

# Perbandingan Jarak Manhattan dan Chebyshev pada K-Medoids berbasis Komputasi Kuantum

## *Comparison of Manhattan and Chebyshev Distance Metrics in Quantum-Based K-Medoids Clustering*

<sup>1</sup>Solikhun\*, <sup>2</sup>Muhammad Rahmansyah Siregar, <sup>3</sup>Lise Pujiastuti, <sup>4</sup>Mochamad Wahyudi, <sup>5</sup>Deny Kurniawan

<sup>1,2</sup>Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>3</sup>Program Studi Sistem Informasi, STMIK Antar Bangsa, Tangerang, Indonesia

<sup>4</sup>Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia

<sup>5</sup>Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia

\*e-mail: [solikhun@amiktunasbangsa.ac.id](mailto:solikhun@amiktunasbangsa.ac.id)

(*received*: 25 March 2025 , *revised*: 29 April 2025, *accepted*: 30 April 2025 )

### Abstrak

Anemia merupakan suatu kondisi jumlah sel darah merah dalam aliran darah menurun. Anemia dapat menyebabkan kelelahan dan menurunkan produktivitas. Clustering merupakan teknik pada data mining untuk mengidentifikasi pola yang dapat dimanfaatkan dalam proses pengambilan keputusan. Clustering dalam kasus anemia sangat penting karena membantu mengidentifikasi berbagai pola seberapa parah kondisi serta memahami faktor-faktor yang berkontribusi terhadap anemia. Komputer kuantum menggunakan mekanika kuantum dalam pemrosesan informasi dan telah mencatat kemajuan signifikan dalam satu dekade terakhir. Komputasi kuantum merupakan metode pemrosesan informasi tingkat lanjut yang memanfaatkan qubit, memungkinkan mereka untuk berada dalam beberapa keadaan secara bersamaan. Teknologi ini berpotensi menyelesaikan masalah dengan kecepatan eksponensial lebih tinggi dibandingkan komputer klasik. Dalam penelitian ini, peneliti menggunakan algoritma K-Medoids yang dihitung menggunakan persamaan kuantum. Penelitian ini membandingkan dua metode perhitungan jarak, yaitu jarak Chebyshev dan jarak Manhattan. Hasil penelitian menunjukkan bahwa algoritma Manhattan lebih unggul dalam konteks medis, terutama untuk deteksi kasus positif, dengan recall 0.57 dan F1-score 0.695, meskipun memiliki precision yang sedikit lebih rendah 0.88. Hal ini membuatnya lebih cocok untuk aplikasi medis di mana false negative berisiko tinggi, seperti deteksi penyakit, meskipun biaya dan MSE lebih tinggi. Sebaliknya, Chebyshev memiliki precision sempurna yaitu 1 dan akurasi lebih tinggi 80%, namun recall yang rendah 0.33 menunjukkan bahwa banyak kasus positif terlewatkan. Oleh karena itu, Manhattan lebih direkomendasikan dalam aplikasi medis yang memerlukan deteksi lebih banyak kasus positif, sementara Chebyshev lebih efisien untuk konteks yang memprioritaskan akurasi dan biaya.

**Kata kunci:** data mining, komputasi kuantum, k-medoids, jarak manhattan, jarak chebyshev

### Abstract

Anemia is a condition characterized by a decrease in the number of red blood cells or hemoglobin levels in the bloodstream. It can lead to fatigue and reduced productivity. Clustering is a technique in data mining used to identify patterns that can support decision-making processes. In the case of anemia, clustering plays a crucial role in identifying various severity patterns and understanding the contributing factors behind the condition. Quantum computers, which utilize the principles of quantum mechanics for information processing, have made significant advancements over the past decade. Quantum computing is an advanced method of information processing that leverages qubits, enabling systems to exist in multiple states simultaneously. This technology offers the potential to solve complex problems at exponentially faster speeds than classical computers. In this study, researchers applied the K-Medoids clustering algorithm, calculated using quantum-based equations. The research compares two distance measurement methods: Chebyshev distance and Manhattan

<http://sistemasi.ftik.unisi.ac.id>

*distance. The results show that the Manhattan algorithm performs better in medical contexts, particularly for detecting positive cases, with a recall of 0.57 and an F1-score of 0.695, although it has a slightly lower precision of 0.88. This makes it more suitable for medical applications where false negatives carry high risks, such as disease detection, despite its higher cost and mean squared error (MSE). On the other hand, Chebyshev distance achieved perfect precision (1.0) and higher accuracy (80%), but its low recall (0.33) indicates that many positive cases were missed. Therefore, Manhattan distance is more recommended for medical applications that require the detection of more positive cases, while Chebyshev is more efficient for scenarios that prioritize accuracy and cost.*

**Keywords:** data mining, quantum computing, clustering, k-medoid, manhattan distance, chebyshev distance.,

## 1 Pendahuluan

Anemia adalah suatu kondisi medis yang ditandai oleh berkurangnya jumlah sel darah merah atau kadar hemoglobin, yang dapat memberikan dampak signifikan terhadap kondisi kesehatan seseorang. Konsentrasi hemoglobin dalam darah dianggap sebagai indikator yang paling berpengaruh terhadap anemia. Anemia dapat menyebabkan kelelahan dan menurunkan produktivitas, serta jika terjadi selama kehamilan, dapat meningkatkan risiko kematian ibu [1]. Dalam penelitian ini, peneliti menggunakan algoritma K-Medoid yang dihitung menggunakan persamaan kuantum untuk melakukan clustering pada anemia. Clustering anemia membantu dalam mengidentifikasi jenis anemia tertentu berdasarkan gejala, penyebab, dan hasil tes laboratorium. Hal ini sangat penting untuk menentukan pendekatan pengobatan yang tepat. Dengan clustering anemia, dokter dapat merancang rencana perawatan yang lebih personal dan terarah.

Komputer kuantum menggunakan mekanika kuantum dalam pemrosesan informasi dan telah mencatat kemajuan signifikan dalam satu dekade terakhir [2]. Komputasi kuantum adalah metode pemrosesan informasi tingkat lanjut yang memanfaatkan qubit, memungkinkan mereka untuk berada dalam beberapa keadaan secara bersamaan. Sementara komputer klasik menggunakan kabel dan gerbang logika, komputer kuantum menggunakan kabel dan gerbang kuantum elementer. Tidak seperti bit klasik yang hanya dapat merepresentasikan nilai 0 atau 1, qubit dapat berada dalam keadaan selain  $|0\rangle$  atau  $|1\rangle$ , menjadikannya berbeda dari bit konvensional [3]–[5]. Komputasi kuantum memiliki potensi untuk menyelesaikan masalah dengan kecepatan eksponensial lebih tinggi dibandingkan dengan komputer klasik [6]. Data mining merupakan tahap krusial dalam proses Knowledge Discovery in Database, yang mencakup langkah-langkah seperti pemilihan data, pemrosesan, transformasi, proses data mining itu sendiri, serta evaluasi hasil [7]. Data mining dapat mengungkap pengetahuan tersembunyi dalam basis data yang besar [8]. Proses ini mengidentifikasi pola yang dapat dimanfaatkan untuk mendukung proses pengambilan keputusan [9]. Clustering melibatkan penggunaan algoritma untuk membagi data ke dalam kelompok tertentu yang disebut kluster [10], [11]. Prinsip dasar dari pengelompokan adalah mengelompokkan beberapa objek ke dalam satu kluster, di mana kualitas kluster diukur berdasarkan tingkat kesamaan yang tinggi antara objek dalam kluster yang sama serta tingkat perbedaan yang tinggi antara kluster yang berbeda [12], [13]. Salah satu algoritma yang umum digunakan dalam teknik clustering adalah K-Medoid, yang juga dikenal dengan nama Partitioning Around Medoids (PAM). Algoritma ini mirip dengan K-Means karena keduanya merupakan algoritma partisi yang membagi data menjadi beberapa kelompok. Perbedaannya terletak pada metode penentuan pusat kluster. Algoritma K-Means menghitung rata-rata (mean) dari semua titik dalam suatu kluster untuk menentukan pusatnya, sedangkan pendekatan K-Medoids memilih salah satu objek data asli, yang disebut medoid, untuk mewakili pusat kluster [14], [15].

Tujuan penelitian yaitu menerapkan algoritma K-medoid dengan komputasi kuantum pada data anemia untuk diagnosis pasien penyakit anemia. Manfaat dari penelitian ini yaitu algoritma ini dapat digunakan oleh klinisi dan peneliti untuk memprediksi berbagai hasil pada pasien anemia, seperti risiko komplikasi, kecepatan pemulihan, atau kebutuhan transfusi darah. Dengan meningkatkan akurasi pengelompokan data pasien, algoritma ini memungkinkan pengembangan rencana perawatan yang lebih efektif. Selain itu, dengan efisiensi dalam mengelola dan menganalisis dataset besar, penelitian ini juga mendukung pengambilan keputusan klinis serta perumusan kebijakan kesehatan.

## 2 Tinjauan Literatur

Penelitian sebelumnya oleh [16], menggunakan metode K-Medoids dan CLARA untuk mengelompokkan kabupaten/kota pada provinsi Sulawesi Selatan dan Barat berdasarkan indikator yang membentuk Angka Partisipasi Pendidikan. Tujuan penelitian ini adalah memetakan wilayah di kedua provinsi tersebut ke dalam beberapa kelompok. Hasil penelitian menunjukkan bahwa nilai silhouette yang dihasilkan oleh metode K-Medoids dan CLARA dengan  $k=4$  adalah 0,49 untuk masing-masing metode.

Penelitian sebelumnya oleh [17] Penelitian ini mengklasifikasikan kabupaten dan kota di Provinsi Sulawesi Selatan beserta empat sektornya berdasarkan indikator yang membentuk Indeks Kualitas Udara (IKU) tahun 2019 dan 2020. Klasifikasi dilakukan menggunakan analisis cluster non-hierarkis dengan metode K-Medoids. Hasil penelitian menunjukkan bahwa metode K-Medoids dengan  $k=2$  menghasilkan pengelompokan terbaik dengan nilai koefisien silhouette sebesar 0,56. Data passive sampler tahun 2019 menunjukkan bahwa kelompok 1 memiliki IKLH sangat baik (IKU 84,14), sedangkan kelompok 2 memiliki IKLH buruk (IKU 60,04). Untuk data tahun 2020, kelompok 1 masuk dalam kategori IKLH baik (IKU 80,68), sedangkan kelompok 2 tetap dalam kategori IKLH buruk (IKU 61,53).

Penelitian sebelumnya oleh [18] Penelitian ini menggunakan jarak Euclidean dan jarak Chebyshev untuk mengelompokkan data Gaikindo Wholesales tahun 2021, yang merupakan data spesifikasi produk yang dijual oleh Gaikindo. Tujuan penelitian ini adalah membandingkan jarak Euclidean dan Chebyshev dalam clustering K-Medoids untuk menentukan hasil yang paling optimal. Evaluasi dilakukan dengan menggunakan indeks koefisien silhouette. Hasil clustering menunjukkan bahwa penggunaan jarak Euclidean menghasilkan klaster yang lebih optimal dibandingkan dengan jarak Chebyshev.

Penelitian sebelumnya oleh [19] menganalisis cluster komparatif dengan K-Means dan K-Medoids untuk mengkaji distribusi virus COVID-19. Perbandingan kedua algoritma divalidasi menggunakan nilai Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa DBI untuk K-Means adalah 0,9762331449809145, sedangkan untuk K-Medoids adalah 0,9809235412405508. Maka dapat disimpulkan bahwa K-Means menghasilkan clustering yang lebih baik.

Penelitian sebelumnya oleh [20] Penelitian ini menghitung jarak antara koordinat karyawan dari smartphone dan pusat Iteas menggunakan metode Euclidean, Manhattan, dan Haversine. Kinerja masing-masing metode dievaluasi berdasarkan ketepatan perhitungan jarak, akurasi keputusan, dan waktu pemrosesan data. Hasil penelitian menunjukkan bahwa metode Manhattan adalah yang tercepat dalam memproses seratus data, metode Haversine adalah yang paling akurat dengan selisih jarak kecil dan akurasi 98,66%, serta paling presisi dalam menentukan lokasi dengan akurasi 90%. Kesimpulan dari penelitian ini adalah metode Euclidean dan Haversine dapat diterapkan dalam sistem identifikasi lokasi karyawan karena memiliki selisih jarak kurang dari 0,5 meter, sedangkan metode Manhattan, dengan selisih 6,67 meter, tidak cocok untuk penelitian ini.

Penelitian sebelumnya oleh [21] Penelitian membandingkan K-Means dan K-Medoids dalam menganalisis peningkatan kasus HIV di Jawa Barat pada tahun 2022. Hasil penelitian menunjukkan bahwa algoritma K-Means lebih unggul dibandingkan K-Medoids dalam pengelompokan data, dengan K-Means mencapai nilai DBI terbaik sebesar 0,102 pada  $k=4$ , sedangkan K-Medoids mencatat DBI optimal sebesar 0,130 pada  $k=3$ . Dari segi kecepatan komputasi, K-Means membutuhkan waktu 5 detik, sedangkan K-Medoids memerlukan 5 menit.

Penelitian sebelumnya oleh [22] melakukan analisis pasien COVID-19 yang mengalami obesitas, non-obesitas, serta obesitas dengan penyakit kardiovaskular menggunakan metode K-Medoids. Penelitian ini membandingkan dan mengukur pengaruh atribut dalam dataset serta menentukan persentase pasien yang terinfeksi dan tidak terinfeksi COVID-19. Hasil penelitian menunjukkan bahwa metode K-Medoids efektif dalam clustering data. Dari dataset mentah, 62,62% pasien tidak terinfeksi COVID-19, sedangkan 37,38% terinfeksi. Dari populasi yang diuji, 74,54% termasuk kategori non-obesitas, 25,46% obesitas, dan 0,57% obesitas dengan penyakit kardiovaskular. Model clustering dengan K-Medoids menunjukkan pengaruh atribut yang signifikan. Pada dataset obesitas, atribut pneumonia dan hipertensi memiliki pengaruh sebesar 150,15% dan 172,04%. Pada dataset non-obesitas, atribut jenis kelamin dan hipertensi memiliki pengaruh sebesar

39,50% dan 106,61%. Sementara itu, pada pasien dengan obesitas dan penyakit kardiovaskular, atribut jenis kelamin dan pneumonia memiliki pengaruh sebesar 159,07% dan 300%.

Penelitian sebelumnya yang menggunakan algoritma K-Medoids adalah [23] Dalam penelitian ini, K-Medoids digunakan untuk mengelompokkan virus MERS-CoV berdasarkan urutan DNA. Tujuan penelitian ini adalah memberikan informasi mengenai karakteristik virus MERS-CoV di beberapa negara. Para peneliti menggunakan perangkat lunak RStudio dan Notepad untuk melakukan perhitungan. Hasil penelitian menghasilkan dua klaster, dengan klaster nomor 2 memiliki nilai koefisien silhouette tertinggi, yaitu 0,61534. Klaster 1 terdiri dari 84 urutan DNA MERS-CoV, sedangkan Klaster 2 terdiri dari 16 urutan DNA MERS-CoV.

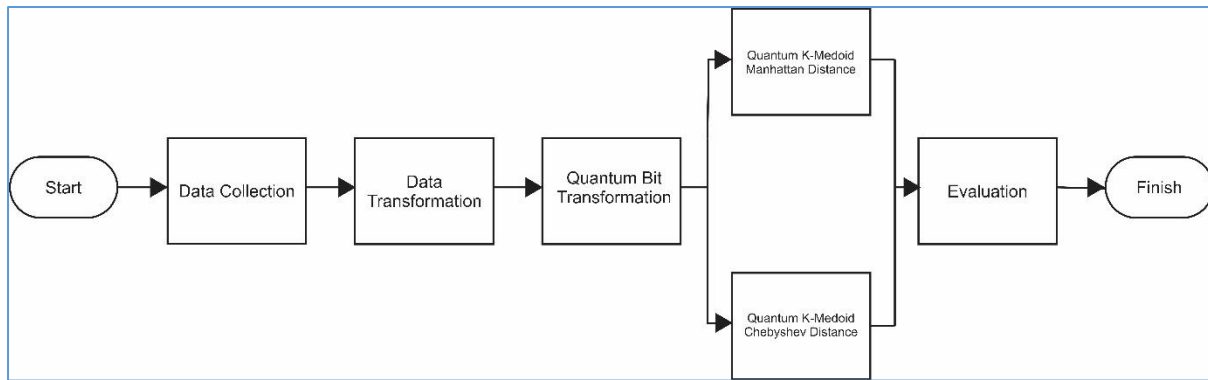
Penelitian lain oleh [24] Penelitian ini mengelompokkan data pasien COVID-19 menggunakan algoritma K-Means dan K-Medoids berdasarkan jenis transmisi, usia, jenis kelamin, fasilitas kesehatan, dan kecamatan. Studi ini membandingkan efektivitas algoritma clustering K-Medoids dan K-Means dalam pengelompokan pasien COVID-19. Hasil penelitian menunjukkan bahwa algoritma K-Means lebih optimal dibandingkan K-Medoids dalam tugas ini, dibuktikan dengan nilai Davies-Bouldin Index terbaik sebesar 0,139 pada eksperimen dengan  $k = 4$ .

Referensi utama untuk penelitian ini adalah [25] Penelitian ini membandingkan kinerja algoritma clustering K-Means dan K-Medoids dalam menganalisis dataset diabetes dan memprediksi penyakit. Studi ini berfokus pada penggunaan data resep biokimia akhir pasien diabetes untuk identifikasi penyakit. Hasil penelitian menunjukkan bahwa algoritma K-Means mencapai akurasi 87%, sedangkan algoritma K-Medoids mencapai akurasi 80%, menjadikan K-Means sebagai salah satu teknik terbaik untuk prediksi penyakit.

Dalam penelitian [25], ditemukan bahwa akurasi algoritma K-Medoids adalah 80%, lebih rendah dibandingkan dengan algoritma K-Means yang memiliki akurasi 87%. Berdasarkan hal ini, para peneliti mencari model alternatif untuk K-Medoids dengan pendekatan kuantum. Penelitian ini melakukan analisis komparatif terhadap dua metode perhitungan jarak: Quantum K-Medoids dengan Chebyshev dan Quantum K-Medoids dengan Manhattan. Kedua pendekatan ini akan dievaluasi berdasarkan akurasi hasil clustering dalam pengelompokan anemia. Dalam penelitian ini, para peneliti membandingkan perhitungan jarak Chebyshev dengan perhitungan jarak Manhattan menggunakan Quantum K-Medoids. Perbedaan utama dengan penelitian sebelumnya adalah penerapan clustering pada anemia berdasarkan data medis pasien anemia. Penelitian ini menghadirkan kebaruan dengan memanfaatkan komputasi kuantum bersama algoritma K-Medoids untuk clustering data anemia. Penelitian ini dapat digunakan sebagai salah satu metode dalam pengelompokan penyakit anemia, khususnya di bidang medis, serta memberikan wawasan mengenai pendekatan algoritma kuantum dalam clustering.

### **3 Metode Penelitian**

Data pada penelitian ini diperoleh dari website Kaggle, yang berisi dataset terkait anemia, mencakup parameter medis seperti kadar hemoglobin, jumlah sel darah merah, hematokrit, serta faktor-faktor lain yang berkontribusi dalam diagnosis anemia. Penelitian ini akan dilaksanakan berdasarkan rancangan alur penelitian yang disajikan dalam flowchart berikut, guna memastikan setiap tahapan berjalan secara sistematis dan terstruktur. Penelitian ini termasuk dalam penelitian komparatif, karena membandingkan kinerja algoritma Quantum K-Medoids dengan perhitungan jarak Manhattan dan chebyshev. Perbandingan ini dilakukan untuk mengevaluasi perbedaan hasil clustering, akurasi, dan efisiensi dalam mengelola dataset anemia.



**Gambar 1. Rancangan penelitian**

Gambar 1 adalah rancangan penelitian yang dimulai dari mengumpulkan data, transformasi data, transformasi qubit, implementasi kuantum k-medoid dan evaluasi hasil. Langkah awal yang dilakukan dalam penelitian ini adalah mengumpulkan data, tahap ini melibatkan pengumpulan data yang diperlukan untuk memprediksi anemia. Dataset diperoleh dari situs Kaggle, dengan memanfaatkan 20 data awal dari total 300 data yang tersedia. Dataset ini mencakup rekam medis pasien anemia, termasuk jenis kelamin, kadar hemoglobin, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), dan mean corpuscular volume (MCV). Tahapan kedua adalah melakukan transformasi data ke biner. Pada tahap ini dataset yang dikumpulkan dikonversi ke dalam representasi biner (1 atau 0) untuk mempermudah pemrosesan lebih lanjut. Proses ini mencakup normalisasi atau pengkodean variabel kategorikal menjadi variabel biner agar data dapat dikelompokkan dengan lebih efektif. Setelah semua data telah di-transformasikan ke biner, selanjutnya adalah tahapan ketiga yaitu transformasi ke dalam bentuk qubit (Quantum bit), dimana qubit berfungsi sebagai unit dasar data dalam komputasi kuantum, memungkinkan penyimpanan dan manipulasi informasi berdasarkan prinsip kuantum. Transformasi ini mendukung penerapan algoritma K-Medoids dalam lingkungan komputasi kuantum. Tahapan keempat adalah melakukan implementasi quantum dengan algoritma K-Medoid menggunakan Manhattan Distance dan Chebyshev Distance yang masing masing menggunakan rumus (1) untuk K-medoid Manhattan dan (2) untuk K-Medoid Chebyshev.

$$d(a_x, b_y) = \sum_{z=1}^n ||a_{az} > - |b_{bz} >| \quad (1)$$

Notasi:

$a_x$  = titik data pertama

$b_y$  = titik data kedua

$z$  = indeks fitur

$a_{az}$  = Nilai fitur ke-z dari vektor  $a_x$

$b_{bz}$  = Nilai fitur ke-z dari vektor  $b_y$

$$d_{Chebyshev}(x, y) = \text{Max}||x_i > - |y_i >| \quad (2)$$

Notasi:

$x, y$  = vector data

$x_i, y_i$  = Komponen ke-i dari vektor  $x$  dan  $y$

Pada tahap ini, hasil pengelompokan data dianalisis menggunakan metode K-Medoids dengan komputasi kuantum Manhattan dan K-Medoids dengan komputasi kuantum Chebyshev. Analisis ini memberikan wawasan mengenai struktur dan pola data yang ditemukan oleh algoritma clustering. Tahapan terakhir adalah evaluasi, tahap evaluasi mencakup penilaian kinerja metode klasik dan K-Medoids dengan komputasi kuantum dalam memprediksi anemia menggunakan perhitungan Jarak Chebyshev. Evaluasi ini menggunakan metrik akurasi untuk menentukan sejauh mana model yang dihasilkan dapat mengidentifikasi anemia dengan baik.

#### 4 Hasil dan Pembahasan

Penelitian ini membandingkan dua perhitungan jarak K-Medoids dengan komputasi kuantum, yaitu perhitungan Jarak Chebyshev dan perhitungan Jarak Manhattan. Tujuan dari penelitian ini adalah untuk mengetahui perhitungan jarak mana yang lebih optimal dalam mengelompokkan data rekam medis anemia dengan pendekatan kuantum. Tahapan pertama yang dilakukan pada penelitian ini adalah melakukan transformasi data dengan aturan:

- Gender memengaruhi ukuran rata-rata hemoglobin dalam darah, 1 untuk pria dan 0 untuk wanita.
- Hemoglobin diukur dengan rentang normal untuk pria dewasa yaitu 13 - 17 g/dL, sedangkan untuk wanita dewasa 12 - 15 g/dL. Jika nilai hemoglobin berada dalam rentang normal, maka diberi nilai 0; jika di luar batas normal, maka nilainya 1.
- Mean Corpuscular Hemoglobin (MCH) memiliki rentang normal 27 - 31 pg/sell dan mengukur rata-rata kandungan hemoglobin dalam setiap sel darah merah. Jika nilainya di luar batas normal, maka diberi nilai 1; jika dalam batas normal, nilainya 0.
- Mean Corpuscular Hemoglobin Concentration (MCHC) menunjukkan konsentrasi rata-rata hemoglobin dalam setiap unit volume sel darah merah dengan rentang normal 32 - 36 g/dL. Jika nilainya berada di luar rentang normal, maka nilainya 1; jika dalam batas normal, maka nilainya 0.
- Mean Corpuscular Volume (MCV) mengukur volume rata-rata sel darah merah dengan nilai normal antara 80 - 100 fL. Jika nilainya di luar batas normal, maka diberikan nilai 1, sedangkan jika dalam batas normal, nilainya 0.

Aturan ini memastikan bahwa data rekam medis anemia terstandarisasi dan dapat digunakan secara efektif untuk analisis dan pemrosesan lebih lanjut dalam format yang konsisten. Hasil pengkodean biner disajikan dalam Tabel 1.

**Tabel 1. Transformasi biner dataset rekam medis anemia**

| No  | Gender | Hemoglobin | MCH | MCHC | MCV | Anemia |
|-----|--------|------------|-----|------|-----|--------|
| 1   | 1      | 0          | 1   | 1    | 0   | 0      |
| 2   | 0      | 1          | 1   | 1    | 1   | 0      |
| 3   | 0      | 1          | 1   | 1    | 1   | 1      |
| ... | ...    | ...        | ... | ...  | ... | ...    |
| 20  | 0      | 0          | 1   | 1    | 1   | 0      |

Tahapan selanjutnya adalah mengubah data kedalam bentuk qubit. Sebagai contoh, sampel diambil dari dataset nomor 1 dengan kode biner 101100. Kode ini menunjukkan bahwa untuk kondisi berikut: atribut usia  $\geq 45$ , hemoglobin normal, MCH abnormal, MCHC abnormal, dan MCV normal, hasil klasifikasi anemia adalah "tidak". Data dalam Tabel 1 di atas kemudian dikonversi ke dalam bentuk qubit sebagaimana ditampilkan dalam Tabel 2 berikut:

**Tabel 2. Transformasi qubit dataset anemia**

| No  | X1                                     | X2                                     | X3                                     | X4                                     | X5                                     | X6                                     |
|-----|----------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|
| 1   | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ |
| 2   | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ |
| 3   | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ |
| ... | ...                                    | ...                                    | ...                                    | ...                                    | ...                                    | ...                                    |
| 20  | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ |

Tahapan berikutnya adalah melakukan implementasi algoritma K-Medoid, pertama adalah implementasi Kuantum k-medoid dengan menggunakan perhitungan jarak Chebyshev. Medoid awal dan medoid ke-2 yang diberikan pada kedua metode memiliki data yang sama. Hasil yang didapatkan

<http://sistemasi.ftik.unisi.ac.id>



dari pengujian kedua algoritma akan dibandingkan dengan akurasi dalam mencluster anemia. Tabel 3 berisi hasil dari perhitungan dengan Quantum K-Medoid Chebyshev Distance pada Epoch 1.

**Tabel 3. Hasil uji quantum k-medoids epoch-1 menggunakan jarak chebyshev**

| C1                                     | C2                                     | Cluster | Data Real | Description |
|----------------------------------------|----------------------------------------|---------|-----------|-------------|
| $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | 1       | 0         | true        |
| $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 2       | 0         | false       |
| $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 2       | 1         | true        |
| ...                                    | ...                                    | ...     | ...       | ...         |
| $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | 1       | 0         | true        |
| Precision                              |                                        |         |           | 0.4         |
| Recall                                 |                                        |         |           | 0.333333333 |
| F-1 Score                              |                                        |         |           | 0.363636364 |
| accuracy                               |                                        |         |           | 70 %        |

Pada tabel 3, hasil yang diperoleh pada Epoch-1 menunjukkan akurasi sebesar 70% dengan presisi 0.4, recall 0.3333 dan F-1 Score 0.3636. Selanjutnya, perhitungan dilakukan pada Epoch-2 dengan medoid yang diperbarui. Hasil yang diperoleh dari medoid baru dapat dilihat pada Tabel 4.

**Tabel 4. Hasil uji quantum k-medoids epoch-2 menggunakan jarak chebyshev**

| C1                                     | C2                                     | Cluster | Data Real | Description |
|----------------------------------------|----------------------------------------|---------|-----------|-------------|
| $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | 1       | 0         | true        |
| $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 1       | 0         | true        |
| $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 1       | 1         | false       |
| ...                                    | ...                                    | ...     | ...       | ...         |
| $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ | 1       | 0         | true        |
| Precision                              |                                        |         |           | 1           |
| Recall                                 |                                        |         |           | 0.333333333 |
| F-1 Score                              |                                        |         |           | 0.5         |
| Accuracy                               |                                        |         |           | 80 %        |

Pada tabel 4, hasil pengujian menggunakan algoritma K-Medoids kuantum dengan perhitungan Jarak Chebyshev menunjukkan bahwa akurasi yang diperoleh pada Epoch-2 mencapai 80% dengan presisi 1, recall 0.3333 dan F-1 Score 0.5 . Selanjutnya adalah pengujian dengan Quantum K-Medoid dengan Manhattan Distance. Hasil simulasi pengujian algoritma K-Medoids dengan perhitungan Jarak Manhattan menggunakan komputasi kuantum menunjukkan akurasi sebesar 55% pada epoch-1 dan 65% pada epoch-2. Hasil dari pengujian K-medoid Manhattan distance pada epoch-1 dapat dilihat pada tabel 5.

**Tabel 5. Hasil uji quantum k-medoids epoch-1 menggunakan jarak manhattan**

| C1                                     | C2                                     | Cluster | Data Real | Description |
|----------------------------------------|----------------------------------------|---------|-----------|-------------|
| $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ | $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | 2       | 1         | true        |
| $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ | 1       | 1         | false       |
| $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ | 1       | 0         | true        |
| ...                                    | ...                                    | ...     | ...       | ...         |
| $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ | 1       | 1         | false       |
| Precision                              |                                        |         |           | 0.777777778 |
| Recall                                 |                                        |         |           | 0.5         |
| F-1 Score                              |                                        |         |           | 0.608695652 |
| Accuracy                               |                                        |         |           | 55 %        |

Pada tabel 5, hasil dari uji Quantum K-Medoid dengan Manhattan Distance pada Epoch-1 menunjukkan akurasi sebesar 55%, presisi 0.777, recall 0.5 dan F-1 score 0.6086 selanjutnya adalah melakukan ujicoba Epoch -2 dengan medoid baru yang sama dengan medoid Chebyshev Distance yang dapat dilihat pada Tabel 6.

**Tabel 6. Hasil uji quantum k-medoids epoch-2 menggunakan jarak manhattan**

| C1                                     | C2                                     | Cluster | Data Real | Description |
|----------------------------------------|----------------------------------------|---------|-----------|-------------|
| $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 2       | 1         | true        |
| $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ | 1       | 1         | false       |
| $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 1       | 0         | true        |
| ...                                    | ...                                    | ...     | ...       | ...         |
| $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ | 1       | 1         | false       |
| Precision                              |                                        |         |           | 0.888888889 |
| Recall                                 |                                        |         |           | 0.571428571 |
| F-1 Score                              |                                        |         |           | 0.695652174 |
| Accuracy                               |                                        |         |           | 65 %        |

Pada Tabel 6, hasil uji coba dengan Epoch 2 Manhattan menunjukkan presisi 0.888, recall 0.5714, F-1 Score 0.6956 dan akurasi sebesar 65. Untuk nilai cost yang didapatkan oleh kedua pendekatan masing masing adalah Manhattan = 72.629 dan Chebyshev = 41.152. MSE yang didapatkan oleh kedua pendekatan masing masing adalah Manhattan = 0.35 dan Chebyshev = 0.2. Untuk perbandingan lebih lanjut kedua pendekatan yang digunakan dapat dilihat pada Tabel 7.



**Tabel 7. Komparasi pendekatan manhattan dan chebyshev**

| Metrik    | Manhattan | Chebyshev | Performa Algoritma |
|-----------|-----------|-----------|--------------------|
| Akurasi   | 65%       | 80%       | Chebyshev          |
| Presisi   | 0.88      | 1         | Chebyshev          |
| Recall    | 0.57      | 0.33      | Manhattan          |
| F-1 Score | 0.69      | 0.5       | Manhattan          |
| MSE       | 0.35      | 0.2       | Chebyshev          |
| Cost      | 72.629    | 41.152    | Chebyshev          |

Dalam perbandingan antara algoritma Manhattan dan Chebyshev, masing-masing memiliki keunggulan dan kelemahan yang berbeda, baik dari sisi performa maupun aplikasi medis. Dari segi performanya, Chebyshev unggul dalam hal precision dengan nilai 1, yang menunjukkan bahwa semua prediksi positif yang dibuat benar, tanpa ada kesalahan positif (false positive). Ini menunjukkan bahwa Chebyshev sangat efisien dalam memberikan prediksi positif yang tepat, tetapi ini datang dengan penurunan recall yang signifikan (0.33), yang berarti banyak kasus positif yang sebenarnya terlewatkan (false negative). F1-score Chebyshev pun lebih rendah (0.5), menandakan ketidakseimbangan antara precision dan recall. Meskipun Chebyshev memiliki akurasi yang lebih tinggi (80%) dan cost yang lebih rendah, model ini mungkin kurang efektif dalam kasus data yang tidak seimbang, di mana mengabaikan banyak kasus positif bisa merugikan. Di sisi lain, Manhattan menawarkan recall yang lebih tinggi (0.57), artinya model ini lebih baik dalam menemukan kasus positif meskipun dengan sedikit penurunan precision (0.88), yang berarti ada sedikit kesalahan dalam prediksi positif. F1-score Manhattan yang lebih tinggi (0.695) mencerminkan keseimbangan yang lebih baik antara precision dan recall, yang lebih menguntungkan untuk aplikasi di mana false negative lebih berisiko, seperti dalam deteksi penyakit. MSE yang lebih tinggi pada Manhattan menunjukkan bahwa model ini mungkin kurang tepat dalam hal akurasi perhitungan numerik, namun, dalam konteks medis, fokusnya pada recall yang lebih tinggi lebih penting daripada kesalahan perhitungan numerik kecil. Cost Manhattan juga lebih tinggi, menunjukkan bahwa perhitungan jarak lebih mahal secara komputasi, meskipun ini bisa menjadi risiko yang wajar untuk mendapatkan deteksi lebih banyak kasus positif. Dalam aplikasi medis, Manhattan lebih cocok karena recall-nya yang lebih tinggi, yang berarti lebih banyak pasien sakit dapat terdeteksi, meskipun mungkin ada kesalahan dalam mendeteksi beberapa pasien yang sehat. Di dunia medis, lebih baik mendeteksi lebih banyak kasus positif untuk memastikan bahwa tidak ada pasien yang terlewatkan, meskipun ada sedikit kesalahan dalam diagnosis. Di sisi lain, Chebyshev, dengan precision sempurna, lebih cocok untuk aplikasi di mana false positive harus sangat diminimalkan, tetapi dalam konteks medis, ini bisa berbahaya jika terlalu banyak kasus positif yang terlewatkan.

## 5 Kesimpulan

Penelitian ini membandingkan dua perhitungan jarak K-Medoids dengan komputasi kuantum, yaitu perhitungan Jarak Chebyshev dan Jarak Manhattan. Tujuan dari penelitian ini adalah untuk mengetahui perhitungan jarak mana yang lebih optimal dalam mengelompokkan data rekam medis anemia dengan pendekatan kuantum. Secara keseluruhan, pilihan antara Manhattan dan Chebyshev tergantung pada prioritas aplikasi. Untuk aplikasi medis yang membutuhkan deteksi kasus positif sebanyak mungkin, Manhattan adalah pilihan yang lebih ideal, meskipun lebih mahal secara komputasi. Sedangkan Chebyshev lebih efisien dan tepat untuk konteks di mana akurasi dan biaya menjadi prioritas, dan ketika dataset lebih seimbang. Saran untuk penelitian selanjutnya adalah untuk meningkatkan ukuran dataset yang digunakan, karena penelitian ini hanya menggunakan 20 data, yang memiliki proporsi nilai 1 (kasus positif) yang sangat sedikit. Penggunaan dataset yang lebih besar dan lebih seimbang dapat memberikan gambaran yang lebih akurat tentang performa algoritma K-Medoids dengan komputasi kuantum dalam mengelompokkan data rekam medis anemia. Penelitian

<http://sistemasi.ftik.unisi.ac.id>

lebih lanjut juga dapat mengeksplorasi penggunaan algoritma lain atau pendekatan lain dalam komputasi kuantum, seperti Quantum K-Means atau algoritma klasifikasi lainnya, untuk membandingkan hasil yang lebih variatif dan memastikan keandalan model dalam aplikasi medis.

## Referensi

- [1] M. Jaiswal, A. Srivastava, and T. J. Siddiqui, “Machine Learning Algorithms for Anemia Disease Prediction,” in *Lecture Notes in Electrical Engineering*, 2019, pp. 463–469. doi: 10.1007/978-981-13-2685-1\_44.
- [2] C. Bauckhage, R. Sifa, and S. Wrobel, “Adiabatic Quantum Computing for Max-Sum Diversification,” in *Proceedings of the 2020 SIAM International Conference on Data Mining*, Philadelphia, PA: Society for Industrial and Applied Mathematics, 2020, pp. 343–351. doi: 10.1137/1.9781611976236.39.
- [3] A. Naik, E. Yeniaras, G. Hellstern, G. Prasad, and S. K. L. P. Vishwakarma, “From Portfolio Optimization to Quantum Blockchain and Security: A Systematic Review of Quantum Computing in Finance,” *arXiv*, pp. 1–64, Jun. 2023, doi: 10.48550/arXiv.2307.01155.
- [4] U. Awan, L. Hannola, A. Tandon, R. K. Goyal, and A. Dhir, “Quantum Computing Challenges in the Software Industry. A fuzzy AHP-based Approach,” *Information and Software Technology*, Vol. 147, No. February, p. 106896, Jul. 2022, doi: 10.1016/j.infsof.2022.106896.
- [5] T. Baidawi and Solikhun, “A Comparison of Madaline and Perceptron Algorithms on Classification with Quantum Computing Approach,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, Vol. 8, No. 2, pp. 280–287, Apr. 2024, doi: 10.29207/resti.v8i2.5502.
- [6] F. F. C. Silva, P. M. S. Carvalho, and L. A. F. M. Ferreira, “A Quantum Computing Approach for Minimum Loss Problems in Electrical Distribution Networks,” *Scientific Reports*, Vol. 13, No. 1, p. 10777, Jul. 2023, doi: 10.1038/s41598-023-37293-9.
- [7] U. S., “Penerapan Data Mining dengan mengimplementasikan Algoritma K-Means dalam Proses Clustering untuk Pengelompokan Mahasiswa Calon Penerima Beasiswa KIP,” *Building of Informatics, Technology and Science (BITS)*, Vol. 5, No. 1, p. 70–85, Jun. 2023, doi: 10.47065/bits.v5i1.3411.
- [8] B. Hartono, S. Eniyati, and K. Hadiono, “Perbandingan Metode Perhitungan Jarak pada Nilai Centroid dan Pengelompokan Data menggunakan K-Means Clustering,” *Jurnal Sistem Komputer dan Informatika (JSON)*, Vol. 4, No. 3, p. 503, Mar. 2023, doi: 10.30865/json.v4i3.6021.
- [9] M. R. Muttaqin and M. Defriani, “Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa,” *ILKOM Jurnal Ilmiah*, Vol. 12, No. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [10] . F., felix toknady Kesuma, and S. P. Tamba, “Penerapan Data Mining untuk menentukan Penjualan Sparepart Toyota dengan Metode K-Means Clustering,” *Jurnal Sistem Informasi dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, Vol. 2, No. 2, pp. 67–72, Aug. 2020, doi: 10.34012/jusikom.v2i2.376.
- [11] V. A. Ekasetya and A. Jananto, “Klusterisasi Optimal dengan Elbow Method untuk Pengelompokan Data Kecelakaan Lalu Lintas di Kota Semarang,” *Jurnal Dinamika Informatika*, Vol. 12, No. 1, pp. 20–28, Aug. 2020, doi: 10.35315/informatika.v12i1.8159.
- [12] I. N. M. Adiputra, “Clustering Penyakit DBD pada Rumah Sakit Dharma Kerti menggunakan Algoritma K-Means,” *INSERT: Information System and Emerging Technology Journal*, Vol. 2, No. 2, pp. 99–105, Jan. 2022, doi: 10.23887/insert.v2i2.41673.
- [13] Y. Wibowo, “Penerapan Data Mining Dalam Pengelompokan Data Member Card Mitra10 untuk meningkatkan Rewards terhadap Konsumen dengan Metode Fuzzy Subtractive Clustering,” *TIN: Terapan Informatika Nusantara*, Vol. 2, No. 8, pp. 471–475, Jan. 2022, doi: 10.47065/tin.v2i8.993.
- [14] K. Rizal; Aidilof, Hafizh Al Kautsar; , Mukhlis; Nur, “Penerapan Algoritma K-Medoid dalam Perbandingan Daya Serap Akademik Siswa Sekolah Perkotaan dan Sekolah Pedesaan selama Masa Pandemi,” *Jurnal Tekno Kompak*, Vol. 1, No. 1, pp. 6–13, 2021.
- [15] A. Rohmah Zaidah, C. Indira Septiarani, M. Sholikhhatun Nisa, A. Yusuf, and N. Wahyudi, “Komparasi Algoritma K-Means, K-Medoid, Agglomeartive Clustering terhadap Genre

- Spotify,” *Jurnal Ilmiah Ilmu Komputer*, Vol. 7, No. 1, pp. 49–54, 2021, doi: 10.35329/jiik.v7i1.186.
- [16] W. AS, M. K. Aidid, and M. Nusrang, “Pengelompokan Kabupaten/Kota Provinsi Sulawesi Selatan dan Barat berdasarkan Angka Partisipasi Pendidikan SMA/SMK/MA menggunakan K-Medoid dan CLARA,” *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, Vol. 1, No. 3, p. 48, Dec. 2019, doi: 10.35580/variasiunm12899.
- [17] I. Irwan, M. S. Wahyuni, S. Sulaiman, and A. M. Mu’adz, “Analisis K-Medoid untuk Pemetaan Tingkat Pencemaran Udara di Provinsi Sulawesi Selatan,” *Journal of Mathematics Computations and Statistics*, Vol. 5, No. 2, p. 106, Oct. 2022, doi: 10.35580/jmathcos.v5i2.38215.
- [18] G. P. I.R, A. Aziz, and M. P. T.S, “Implementasi Euclidean dan Chebyshev Distance pada K-Medoids Clustering,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 6, No. 2, pp. 710–715, 2022, doi: 10.36040/jati.v6i2.5443.
- [19] F. Fathurrahman, S. Harini, and R. Kusumawati, “Evaluasi Clustering K-Means dan K-Medoid pada Persebaran Covid-19 di Indonesia dengan Metode Davies-Bouldin Index (DBI),” *Jurnal Mnemonic*, Vol. 6, No. 2, pp. 117–128, Oct. 2023, doi: 10.36040/mnemonic.v6i2.6642.
- [20] Y. Miftahuddin, S. Umaroh, and F. R. Karim, “Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, dan Manhattan dalam Penentuan Posisi Karyawan,” *Jurnal Tekno Insentif*, Vol. 14, No. 2, pp. 69–77, Aug. 2020, doi: 10.36787/jti.v14i2.270.
- [21] M. Soni, N. Rahaningsih, and R. Danar Dana, “Komparasi Algoritma K-Means dan K-Medoids Clustering pada Data Penyebaran Kasus HIV di Provinsi Jawa Barat,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 7, No. 6, pp. 3766–3772, Feb. 2024, doi: 10.36040/jati.v7i6.8274.
- [22] H. Ramdan, A. Gunawan, and G. Gunawan, “Analisis Pengaruh Kardiovaskular dalam Kasus Covid-19 terhadap Obesitas menggunakan Metode K-Medoid,” *Indonesian Journal Computer Science*, Vol. 3, No. 1, pp. 16–24, Apr. 2024, doi: 10.31294/ijcs.v3i1.2558.
- [23] S. Wulandari and N. Dwitianti, “Implementasi Algoritma Clustering Partitioning Around Medoid (PAM) dalam Clustering Virus MERS-Cov,” *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, Vol. 5, No. 1, p. 70, Aug. 2020, doi: 10.30998/string.v5i1.6469.
- [24] U. R. Gurning and M. Mustakim, “Penerapan Algoritma K-Means dan K-Medoid untuk Pengelompokkan Data Pasien Covid-19,” *Building of Informatics, Technology and Science (BITS)*, Vol. 3, No. 1, pp. 48–55, Jun. 2021, doi: 10.47065/bits.v3i1.1003.
- [25] K. Saravananathan and T. Velmurugan, “Quality Based Analysis of Clustering Algorithms using Diabetes Data for the Prediction of Disease,” *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, No. 11S2, pp. 448–452, Oct. 2019, doi: 10.35940/ijitee.K1072.09811S219.