

Komparasi Algoritma K-Nearest Neighbor dengan Linear Regression untuk Memprediksi Total Kasus Positif Covid-19 di Pamekasan

Luluk Suhartini^{1*}, Moh. Wasil²

¹ Institut Sains dan Teknologi Annuqayah, Indonesia

Info Artikel	ABSTRAK
<u>Riwayat Artikel</u> Diterima: 23-05-2023 Disetujui: 18-06-2023 <u>Kata Kunci</u> COVID-19; Data Mining; kNN; Linear Regression; RapidMiner;	COVID-19 telah ditetapkan sebagai pandemic sejak 30 Januari 2020 oleh WHO. Hal ini mengacu pada tingkat penyebar-an virus yang sudah mewabah ke belahan dunia, tak terkecuali Indonesia. Jumlah kasus baru tersebar di 34 provinsi dan hampir di semua kabupaten di Indonesia, termasuk Kabupaten Pamekasan. Dengan jumlah kasus COVID-19 yang selalu berubah, maka diperlukan kajian untuk memprediksi jumlah kasus COVID-19 di masa yang akan datang. Pada penelitian ini dilakukan perbandingan algoritma k-Nearest Neighbor dengan Linear Regression dalam memprediksi total kasus COVID-19. Dataset yang digunakan adalah dataset pasien kasus positif COVID-19 di Kabupaten Pamekasan. Pada penelitian ini digunakan model validasi 10 Fold Cross Validation untuk mengukur performa algoritma kNN dan Linear Regression sedangkan model evaluasi menggunakan Root Mean Square Error (RMSE) dengan menggunakan tools RapidMiner. Berdasarkan hasil penelitian, algoritma kNN memperoleh nilai RMSE lebih kecil dari pada Linear Regression dan dapat diterapkan untuk memprediksi kasus Covid-19 di Pamekasan dengan nilai RMSE sebesar 0.057.

*lulukkhafi@gmail.com

1. PENDAHULUAN

Covid-19 atau Corona Virus Disease 2019 adalah penyakit baru yang muncul pada akhir tahun 2019[1]. Hal ini berawal dari kasus gejala pneumonia aneh yang dijumpai pada beberapa pasien di kota Wuhan Provinsi Hubei China, 31 Desember 2019 lalu. Di kemudian hari diketahui penyebab gangguan pernafasan ini adalah virus jenis SARS-CoV-2[2] yang dapat menyebabkan kematian[3]. COVID-19 ditetapkan sebagai pandemi sejak 30 Januari 2020[2]. Hal ini mengacu pada tingkat penyebaran virus yang sudah mewabah ke seluruh dunia tanpa terkendali[4]. Tingkat penyebaran yang signifikan ini sampai ke Indonesia setelah ditemukannya kasus pertama yang menginfeksi seorang perempuan asal Jakarta Selatan pasca terjadinya kontak langsung dengan WN Jepang di Malaysia yang terkonfirmasi positif COVID-19[5]. Jumlah kasus baru yang tersebar di 34 provinsi dan hampir di semua kabupaten di Indonesia terus diumumkan oleh Pemerintah Indonesia[5]. Tak terkecuali di Kabupaten Pamekasan, yang terus mengalami peningkatan dan sampai saat ini menunjukkan update data terakhir kasus positif COVID-19 yang terkonfirmasi mencapai 2614 orang, 2408 kasus sembuh dan 201 kasus meninggal dunia pada tanggal 07 Februari 2022[6]. Dengan jumlah kasus COVID-19 yang terus meningkat, maka diperlukan kajian lebih lanjut untuk memprediksi dan memperkirakan jumlah kasus positif COVID-19 di waktu mendatang [7]. Hal ini dimaksudkan agar dijadikan bahan kajian untuk mengambil keputusan dalam melakukan pencegahan dan mengurangi jumlah kasus positif COVID-19.

Untuk melakukan prediksi, dapat dilakukan proses Data Mining yaitu proses pencarian informasi dari kumpulan data yang berukuran besar dengan cara mencari pola atau informasi menarik yang ada pada database yang dapat diuraikan melalui temuan pengetahuan dengan penggunaan teknik tertentu [8]. Beberapa jenis prediksi yang menggunakan teknik data mining

diantaranya adalah Klasifikasi dan Regresi [9]. Khalid Alkhatib [10] dalam penelitiannya menyebutkan metode kNN merupakan suatu metode Machine Learning dimana banyak digunakan untuk melakukan prediksi dalam beberapa tahun terakhir. Dalam penelitian tersebut digunakan algoritma kNN untuk memprediksi harga saham, dan menyimpulkan bahwa hasil prediksi hampir sejajar dengan harga saham actual. Selanjutnya penelitian yang dilakukan oleh Atmaja [11] yang menerapkan algoritma kNN untuk memprediksi pengelompokan tingkat risiko penyebaran COVID-19 di Jawa Barat, menghasilkan angka 85% sebagai nilai Accuracynya dengan angka Deviasi Standar $\pm 19,95\%$, 60% untuk nilai Precisionnya, dan 60% untuk nilai Recall. Li Kuang [12] dalam penelitiannya menggunakan pembobotan Bayesian Network model kNN untuk memprediksi durasi kecelakaan lalu lintas.

Hasil eksperimen tersebut menunjukkan bahwa algoritma dapat meningkatkan akurasi prediksi. Ching-Hsue Cheng [13] juga menggunakan metode kNN untuk memprediksi kesulitan keuangan. Mirko Kück [14] menggunakan model K-Nearest Neighbor dalam memprediksi permintaan pelanggan untuk perencanaan produksi. Dalam penelitian tersebut disimpulkan bahwa terdapat kesesuaian model KNN untuk meramalkan time series data industry. Francesco Borghesan [15] dalam penelitiannya menggunakan kNN untuk melakukan peramalan gangguan pada mesin pengontrol pabrik dan menyimpulkan bahwa metode tersebut mampu memberikan prediksi yang baik dan secara komputasi menjadi lebih cepat. Sedangkan M, Lestandy [16] menyimpulkan hasil penelitiannya yang menerapkan metode kNN untuk memprediksi kasus aktif Covid-19 menghasilkan MSE sebesar 0,007 dengan nilai akurasi 72,3337%.

Pada penelitian lain dilakukan analisa dan perbandingan beberapa algoritma Data Mining untuk memprediksi peningkatan jumlah kasus COVID-19 di Indonesia. Dari beberapa algoritma tersebut, menunjukkan Linear Regression sebagai algoritma yang memperoleh nilai Root Mean Square Error (RMSE) terkecil dibanding algoritma yang lain, dengan nilai RMSE sebesar 1625.580 ± 414.224 (micro average: 1671.084 ± 0.000) [7]. Penelitian yang dilakukan oleh Zhi-yong Huang untuk memprediksi angka penyakit paru obstruktif kronik berdasarkan beberapa model Linear Regression berbobot lokal dengan pengelompokan K-means [17].

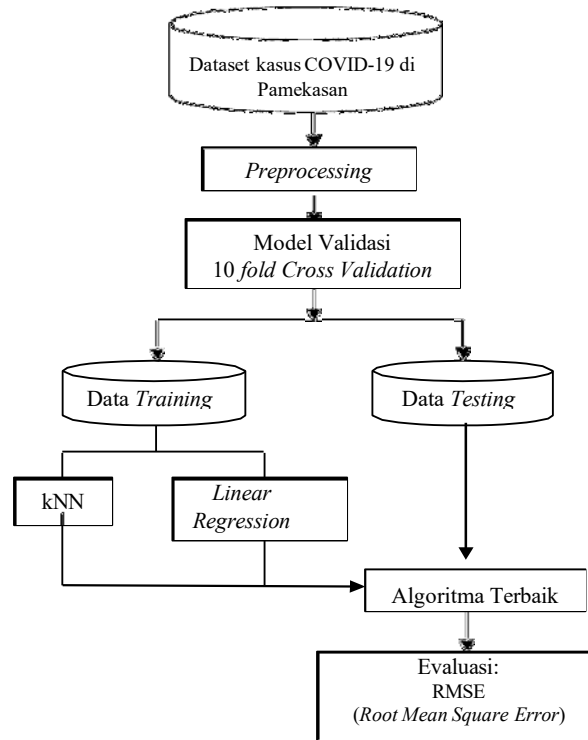
Chien-Hsun Huang [18] memprediksi biaya tenaga kerja BIM dengan Random Forest dan Regresi Linier sederhana. Penelitian ini mencoba mengembangkan metodologi baru dan mengusulkan pendekatan hibrida dengan menggabungkan Random Forest (RF) dan Simple Linear Regression (SLR) untuk meningkatkan akurasi prediksi pada proyek Biaya tenaga kerja BIM dalam tahap konstruksi. Studi kasus dilakukan untuk mendemonstrasikan dan memvalidasi hasil prediksi menggunakan sembilan belas proyek BIM yang diselesaikan dari perusahaan konstruksi terkemuka di Taiwan. Analisis pengelompokan juga diadopsi dan dikonfirmasi untuk meningkatkan kinerja model. Melalui perbandingan penelitian, pendekatan hibrida dari metodologi prediksi yang diusulkan dalam penelitian ini telah terbukti efektif dalam mengurangi risiko prediksi biaya tenaga kerja BIM. Dari penjabaran diatas, menunjukkan bahwa kedua metode tersebut sama-sama mempunyai kinerja yang baik dalam memprediksi. Pada penelitian ini, mengusulkan untuk membandingkan algoritma kNN dengan Linear Regression dalam memprediksi total kasus COVID-19 di Pamekasan

2. METODE

Langkah awal yang dilakukan sebelum penelitian ini adalah tahap pengumpulan data. Data yang dikumpulkan sesuai dengan topik penelitian. Dalam penelitian ini dataset yang digunakan adalah dataset kasus COVID-19 di Kabupaten Pamekasan. Dataset dibagi menjadi data training dan data testing dengan perbandingan 80% dan 20%. Hal ini mengacu pada hasil penelitian A. Tri et al [19] yang menyatakan nilai Mean Absolute Percentage Error (MAPE) minimum diperoleh pada proporsi tersebut.

Pada tahap preprocessing, dataset akan dinormalisasi secara manual menggunakan tool Microsoft Excel untuk meningkatkan waktu komputasi. Setelah data dinormalisasi data akan diubah dari data univariate menjadi data multivariate, dengan tujuan untuk mendapatkan

parameter dan model terbaik untuk prediksi [20]. kemudian data akan divalidasi dengan menggunakan model 10 Fold Cross validation yaitu suatu tindakan membagi data menjadi 10 bagian secara random. Sembilan bagian dijadikan data training, satu bagian menjadi data testing. Selanjutnya dilakukan pengujian kNN dan Linear Regression menggunakan tool RapidMiner. Tahap terakhir melakukan evaluasi dengan membandingkan RMSE dari kedua algoritma tersebut untuk kemudian dilakukan prediksi dengan model algoritma terbaik. Tahapan penelitian ditunjukkan pada gambar 1 berikut.



Gambar 1 Tahapan Penelitian

C. Linear Regression

Dalam penerapan Linear Regression dibutuhkan pengetahuan tentang kombinasi variable/objek yang tepat, target, serta jumlah permintaan terhadap variable ataupun objek yang sedang diteliti. Rumus untuk Linear Regression dengan metode kuadrat terkecil adalah [23]:

$$aa = (\sum yy)(\sum x^2) - (\sum x)(\sum yx) \quad n(\sum x^2) - (\sum x)^2 \quad (1)$$

$$bb = (\sum xy) - (\sum x)(\sum yy) \quad n(\sum x^2) - (\sum x)^2 \quad (2)$$

$$yy = aa + b. x1 + bb. xx^2 \dots \quad (3)$$

D. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) adalah metode untuk menghitung kesalahan prediksi untuk mengevaluasi data sebenarnya dan mengetahui hasil akurasi prediksi. Metode ini bekerja dengan cara mengetahui eror yang ditunjukkan untuk mengetahui perbedaan data antara hasil prediksi dengan data aktual [20]. RMSE adalah hasil akar dari nilai MSE sehingga menghasilkan akurasi data. Akurasi data akan bernilai baik jika nilai RMSE nya kecil. Berikut ini Rumus menghitung RMSE [24].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

Ket:

y_t = Nilai indeks actual

$\hat{y}_t$ = Nilai indeks prediksi

n = Jumlah sampel MS

3. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada Penelitian ini digunakan dataset univariate kasus positif COVID-19 harian di Kabupaten Pamekasan dari tanggal 01 Januari sampai tanggal 29 September 2021 yang terdiri dari 204 record data pada Tabel 1. Dataset dibagi menjadi data training dan data testing.

Tabel 1. Data kasus Covid-19 di Kabupaten Pamekasan

Tanggal	Total Kasus
01/09/2021	11
02/09/2021	15
03/09/2021	24
04/09/2021	27
05/09/2021	34
06/09/2021	39
07/09/2021	45
08/09/2021	59
...	...
29/09/2021	1807

B. Preprocessing Data

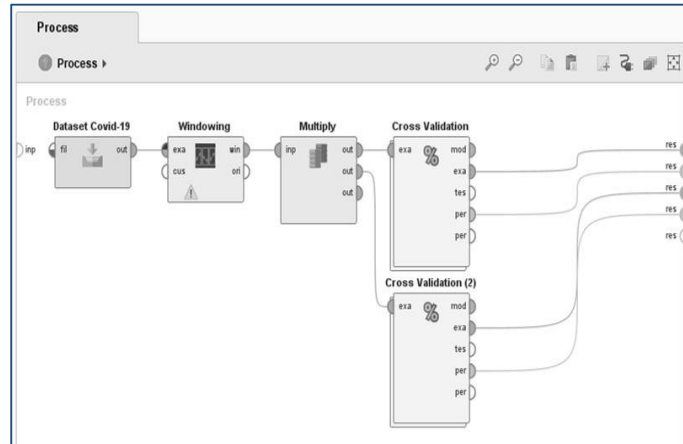
Sebelum diinput ke RapidMiner, data tersebut dinormalisasi menggunakan normalisasi min-max terlebih dahulu dengan tool Microsoft Excel untuk mendapatkan data dengan nilai yang lebih kecil. Hasil normalisasi ditunjukkan pada Tabel 2.

Tabel 2 Normalisasi data dengan min-max normalization

Tanggal	Total Kasus
01/09/2021	0
02/09/2021	0.00219539
04/09/2021	0.008781559
05/09/2021	0.012623491
06/09/2021	0.015367728
07/09/2021	0.018660812
08/09/2021	0.026344676
...	...
29/09/2021	1

C. Pengujian Model

Dataset yang sudah dinormalisasi diubah menjadi data multivariate, dengan tujuan untuk mendapatkan parameter dan model terbaik untuk prediksi. Pada penelitian ini data awal yang hanya terdiri dari 1 variabel akan diubah menjadi data multivariate dengan 10 variabel independent, yakni dari 1 periode sampai dengan 10 periode yang masing-masing mempunyai 1 variabel dependent menggunakan operator windowing, kemudian dilakukan validasi dengan teknik 10 fold cross validation dan pengujian model algoritma kNN dengan Linear Regression menggunakan tool RapidMiner. Proses pengujian algoritma ditunjukkan pada gambar 2 berikut.



Gambar 2 Proses pengujian algoritma

Dari proses pengujian model algoritma pada gambar 2, selanjutnya dilakukan perbandingan hasil pengujian algoritma terbaik dengan membandingkan nilai RMSE dari algoritma kNN dan Linear Regression.

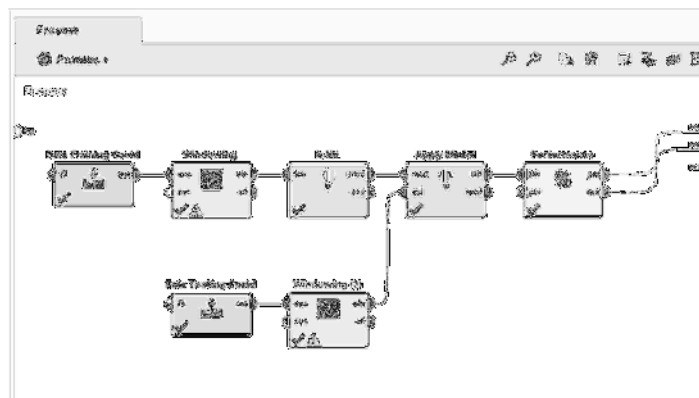
Tabel 3 Perbandingan RMSE algoritma kNN dengan Linear Regression

Algoritma	RMSE
kNN	0.007 +/- 0.005 (micro average 0.008 +/- 0.000)
Linear Regression	0.348 +/- 0.027 (micro average 0.349 +/- 0.000)

Pada Tabel 3 perbandingan algoritma kNN dengan Linear Regression, nilai RMSE terkecil ditunjukkan oleh algoritma kNN pada periode 4 dengan nilai RMSE : 0.007 +/- 0.005 (micro average: 0.008 +/- 0.000) sehingga algoritma kNN akan digunakan untuk melakukan prediksi total kasus COVID-19 di Kabupaten Pamekasan.

D. Prediksi dengan kNN

Gambar 3 adalah gambaran Skema model untuk melakukan prediksi dengan algoritma kNN.



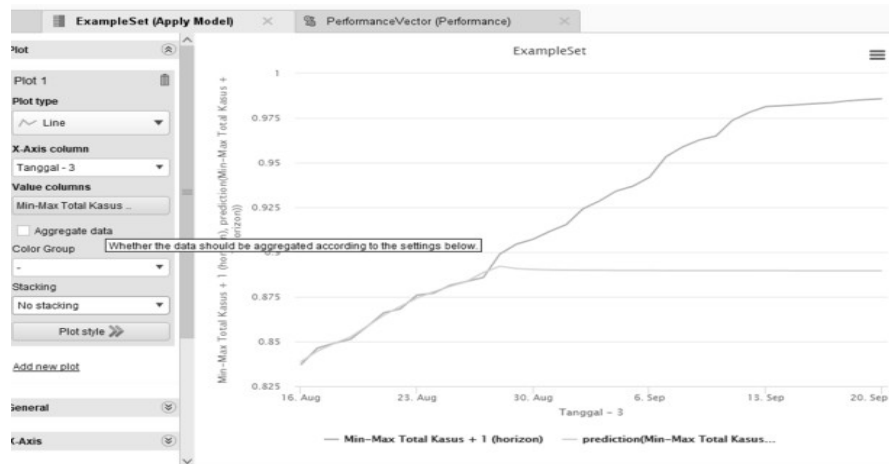
Gambar 3 Model prediksi total kasus Covid-19 dengan kNN

E. Evaluasi

Hasil prediksi kasus Covid-19 harian per tanggal 01 Januari sampai dengan 29 September 2021 ditunjukkan pada gambar 4.

User No.	Window Id	Min-Max Tot...	prediction(Min-Max Total Kasus + 1 (horizon))	Tanggal - 3	Tanggal - 2	Tanggal - 1	Tanggal - 8
1	0	0.937	0.939	Aug 16, 2021	Aug 17, 2021	Aug 18, 2021	Aug 19, 2021
2	1	0.946	0.948	Aug 17, 2021	Aug 18, 2021	Aug 19, 2021	Aug 20, 2021
3	2	0.949	0.949	Aug 18, 2021	Aug 19, 2021	Aug 20, 2021	Aug 21, 2021
4	3	0.951	0.952	Aug 19, 2021	Aug 20, 2021	Aug 21, 2021	Aug 22, 2021
5	4	0.956	0.959	Aug 20, 2021	Aug 21, 2021	Aug 22, 2021	Aug 23, 2021
6	5	0.958	0.958	Aug 21, 2021	Aug 22, 2021	Aug 23, 2021	Aug 24, 2021
7	6	0.959	0.970	Aug 22, 2021	Aug 23, 2021	Aug 24, 2021	Aug 25, 2021
8	7	0.976	0.974	Aug 23, 2021	Aug 24, 2021	Aug 25, 2021	Aug 26, 2021
9	8	0.977	0.979	Aug 24, 2021	Aug 25, 2021	Aug 26, 2021	Aug 27, 2021
10	9	0.981	0.981	Aug 25, 2021	Aug 26, 2021	Aug 27, 2021	Aug 28, 2021
11	10	0.984	0.984	Aug 26, 2021	Aug 27, 2021	Aug 28, 2021	Aug 29, 2021
12	11	0.995	0.999	Aug 27, 2021	Aug 28, 2021	Aug 29, 2021	Aug 30, 2021
13	12	0.999	0.992	Aug 28, 2021	Aug 29, 2021	Aug 30, 2021	Aug 31, 2021
14	13	0.995	0.991	Aug 29, 2021	Aug 30, 2021	Aug 31, 2021	Aug 31, 2021
15	14	0.997	0.990	Aug 30, 2021	Aug 31, 2021	Aug 31, 2021	Aug 31, 2021

Gambar 4 Hasil prediksi total kasus Covid-19.



Gambar 5 grafik perbandingan prediksi dan data aktual kasus covid-19.

Gambar 5 menunjukkan grafik perbandingan hasil prediksi dengan data actual. Berdasarkan grafik tersebut, pola data aktual dan prediksinya menunjukkan jarak yang hampir sama pada awal grafik dan jarak yang tidak terlalu jauh pada pertengahan sampai akhir grafik sehingga dapat dikatakan bahwa hasil prediksi cukup baik dengan RMSE 0.057 +/- 0.000 seperti yang ditunjukkan pada gambar 6



Gambar 6 RMSE Prediksi Kasus Covid-19 di Pamekasan

4. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian model, performa algoritma kNN lebih baik dibandingkan dengan algoritma Linear Regression. Hal ini mengacu pada evaluasi nilai RMSE antara kedua algoritma yang menunjukkan bahwa algoritma kNN memperoleh nilai RMSE lebih kecil dari pada Linear Regression. Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa Algoritma kNN dapat diterapkan untuk memprediksi kasus Covid-19 di Pamekasan dengan nilai RMSE sebesar 0.057.

Penelitian ini tentunya memerlukan pengembangan untuk bahan kajian riset. Pada penelitian berikutnya, dapat menggunakan dataset dalam jumlah yang lebih banyak dengan mempertimbangkan macam-macam metode dan teknik prediksi untuk mendapatkan hasil yang lebih baik. Selain itu penting pula untuk mempelajari macam-macam metode riset terkini terkait COVID-19 dan membandingkan dengan kondisi nyata di lapangan.

5. DAFTAR PUSTAKA

- [1] C. Long et al., “Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?,” *Eur. J. Radiol.*, vol. 126, no. February, p. 108961, 2020, <https://doi.org/10.1016/j.ejrad.2020.108961>.
- [2] F. Ye et al., “Delivery of infection from asymptomatic carriers of COVID-19 in a familial cluster,” *Int. J. Infect. Dis.*, vol. 94, pp. 133–138, 2020, doi: 10.1016/j.ijid.2020.03.042.
- [3] S. Sanche, Y. T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, and R. Ke, “RESEARCH High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2,” *Emerg. Infect. Dis.*, vol. 26, no. 7, pp. 1470–1477, 2020, <https://doi.org/10.3201/eid2607.200282>.
- [4] Kementerian Kesehatan Republik Indonesia, “Situasi Terkini Perkembangan (COVID-19) (21 Mei 2020),” *Kemenkes*, no. September, pp. 17–19, 2020, [Online]. Available: https://covid19.kemkes.go.id/download/Situasi_Ter_kini_050520.pdf.
- [5] D. Susanna, “When will the COVID-19 pandemic in indonesia end?,” *Kesmas*, vol. 15, no. 4, pp. 160–162, 2020, <https://doi.org/10.21109/KESMAS.V15I4.4361>.
- [6] Kompas.com, “Kasus Covid-19 di Pamekasan Madura Naik , Salah Satunya Ibu Hamil,” 2022. <https://jatim.liputan6.com/read/4879634/kasus-covid-19-di-pamekasan-madura-naik-salah-satunya-ibu-hamil>.
- [7] D. Wulan Sari, “Analisis Dan Perbandingan Algoritma Prediksi Dalam Mengetahui Perkiraan Peningkatan Jumlah Kasus Covid-19 Di,” no. June, 2021.
- [8] F. Salsabila and S. M. Intani, “Implementasi Algoritma K-Means Dan C4.5 Dalam Menentukan Tingkat Penyebaran Covid-19 Di Indonesia,” *Siliwangi*, vol. 7, no. 1, pp. 25–30, 2021.
- [9] S. B. Imandoust and M. Bolandraftar, “Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events : Theoretical Background,” *Int. J. Eng. Res. Appl.*, vol. 3, no. 5, pp. 605–610, 2013.
- [10] K. Alkhotib, “Stock Price Prediction Using K - Nearest Neighbor (k NN) Algorithm,” *Int. J. Business, Humanit. Technol.*, vol. 3, no. 3, pp. 32–44, 2013.
- [11] D. M. umi Atmaja, A. R. Hakim, D. Haryadi, and N. Suwaryo, “Penerapan Algoritma K-Nearest Neighbor Untuk Prediksi Pengelompokan Tingkat Risiko Penyebaran COVID-19 Jawa Barat,” *SNTEM*, vol. 1, no. November, pp. 1218–1226, 2021.
- [12] L. Kuang, H. Yan, Y. Zhu, S. Tu, and X. Fan, “Predicting duration of traffic accidents based on cost-sensitive Bayesian network and weighted K-nearest neighbor,” *J. Intell. Transp.*

Syst. Technol. Planning, Oper., vol. 23, no. 2, pp. 161–174, 2019, [https://doi: 10.1080/15472450.2018.1536978](https://doi.org/10.1080/15472450.2018.1536978).

[13] C. H. Cheng, C. P. Chan, and Y. J. Sheu, “A novel purity-based k nearest neighbors imputation method and its application in financial distress prediction,” *Eng. Appl. Artif. Intell.*, vol. 81, no. March, pp. 283–299, 2019, [https://doi: 10.1016/j.engappai.2019.03.003](https://doi.org/10.1016/j.engappai.2019.03.003).

[14] M. Kück and M. Freitag, “Forecasting of customer demands for production planning by local k-nearest neighbor models,” *Int. J. Prod. Econ.*, vol. 231, p. 107837, 2021, [https://doi: 10.1016/j.ijpe.2020.107837](https://doi.org/10.1016/j.ijpe.2020.107837).

[15] F. Borghesan, M. Chioua, and N. F. Thornhill, “Forecasting of process disturbances using k-nearest neighbours, with an application in process control,” *Comput. Chem. Eng.*, vol. 128, no. 675215, pp. 188–200, 2019, [https://doi: 10.1016/j.compchemeng.2019.05.009](https://doi.org/10.1016/j.compchemeng.2019.05.009).

[16] M. Lestandy and L. Syafa’ah, “Prediksi Kasus Aktif Covid-19 Menggunakan Metode K-Nearest Neighbors,” in *Seminar Nasional Teknologi dan Rekayasa (SENTRA) 2020*, 2020, pp. 45–48.

[17] Z. yong Huang et al., “Predicting the morbidity of chronic obstructive pulmonary disease based on multiple locally weighted linear regression model with K-means clustering,” *Int. J. Med. Inform.*, vol. 139, no. April, p. 104141, 2020, [https://doi: 10.1016/j.ijmedinf.2020.104141](https://doi.org/10.1016/j.ijmedinf.2020.104141).

[18] C. H. Huang and S. H. Hsieh, “Predicting BIM labor cost with random forest and simple linear regression,” *Autom. Constr.*, vol. 118, no. May, p. 103280, 2020, [https://doi: 10.1016/j.autcon.2020.103280](https://doi.org/10.1016/j.autcon.2020.103280).

[19] A. Tri et al., “Peramalan Jumlah Titik Panas Provinsi Kalimantan Timur Menggunakan Metode Radial Basis Function Neural Network,” *JAMBURAJ. Probab. Stat.*, vol. 1, no. November, 2020.

[20] L. Suhartini, M. B. Sulthan, I. Wahyudi, T. Informasi, and F. Teknik, “Optimasi K-Nearest Neighbor Dengan Particle Swarm,” vol. 2, no. 2, 2021.

[21] M. S. Bhullar and A. Kaur, “Use of data mining in education sector,” *Lect. Notes Eng. Comput. Sci.*, vol. 1, no. April, pp. 513–516, 2012.

[22] L. Zhang, Q. Liu, W. Yang, N. Wei, and D. Dong, “An Improved K -nearest Neighbor Model for Short-term Traffic Flow Prediction,” *Procedia - Soc. Behav. Sci.*, vol. 96, no. Cictp, pp. 653–662, 2013, [https://doi: 10.1016/j.sbspro.2013.08.076](https://doi.org/10.1016/j.sbspro.2013.08.076).

[23] T. Indarwati, T. Irawati, and E. Rimawati, “Penggunaan Metode Linear Regression Untuk Prediksi Penjualan Smartphone,” *J. Teknol. Inf. dan Komun.*, vol. 6, no. 2, pp. 2–7, 2019, [https://doi: 10.30646/tikomsin.v6i2.369](https://doi.org/10.30646/tikomsin.v6i2.369).

[24] M. E. Lasulika, “Prediksi Harga Komoditi Jagung Menggunakan K-Nn Dan Particle Swarm Optimazation,” vol. 9, pp. 233–238, 2017.dari keseluruhan proses, dan bukan sebagai dokumen independen.