



Analisis Perbandingan Algoritma Supervised Learning untuk Prediksi Kasus Covid-19 di Jakarta

Angeline Septhiani¹, Hendry²

^{1,2}Universitas Kristen Satya Wacana, Indonseia

e-mail: ¹1672019026@student.uksw.edu, ²hendry@uksw.edu

Abstract

Coronavirus disease or called COVID-19 is a pandemic according to World Health Organization (WHO) in February. The virus gives several symptoms, such as cough, asthma, and fever. The data and information are the important part of making a good decision. Those data need to be processed and analyzed to be useful information. In this research, the data will be used to predict the COVID-19 issue in Jakarta, using several supervised learning algorithm models, such as K-Nearest Neighbors, Neural Network, Linear Regression, Support Vector Machine, and Random Forest. Using 10 Fold Cross Validation in model testing and T-Test to get the model with the best accuracy. According to this research, the algorithm that has the best accuracy is K-Nearest Neighbors with the lowest RMSE, 1096.188 +/- 365.077 (micro average: 1149.601 +/- 0.000).

Keywords: Covid-19, supervised learning, prediction, RapidMiner

Abstrak

Coronavirus disease atau yang biasa yang disebut dengan COVID-19 merupakan pandemi yang ditetapkan World Health Organization (WHO) pada bulan Februari. Virus tersebut dapat menimbulkan berbagai gejala seperti batuk kering, sesak nafas, dan demam. Data dan informasi merupakan bagian penting dalam pengambilan keputusan. Data tersebut perlu diolah dan dianalisis sehingga menjadi data yang bermanfaat. Dalam penelitian ini data tersebut akan digunakan untuk memprediksi kasus COVID-19 di Jakarta dengan melakukan perbandingan terhadap beberapa model algoritma supervised learning seperti K-Nearest Neighbors, Linear Regression, Random Forest, Support Vector Machine, dan Neural Network. Dalam pengujian model akan menggunakan dua tahap yaitu 10 Fold Cross Validation dan menggunakan T-Test untuk mengetahui tingkat akurasi yang paling baik. Berdasarkan penelitian ini, algoritma yang memiliki akurasi paling baik adalah K-Nearest Neighbors dengan hasil RMSE paling kecil 1096.188 +/- 365.077 (micro average: 1149.601 +/- 0.000).

Kata kunci: Covid-19, supervised learning, prediksi, RapidMiner

1. PENDAHULUAN

Pada akhir bulan desember tahun 2019, virus yang belum teridentifikasi sebelumnya muncul di Wuhan, Tiongkok, dan menyebar ke beberapa kota di China. Pada bulan Januari virus tersebut menyebar ke seluruh dunia termasuk Thailand, Korea, Jepang, Amerika Serikat, Singapura, Vietnam, Indonesia, dan negara-negara lainnya [1]. Masyarakat yang terinfeksi atau terpapar virus akan mengalami berbagai gejala mulai dari gejala ringan, hingga kondisi yang membutuhkan perawatan khusus seperti kegagalan respirasi akut. Gejala umum yang sering dialami antara lain batuk kering, sesak nafas, dan demam. Virus ini dapat menyebar melalui kontak fisik seperti batuk, bersin, atau berbicara, serta menyentuh permukaan yang terkontaminasi [2]. World Health Organization (WHO) secara resmi menamai virus tersebut sebagai coronavirus disease (COVID-19) pada tanggal 11 Februari 2020. Pada tanggal 2 Maret 2020, Indonesia mencatat kasus pertama dan kedua yang diawali dengan dua orang terkonfirmasi



positif COVID-19 [3]. Setelah adanya pengumuman mengenai hal tersebut, tidak dapat diperkirakan peningkatan kasus setiap harinya. Indonesia mengkonfirmasi kasus positif sebanyak 6.727.346 jiwa, 6.559.736 pasien sembuh dan 160.756 pasien meninggal, dimana provinsi DKI Jakarta memiliki total kasus yang lebih tinggi dibandingkan dengan provinsi lainnya [4].

Berdasarkan keadaan tersebut, diperlukan sebuah prediksi untuk mengetahui peningkatan jumlah kasus COVID-19 di Indonesia. Di era globalisasi yang semakin maju perkembangan teknologi menduduki peringkat pertama dalam membawa perubahan pada masyarakat. Kemajuan teknologi memberikan kemudahan pada berbagai aspek kehidupan manusia serta mempermudah mengakses informasi di seluruh dunia [5],[6]. Kemajuan teknologi yang paling berdampak dan bermanfaat bagi kehidupan masyarakat adalah bidang kecerdasan buatan [7]. Salah satu cabang ilmu kecerdasan buatan adalah *Machine Learning*. *Machine Learning* dapat belajar sendiri untuk membuat sebuah keputusan tanpa harus diprogram berulang kali oleh manusia. *Machine Learning* dapat mengidentifikasi pola tertentu saat data tersedia sebagai input untuk menganalisis kumpulan data. *Machine Learning* menggunakan dua teknik pembelajaran yaitu *supervised learning* (pembelajaran yang terawasi) dan *unsupervised learning* (pembelajaran tidak terawasi) untuk mengolah data yang besar[8]. Berdasarkan latar belakang yang telah dijelaskan sebelumnya, penelitian ini dilakukan untuk mengetahui akurasi yang paling sesuai untuk memprediksi jumlah kasus COVID-19 di Indonesia dengan menggunakan metode *supervised learning*.

2. METODOLOGI PENELITIAN

2.1. Tinjauan Pustaka

Penelitian ini dibuat menurut penelitian terdahulu yang masih memiliki keterkaitan di dalam pengkajiannya. Dalam jurnal penelitian dengan judul “Prediksi Jumlah Kasus COVID-19 Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) (Studi Kasus Kabupaten Sidoarjo)” membahas mengenai metode ARIMA yang digunakan untuk memperkirakan jumlah kasus COVID-19 di lokasi tersebut[9].

Penelitian terdahulu dengan judul “Prediksi Kesembuhan Pasien COVID-19 di Indonesia Melalui Terapi Menggunakan Metode Naive Bayes” membahas mengenai metode naïve bayes yang digunakan untuk memperkirakan kesembuhan pasien COVID-19 di Indonesia melalui perhitungan *data mining*. Akurasi yang diperoleh adalah akurasi keberhasilan 96,51% dengan presisi sukses 100%, kegagalan 95,71%, sensitivitas keberhasilan 84,21%, dan sensitivitas kegagalan 100% [10].

Penelitian yang berjudul “Prediksi Kasus Aktif COVID-19 menggunakan Metode K-Nearest Neighbors” membahas mengenai metode KNN yang akan digunakan untuk memperkirakan penyebaran dengan menggunakan 260 total data dengan membagi 80% menjadi data training dan 20% menjadi data testing [11].

Penelitian yang berjudul “Prediksi Kasus Aktif Kumulatif COVID-19 Indonesia Menggunakan Model Regresi Linier Berganda” membahas mengenai metode regresi linier berganda yang akan digunakan untuk memperkirakan kasus aktif kumulatif dengan variabel prediktor dan variabel respon [12].



Penelitian yang berjudul “Analisis dan Perbandingan Algoritma Prediksi dalam Mengetahui Perkiraan Peningkatan Jumlah Kasus COVID-19 di Indonesia Dengan Metodologi CRISP-DM” menggunakan metodologi Cross Industry Standard Process for Data Mining (CRISP-DM) dan enam algoritma yaitu jaringan syaraf tiruan, *Random Forest*, *Support Vector Machine*, *Linear Regression*, *Polynomial Regression* dan *Gaussian Process* yang akan dibandingkan untuk memperkirakan jumlah kasus COVID-19 [13].

2.1.1. Prediksi

Prediksi merupakan sebuah proses peramalan suatu variabel dimasa yang akan datang dengan mendasarkan pada pertimbangan daripada data yang berasal dari masa lampau [14]. Prediksi menurut Kamus Besar Bahasa Indonesia (KBBI) merupakan kegiatan untuk menduga sesuatu hal yang akan terjadi. Menurut Makridakis dalam jurnal Laras Purwati mengartikan bahwa prediksi merupakan bagian internal dari kegiatan untuk mengambil suatu *captain*. Menurut Buffa dalam jurnal Laras Purwati Ayuningtias mengartikan bahwa prediksi merupakan penggunaan teknik-teknik statistik berdasarkan pengolahan angka historis dalam bentuk gambaran masa depan. Menurut Imami dalam jurnal Laras Purwati Ayuningtias mengartikan bahwa prediksi merupakan suatu kegiatan untuk memperkirakan sesuatu yang akan terjadi di masa yang akan datang [15].

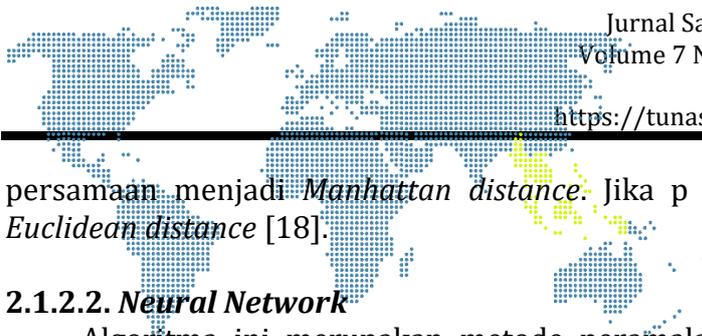
2.1.2. Supervised Learning

Supervised Learning merupakan pembelajaran yang terarah atau diawasi dengan mempelajari *data training* sebagai label. Dalam *supervised learning* pembelajaran terarah dapat berupa klasifikasi maupun regresi. Terdapat beberapa algoritma yang merupakan kategori *supervised learning* antara lain : [16]

2.1.2.1. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) diperkenalkan pertama kali pada tahun 1967 oleh T.Cover dan P.Hart dimana KNN merupakan salah satu algoritma *machine learning* yang paling sederhana. KNN mempelajari dan mengklasifikasi data tanpa membangun sebuah model serta memerlukan pelatihan data dalam mengklasifikasikan objek dengan mencari jarak terdekat. KNN juga harus mengingat semua data *training* dalam memori [17]. Proses pelatihan pada metode ini menggunakan dokumen yang telah dikategorikan. Selanjutnya dilakukan tahap *preprocessing* untuk mendapatkan bobot dari *training* yang dilakukan. Algoritma ini menghitung jarak dari setiap titik pada data uji dengan data latih setiap kelas, dan mengurutkan dari jarak terdekat ke jarak terjauh. Data yang terpilih adalah data yang memiliki jarak terdekat antara data uji dengan data latih sejumlah k. Jika lebih dari 1 tetangga yang diambil, maka kelas data uji akan ditentukan oleh kelas mayoritas. KNN secara *default* menggunakan *Minkowski distance*. *Minkowski distance* dapat ditulis dengan persamaan sebagai berikut :

$$D(x,y) = \sqrt[p]{\sum_{i=1}^n (X_i - Y_i)^p}$$
, dengan keterangan d merupakan jarak antara titik x dengan titik y, n merupakan jumlah fitur, dan jika $p = 1$ maka



persamaan menjadi *Manhattan distance*. Jika $p = 2$ maka persamaan menjadi *Euclidean distance* [18].

2.1.2.2. *Neural Network*

Algoritma ini merupakan metode peramalan dengan tingkat *error* cukup rendah serta dalam proses generalisasi cukup baik karena didukung oleh data *training* dan proses pembelajaran menyesuaikan bobot sehingga model ini dapat meramalkan data *time series* untuk beberapa waktu ke depan [19]. *Neural network* juga merupakan paradigma dalam mengolah informasi yang terinspirasi dari sistem saraf, dimana algoritma ini tersusun dari beberapa lapisan [20].

2.1.2.3. *Linear Regression*

Linear regression merupakan algoritma prediksi yang digunakan untuk mendapatkan model jenis keterkaitan antara variabel dependen (pemberi pengaruh) dengan satu atau banyak variabel independen (terpengaruh). Terdapat dua jenis *linear regression* antara lain *simple linear regression* dan *multiple linear regression* [21],[22]. *Simple linear regression* adalah suatu algoritma untuk memodelkan hubungan antara satu variabel terikat dan satu variabel bebas [23]. *Multiple linear regression* merupakan algoritma yang berguna untuk mengidentifikasi hubungan satu variabel terikat dengan dua atau lebih variabel bebas [12].

2.1.2.4. *Support Vector Machine (SVM)*

Pada tahun 1992 teori mengenai *support vector machine* (SVM) diperkenalkan oleh Vapnik, Boser, dan Guyon, sehingga berkembang dengan pesat. SVM merupakan salah satu metode *supervised learning* yang menghasilkan pemetaan dari fungsi input dan output dari sekumpulan data *training* yang didasarkan pada teori optimasi yang berasal dari teori pembelajaran statistik. Fungsi pemetaan tersebut dapat berupa fungsi regresi maupun fungsi klasifikasi [24].

2.1.2.5. *Random Forest*

Algoritma *Random forest* merupakan algoritma yang dapat meningkatkan akurasi dengan cara dilakukan pemilihan secara acak dan membangkitkan simpul anak untuk setiap node [25],[26]. *Random forest* terdiri dari root node, internal node, dan leaf node. Root node atau yang biasa disebut dengan akar dari pohon keputusan yang letaknya di paling atas. Internal node mempunyai minimal dua output dengan satu input. Leaf node terletak di bagian akhir yang hanya memiliki satu input dan tidak mempunyai output [26].

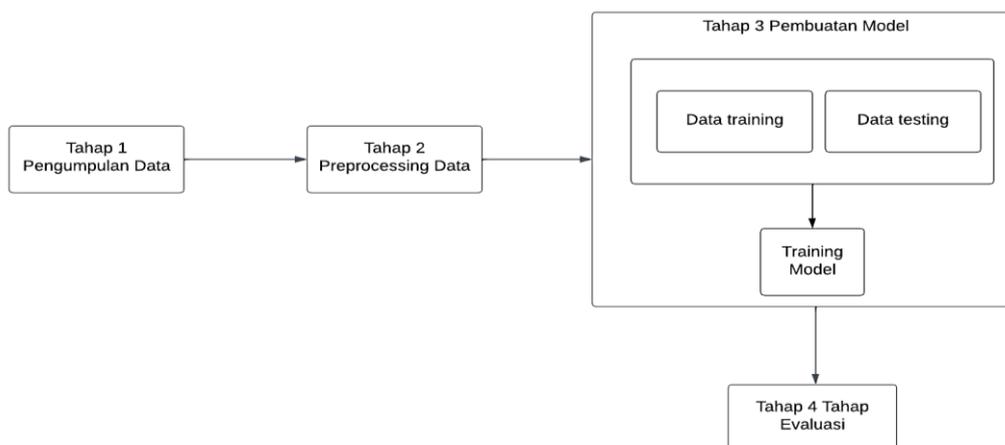
2.1.2.6. *RapidMiner*

RapidMiner merupakan *software* yang berguna untuk menganalisa data serta dapat mengintegrasikan berbagai bahasa pemrograman dengan mudah. Aplikasi ini mendukung beberapa format seperti .csv, .xls, dan format lainnya. RapidMiner ditulis dengan bahasa pemrograman Java serta bekerja dengan banyak sistem

operasi dapat bekerja dengan banyak sistem operasi karena ditulis dengan bahasa pemrograman Java [27]. RapidMiner dapat digunakan dengan mudah oleh pengguna karena menyediakan tampilan yang *user friendly*. Tampilan atau *perspective* dari aplikasi ini ada tiga antara lain *welcome perspective*, *design perspective*, dan *result perspective* [27].

2.2. Metode Penelitian

Dalam penelitian tersebut, peneliti melakukan beberapa langkah. Penelitian ini dilakukan untuk mengidentifikasi dan memecahkan permasalahan yang ada. Tahapan penelitian yang dilakukan adalah sebagai berikut :



Gambar 1. Alur Penelitian

1. Pengumpulan Data

Penelitian ini menggunakan data sudah dikumpulkan sebelumnya atau disebut dengan data sekunder yang didapatkan dari website <https://corona.jakarta.go.id/id>. Pada website tersebut terdapat data masyarakat yang sembuh, meninggal, dan dirawat dari tahun 2021 sampai tahun 2023. Data tersebut dikelola terlebih dahulu pada tahap *preprocessing data* sebelum digunakan untuk pembuatan model.

2. Preprocessing Data

Peneliti melakukan *cleaning* terhadap data yang sudah disiapkan. Proses ini dilakukan secara otomatis menggunakan RapidMiner. Setelah data siap, data tersebut akan digunakan dalam pembuatan model. Data tersebut dikelompokkan ke dalam 2 kategori, yaitu data training dan data testing.

3. Pembuatan Model

Dalam tahap ini, peneliti membuat model untuk setiap algoritma, yaitu *K-Nearest Neighbors*, *Linear Regression*, *Random Forest*, *Support Vector Machine*, dan *Neural Network*. Proses ini dilakukan dengan menggunakan RapidMiner. RapidMiner akan menampilkan informasi mengenai model yang sudah dibuat tersebut, seperti tingkat akurasi, tingkat error dan beberapa informasi lainnya yang diperlukan untuk melakukan tahap selanjutnya.

4. Tahap Evaluasi

Pada tahap akhir ini, peneliti mengevaluasi model-model yang sudah dibuat sebelumnya menggunakan informasi yang diberikan oleh RapidMiner. Setelah itu, peneliti mengambil satu model yang menghasilkan prediksi yang paling baik.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Tahap pertama yang dilakukan peneliti yaitu mengumpulkan data. Gambar 2 merupakan *dataset* yang akan digunakan.

Tanggal	Meninggal	Sembuh	Self Isolation	Masih Pera...	Total Pasien	Positif Harian	Positif Aktif	Sembuh Har...
Mar 1, 2020	0	0	0	0	0	0	0	0
Mar 2, 2020	0	0	0	0	0	0	0	0
Mar 3, 2020	1	0	0	2	3	3	2	0
Mar 4, 2020	1	0	0	2	3	0	2	0
Mar 5, 2020	3	0	0	4	7	4	4	0
Mar 6, 2020	3	0	0	4	7	0	4	0
Mar 7, 2020	3	0	0	4	7	0	4	0
Mar 8, 2020	3	0	0	4	7	0	4	0
Mar 9, 2020	3	0	0	31	34	27	31	0
Mar 10, 2020	3	0	0	31	34	0	31	0
Mar 11, 2020	3	0	0	33	36	2	33	0
Mar 12, 2020	5	0	0	57	62	26	57	0
Mar 13, 2020	7	0	0	65	72	10	65	0
Mar 14, 2020	9	0	0	70	79	7	70	0

Gambar 2. Dataset Covid-19

3.2. Preprocessing Data

Tahap selanjutnya adalah preprocessing data. Pada tahap ini data yang telah dikumpulkan sebelumnya akan dibersihkan sehingga tidak terdapat *missing value* pada data yang digunakan.

Name	Type	Missing	...	Filter (2 / 2 attributes):	Search for Attributes	
▼ Tanggal	Date	0	Earliest date	Mar 1, 2020	Latest date	Feb 4, 2023
▼ Positif Aktif	Integer	0	Min	0	Max	113138

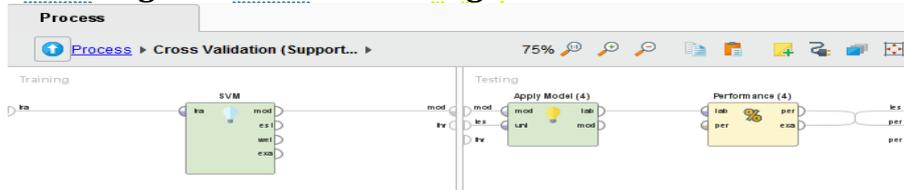
Gambar 3. Hasil Preprocessing Data

3.3. Modeling

Pada tahap pemodelan penulis membuat model untuk setiap algoritma yaitu *K-Nearest Neighbors*, *Linear Regression*, *Random Forest*, *Support Vector Machine*, dan *Neural Network*. Setiap model algoritma yang sudah terbentuk akan dilakukan perbandingan pada tahap evaluasi.

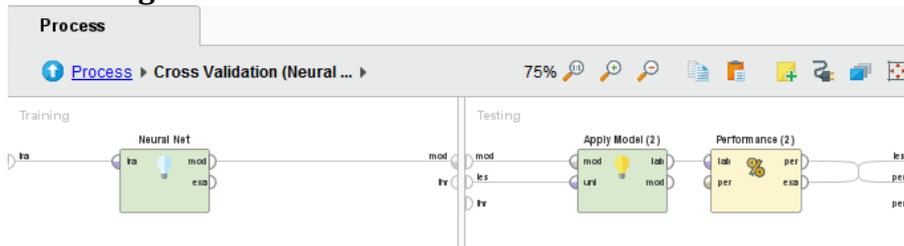


3.3.1. Model Algoritma K-Nearest Neighbors



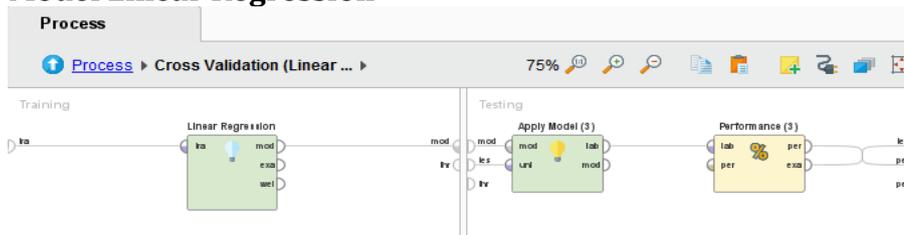
Gambar 4. Model K-Nearest Neighbors

3.3.2. Model Algoritma Neural Network



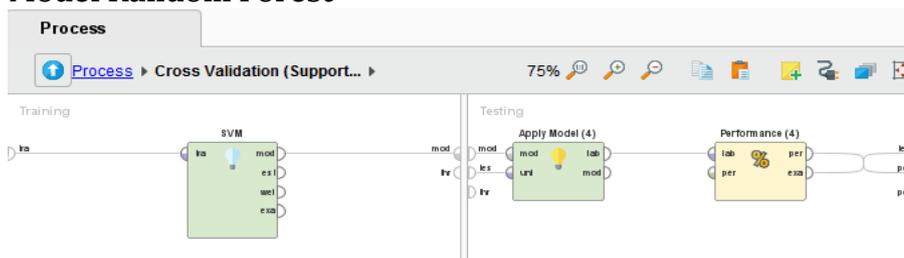
Gambar 5. Model Neural Network

3.3.3. Model Linear Regression



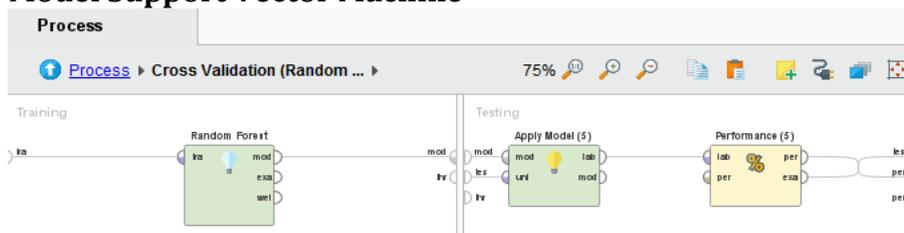
Gambar 6. Model Linear Regression

3.3.4. Model Random Forest



Gambar 7. Model Random Forest

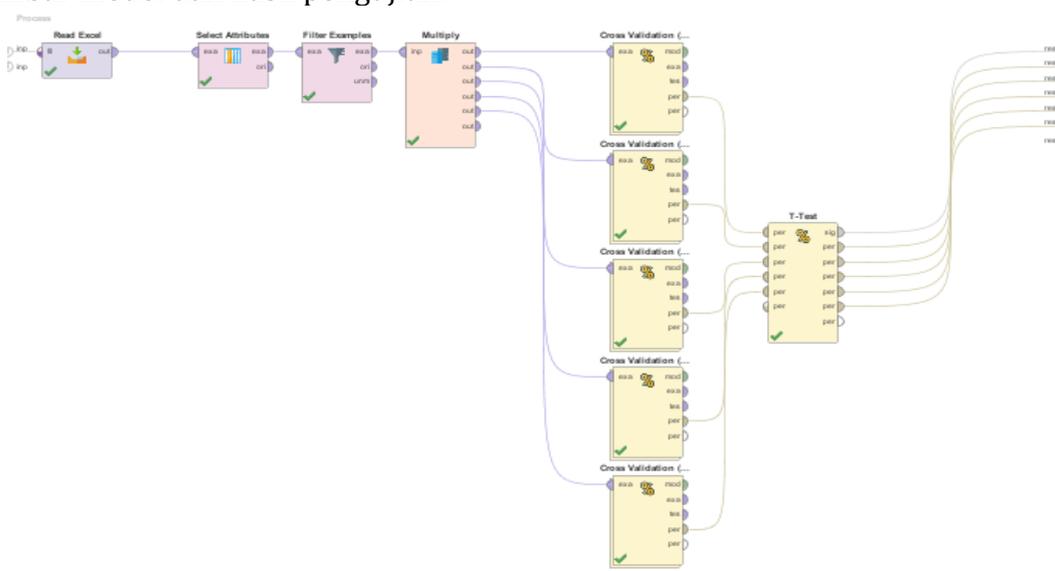
3.3.5. Model Support Vector Machine



Gambar 8. Model Support Vector Machine

3.4. Tahap Evaluasi

Pada tahap evaluasi akan dilakukan perbandingan hasil pengujian algoritma *K-Nearest Neighbors*, *Linear Regression*, *Random Forest*, *Support Vector Machine*, dan *Neural Network* akan menggunakan *cross validation 10-fold cross validation* dan akan dilanjutkan menggunakan T-test dengan $\alpha = 0,05$. Berikut merupakan gambar model dan hasil pengujian.



Gambar 9. Desain Model Pengujian

A	B	C	D	E	F
	1096.188 +/- 365.077	16801.456 +/- 3136....	17459.014 +/- 2816....	17436.848 +/- 3222....	1338.730 +/- 344.710
1096.188 +/- 365.077		0.000	0.000	0.000	0.144
16801.456 +/- 3136....			0.628	0.660	0.000
17459.014 +/- 2816....				0.987	0.000
17436.848 +/- 3222....					0.000
1338.730 +/- 344.710					

Gambar 10. Hasil Pengujian Menggunakan T-Test

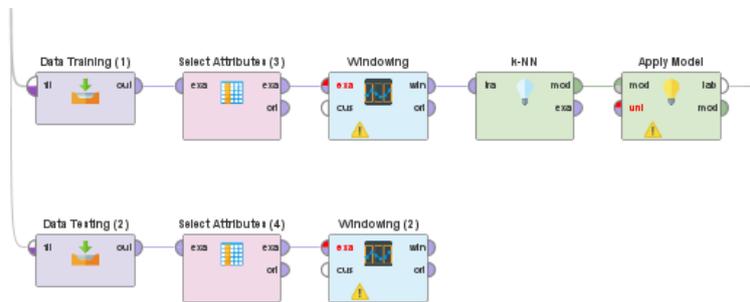
Berdasarkan Gambar 10, algoritma yang paling dominan dibandingkan algoritma lain adalah algoritma *K-Nearest Neighbors* merupakan algoritma yang. Hal tersebut dikarenakan algoritma *K-Nearest Neighbors* memiliki nilai yang lebih kecil dari $\alpha = 0.050$ dibandingkan algoritma lainnya.

Tabel 1. Hasil Pengujian Menggunakan Cross Validation

Algoritma	RMSE
K-Nearest Neighbors	1096.188 +/- 365.077 (micro average: 1149.601 +/- 0.000)
Neural Network	16801.456 +/- 3136.579 (micro average: 17062.920 +/- 0.000)
Linear Regression	17459.014 +/- 2816.132 (micro average: 17662.240 +/- 0.000)

Algoritma	RMSE
Support Vector Machine	17436.848 +/- 3222.548 (micro average: 17702.825 +/- 0.000)
Random Forest	1338.730 +/- 344.710 (micro average: 1378.093 +/- 0.000)

Dari hasil pengujian setiap algoritma, algoritma yang memiliki RMSE paling kecil atau mendekati nol (0) adalah algoritma *K-Nearest Neighbors* dengan hasil akurasi sebesar 1096.188 +/- 365.077 (micro average: 1149.601 +/- 0.000), sehingga algoritma *K-Nearest Neighbors* akan digunakan pada penelitian ini. Berikut merupakan model dan hasil prediksi positif covid-19.

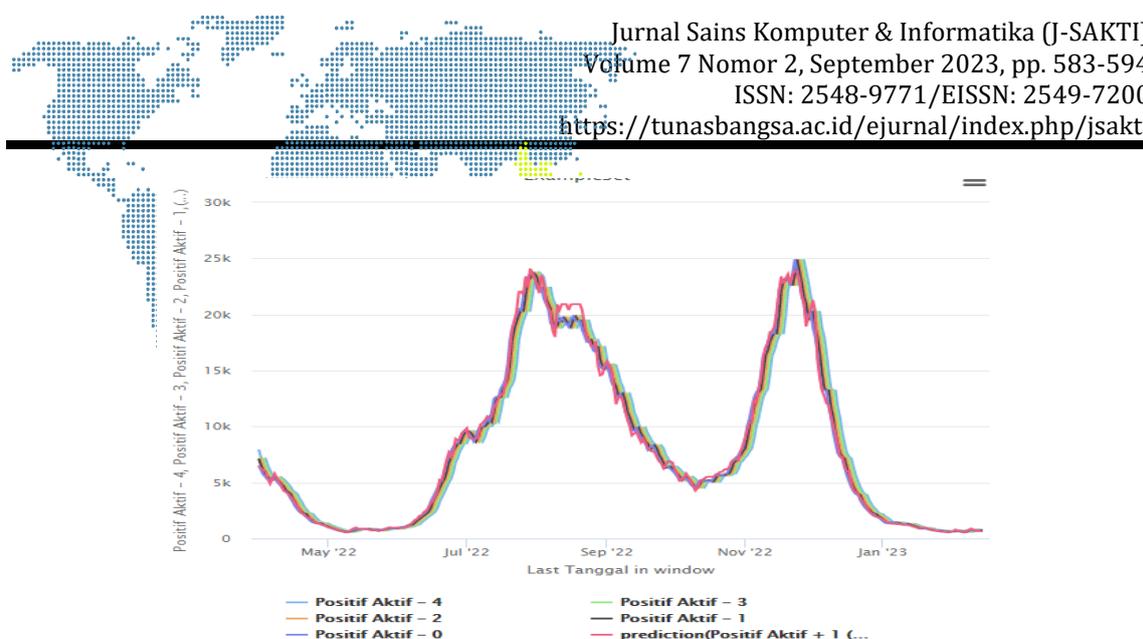


Gambar 11. Desain Model Prediksi

Row No.	Last Tanggal...	prediction(Po...	Positif Aktif - 4	Positif Aktif - 3	Positif Aktif - 2	Positif Aktif - 1	Positif Aktif - 0
1	Mar 31, 2022 ...	6457.848	7948	7217	7057	7120	6529
2	Apr 1, 2022 1...	6269.738	7217	7057	7120	6529	6094
3	Apr 2, 2022 1...	6089.115	7057	7120	6529	6094	5706
4	Apr 3, 2022 1...	6029.561	7120	6529	6094	5706	5502
5	Apr 4, 2022 1...	5312.843	6529	6094	5706	5502	5246
6	Apr 5, 2022 1...	4888.783	6094	5706	5502	5246	5238
7	Apr 6, 2022 1...	5326.405	5706	5502	5246	5238	5455
8	Apr 7, 2022 1...	5795.875	5502	5246	5238	5455	5297
9	Apr 8, 2022 1...	5384.768	5246	5238	5455	5297	5103
10	Apr 9, 2022 1...	5046.723	5238	5455	5297	5103	4871
11	Apr 10, 2022 ...	4726.401	5455	5297	5103	4871	4642
12	Apr 11, 2022 ...	4272.881	5297	5103	4871	4642	4297
13	Apr 12, 2022 ...	4168.379	5103	4871	4642	4297	4150
14	Apr 13, 2022 ...	3986.056	4871	4642	4297	4150	4003

ExampleSet (321 examples, 2 special attributes, 5 regular attributes)

Gambar 12. Hasil Prediksi Positif Covid-19



Gambar 13. Grafik Hasil Prediksi Positif Covid-19

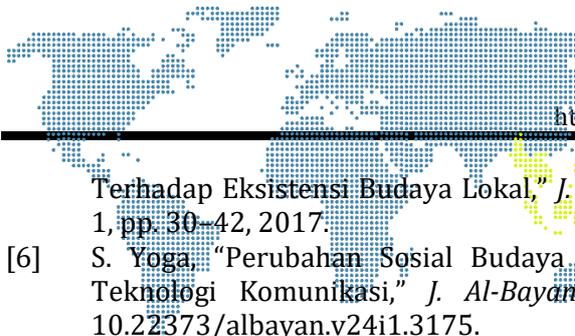
Hasil prediksi menunjukkan bahwa jumlah positif covid-19 di Jakarta cenderung menurun.

4. SIMPULAN

Berdasarkan hasil perbandingan antara algoritma *K-Nearest Neighbors*, *Linear Regression*, *Random Forest*, *Support Vector Machine*, dan *Neural Network* untuk melakukan prediksi terhadap jumlah positif covid-19 di Jakarta, algoritma yang paling cocok digunakan adalah *K-Nearest Neighbors* karena memiliki tingkat kesalahan atau RMSE yang kecil atau mendekati nol (0) dibandingkan dengan algoritma yang lainnya. Hasil RMSE *K-Nearest Neighbors* sebesar 1096.188 +/- 365.077 (micro average: 1149.601 +/- 0.000), RMSE *Neural Network* sebesar 16801.456 +/- 3136.579 (micro average: 17062.920 +/- 0.000), model *Linear Regression* sebesar 17459.014 +/- 2816.132 (micro average: 17662.240 +/- 0.000), model *Support Vector Machine* sebesar 17436.848 +/- 3222.548 (micro average: 17702.825 +/- 0.000), sedangkan model *Support Vector Machine* sebesar 1338.730 +/- 344.710 (micro average: 1378.093 +/- 0.000). Selain itu hasil T-test menunjukkan bahwa *K-Nearest Neighbors* lebih dominan dibandingkan algoritma lainnya.

DAFTAR PUSTAKA

- [1] Y. C. Wu, C. S. Chen, and Y. J. Chan, "The outbreak of COVID-19: An overview," *J. Chinese Med. Assoc.*, vol. 83, no. 3, pp. 217–220, 2020, doi: 10.1097/JCMA.0000000000000270.
- [2] A. Aditia, "Covid-19: Epidemiologi, Virologi, Penularan, Gejala Klinis, Diagnosa, Tatalaksana, Faktor Risiko Dan Pencegahan," *J. Penelit. Perawat Prof.*, vol. 3, no. November, pp. 653–660, 2021, [Online]. Available: <http://jurnal.globalhealthsciencegroup.com/index.php/JPPP%0ACOVID-19>
- [3] A. I. Almuttaqi, "Kekacauan respon terhadap covid-19 di indonesia," *Habibie Cent. Insights*, vol. 1, no. 13, pp. 1–7, 2020.
- [4] antaranews.com, "Data Perkembangan COVID-19," 2023.
- [5] R. D. Nasution, "Pengaruh Perkembangan Teknologi Informasi Komunikasi



- Terhadap Eksistensi Budaya Lokal," *J. Penelit. Komun. dan Opini Publik*, vol. 21, no. 1, pp. 30–42, 2017.
- [6] S. Yoga, "Perubahan Sosial Budaya Masyarakat Indonesia Dan Perkembangan Teknologi Komunikasi," *J. Al-Bayan*, vol. 24, no. 1, pp. 29–46, 2019, doi: 10.22373/albayan.v24i1.3175.
- [7] Y. Devianto and S. Dwiasnati, "Kerangka Kerja Sistem Kecerdasan Buatan dalam Meningkatkan Kompetensi Sumber Daya Manusia Indonesia," *J. Telekomun. dan Komput.*, vol. 10, no. 1, p. 19, 2020, doi: 10.22441/incomtech.v10i1.7460.
- [8] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.
- [9] L. Ainiyah and M. Bansori, "Prediksi Jumlah Kasus COVID-19 Menggunakan Metode Autoregressive Integrated Moving Average," *J. Sains Dasar*, vol. 10, no. 2, pp. 62–68, 2021.
- [10] P. T. A. Barus Okky, "Prediksi kesembuhan pasien COVID-19 di Indonesia melalui terapi menggunakan metode Naïve Bayes," *J. Inf. Syst. Dev.*, vol. 6, no. 2, pp. 59–66, 2021, [Online]. Available: <https://ejournal.medan.uph.edu/index.php/isd/article/view/460>
- [11] M. Lestandy and L. Syafa'ah, "Prediksi Kasus Aktif Covid-19 Menggunakan Metode K-Nearest Neighbors," *Semin. Nas. Teknol. dan Rekayasa 2020*, pp. 45–48, 2020.
- [12] Elen Riswana Safila Putri, Fahriza Novianti, Yasirah Rezaqita Aisyah Yasmin, and Dian Candra Rini Novitasari, "Prediksi Kasus Aktif Kumulatif Covid-19 Di Indonesia Menggunakan Model Regresi Linier Berganda," *Transform. J. Pendidik. Mat. dan Mat.*, vol. 5, no. 2, pp. 567–577, 2021, doi: 10.36526/tr.v5i2.1231.
- [13] M. Crisp-dm, "Analisis Dan Perbandingan Algoritma Prediksi Dalam Mengetahui Perkiraan Peningkatan Jumlah Kasus Covid-19 Di," no. June, 2021.
- [14] M. Metode, E. Smoothing, and U. K. Malang, "Sistem Informasi Prediksi Jumlah Kebutuhan Bahan Produksi Pada PT Agaricus Sido Makmur Sentosa Menggunakan Metode Exponential Smoothing," vol. 67, [Online]. Available: <https://media.neliti.com/media/publications/184428-ID-sistem-pendukung-keputusan-prediksi-pema.pdf>
- [15] Supono dan Putratama, "Penerapan Algoritma Naïve Bayes Pada Prediksi Kelulusan Mahasiswa (Studi Kasus Institut Medika Drg. Suherman)," vol. 24, no. 2, p. 3, 2016.
- [16] I. D. Id, *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*, 1st ed. UR PRESS, 2021.
- [17] M. S. Fajri, N. Septian, and E. Sanjaya, "Evaluasi Implementasi Algoritma Machine Learning K-Nearest Neighbors (kNN) pada Data Spektroskopi Gamma Resolusi Rendah," *Al-Fiziya J. Mater. Sci. Geophys. Instrum. Theor. Phys.*, vol. 3, no. 1, pp. 9–14, 2020, doi: 10.15408/fiziya.v3i1.16180.
- [18] E. R. A. Revolusi, E. A. Sari, and U. N. Malang, "Peran Pustakawan AI (Artificial Intelligent) Sebagai Strategi Promosi Perpustakaan Perguruan Tinggi Di Era Revolusi 4.0," vol. 3, pp. 64–73, 2019.
- [19] D. Agus Perdana Windarto, Darmeli Nasution, Anjar Wanto, *Jaringan Saraf Tiruan Algoritma Prediksi & Implementasi*. Yayasan Kita Menulis, 2020.
- [20] G. S. Mustika, U. Budiyanto, and S. Subandi, "Prediksi Jumlah Pasien Sembuh Covid-19 Menggunakan Jaringan Syaraf Tiruan," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 18, no. 2, p. 63, 2021, doi: 10.36080/bit.v18i2.1667.
- [21] Lisnawati et al., "Algoritma Linear Regression dalam Memprediksi Pertumbuhan Jumlah Penduduk," *J. Ilm. Wahana Pendidik.*

- <https://jurnal.unibrah.ac.id/index.php/JIWP>, vol. 8, no. 3, pp. 178–183, 2022, doi: 10.5281/zenodo.6408866.
- [22] W. H. Nugroho, Y. Widiastiwi, and I. N. Isnainiyah, “Sistem Prediksi Jumlah Kasus Covid-19 di Jakarta Menggunakan Metode Linear Regression,” vol. 19, pp. 154–164, 2022.
- [23] Graphpad.com, “Kalkulasi Regresi Linier,” 2023. <https://www.graphpad.com/quickcalcs/linear1/>
- [24] Obey Al Farobi, “Implementasi Metode Support Vector Machine (Svm) Untuk Mengetahui Respon Masyarakat Indonesia Terhadap Implementasi Metode Support Vector Machine (Svm) Untuk Mengetahui Respon Masyarakat Indonesia,” p. 82, 2021.
- [25] W. Apriliah, I. Kurniawan, M. Baydhowi, and T. Haryati, “Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest,” *Sistemasi*, vol. 10, no. 1, p. 163, 2021, doi: 10.32520/stmsi.v10i1.1129.
- [26] G. A. Sandag, “Prediksi Rating Aplikasi App Store Menggunakan Algoritma Random Forest,” *Cogito Smart J.*, vol. 6, no. 2, pp. 167–178, 2020, doi: 10.31154/cogito.v6i2.270.167-178.
- [27] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, “Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 1, pp. 8–17, 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.