



PENERAPAN ALGORITMA K-MEDOIDS CLUSETERING UNTUK REKOMENDASI MENU DAN STRATEGI STOK BAHAN BAKU

Rinci Kembang Hapsari¹⁾, M Safi Anwar Anas²⁾, Reza Zulkifli Ferdiansyah³⁾, Hanif Prasetyo⁴⁾, Mochamad Muhajir⁵⁾

^{1,2,3,4} Teknik Informatika, Institut Teknologi Adhi Tama Surabaya

email: ¹rincikembang@itats.ac.id, ²anassafi375@gmail.com, ³zulkifliferdiansyah@gmail.com,

⁴hanifprasetyo08@gmail.com, ⁵ 062017106906@mhs.itats.ac.id

ARTICLE INFO

Article History:

Recieved : 26 Maret 2024

Accepted : 1 Juni 2024

Published : 14 Juni 2024

Keywords:

Clusterization

Coffee shop

K-medoids

Sales data

IEEE style in citing this article:

R. K. Hapsari, M. S. A. Anas, R. Z. Ferdiansyah, H. Prasetyo, "Penerapan Algoritma K-Medoids Clustering untuk Rekomendasi Menu dan Strategi Stok Bahan Baku", *jurnal.ilmiah.informatika*, vol. 9, no. 1, pp. 30-xx, Jun. 2024.

ABSTRACT

A coffee shop is a place that provides coffee and other hot drinks. Many customers, especially young people, visit coffee shops to enjoy food and drinks while relaxing as the growth of coffee shops continues to increase due to modern places and affordable food and drink prices. In competing for coffee shop services to customers, many coffee shops have developed minimum variants/types of food sold in coffee shops. The large number of food and drink variants means that customers need a long time to choose a menu, making it difficult for the purchasing department to provide raw material stock. So, this research aims to group the menus in coffee shops into 2 clusters, namely the menu cluster that sells well and the menu cluster that sells quite well. In this research, sales data from coffee shop menus were grouped by implementing the k-medoids algorithm. Every member of Cluster 1 and every member of Cluster 2 were grouped. From the tests that have been carried out it can help customers and coffee shop entrepreneurs to support purchasing strategies. By looking at the cluster 1 menu it can be used as menu recommendation information so that it is easier for consumers to choose the drink and food menu at the coffee shop. It can also be used as a basis for purchasing raw food and beverage materials

1. PENDAHULUAN

Awal perkembangan usaha kedai kopi memiliki kesan yang kurang nyaman dan sarana yang minim. Seiring perkembangan bisnis, kedai kopi sekarang merupakan tempat yang nyaman. Hal ini dikarenakan banyak kedai kopi yang memiliki fasilitas WIFI, ruang berAC serta tempat duduk yang nyaman dilengkapi dengan desain interior yang khas. Hal ini membuat peminat kedai kopi semakin banyak. Semakin banyaknya peminat kedai kopi, membuat pengusaha kedai kopi harus mampu bersaing dengan pengusaha kedai kopi lainnya. Dalam upaya meningkatkan kemampuan bersaing dengan pengusaha kedai kopi yang lain, seorang pengusaha diharuskan untuk dapat mengambil keputusan yang tepat untuk menentukan strategi usahanya. Dimana data transaksi penjualan merupakan salah satu sumber yang dapat digunakan sebagai dasar untuk menentukan strategi penjualan.

Dengan menggunakan teknik data mining, pengusaha bisa menggali informasi implisit yang terdapat dalam data transaksi penjualan, untuk mendukung pengambilan strategi penjualannya. Telah banyak diterapkan teknik data mining dalam dunia usaha, salah satunya adalah teknik klasterisasi. Pada permasalahan klasterisasi telah banyak dilakukan penelitian, antara lain yaitu penerapan teknik klasterisasi dengan menggunakan algoritma klasterisasi lain yaitu Agglomerative Hierarchical Clustering digunakan dalam pengelompokan pelanggan untuk menentukan strategi pemasaran[1].

Pada penelitian lain Teknik klasterisasi digunakan sebagai strategi pemasaran penjualan buku pada marketplace, yaitu Tokopedia dan Shopee dengan menggunakan algoritma K-means dan k-medoids. Dimana transaksi

penjualan terbagi menjadi tiga cluster, yaitu pada cluster pertama beranggotakan produk-produk yang banyak diminati, pada cluster kedua beranggotakan produk-produk yang cukup diminati dan pada cluster ketiga beranggotakan produk-produk yang kurang diminati. Pada kedua algoritma ini proses evaluasi diukur dengan menghitung nilai Indeks Davies Bouldin (DBI), untuk memaksimalkan jarak antar cluster dan meminimalkan jarak intra cluster. Berdasarkan nilai DBI tersebut diperoleh hasil bahwa algoritma k-medoids memiliki kinerja lebih baik daripada algoritma k-means[2].

Algoritma K-Medoids merupakan algoritma klasterisasi analitik yang memiliki tujuan untuk menghasilkan sekumpulan k-cluster diantara data yang membutuhkan suatu objek data dalam pengumpulan datanya. Algoritma K-Medoids dapat digunakan untuk mengetahui pola sebaran penyebaran COVID-19, yang dapat dimanfaatkan untuk masyarakat. Hasil pengelompokan riset baru COVID-19 menunjukkan dapat menghasilkan kelompok dari berbagai daerah di Indonesia. Salah satu ciri gejala COVID-19, Dimana pasien mengalami suhu badan lebih dari 36,90C yang diikuti dengan demam dan batuk [3].

Dalam penggalan pola pemilihan program studi mahasiswa baru tahun 2018 di universitas Kanjuruhan, algoritma k-medoid dengan jumlah cluster sebanyak tiga dan jumlah data sebanyak 15 dapat menghasilkan nilai Silhouette Coefficient terbaik yaitu 0.690754 [4]. Algoritma K-medoids juga digunakan dalam segmentasi pelanggan untuk menentukan strategi pemasaran berdasarkan karakteristik dan tipe pelanggan. Dari hasil pengujian menunjukkan bahwa nilai Silhouette Index maksimum adalah 0,375, sedangkan nilai Davies Doulbin Index

minimum adalah 1,030 pada saat jumlah cluster optimalnya adalah tiga cluster [5].

Penelitian tentang pengelompokan wilayah sebaran anak yang mengalami cacat, algoritma K-Means dan algoritma K-Medoids telah mampu mengelompokkan data sebaran anak cacat yang terdapat di Provinsi Riau dengan menghasilkan tiga cluster. Nilai validitas Silhouette Coefficient yang dihasilkan pada algoritma K-Means adalah 0.1443, sedangkan untuk Nilai validitas Silhouette Coefficient adalah 0.5009. Nilai ini menunjukkan bahwa algoritma K-Medoids lebih baik dari algoritma K-Means dalam melakukan pengelompokan pada data sebaran Anak Cacat [6]. Algoritma K-Medoids merupakan algoritma yang diciptakan untuk mengatasi kelemahan algoritma K-Means. Dimana algoritma K-Means bersifat sensitif terhadap outlier, karena nilai yang

sangat besar dapat secara substansial mendistorsi distribusi data [7].

Berdasarkan penelitian yang terdahulu, maka artikel ini mengelompokkan menu makanan dan minuman yg ada di kedai kopi dengan menggunakan metode, yaitu K-Medoids clustering.

2. METODE PENELITIAN

Pada bagian ini menjelaskan tahapan yang dilakukan dalam melakukan penelitian. Metode K-Medoids clustering dipilih untuk melakukan pengelompokan data menu makanan. Gambar 1 merupakan diagram alir tahapan penelitian yang terdapat tiga proses, yaitu pre-proses data, proses klasterisasi dengan menggunakan algoritma K-Medoids dan yang terakhir adalah pengukuran performa system.



Gambar 1. Alur tahapan penelitian

2.1. Dataset

Data yang digunakan merupakan data primer. Data yang didapatkan oleh peneliti merupakan data penjualan atau jumlah setiap makanan dan minuman yang terjual di coffee break dalam kurun waktu 4 minggu, dan terdapat 53 data transaksi penjualan. Data dibagi menjadi 4 parameter yaitu, pekan pertama, pekan kedua, pekan ketiga, dan pekan keempat.

2.2 Pre-proses Data

Preprocessing merupakan proses untuk mengubah data mentah/asli menjadi data yang lebih mudah dipahami. Banyak data mentah sering kali memiliki format yang tidak konsisten, tidak lengkap, dan memiliki kesalahan, sehingga pre-processing data penting

dilakukan. Preprocessing data seringkali diperlukan untuk memastikan validitas dan keandalan hasil analisis data [8]. Misalnya, data transaksi penjualan biasanya memiliki banyak nilai dan outlier yang hilang karena kesalahan dalam pengumpulan, transmisi, dan penyimpanan data [9]. Langkah pemrosesan awal data dapat diterapkan untuk menghilangkan outlier dan mengisi nilai yang hilang untuk analisis data yang lebih andal [10].

Pada proses ini melibatkan serangkaian teknik untuk membersihkan, mentransformasi, dan mempersiapkan data mentah agar sesuai untuk analisis lebih lanjut. Data preprocessing merupakan proses paling awal sebelum melakukan data mining. Tabel 1 merupakan

informasi data penjualan makan dan minuman selama 4 minggu berturut-turut.

Tabel 1. Data Transaksi Penjualan

No	Nama Menu	Pekan			
		I	II	III	IV
1	Nasi goreng	7	5	12	5
2	Nasi ayam penyet	19	20	18	14
...
51	Burger	3	5		4

Terdapat 51 menu dan tiap menu memiliki jumlah penjualannya untuk setiap pekan, yang mana pada data ke-51 terdapat informasi yang tidak lengkap, yaitu pada pekan ke-3. Sehingga perlu dilakukan preprocessing dengan melakukan proses pembersihan data atau data cleaning. Pada proses ini data mentah yang telah diperoleh perlu dilakukan seleksi kembali. Dimana dilakukan penghapusan atau menghilangkan data-data yang tidak relevan, tidak lengkap, dan tidak akurat[11]. Maka dari itu dilakukanlah penghapusan data ke-51. Proses itu digunakan untuk menghindari adanya noise pada dataset.

5.3 Klasterisasi (K-Medoids)

Klasterisasi merupakan suatu proses pengelompokkan suatu record, hasil observasi, maupun pengelompokkan kelas yang memiliki kemiripan objek [3]. Dalam klasterisasi melakukan proses pengelompokan data yang didasarkan pada kemiripan objek, sehingga klasterisasi digolongkan sebagai metode unsupervised learning [12].

Dalam penelitian ini proses klasterisasi menggunakan algoritma K-Medoids. Algoritma K-Medoids juga disebut dengan partitioning around medoids, yang merupakan varian algoritma K-Means. Pada K-Medoids

pusat cluster didasarkan pada penggunaan medoids bukan dari hasil pengamatan yang terdapat pada setiap cluster. Hal ini bertujuan untuk mereduksi sensitivitas dari partisi yang dihasilkan terkait dengan nilai-nilai ekstrim yang terdapat pada dataset [13]. Algoritma K-Medoids adalah suatu algoritma yang memperbaiki kelemahan Algoritma K-Means yang sensitif terhadap outlier karena objek dengan nilai yang besar bisa jadi menyimpang terhadap distribusi data [5][14].

Alur proses algoritma K-Medoids digambarkan dalam flowchart Gambar 2. Pada input menentukan jumlah k, merupakan proses untuk memasukkan jumlah kelompok yang diinginkan. Pada proses perhitungan jarak masing-masing objek data dengan medoid yang digunakan untuk menentukan anggota cluster, digunakan persamaan (1)

$$D(i, j) = \sqrt{(x_{1i} - y_{1j})^2 + (x_{2i} - y_{2j})^2 + \dots + (x_{ni} - y_{nj})^2} \quad (1)$$

Dengan

$D(i, j)$: Jarak data ke i terhadap data pusat (medoids) ke j

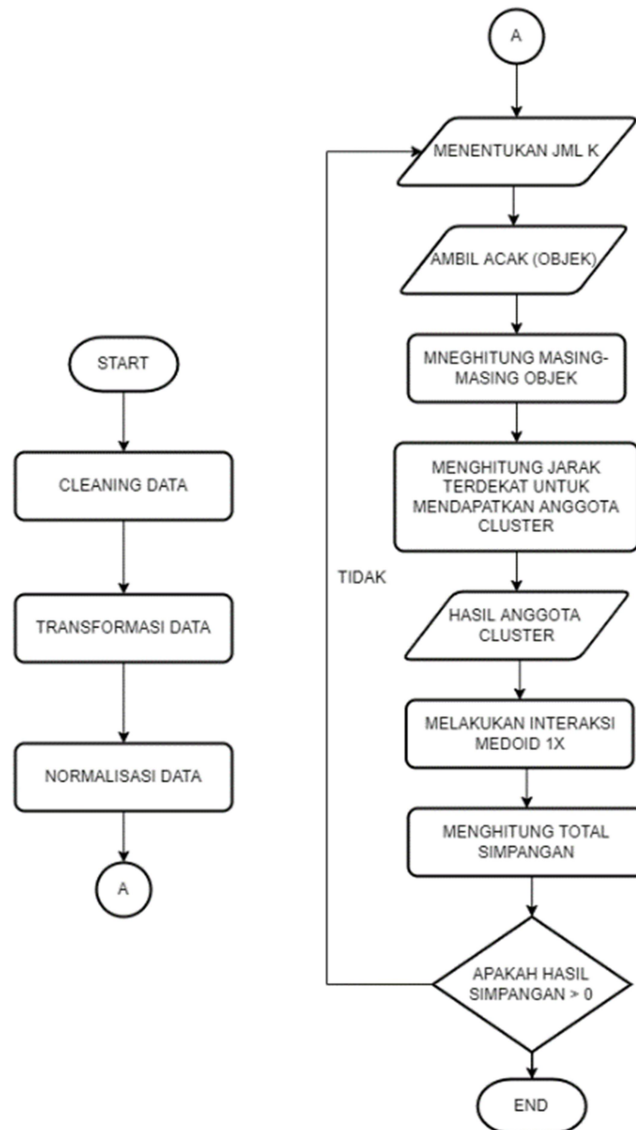
x_{ni} : Titik pusat (medoids) ke j pada attribute data ke n

y_{nj} : Data ke-i pada attribute data ke n

Setelah didapatkan jarak setiap obyek data terhadap medoids di masing-masing cluster dilakukan penentuan cluster

setiap obyek data. Kemudian menghitung nilai total simpangan (S) dengan mengurang-kan nilai total distance baru

dikurangi dengan nilai total distance lama.



Gambar 2. Flowchart Algoritma K-Medoids

Jika $S < 0$, maka objek data ditukar dengan data cluster untuk membentuk sekelompok k objek baru sebagai medoids. Diulang proses menghitung jarak setiap obyek data dengan medoids. Dan menghitung total simpangan (S) jika $S > 0$, proses berhenti.

3. HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini diambil dari transaksi penjualan. Dimana total jumlah dataset valid adalah 50 data, berisi 4 parameter yang di mana setiap parameter berisi total penjualan makanan dan minuman selama 1 periode. Dan dilakukan pengelompokan menjadi 2 kelompok, yaitu kelompok satu adalah laku dimana menunjukkan menu

makanan atau minuman yang laku terjual, dan kelompok dua adalah cukup, yaitu menu makanan atau minuman yang cukup laku terjual.

Dengan menggunakan dataset yang telah dilakukan proses normalisasi, tahapan proses klusterisasi menggunakan algoritma k-Medoids, adalah sebagai berikut :

- a. Menentukan jumlah cluster yang akan dibentuk adalah dua cluster. Dan menentukan centroid (pusat cluster) awal sebagai medoid yang ditunjukkan pada Tabel 2. Pemilihan medoids dilakukan secara random, diambil 2 data secara acak.

Tabel 2. Medoids Awal Cluster (Iterasi pertama)

Nama Cluster (data ke-)	Pekan			
	I	II	III	IV
Cluster 1 (data ke-4)	5	7	12	13
Cluster 2 (data ke-10)	7	5	10	9

- b. Mengelompokkan objek-objek data non medoids ke dalam cluster yang paling dekat dengan medoids berdasarkan nilai jarak Euclidean, yang dirumuskan pada persamaan (1). Hasil perhitungan jarak dari data ke 1 dengan nilai (7, 5, 12, 5) didapatkan nilai sebagai berikut :

$$D(x_1, c_1) = \sqrt{(5-7)^2 + (7-5)^2 + (12-12)^2 + (13-5)^2}$$

$$= 8,4852$$

$$D(x_1, c_2) = \sqrt{(7-7)^2 + (5-5)^2 + (10-12)^2 + (9-5)^2}$$

$$= 4,472$$

Tabel 3. Medoids Awal Cluster (Iterasi pertama)

NoId	C1	C2	Kedekatan	Anggota Cluster
1	8,4852	4,4721	4,4721	2
2	20,0499	21,4009	20,0499	1
...
50	27,6586	29,8161	27,6586	1
Total kedekatan =			587,2030	

Perhitungan jarak dengan medoids dilakukan untuk semua data dalam dataset, yaitu sampai data yang ke-50. Hasil keseluruhan perhitungan jaraknya ditunjukkan pada Tabel 3.

$$\text{Simpanan} = 556,6 - 431,9$$

$$= 124,7 > 0$$

- c. Perhitungan simpangan atau selisih antara total kedekatan baru dikurangi total kedekatan lama, berhenti jika simpangan lebih besar dari pada 0. Dan itu terjadi pada iterasi ke-3, Dimana nilai simpangannya adalah :

Hasil penelitian diperoleh anggota dari cluster 1 dan cluster 2. Data anggota ini, dikelompokkan menjadi “laku” dan “tidak laku”. Pengelompokan ini berguna untuk membuktikan perhitungan berhasil atau tidak. Jika semua anggota pada cluster 1 ada dalam “laku” dan tidak ada anggota yang berada di dalam “tidak laku”, maka perhitungan benar.

Tabel 4. Hasil Perhitungan Jarak ke Setiap Medoids

NoId	Average	Anggota Cluster
43	23	1
50	22,5	1
...
34	11,75	1
16	10,75	1
40	9,75	2
6	9,5	2
...
38	1,75	2
45	1,75	2

Pada Tabel 4, average adalah rata-rata dari penjualan minggu pertama sampai minggu keempat. Setelah itu diurutkan dari yang terbesar sampai terkecil. Nomer pada tabel 5 sudah teracak menurut urutan average. Dari urutan data 1 samapi 20, semuanya berisi anggota cluster 1, tidak ada anggota cluster 2.

4. UCAPAN TERIMA KASIH

Terimakasih kami ucapkan kepada Bapak Pembina, Bapak Ketua YPTS, Bapak Pinpinan Institut Teknologi Adhi Tama Surabaya karena telah mensupport kami. Sehingga kolaborasi penelitian dosen dan mahasiswa dapat diselesaikan dan dipublikasikan.

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik kesimpulan, bahwa sistem rekomendasi menu makanan dan minuman dibuat dengan mengimplemantasikan algoritma Clustering K-Medoids. Terdapat lima tahap pada proses cluster yaitu cleaning data, menentukan jumlah Cluster, menentukan masing-masing medoids dari

setiap cluster, menghitung jarak, menghitung simpangan/selisih kedekatan. Sehingga didapatkan output dua cluster beserta setiap anggotanya, yang berupa menu makanan dan minuman. Hasil pengamatan menunjukkan cluster satu merupakan manu makanan dan minuman yang paling laku dan cluster dua beranggotakan menu makanan atau minuman yang laku.

Anggota cluster satu yang merupakan menu makanan dan minuman yang paling laku, dapat digunakan sebagai masukan rekomendasi menu untuk pelanggan dan juga untuk memudahkan pihak kedai kopi dalam melakukan strategi stok bahan baku.Strategi stok bahan baku yaitu dari hasil perhitungan clustering digunakan sebagai patokan dalam proses pembelian dan pengadaan bahan baku makanan dan minuman. Dimana sistem rekomendasi menu berdasarkan hasil clustering dapat digunakan dalam mempermudah konsumen untuk memilih menu makanan dan minuman di kedai kopi.

6. REFERENSI

- [1] W. Widyawati, W. L. Y. Saptomo, and

- Y. R. W. Utami, "Penerapan Agglomerative Hierarchical Clustering Untuk Segmentasi Pelanggan," *J. Ilm. SINUS*, vol. 18, no. 1, p. 75, 2020, doi: 10.30646/sinus.v18i1.448.
- [2] W. M. Baihaqi, K. Indartono, and S. Banat, "Penerapan Teknik Clustering Sebagai Strategi Pemasaran pada Penjualan Buku Di Tokopedia dan Shopee," *Paradig. - J. Komput. dan Inform.*, vol. 21, no. 2, pp. 243–248, 2019, doi: 10.31294/p.v21i2.6149.
- [3] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 166–173, 2020, doi: 10.36294/jurti.v4i1.1296.
- [4] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang," *RAINSTEK J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 53–68, 2019, doi: 10.21067/jtst.v1i3.3046.
- [5] A. A. D. Sulistyawati and M. Sadikin, "Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan," *Sistemasi*, vol. 10, no. 3, p. 516, 2021, doi: 10.32520/stmsi.v10i3.1332.
- [6] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [7] N. Pulungan, S. Suhada, and D. Suhendro, "Penerapan Algoritma K-Medoids Untuk Mengelompokkan Penduduk 15 Tahun Keatas Menurut Lapangan Pekerjaan Utama," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, 2019, doi: 10.30865/komik.v3i1.1609.
- [8] R. K. Hapsari, E. Purwanti, W. Widyanto, R. Gunawan, F. Nurlaily, and A. H. Salim, "Optimization Based Random Forest Algorithm Modification for Detecting Monkeypox Disease," in *2023 Sixth International Conference on Vocational Education and Electrical Engineering (ICVEE)*, IEEE, 2023, pp. 340–346. doi: 10.1109/icvee59738.2023.10348223.
- [9] I. Khan, A. Capozzoli, S. P. Corgnati, and T. Cerquitelli, "Fault detection analysis of building energy consumption using data mining techniques," *Energy Procedia*, vol. 42, no. April 2014, pp. 557–566, 2013, doi: 10.1016/j.egypro.2013.11.057.
- [10] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data," *Front. Energy Res.*, vol. 9, no. March, pp. 1–17, 2021, doi: 10.3389/fenrg.2021.652801.
- [11] R. K. Hapsari, A. H. Salim, B. D. Meilani, T. Indriyani, and A. Rachman, "Comparison of the Normalization Method of Data in Classifying Brain Tumors with the k-NN Algorithm," in *The 2nd International Conference on Neural Networks and Machine Learning*, Atlantis Press International BV, 2023, pp. 21–29. doi: 10.2991/978-94-6463-174-6_3.
- [12] D. D. C. Nugraha, Z. Naimah, M.

- Fahmi, and N. Setiani, "Klasterisasi Judul Buku dengan Menggunakan Metode K-Means," *Semin. Nas. Apl. Teknol. Inf. Yogyakarta*, vol. 21, no. 1, pp. 1907–5022, 2014.
- [13] B. Riyanto, "Penerapan Algoritma K-Medoids Clustering Untuk Pengelompokan Penyebaran Diare Di Kota Medan (Studi Kasus: Kantor Dinas Kesehatan Kota Medan)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 562–568, 2019, doi: 10.30865/komik.v3i1.1659.
- [14] S. Akter, F. Reza, and M. Ahmed, "Convergence of Blockchain, k-medoids and homomorphic encryption for privacy preserving biomedical data classification," *Internet Things Cyber-Physical Syst.*, vol. 2, no. March, pp. 99–110, 2022, doi: 10.1016/j.iotcps.2022.05.006.