunakan, yaitu:

2.8.1. Nodes pada Hidden Layer

Nodes pada *hidden layer* akan mempengaruhi proses *training*. Jumlah *nodes* pada *hidden layer* yang terlalu sedikit akan menyebabkan model mengalami *underfitting*. Jumlah *nodes* pada *hidden layer* yang terlalu banyak akan menyebabkan model mengalami *overfitting*. Berikut beberapa cara untuk menentukan jumlah *nodes* pada *hidden layer* (Heaton, 2017):

- a. Jumlah *nodes* pada *hidden layer* berada di antara jumlah *nodes* pada *input layer* dan *output layer*.
- b. Jumlah *nodes* pada *hidden layer* bernilai $2/3$ dari jumlah *nodes* pada *input layer* ditambah *output layer*.
- c. Jumlah *nodes* pada *hidden layer* kurang dari dua kali jumlah *nodes* pada *input layer*.

2.8.2. Dropout

Menurut Srivastava, dkk. (2014), *dropout* merupakan pengabaian secara acak *nodes* tertentu dalam lapisan selama proses *training*. *Dropout* dapat dilakukan untuk mencegah terjadinya *overfitting* yang terjadi selama proses *training*.

2.8.3. Epochs

Epochs adalah jumlah iterasi atau perulangan selama proses *training* yang memberikan *input* dari jaringan dan juga memperbarui bobot jaringan. Proses pelatihan tersebut berlangsung pada *neural network* sampai ke awal lagi (Vijayalakshmi dan Venkatachalapathy, 2019).

2.8.4. Batch Size

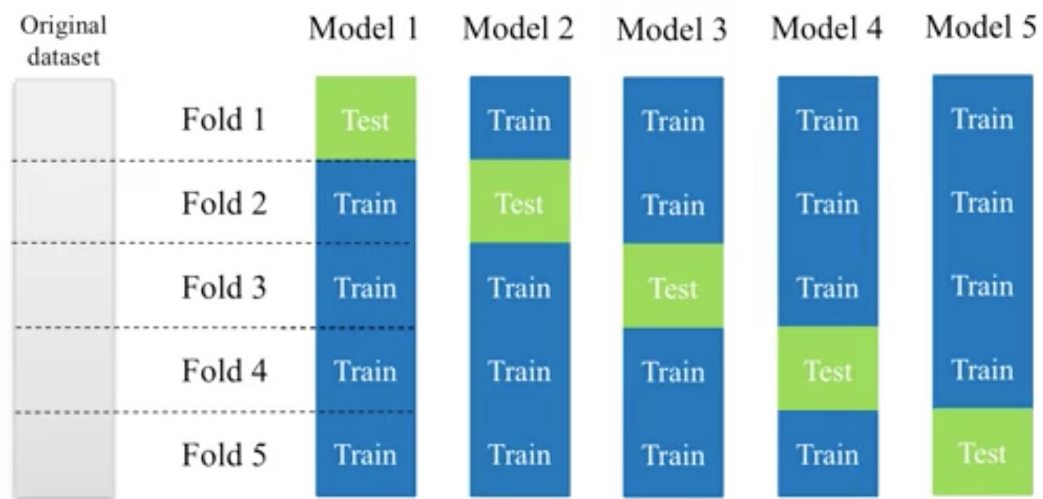
Menurut Rochmawati, dkk. (2021), *batch size* merupakan salah satu *hyperparameter* penting yang digunakan dalam pembelajaran mesin dan mengacu pada jumlah contoh pelatihan yang digunakan dalam satu iterasi.

2.8.5. Learning Rate

Learning rate adalah parameter yang sangat penting pada kinerja suatu jaringan terhadap waktu yang dibutuhkan agar tercapainya suatu target yang optimal. Suatu target dapat dikatakan optimal apabila nilai perubahan bobot serta *error* yang dihasilkan kecil. Proses pelatihan pada beberapa kasus memerlukan banyak iterasi, sehingga membutuhkan waktu yang cukup lama. Oleh karena itu, untuk mempercepat iterasi digunakannya suatu parameter yaitu *learning rate*. Nilai *learning rate* yang umum digunakan, yaitu 0.1, 0.01, 0.001 (Aina, 2018).

2.9. *K-Fold Cross Validation*

K-fold cross validation merupakan metode statistik untuk menganalisis suatu model yang memiliki generalisasi yang baik. Pada *k-fold cross validation* data dibagi menjadi k lipatan dengan ukuran yang sama atau hampir sama, kemudian $k-1$ lipatan digunakan sebagai data *training*, dan sisanya menjadi data *testing*. Hal ini diulang sebanyak k kali agar diperoleh nilai di setiap bagiannya dan nilai rata-rata dari semua perulangan (Putra, 2020).



Gambar 6. Contoh *K-Fold Cross Validation* dengan $k=5$. (Sumber: Bronshtein, 2017)

2.10. Covid-19

WHO China *Country Office* melaporkan kasus pneumonia yang tidak diketahui etiologinya di Kota Wuhan, Provinsi Hubei, China pada 31 Desember 2019. China mengidentifikasi pneumonia yang tidak diketahui penyebabnya sebagai jenis baru *coronavirus* pada 7 Januari 2020. *Coronavirus* merupakan virus yang disebabkan oleh penyakit ringan sampai berat, seperti *common cold* atau pilek, MERS, dan SARS (Safrizal, dkk, 2020). Covid-19 (*Coronavirus Disease 2019*) merupakan

penyakit yang disebabkan oleh jenis coronavirus yaitu Sars-CoV-2 yang menyerang sistem pernapasan (Rahmalina dan Puspita, 2021).

III. METODOLOGI PENELITIAN

3.1. Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2021/2022 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data harian kasus Covid-19 di Provinsi Lampung yang diperoleh dari situs kementerian kesehatan RI <https://infeksiemerging.kemkes.go.id/category/situasi-terkini> yang terhitung sejak 19 Maret 2020 sampai dengan 24 Januari 2022 dan diunduh dengan cara satu per satu dalam bentuk *pdf* kemudian dimasukkan ke dalam *software excel*. Data Kasus Covid-19 yang didapatkan memuat 6 variabel, yaitu kasus baru, kasus sembuh, kasus kematian, jumlah kasus baru, jumlah kasus sembuh, dan jumlah kasus kematian. Variabel data kasus Covid-19 yang digunakan dalam penelitian ini hanya 3 variabel, yaitu kasus baru, kasus sembuh, dan kasus kematian yang dimulai pada tanggal 19 Maret 2020 sampai dengan 26 Oktober 2021. Kemudian, data kasus baru, kasus sembuh, dan kasus kematian pada tanggal 27 Oktober sampai 24 Januari 2022 akan dijadikan perbandingan dengan data hasil peramalan yang didapatkan.

Tabel 1. Data Harian Kasus Covid-19 di Provinsi Lampung

Tanggal	Kasus Baru	Kasus Sembuh	Kasus Kematian	Jumlah Kasus Baru	Jumlah Kasus Sembuh	Jumlah Kasus Kematian
19/03/2020	1	0	0	1	0	0
20/03/2020	0	0	0	1	0	0
21/03/2020	0	0	0	1	0	0
22/03/2020	0	0	0	1	0	0
23/03/2020	0	0	0	1	0	0
...
20/01/2022	1	2	0	49654	45149	3821
21/01/2022	3	1	0	49657	45150	3821
22/01/2022	6	0	1	49663	45150	3822
23/01/2022	2	1	0	49665	45151	3822
24/01/2022	3	1	0	49668	45152	3822

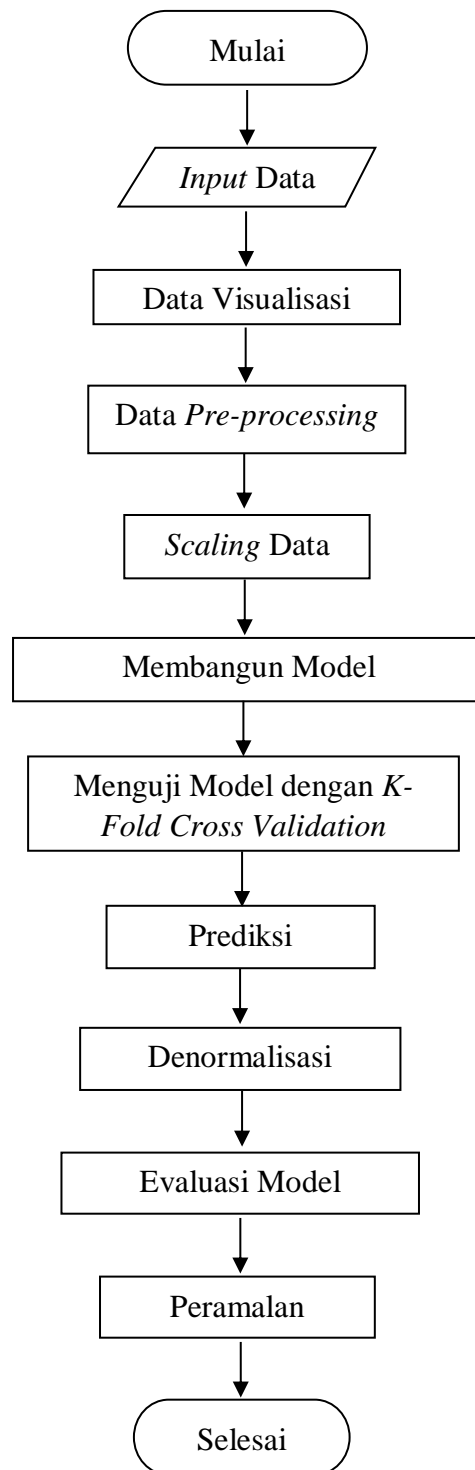
3.3. Metode Penelitian

Penelitian ini menggunakan metode *Artificial Neural Networks* (ANN) dengan *K-Fold Cross Validation* untuk memprediksi kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung dengan bantuan *software python* yang didukung oleh Google Colab.

Adapun langkah-langkah pada penelitian ini yaitu:

- Menginput data kasus Covid-19 di Provinsi Lampung yang di mulai pada tanggal 19 Maret 2020 sampai 24 Januari 2022.
- Mengambil data kasus Covid-19 yang akan diteliti yaitu kasus baru, kasus sembuh, dan kasus kematian yang di mulai pada tanggal 19 Maret 2020 sampai 26 Oktober 2021 dengan menggunakan *iloc function*.
- Melakukan visualisasi data dengan membuat plot untuk melihat pola data kasus baru, kasus sembuh, dan kasus kematian.
- Melakukan *preprocessing* data dengan melihat ada atau tidaknya *missing value* pada data kasus baru, kasus sembuh, dan kasus kematian.

- e. Melakukan transformasi data dengan menggunakan *Standard Scaler*.
- f. Melakukan *hypertuning* untuk mendapatkan parameter terbaik, yaitu *batch size*, *epoch*, *learning rate*, dan *dropout*.
- g. Menguji model *artificial neural network* dengan menggunakan *K-Fold Cross Validation* untuk parameter terbaik yang didapatkan dari *hypertuning*.
- h. Melakukan prediksi pada data kasus baru, kasus sembuh, dan kasus kematian.
- i. Melakukan denormalisasi untuk mengembalikan nilai data agar bisa dibandingkan dengan data aktual.
- j. Melakukan evaluasi terhadap model *artificial neural network* yang dibangun berdasarkan nilai RMSE, MAPE, dan akurasi yang peroleh.
- k. Melakukan peramalan untuk menentukan nilai kasus baru, kasus sembuh, dan kasus kematian pada Covid-19 di Provinsi Lampung dalam jangka waktu 90 hari ke depan.



Gambar 7. Flowchart proses metode ANN

V. KESIMPULAN

Berdasarkan hasil dan pembahasan pada bab sebelumnya, dapat diambil kesimpulan sebagai berikut:

1. Metode *artificial neural network* dengan menggunakan *k-fold cross validation* dapat digunakan untuk meramalkan kasus Covid-19 di Provinsi Lampung.
 - a. Metode *artificial neural network* pada kasus baru menggunakan 1 *hidden layer* (30 units), *dropout* sebesar 0.2, *batch size* sebesar 16, *epoch* sebesar 100, dan *learning rate* sebesar 0.01.
 - b. Metode *artificial neural network* pada kasus sembuh menggunakan 1 *hidden layer* (30 units), *dropout* sebesar 0.2, *batch size* sebesar 32, *epoch* sebesar 50, dan *learning rate* sebesar 0.01.
 - c. Metode *artificial neural network* pada kasus kematian menggunakan 3 *hidden layer* (30, 20, dan 13 units), *dropout* sebesar 0.2, *batch size* sebesar 32, *epoch* sebesar 100, dan *learning rate* sebesar 0.01.
2. Hasil prediksi menggunakan metode *artificial neural network* dengan *k-fold cross validation* pada kasus baru, kasus sembuh, dan kasus kematian menghasilkan nilai MAPE sebesar 1.78%, 1.92% dan 0.95%, sehingga dapat diartikan bahwa hasil tersebut memiliki kemampuan prediksi yang baik dengan nilai akurasi sebesar 98.22% untuk kasus baru, sebesar 98.08% untuk kasus sembuh, dan sebesar 99.05% untuk kasus kematian. Oleh karena itu, metode ini dapat digunakan untuk melakukan peramalan pada periode ke depan.
3. Peramalan untuk kasus baru, kasus sembuh, kasus kematian untuk periode 27 Oktober 2021 sampai dengan 24 Januari 2022 mengalami penurunan

DAFTAR PUSTAKA

- Aina, I. Q. 2018. Implementasi Artificial Neural Netwrok (ANN) dengan Algoritma Backpropagation Untuk Memprediksi Volume Penjualan di Bukalapak. Tugas Akhir. Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia, Yogyakarta.
- Berry, M. J. A. dan Linoff, G.S. 2004. *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management*. 2nd edition. John Wiley dan Sons, Indiana.
- Bronshtein, A. 2017. Train / Test Split and Cross Validation in Python. <https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6>. Diakses pada 12 Juli 2022.
- Dewi, N. F., Hindami, R. G., Sabila, S. K., Salsabila, T., Cortina, V. B., dan Kariyam. 2020. Perbandingan Metode Double Exponential Smoothing dan Artificial Neural Network Untuk Meramalkan Perkembangan Covid-19 di Indonesia, hlm. 312-318. *Seminar Nasional Matematika dan Pendidikan Matematika (5th SENATIK)*, Semarang.
- Gullo, F. 2015. From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. *Science Direct Journal*. **62**(3): 18-22.
- Halim, K. N. A., Jaya, A. S. M., dan Fadzil, A. F. A. 2020. Data Pre-Processing Algorithm for Neural Netwrok Binary Classification Model in Bank Tele-Marketing. *International Journal of Innovative Technology and Exploring Engineering*. **9**(3): 272-277.
- Han, J. dan Kamber, M. 2006. *Data Mining: Concept and Techniques*. 2nd edition. Morgan Kaufmann Publishers, San Francisco.

- Haykin, S. 2009. *Neural Networks and Learning Machines*. 3th edition. Prentice Hall, Canada.
- Hermawati, F.A. 2013. *Data Mining*. Andi, Yogyakarta.
- Heaton, J.T. 2008. *Introduction to Neural Network for Java*. 2nd Ed. Heaton Research Inc. Florida.
- Jumarwanto, A., Hartanto, R., dan Prastiyanto, D. 2009. Aplikasi Jaringan Saraf Tiruan Backpropagation Untuk Memprediksi Penyakit THT di Rumah Sakit Mardi Rahayu Kudus. *Jurnal Teknik Elektro*. **1**(1): 11-21.
- Kemenkes. 2021. *Positif Covid-19 Usai di Vaksinasi*. <https://www.kemkes.go.id/article/view/21022300001/positif-covid-19-usai-divaksinasi-begini-penjelasan-komnas-kipti-dan-kemenkes.html>. Diakses pada 12 Januari 2021.
- Larose, D.T. dan Larose, C.D. 2014. *Discovering Knowledge in Data – An Introduction to Data Mining*. 2nd edition. John Wiley dan Sons, New Jersey.
- Michelucci, U. 2018. *Applied Deep Learning : A Case - Based Approach to Understanding Deep Neural Network*. Apress Berkeley. California.
- Putra, J. W. G. 2020. *Pengenalan Konsep Pembelajaran Mesin dan Deep learning*. Ed. ke-1.4. Tokyo.
- Rahmalina, W. dan Puspita, S. 2021. Pemodelan Seasonal Autoregressive Integrated Moving Average Untuk Memprediksi Jumlah Kasus Covid-19 di Padang. *Jurnal Matematika Integratif*. **17**(1): 23-31.
- Ramadhona, G., Setiawan, B.D., dan Bachtiar, F.A. 2018. Prediksi Produktivitas Padi Menggunakan Jaringan Syaraf Tiruan Backpropagation. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. **2**(12): 6048-6057.
- Rochmawati, N., Hidayati, H.B., Yamasari, Y., Tjahyaningtjas, H.P.A., Yustanti,

- W., dan Prihanto, A. 2021. Analisa Learning Rate dan Batch size Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam. *Journal Information Engineering and Educational Technology*. **5**(2):44-48.
- Safrizal, Putra, D.I., Sofyan, S., dan Bimo. 2020. Pedoman Umum Menghadapi Pandemi Covid-19 Bagi Pemerintah Daerah.
- Santoni, M.M., Chamidah, N., dan Matondang, N. 2020. Prediksi Hipertensi Menggunakan *Decision Tree*, *Naïve Bayes*, dan *Artificial Neural Network* pada *Software KNIME*. *Jurnal Techno.COM*. **19**(4): 353-363.
- Santoso, J. F. N., Setaiawan, A., dan Rostianingsih, S. 2021. Perbandingan dan Analisis Metode *Artificial Neural Network* dan SIRD Pada Kasus Covid-19 di Surabaya. *Jurnal Infra*. **9**(1): 1-6.
- Sarkar, K. 2018. ReLU : Not a Differentiable Function: Why used in Gradient Based Optimization? and Other Generalizations of ReLU. <https://medium.com/@kanchansarkar/relu-not-a-differentiable-function-why-used-in-gradient-based-optimization-7fef3a4cecec>. Diakses pada 12 Juli 2022.
- Satya, R. dan Abraham, A. 2013. Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal of Advanced Research in Artificial Intelligence*. **2**(2): 34-38.
- Sitepu, A. C. dan Sigiro, M. 2021. Analisis Fungsi Aktivasi Relu dan Sigmoid Menggunakan Optimizer SGD dengan Representasi MSE Pada Model Backpropagation. *Jurnal Teknik Informatika Komputer Universal*. **1**(1): 12-25.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., dan Salakhutdinov, R. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. **15**: 1929-1958.
- Sulastrri dan Nugroho, Y. S. 2017. Penerapan Data Mining Untuk Prediksi Rating Penjualan Buku Menggunakan Metode Naive Bayes. *Jurnal Dutacom*. **12**(2): 57-72.
- Vijayalakshmi, V dan Venkatachalapathy, K. 2019. Deep Neural Network for

Multi-Class Prediction of Student Performance in Educational Data. *International Journal of Recent Technology and Engineering (IJRTE)*. **8(2)**: 5073-5081.

Virtyani, M.Z., Muljaningsih, S., dan Asmara, K. 2021. Studi Peristiwa Penetapan Covid-19 Sebagai Pandemi Oleh World Health Organization Terhadap Saham Sektor Healthcare di Bursa Efek Indonesia. *Jurnal Sekuritas*. **4(3)**: 240-252.