**BULLETIN OF INFORMATION TECHNOLOGY (BIT)**
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT

# *Optimasi* Algoritma K- Nearest Neighbor Berbasis *Particle Swarm Optimization* Untuk Meningkatkan Kebutuhan Barang

**Taofik Safrudin, Gatot Tri Pranoto, Wahyu Hadikristanto**
Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia
Email: [1]taofik.99@mhs.pelitabangsa.ac.id. [2]gatot.pranoto@pelitabangsa.ac.id. [3]wahyu.hadikristanto@pelitabangsa.ac.id.
Email Penulis Korespondensi: taofiksafrudin4@gmail.com

**Abstrak**− Penerapan algoritma K-Nearest Neighbour dapat diimplementasikan dimana hasilnya juga menunjukkan sebuah wawasan baru, yaitu memprediksi tingkat kebutuhan. Dengan rasio 90%:10%, dimana disini ada 50 objek data yang diuji untuk memprediksi tingkat kebutuhan dalam 2 kelompok yaitu kebutuhan Rendah ataupun kebutuhan Tinggi. Hasil skenario model menunjukkan ada 2 objek masuk ke kelompok kebutuhan Rendah dan 1 objek masuk kedalam kelompok kebutuhan Tinggi. Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **82% ,** kemudian nilai **Precision** sebesar **87.50%**, dan nilai **Recall** sebesar **80%**. Dengan mengukur kinerja model dengan *Cross Validation* maka akurasi yang dihasilkan memiliki nilai standar atau simpangan baku, yang bertujuan untuk melihat jarak antara rata-rata akurasi dengan akurasi setiap percobaan. Sedangankan Hasil Pengujian menggunkan PSO Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **100% ,** kemudian nilai **Precision** sebesar **100%**, dan nilai **Recall** sebesar **100%** hasil pengujian mengalami peningkatkan secara segenefikan

**Kata Kunci :** Data Mining, PSO, *K-NN,Barang Pokok*

***Abstract***− *Abstract− The application of the K-Nearest Neighbor algorithm can be implemented where the results also show a new insight, namely predicting the level of need. With a ratio of 90%:10%, where there are 50 data objects tested to predict the level of needs in 2 groups, namely low needs or high needs. The results of the model scenario show that there are 2 objects in the Low needs group and 1 object in the High needs group. In evaluating this model, it was obtained from 10 fold Cross Validation that the Accuracy value was 82%, then the Precision value was 87.50%, and the Recall value was 80%. By measuring the performance of the model with Cross Validation, the resulting accuracy has a standard value or standard deviation, which aims to see the distance between the average accuracy and the accuracy of each experiment. While the Test Results using PSO In the evaluation of this model, it is obtained from 10 fold Cross Validation the Accuracy value is 100%, then the Precision value is 100%, and the Recall value is 100%, the test results have increased significantly*

***Keywords:*** *Data Mining, PSO, K-NN, Staple Goods*

## 1. PENDAHULUAN

Stock barang pada suatu gudang merupakan suatu elemen yang sangat pentin yang dapat memproses persediaan barang sangat diperlukan untuk menghindari kekurangan dalam produksi. Proses untuk mengatur ketersedian barang yang akan di proses da untuk memaksimalkan kebutuhan barang maka diperlukan pengiriman barang tertentu, perusahaan yang bergerak di bidang industri bahan pokok. Barang yang di proses atau perminggu dan perbulanya selalu bertambah barang tersebut sehingga perlu penempatan yang tepat dan pengelompokan barang sesuai kebutuhan.

Barang yang ada di dalam sebuah gudang merupakan stock kebutuhan barang pengelompokan barang pokok mengungkapkan adanya suatu pendataan barang kurang efektif dan barang over atau kelebihan barang sehingga perlunya ada pendataan kebutuhan barang berdasarkan jumlah data barang yang memudahkan dalam pemetaan barang. Data informasi yang didapat mengontrol hasil kebutuhan pokok setiap harinya dan selalu bermasalah dalam kegiatan stock yang dilakukan dan tidak teraturnya penempatan produk sehingga menyulitkan ketika akan mengambil produk satu dengan yang lainnya. Kebutuhan barang harus dipenuhi sehingga mengurangi kerugian yang ditimbulkannya dan tidak sampai berpengaruh secara signifikan.

Klasifikasi adalah proses menemukan model atau fungsi yang menggambarkan dan membedakan kelas atau konsep data. Namun terdapat beberapa masalah pada algoritma K-NN diantaranya adalah penentuan nilai k untuk pemilihan jumlah tetangga terdekatnya sangat sulit, karena nilai k sangat peka atau sensitif terhadap hasil klasifikasi. Pada penelitian ini, akan dilakukan pemodelan klasifiasi dengan menggunakan algoritma K-NN yang difokuskan pada proses penentuan nilai k terbaik pada dataset IKG (Indeks Kesulitas Geografis) desa. Pada penelitian ini akan melakukan integrasi algoritma K-NN dengan menentukan nilai k optimal dengan optimize parameters berdasar algoritma genetika[1].

Komoditi dan harganya karet mengalami perubahan yang fluktuatif dan menunjukkan pola yang tidak stasioner, di sisi lain pengambilan keputusan bisnis memerlukan data yang akurat dan terukur. Algoritma k-NN merupakan algoritma yang merupakan algoritma unsupervised, dan terbukti baik pada data mining. Sedangkan Particle Swarm Optimization (PSO) menunjukkan performa optimasi yang lebih baik dibandingkan dengan metode yang lain. Penelitian ini bertujuan untuk merancang metode prakiraan yang dapat memperkirakan tingkat harga dan volume permintaan untuk TSR 20. Hasil penelitian dari prediksi harga komoditi karet dengan menggunakan model k-NN mendapatkan nilai RMSE sebesar 0,087 sedangkan bila menggunakan k-NN yang dioptimasi dengan menggunakan Particle Swarm Optimization didapatkan nilai RMSE sebesar 0,082 lebih baik dibandingkan dengan hanya menggunakan k- NN saja[2].

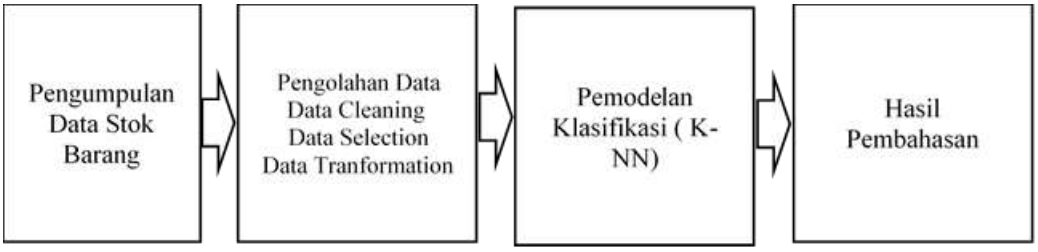
BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT

Data Mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut, untuk menentukan kebutuhan barang untuk mengatasi permasalahan yang terjadi diatas. Algoritma Nearest Neighbor merupakan salah satu metode yang digunakan untuk pemecahan masalah pada bidang Data Mining dan algoritma ini memiliki ciri yaitu dengan pendekatan untuk mencari kasus dengan menghitung kedekatan kasus yang baru dengan kasus yang lama.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Dalam melakukan penelitian seorang peneliti memiliki pedoman yang secara bertahap yang akan dilakukan sebagai berikut:




**Gambar 1.** Tahapan Penelitian

1. Pengumpulan Data

   Pada tahap ini penulis mengumpulkan data sampel yang berkaitan dengan subyek penelitian dan dilanjutkan dengan proses pencarian data stok barang, serta mengidentifikasi masalah kualitas datanya untuk mendapatkan bagian yang menarik dari data yang dapat digunakan dalam penelitian untuk mendapatkan informasi.

2. Pengolahan Data

   Tahap ini merupakan tahap pengelolaan data berkaitan dengan pencarian subyek penelitian, dengan melakukan pemilihan tabel, record, dan atribut-atribut data, termasuk juga proses cleaning, selection dan transformasi data yang kemudian akan dijadikan masukan pada tahap *data mining*.

3. Pemodelan

   Pemodelan pada penelitian ini dilakukan dengan data mining dan algoritma K-NN. Teknik ini dipilih karena merupakan metode yang umum dipakai pada penelitian data mining untuk mencari pendekatan terhadap kasus untuk di prediksi.

4. Hasil Dan Pembahasan

   Pada tahap ini merupakan evaluasi dan hasil penelitian yang telah dilakukan sehingga dapat membantu dalam pemecahan dari suatu data dalam pengambilan keputusan yang akan dijadikan informasi dan pengetahuan.

### 2.2 Pengujian Metode

Pemodelan pada penelitian ini dilakukan dengan data mining dan metode K Nearest Neighbor (KNN) untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Untuk pemilihan atribut terdiri dari n neighbors (biasa disebut k). parameter k pada testing ditentukan berdasarkan nilai k optimum pada saat training. Nilai k optimum diperoleh dengan mencoba-coba. Menghitung kuadrat jarak euclid (euclidean distance) masing-masing obyek terhadap data sampel yang diberikan.

$$\mathrm{D(a,b)} = \sqrt{\sum_{k=1}^{d} (a_k - b_k)^2}$$

Berdasarkan rumus 4 dimana matriks D(a,b) adalah jarak skalar dari kedua vektor a (data latih) dan b (data uji) darimatriks denganukurand dimensi. Mengurutkanhasilobjek-objek tersebut ke dalam kelompok yang mempunyai jarak euclidian terkecil. Mengumpulkan kategori y (klasifikasi nearest neighbor berdasarkan nilai k) Dengan menggunakan kategori nearest neighbor yang paling mayoritas, maka dapat diprekdisikan kategori objek tersebut.

BULLETIN OF INFORMATION TECHNOLOGY (BIT)
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT

# 3. HASIL DAN PEMBAHASAN

## 3.1 Implementasi Algoritma K-Nearest Neighbour

Perhitungan algoritma K-NN dari penulis menggunakan data uji *(testing)* sebanyak 12 record data yang didapat pada dataset sebelumnya. Selanjutnya dalam menerapkan algoritma K-NN pada data uji tersebut ditentukan dengan langkah-langkah seperti dibawah berikut

1. Menentukan jumlah tetangga yang akan diperhitungkan, dinotasikan dalam **(k)**, dalam hal ini penulis menentukan 3 tetangga terdekat atau dengan kata lain **(k = 3)**
2. Menghitung jarak dengan perhitungan *Euclidean Distance* dari masing-masing setiap tetangga terhadap objek target data uji, kemudian diurutkan dari yang terkecil sampai terbesar.
3. Menentukan objek target data uji masuk dalam suatu kelompok dengan melihat nilai mayoritas sebagai hasil prediksi dari data uji tersebut.

Dengan langkah yang ditetapkan tersebut, maka berikut adalah perhitungan dari masing – masing objek data uji yang terdapat 12 record data.

1. Kebtuhan Stok A **(x1 = 12, y1 = 5, z1= 16)**

Dengan menerapkan rumus perhitungan jarak *Euclidean Distance* berikut :

$$D = \sqrt{(x1 - x2)^2 + (y1 - y2)^2 + (z1 - z2)^2}$$

Keterangan :
D = Euclidean Distance
x1 = Angka kasus suspect dari objek
y1 = Angka kasus meninggal dari objek
x2 = Angka kasus suspect dari tiap tetangga objek (data training)
y2 = Angka kasus meninggal dari tiap tetangga objek (data training)

Didapatkan hasil dari perhitungan jarak *Euclidean Distance* antara objek stok kebtuhan akhir dengan tiap-tiap tetangga dapat dilihat dari Tabel 1 berikut.

**Tabel 1.** Euclidean Distance Terhadap Objek

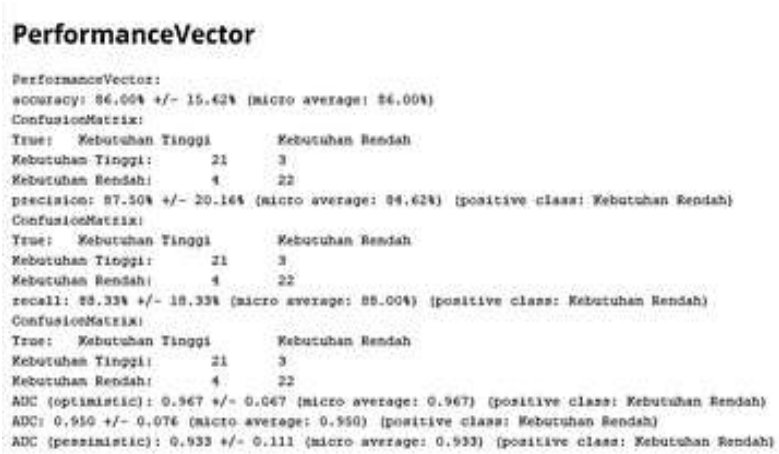
| 12 | 5 | 16 | Tinggi | | |
|---|---|---|---|---|---|
| Stok Awal | Sale | Stok Akhri | Tingkat Kebutuhan | Jarak Terhadap Kebutuhan Stok | K=3 |
| 2 | 7 | 3 | Kebutuahn Tinggi | 17 | |
| 10 | 20 | 3 | Kebutuhan Rendah | 20 | |
| 7 | 16 | 3 | Kebutuhan Rendah | 18 | |
| 11 | 5 | 4 | Kebutuhan Rendah | 12 | 3 |
| 4 | 1 | 3 | Kebutuhan Rendah | 16 | |
| 35 | 1 | 86 | Kebutuahn Tinggi | 74 | |
| 6 | 9 | 10 | Kebutuahn Tinggi | 9 | 1 |
| 3 | 8 | 4 | Kebutuahn Tinggi | 15 | |
| 21 | 24 | 11 | Kebutuhan Rendah | 22 | |
| 21 | 5 | 62 | Kebutuahn Tinggi | 47 | |
| 20 | 4 | 19 | Kebutuhan Rendah | 9 | 2 |
| 3 | 8 | 231 | Kebutuahn Tinggi | 215 | |

Berdasar urutan dari yang terkecil sampai dengan yang terbesar, dari 3 tetangga terdekat, dapat dilihat terdapat 3 tetangga dengan label Kebutuhan Rendah dan 1 tetangga dengan label Kebutuhan Tinggi, sehingga secara mayoritas dapat disimpulkan bahwa Objek A

## 3.2 Evaluasi Data Uji pada RapidMiner

Pengujian model dilakukan dengan maksud mengetahui hasil perhitungan yang dianalisa dan mengukur metode serta algoritma yang digunakan apakah berfungsi dengan baik atau tidak. Penerapan model dengan algoritma K-NN pada penelitian ini menggunakan metode K-Fold Cross Validation yang dapat dimanfaatkan dengan menggunakan *tools* Rapid Miner. Penulis menggunakan metode K-Fold Cross Validation karena fungsi ini tidak hanya membuat beberapa sampel data uji berulang kali, tetapi membagi dataset menjadi bagian terpisah dengan ukuran yang sama dan untuk peningkatkan dari algoritma menggunakan PSO. Model dilatih oleh subset data latih dan divalidasi oleh subset validasi (data uji) sebanyak k (itrerasi). Dengan K-Fold Cross Validation dapat mengurangi waktu komputasi dengan tetap menjaga klasifikasi model. ***Running Process*** pada *tools* Rapid Miner, didapatkan hasil *PerformanceVector* seperti pada Gambar di bawah ini.

**BULLETIN OF INFORMATION TECHNOLOGY (BIT)**
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT




**Gambar 2.** Hasil Evaulasi PerformanceVector K-NN


**Gambar 1**. Hasil Evaulasi PerformanceVector K-NN Model PSO

Tabel Confusion Matrix dari evaluasi model tersebut mendapatkan hasil *Accuracy*, *Precission* dan *Recall* dilihat seperti Tabel 2 berikut.
Perhitungan dari algoritma *K-NN*, akurasi dilakukan dengan cara jumlah TP + TN dibagi jumlah total data testing yang diuji.

**Tabel *2*** Accuracy K-NN

accuracy: 86.00% +/- 15.62% (micro average: 86.00%)

| | true Kebutuhan Tinggi | true Kebutuhan Rendah | class precision |
|---|---|---|---|
| pred. Kebutuhan Tinggi | 21 | 3 | 87.50% |
| pred. Kebutuhan Rendah | 4 | 22 | 84.62% |
| class recall | 84.00% | 88.00% | |

**Tabel 3** Accuracy K-NN Model PSO

accuracy: 100.00%

| | true Kebutuhan Tinggi | true Kebutuhan Rendah | class precision |
|---|---|---|---|
| pred. Kebutuhan Tinggi | 2 | 0 | 100.00% |
| pred. Kebutuhan Rendah | 0 | 3 | 100.00% |
| class recall | 100.00% | 100.00% | |

**BULLETIN OF INFORMATION TECHNOLOGY (BIT)**
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT

Nilai precision dihitung dengan cara membagi jumlah data benar yang bernilai positif (True Positive) dibagi dengan jumlah data benar yang bernilai positif (True Positive) dan data salah yang bernilai positif (False Positive).

**Tabel 4** Percision K-NN

precision: 87.50% +/- 20.16% (micro average: 84.62%) (positive class: Kebutuhan Rendah)

| | true Kebutuhan Tinggi | true Kebutuhan Rendah | class precision |
|---|---|---|---|
| pred. Kebutuhan Tinggi | 21 | 3 | 87.50% |
| pred. Kebutuhan Rendah | 4 | 22 | 84.62% |
| class recall | 84.00% | 88.00% | |

Nilai recall dihitung dengan cara membagi data benar yang bernilai positive (True Positive) dengan hasil penjumlahan dari data benar yang bernilai positif (True Positive) dan data salah yang bernilai negatif (False Negative).

**Tabel 5** Recall K-NN

recall: 88.33% +/- 16.33% (micro average: 88.00%) (positive class: Kebutuhan Rendah)

| | true Kebutuhan Tinggi | true Kebutuhan Rendah | class precision |
|---|---|---|---|
| pred. Kebutuhan Tinggi | 21 | 3 | 87.50% |
| pred. Kebutuhan Rendah | 4 | 22 | 84.62% |
| class recall | 84.00% | 88.00% | |

**Tabel 6** Recall K-NN Model PSO

recall: 100.00% (positive class: Kebutuhan Rendah)

| | true Kebutuhan Tinggi | true Kebutuhan Rendah | class precision |
|---|---|---|---|
| pred. Kebutuhan Tinggi | 2 | 0 | 100.00% |
| pred. Kebutuhan Rendah | 0 | 3 | 100.00% |
| class recall | 100.00% | 100.00% | |

### 3.2 Pembahasan

Dengan menggunakan skenario perbandingan 90:10 untuk data latih *(training)* dan data uji *(testing)* dari dataset yang dimiliki, maka tahapan model pada penerapan algoritma *K-Nearest Neighbour* menghasilkan prediksi terhadap objek – objek untuk mengelompokkan ke dalam label tingkat kebutuhan, kebutuhan Rendah maupun kebutuhan Tinggi. Dalam rasio yang digunakan pada skenario tersebut, sebanyak 50 record data dijadikan data uji *(testing)*, yang masing-masing objek ini kemudian diprediksi berdasarkan ketentuan yang sudah diterapkan pada algoritma K-NN yang digunakan dalam penelitian ini. Pembasahan mengenai pengujian dari penerapan model algoritma K-NN tersebut juga penulis lakukan menggunakan *tools* Rapid Miner. Pada *tools* tersebut penulis memanfaatkan evaluasi model menggunakan fungsi *Cross Validation* karena fungsi ini tidak hanya membuat beberapa sampel data uji berulang kali, tetapi membagi dataset menjadi bagian terpisah dengan ukuran yang sama. Adapun dengan menggunakan tahapan-tahapan proses evaluasi model dengan *Cross Validation* terhadap data latih *(training)* dan data uji *(testing)*, terbentuk *Confusion Matrix* yang dimana dapat menggambarkan nilai *Accuracy, Precision, dan Recall* dari model yang diterapkan yaitu dalam mengevaluasi penerapan algoritma K-NN pada dataset stok produk.
Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **86% ,** kemudian nilai **Precision** sebesar **87.50%**, dan nilai **Recall** sebesar **80.40%**. Dengan mengukur kinerja model dengan *Cross Validation* maka akurasi yang dihasilkan memiliki nilai standar atau simpangan baku, yang bertujuan untuk melihat jarak antara rata-rata akurasi dengan akurasi setiap percobaan. Sedangankan Hasil Pengujian menggunkan PSO Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **100% ,** kemudian nilai **Precision** sebesar **100%**, dan nilai **Recall** sebesar **100%** hasil pengujian mengalami peningkatkan secara segenefikan.

**BULLETIN OF INFORMATION TECHNOLOGY (BIT)**
Vol 4, No 2, September 2023, Hal 281 - 286
ISSN 2722-0524 (media online)
DOI 10.47065/ bit.v3i1. 724
https://journal.fkpt.org/index.php/BIT

# 4. KESIMPULAN

Pemanfaatan data stok untuk pendekatan metode klasifikasi dapat diterapkan dalam menganalisis data guna melakukan prediksi kebutuhan produk antara produk bahan pokok. Penerapan algoritma K-Nearest Neighbour dapat diimplementasikan dimana hasilnya juga menunjukkan sebuah wawasan baru, yaitu memprediksi tingkat kebutuhan. Dengan rasio 90%:10%, dimana disini ada 50 objek data yang diuji untuk memprediksi tingkat kebutuhan dalam 2 kelompok yaitu kebutuhan Rendah ataupun kebutuhan Tinggi. Hasil skenario model menunjukkan ada 2 objek masuk ke kelompok kebutuhan Rendah dan 1 objek masuk kedalam kelompok kebutuhan Tinggi. Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **86% ,** kemudian nilai **Precision** sebesar **87.50%**, dan nilai **Recall** sebesar **80.40%**. Dengan mengukur kinerja model dengan *Cross Validation* maka akurasi yang dihasilkan memiliki nilai standar atau simpangan baku, yang bertujuan untuk melihat jarak antara rata-rata akurasi dengan akurasi setiap percobaan. Sedangankan Hasil Pengujian menggunkan PSO Pada evaluasi model ini didapatkan dari 10 fold *Cross Validation* nilai ***Accuracy*** sebesar **100% ,** kemudian nilai **Precision** sebesar **100%**, dan nilai **Recall** sebesar **100%** hasil pengujian mengalami peningkatkan secara segenefikan

## REFERENCES

[1] J. S. Informasi And F. Teknik, "Optimasi Teknik Klasifikasi Modified K Nearest Neighbor Menggunakan Algoritma Genetika," No. September, Pp. 130–134, 2014.
[2] A. Ristekdikti *Et Al.*, "Optimasi Algoritma Svm Dan K-Nn Berbasis Particle Swarm Optimization Pada Analisis Sentimen Fenomena Tagar," Vol. Vi, No. 1, 2020, Doi: 10.31294/Jtk.V4i2.
[3] Y. Darmi And A. Setiawan, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," *J. Media Infotama Univ. Muhammadiyah Bengkulu*, Vol. 12, No. 2, Pp. 148–157, 2016.
[4] T. Informatika, "Pengelompokkan Loyalitas Pelanggan Dengan Menggunakan Kombinasi Rfm Dan Algoritma K-Means," Vol. 5, No. 1, Pp. 7–13, 2020.
[5] A. F. Watratan, A. P. B, D. Moeis, S. Informasi, And S. P. Makassar, "Journal Of Applied Computer Science And Technology ( Jacost ) Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," Vol. 1, No. 1, Pp. 7–14, 2020.
[6] D. N. Sulistyowati, N. Yunita, S. Fauziah, And R. L. Pratiwi, "Implementation Of Data Mining Algorithm For Predicting Popularity Of Playstore Games In The Pandemic Period Of," Vol. 6, No. 1, Pp. 95–100, 2020, Doi: 10.33480/Jitk.V6i1.1425.
[7] C. Medel-Ramírez, H. Medel-López, U. Veracruzana, I. De Investigaciones, And S. Económicos, "Data Article . Data Mining For The Study Of The Epidemic ( Sars- Cov-2 ) Covid-19 : Algorithm For The Authors Corresponding Author A Bstract Keywords," Pp. 1–8.
[8] M. Peta, P. Pasien, C.-D. Kombinasi, And M. U. Fahri, "Jurnal Teknologi Terpadu Journal Of Integrated Technology," Vol. 6, No. 1, Pp. 25–30, 2020.
[9] S. Hikmawan, A. Pardamean, And S. N. Khasanah, "Sentimen Analisis Publik Terhadap Joko Widodo Terhadap Wabah Covid-19 Menggunakan Metode Machine Learning," Vol. 20, No. 2, Pp. 167–176, 2020.
[10] N. Dwitri, J. A. Tampubolon, S. Prayoga, And P. P. P. A. N. W. F. I. R. H. Zer, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia," Vol. 4, No. 1, Pp. 128–132, 2020.
[11] Retno Tri Vulandari, *Data Mining*. Yogyakarta: Gava Media, 2017.
[12] Suyanto, *Data Mining*. Yogyakarta: Informatika, 2017.
[13] G. Widi N. Dicky Nofriansyah, *Algoritma Data Mining Dan Pengujian*. Yogyakarta: Cv Budi Utama, 2015.
[14] O. Villacampa, "(Weka - Thesis) Feature Selection And Classification Methods For Decision Making: A Comparative Analysis," *Proquest Diss. Theses*, No. 63, P. 188, 2015.