Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 25, No. 1, November 2025, pp. 15∼24
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 10/C/C3/DT.05.00/2025
DOI: 10.30812/matrik.v25i1.4180

New Approach K-Medoids Clustering Based on Chebyshev Distance with Quantum Computing for Anemia Prediction

Mochamad Wahyudi[1], Solikhun[2], Lise Pujiastuti[3], Gerhard-Wilhelm Weber[4]
[1]Universitas Bina Sarana Informatika, Jakarta, Indonesia
[2]STIKOM Tunas Bangsa, Medan, Indonesia
[3]STMIK Antar Bangsa, Jakarta, Indonesia
[4]Poznan University of Technology, Poznan, Poland

Article Info

Article history:

Received June 29, 2024
Revised November 04, 2025
Accepted November 17, 2025

Keywords:

Clustering;
Chebyshev Distance;
Data Mining;
K-Medoids;
Quantum Computing;
Quantum Bit.

ABSTRACT

Anemia is a condition where the number of red blood cells or hemoglobin levels is below normal, reducing the blood's ability to carry oxygen, which can lead to symptoms such as fatigue, weakness, and shortness of breath. This study aims to improve the performance of the K-Medoids method by using a quantum computing approach to calculate the Chebyshev Distance to predict anemia. The method used is K-Medoids clustering, with Chebyshev distance calculations and quantum computing. A comparative analysis of these methods focuses on their performance, particularly the accuracy of their test results. This study used a dataset of medical records from patients with anemia. The dataset was taken from Kaggle. This dataset includes five attributes used to predict anemia disease patterns. The dataset was tested using the classical method and the K-Medoids algorithm, both with a quantum-computing approach based on the Chebyshev distance. The results of this study reveal a new alternative model for the K-Medoids algorithm with the Chebyshev Distance calculation influenced by the integration of the quantum computing framework. Specifically, the simulation test results show the same accuracy as the classical K-Medoids method and the K-Medoids method with a quantum computing approach, with Chebyshev Distance calculations, with an accuracy of 80%. The conclusion of this study highlights that the K-Medoids method, implemented with a quantum computing approach and Chebyshev Distance calculations, can be used to predict anemia using clustering.

Corresponding Author:

Solikhun, 081396586078,
Informatics Engineering Study Program,
STIKOM Tunas Bangsa, Medan, Indonesia,
Email: solikhun@amiktunasbangsa.ac.id.

How to Cite:
M. Wahyudi, S. Solikhun, L. Pujiastuti, and G.-W. . Weber, "New Approach K-Medoids Clustering Based on Chebyshev Distance with Quantum Computing for Anemia Prediction", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 25, No. 1, pp. 15-24, November, 2025.

1. INTRODUCTION

Anemia is a medical condition characterized by a lower-than-normal number of red blood cells or hemoglobin levels in the blood, which reduces the blood's ability to carry oxygen throughout the body. This condition can lead to symptoms such as fatigue, weakness, paleness, shortness of breath, dizziness, and a rapid heartbeat. The causes of anemia vary and may include iron deficiency, vitamin B12 or folic acid deficiency, bone marrow disorders, chronic diseases, or blood loss. Anemia can also be classified by type and cause, such as iron-deficiency anemia, megaloblastic anemia, hemolytic anemia, or aplastic anemia. Diagnosis of anemia is typically performed with a complete blood count (CBC) and additional tests based on the suspected cause. However, no previous studies have explored quantum-enhanced K-Medoids clustering on healthcare datasets, specifically for anemia prediction. However, no previous studies have explored quantum-enhanced K-Medoids clustering on healthcare datasets, specifically for anemia prediction.

Data mining is a data analysis process that aims to uncover patterns, trends, and other meaningful information hidden in large datasets. This method involves applying various statistical, mathematical, and artificial intelligence techniques to extract insights that support decision-making and strategic planning across multiple fields, including business and science [1–3]. Clustering is a data mining method that groups a set of objects or data into clusters based on the level of similarity between objects. Objects in one cluster are highly similar, while objects from different clusters exhibit significant differences. The main goal of clustering is to identify data structures and patterns to support further analysis and more accurate decision-making across fields such as marketing, biology, and pattern recognition [4, 5].

The K-medoids algorithm is a variant of K-means [6, 7] that addresses K-means' vulnerability to noise. The K-means algorithm uses the mean as the cluster center, making it naturally susceptible to noise. In contrast, the K-medoids algorithm is less sensitive to outliers because it uses the median to determine the center value, resulting in a more complex computational procedure than K-means. Nevertheless, K-Medoids [8] remain a more powerful tool for clustering large datasets than hierarchical clustering.

Some research on K-Medoids Clustering, namely on public fuel filling stations, concerns public infrastructure PT. Pertamina provides fuel to the broader community. One area in Brebes that seems separate from its parent district is South Brebes. Problems arise when a motorbike driver runs out of fuel and wants to refuel with petrol or diesel, but is far from a gas station. The shortest-route search process in this research uses the A* algorithm. The shortest route testing process is carried out by finding the g(n) and h(n) values from the A-Star algorithm using several different distance calculation methods, namely Euclidean distance [9], Manhattan distance [10], and Chebyshev distance [11]. Then, the results of the three distance calculation methods were compared with those of the haversine calculations. The results of 20 tests at one gas station location showed that Euclidean Distance accounted for 56%, Manhattan Distance for 33%, and Chebyshev Distance for 11% [12].

Several studies have compared clustering methods and distance measures to optimize clustering results across various datasets. A study [13] evaluated seven distance metrics, finding that Euclidean distance with k=9 yielded optimal clustering (DBI = 0.224), although limited data was a constraint. Another study [14] used K-Means and K-Medoids on socioeconomic and health data, identifying countries in need of aid, with K-Means performing better (DBI = 0.095, k = 5). A study [**?**] compared K-Means and K-Medoids on hotspot data, revealing that K-Medoids outperformed K-Means, with Acq_time as the most influential variable, optimal at six clusters. In [15], K-Means clustering was applied to tourism data from Bangkalan Regency, yielding three clusters based on visitor profiles. Meanwhile, [16] compared Euclidean and Chebyshev distances in K-Medoids, finding Euclidean distance produced better clustering. Finally, [17] clustered health centers in Kudus Regency using K-Means with Euclidean and Chebyshev distances, showing Euclidean-based clustering performed better (silhouette coefficient = 0.3902).

This research is based on a previous study [18] that grouped wheat seed data. The research conducted by S. Suraya, M. Sholeh, and D. Andayati, titled “Comparison of Distance Metric in K-Mean Algorithm for Clustering Wheat Grain Datasheet”, explores the impact of different distance metrics on the performance of the K-Means clustering algorithm when applied to wheat grain data. The wheat grain dataset includes various types of wheat data. The goal of this research is to develop a clustering model using the K-means algorithm and compare it with other distance metric algorithms. The dataset was tested with the K-means algorithm for clustering values (k) ranging from k = 2 to k = 6. Additionally, the clustering results from the K-means algorithm were compared with those obtained using the Euclidean, Manhattan, and Chebyshev distance metrics. All testing processes were evaluated to select the best groupings. The evaluation used the Davis-Bouldin method to assess the quality of each clustering result. The smallest Davis-Bouldin index value indicates the best clustering solution, and if the evaluation result is negative, the value is finalized. The research methodology follows the Knowledge Discovery in Database (KDD) framework. The results indicate that using the K-means algorithm with different distance metrics produces varying evaluation values.

The Euclidean, Manhattan, and Chebyshev distance metrics consistently yielded the best clustering value of k = 2. Consequently, the research concludes that the optimal number of clusters for the wheat seed dataset is k = 2. The study [19] by D. P. Sari, S. Ridmadhanti, R. Erda, N. J. Margiyanti, T. Y. Handayani, and R. A. Tarigan, titled “Deteksi Dini Anemia pada Remaja di Pulau Nguan Kecamatan Galang Kota Batam Tahun 2020”, focuses on early detection of anemia among adolescents in Pulau Nguan, Kecamatan Galang, Batam City, in 2020. The research highlights the prevalence of anemia in this demographic and underscores the importance of early intervention to mitigate health risks. By employing community-based screening and educational approaches, the study contributes to public health initiatives that improve adolescent health outcomes in rural and underserved areas.

The gap in this research is the absence of clustering methods with optimal distance-metric approaches applied to health data, particularly for anemia prediction, as previous clustering studies have been limited to non-health data, such as wheat seeds, and have not been used to support disease analysis. The study by Suraya et al. [13] focused solely on selecting the optimal distance metric for the K-Means algorithm on wheat seed data, without considering applications in healthcare. In contrast, the study by Sari et al. [14] relied solely on community-based screening and education for early anemia detection without data analysis or clustering methods. Therefore, this research makes a novel contribution by applying clustering methods (K-Medoids with the Chebyshev distance and quantum-computing optimization) to health data. This topic has not been thoroughly explored for anemia prediction.

The proposed research on anemia prediction using the K-Medoids method with Chebyshev distance calculation and quantum computation differs significantly from previous studies. This research's problem is finding other alternative grouping methods to obtain optimal clusters from anemia data. This research aims to group anemia data using the K-Medoids grouping method with Chebyshev Distance calculations using a Quantum computing approach. The grouping results will be compared with those of the classical K-Medoids method and with K-Medoids using a quantum computing approach. The grouping results will be validated with accurate data from medical records for anemia. The novelty of this research is a new approach to grouping data using the K-Medoids method with Chebyshev Distance calculations and quantum computing to predict anemia. The findings indicate that the proposed method achieves higher clustering accuracy and efficiency than the classical K-Medoids method, particularly when handling large datasets and complex calculations. The research concludes that integrating quantum computing with Chebyshev Distance enhances the K-Medoids method, making it a viable alternative for clustering medical data. This research contributes significantly to the field of data science and healthcare by providing a robust method for grouping medical data, aiding early detection and intervention for anemia. Furthermore, the results can support public and governmental organizations in developing targeted health programs and policies based on accurate clustering of anemia data.

2. RESEARCH METHOD

This study aims to develop and evaluate an anemia prediction model using a novel approach: the K-Medoids method optimized with Chebyshev distance calculation and quantum computing integration, and to compare its performance with the classical K-Medoids method in terms of clustering accuracy. The research stages are illustrated in Figure 1.

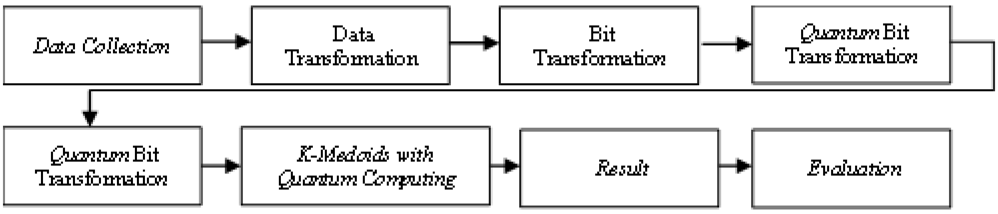

Figure 1. Research Steps

The research begins with data collection, gathering medical records of anemia patients that include key variables such as gender, hemoglobin levels, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and mean corpuscular volume (MCV), all of which are crucial for predicting anemia type and severity. Next, the data are transformed into binary form, with each variable encoded as 0 or 1 depending on whether its value falls within normal ranges, enabling effective grouping. Subsequently, this binary data is converted into quantum bits (qubits) to facilitate processing within a quantum computing framework. The core analysis employs the K-Medoids clustering method using the Chebyshev Distance to group similar data points around medoids as cluster centers. This method is then optimized through quantum computing, applying the same Chebyshev Distance calculation to enhance clustering accuracy. The resulting clusters from both classical and quantum-optimized K-Medoids methods are analyzed to reveal underlying data patterns. Finally, an evaluation phase assesses and compares the performance of both methods in predicting anemia, using accuracy metrics to evaluate the clustering models.

These rules ensure that anemia medical record data is standardized and can be used effectively for further analysis and processing in a consistent format. The results of coding into binary code can be seen in Table 1 below:

Table 1. Anemia Medical Record Dataset

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
4	0	0	1	1	0	0
5	1	0	1	1	0	0
6	0	1	1	1	1	1
7	1	1	1	1	0	1
8	1	1	0	1	0	1
9	0	0	0	1	1	0
10	1	0	1	1	0	0
11	1	0	1	1	0	1
12	0	1	0	1	0	0
13	0	0	1	1	0	0
14	0	0	0	1	1	0
15	0	1	1	1	0	0
16	1	0	1	0	0	0
17	1	0	1	1	0	0
18	1	1	1	1	1	1
19	1	0	1	1	0	0
20	0	0	1	1	1	0

For example, a sample is taken from dataset number 1, namely the binary code 101100; the code has the following meaning: attribute age $>=$ 45, Hemoglobin = normal, MCH = Tidak normal, MCHC = Tidak Normal, MCV = normal, and target (anemia) = no. The data in Table 1 above is changed in qubit form, like in Table 2 below:

Table 2. Qubit Data for Anemia Medical Record

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	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
4	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
5	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
6	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
7	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
8	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
9	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
10	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
11	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
12	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
13	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
14	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
15	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
16	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
17	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
18	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$
19	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
20	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$

The K-Medoids algorithm with Chebyshev distance follows several steps. The first step, Initialization, involves randomly selecting k medoids from the dataset. In the Assignment step, the distance between each data point and each medoid is calculated using the Chebyshev distance. The formula for Chebyshev distance between two points x and y is defined as: $d\_Chebychev(x, y) = Max|x\_i - y\_i|$. Each point is then assigned to the nearest medoid based on this distance. In the Medoid Update step, for each cluster, the total cost (the Chebyshev distance of all points to the medoid) is calculated. A new medoid is selected if the total cost decreases. The process of assignment and medoid update is repeated in the Iteration step until no further changes occur in the medoid, or the maximum number of iterations is reached. The final Output consists of the medoids and the assignment of points to clusters.

To implement the K-Medoids algorithm using Qiskit for Chebyshev distance calculation, the following steps are performed. Environmental Preparation involves installing Qiskit and supporting libraries. In the next step, Define Chebyshev Distance, a function is created to calculate the classical Chebyshev distance. Initialize Medoids randomly selects medoids from the dataset. In the Quantum Assignment step, quantum algorithms are employed to optimize the assignment of points to medoids. Medoid Update is done based on quantum calculations, where the improved formula for Chebyshev distance is: $d_{Chebyshev}(x, y) = Max||x_i > -|y_i > |$. The assignment and medoid update processes are repeated in the Iteration step until convergence is achieved.

3. RESULT AND ANALYSIS

Previous research has identified a gap in the application of clustering methods with optimal distance metrics to health data, particularly for anemia prediction. The study by Suraya et al. [18] focused solely on selecting the optimal distance metric for the K-Means algorithm on wheat seed data, without considering applications in healthcare. Meanwhile, the research by Sari et al. [19] relied only on community-based screening and education for early anemia detection, without utilizing data analysis or clustering methods. Therefore, this study makes a novel contribution by applying the K-Medoids clustering method, optimized using Chebyshev distance calculation and quantum computing, to health data, specifically for anemia prediction, which has not been extensively explored before.

The findings of this research present an alternative model for the K-Medoids clustering method with Chebyshev Distance calculations, namely the K-Medoids method with a quantum computing approach. The attribute values and medoid values are converted into quantum computing. This study clustered anemia medical record data using the K-Medoids and K-Means methods with Chebyshev Distance calculations. The results show the same clustering accuracy, namely 80%. The following are the results of testing epoch-1 and epoch-2 data:

Table 3. K-Medoids Epoch-1 Test Results with Chebyshev Distance

C1	C2	Shortest Distance	Cluster	Data Real	Description
0	1	0	1	0	TRUE
1	0	0	2	0	FALSE
1	0	0	2	1	TRUE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
1	0	0	2	1	TRUE
1	1	1	1	1	FALSE
1	1	1	1	1	FALSE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
0	1	0	1	1	FALSE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
1	0	0	2	0	FALSE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
1	1	1	1	1	FALSE
0	1	0	1	0	TRUE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
Total of Shortest Distance		10	Accuracy		70%

Table 4. K-Medoids Epoch-2 Test Results with Chebyshev Distance

C1	C2	Shortest Distance	Cluster	Data Real	Description
0	1	0	1	0	TRUE
1	1	1	1	0	TRUE
1	1	1	1	1	FALSE
1	1	1	1	0	TRUE

(continued on next page)

Table 2 (continued)

C1	C2	Shortest Distance	Cluster	Data Real	Description
0	1	0	1	0	TRUE
1	1	1	1	1	FALSE
1	0	0	2	1	TRUE
1	1	1	1	1	FALSE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
0	1	0	1	1	FALSE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
1	1	1	1	0	TRUE
0	1	0	1	0	TRUE
1	0	0	2	1	TRUE
0	1	0	1	0	TRUE
1	1	1	1	0	TRUE
Total of Shortest Distance		12	Accuracy		80%

The simulation results for testing the K-Medoids algorithm with quantum computing show 80% accuracy at epoch 2. Following are the results of testing the epoch-1 epoch-2 data:

Table 5. K-Medoids with Quantum Computing Epoch-1 Test Results with Chebyshev Distance

C1	C2	C1 (Decimal)	C2 (Decimal)	Shortest Distance	Cluster	Data Real	Description
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	0	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	0	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	0	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	1	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	0	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	0	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
Total of Shortest Distance				14.14	Accuracy		70%

Table 6. K-Medoids with Quantum Computing Epoch-2 Test Results with Chebychev Distance

C1	C2	C1 (Decimal)	C2 (Decimal)	Shortest Distance	Cluster	Data Real	Description
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE

(continued on next page)

Table 2 (continued)

C1	C2	C1 (Decimal)	C2 (Decimal)	Shortest Distance	Cluster	Data Real	Description
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	1	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	1	FALSE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	1	FALSE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0\\0\end{bmatrix}$	1.41	0	0	2	1	TRUE
$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	0	1.41	0	1	0	TRUE
$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	1.41	1.41	1.41	1	0	TRUE
Total of Shortest Distance				16.97	Accuracy		80%

The findings of this study introduce an alternative model of the K-Medoids method by integrating quantum computing principles into Chebyshev distance calculations. Although both the classical and quantum K-Medoids models achieved 80% accuracy, the quantum-based approach demonstrated superior computational efficiency and scalability, particularly for larger datasets. Visualization graphs comparing classical and quantum K-Medoids clustering results further illustrate the consistency of grouping quality while significantly reducing computation time. These results align with previous studies that highlight the potential of quantum algorithms in optimizing clustering performance. Moreover, compared with other clustering techniques such as K-Means and DBSCAN, the proposed quantum-enhanced K-Medoids model demonstrates superior robustness to non-spherical data distributions, reinforcing its relevance as an efficient and scalable alternative in modern data mining applications.

4. CONCLUSION

This study presents a novel integration of quantum computing principles into the K-Medoids clustering method using Chebyshev distance, offering an alternative computational framework for data mining. Although the experimental results indicate comparable accuracy to the classical K-Medoids method, the proposed quantum-based approach demonstrates the potential for enhanced scalability and computational efficiency. This research advances hybrid quantum–classical data mining models and opens opportunities for further study, such as applying the method to complex health-related datasets or exploring additional quantum optimization techniques to improve clustering performance and stability.

5. ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to STIKOM Tunas Bangsa, Universitas BSI, and STMIK Antar Bangsa for the support. This support has been instrumental in facilitating the successful completion of this research.

6. DECLARATIONS

AI USAGE STATEMENT

During the preparation of this work, the authors used ChatGPT (OpenAI) to improve the language and clarity of the manuscript. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

AUTHOR CONTIBUTION

Conceptualization: Solikhun, Mochamad Wahyudi Methodology: Solikhun, Mochamad Wahyudi Discussion of results: Solikhun, Mochamad Wahyudi, Lise Pujiastuti Original Draft: Solikhun, Mochamad Wahyudi, Lise Pujiastuti Writing Review and Editing: Lise Pujiastuti. Resources: Solikhun, Mochamad Wahyudi, Lise Pujiastuti Supervision: Solikhun, Mochamad Wahyudi Approval of the final text: Solikhun, Mochamad Wahyudi.

FUNDING STATEMENT

This research was supported and funded by STIKOM Tunas Bangsa, Univeritas BSI and STMIK Antar Bangsa. The authors gratefully acknowledge this support, which has contributed to the successful completion of this study.

COMPETING INTEREST

The authors declare that there are no competing interests related to this research, its findings, or its publication.

REFERENCES

[1] X. Gao and G. Li, "A KNN Model Based on Manhattan Distance to Identify the SNARE Proteins," vol. 8, pp. 112 922–112 931, 2020, https://doi.org/10.1109/ACCESS.2020.3003086.

[2] M. Saberi-Karimian, A. Mansoori, M. M. Bajgiran, Z. S. Hosseini, A. Kiyoumarsioskouei, E. S. Rad, M. M. Zo, N. Y. Khorasani, M. Poudineh, S. Ghazizadeh, G. Ferns, H. Esmaily, and M. Ghayour-Mobarhan, "Data mining approaches for type 2 diabetes mellitus prediction using anthropometric measurements," vol. 37, no. 1, p. e24798, January,2023, https://doi.org/10.1002/jcla.24798.

[3] D. Priyanto, B. K. Triwijoyo, D. Jollyta, H. Hairani, and N. G. A. Dasriani, "Data Mining Earthquake Prediction with Multivariate Adaptive Regression Splines and Peak Ground Acceleration," vol. 22, no. 3, pp. 583–592, July,2023, https://doi.org/10.30812/matrik.v22i3.3061.

[4] Y. Liu, X. Yang, S. Zhou, X. Liu, S. Wang, K. Liang, W. Tu, and L. Li, "Simple Contrastive Graph Clustering," vol. 35, no. 10, pp. 13 789–13 800, October,2024, https://doi.org/10.1109/TNNLS.2023.3271871.

[5] J. B. Muñoz, J. Mirocha, S. Furlanetto, and N. Sabti, "Breaking degeneracies in the first galaxies with clustering," vol. 526, no. 1, pp. L47–L55, August,2023, https://doi.org/10.1093/mnrasl/slad115.

[6] M. Faisal, E. M. Zamzami, and Sutarman, "Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance," vol. 1566, no. 1, p. 012112, June,2020, https://doi.org/10.1088/1742-6596/1566/1/012112.

[7] C. Oktarina, K. A. Notodiputro, and I. Indahwati, "Comparison of K-Means Clustering Method and K-Medoids on Twitter Data," vol. 4, no. 1, pp. 189–202, February,2020, https://doi.org/10.29244/ijsa.v4i1.599.

[8] S. Al-Otaibi, V. Cherappa, T. Thangarajan, R. Shanmugam, P. Ananth, and S. Arulswamy, "Hybrid K-Medoids with Energy-Efficient Sunflower Optimization Algorithm for Wireless Sensor Networks," vol. 15, no. 7, p. 5759, March,2023, https://doi.org/10.3390/su15075759.

[9] D. Triyanto, D. Kurniawan, and M. Wahyudi, "Optimization of the Number of Clusters on K-Medoids Using Chebychev and Manhattan on Gold Selling Grouping," vol. 5, no. 3, pp. 2128–2136, 2021.

[10] R. Suwanda, Z. Syahputra, and E. M. Zamzami, "Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K," vol. 1566, no. 1, p. 012058, June,2020, https://doi.org/10.1088/1742-6596/1566/1/012058.

[11] T. Smets, N. Verbeeck, M. Claesen, A. Asperger, G. Griffioen, T. Tousseyn, W. Waelput, E. Waelkens, and B. De Moor, "Evaluation of Distance Metrics and Spatial Autocorrelation in Uniform Manifold Approximation and Projection Applied to Mass Spectrometry Imaging Data," vol. 91, no. 9, pp. 5706–5714, May,2019, https://doi.org/10.1021/acs.analchem.8b05827.

[12] M. A. Fusihan and U. Ghoni, "Perbandingan metode euclidean distance, manhattan distance, chebyshev distance untuk menentukan jarak terpendek spbu di brebes selatan," vol. 3, no. 2, pp. 53–59, December,2023. [Online]. Available: https://www.jurtisi.umbs.ac.id/index.php/jurtisi/article/view/159

[13] R. Buaton and S. Solikhun, "The Application of Numerical Measure Variations in K-Means Clustering for Grouping Data," vol. 23, no. 1, pp. 103–112, November,2023, https://doi.org/10.30812/matrik.v23i1.3269.

[14] D. Hastari, F. Nurunnisa, S. Winanda, and D. D. Aprillia, "Penerapan Algoritma K-Means dan K-Medoids untuk Mengelompokkan Data Negara Berdasarkan Faktor Sosial-Ekonomi dan Kesehatan: Application of K-Means and K-Medoids Algorithms for Grouping Country Data Based on Socio-Economic and Health Factors," pp. 274–281, August,2023. [Online]. Available: https://journal.irpi.or.id/index.php/sentimas/article/view/542

[15] A. Jauhari, D. R. Anamisa, and F. A. Mufarroha, "Analysis of Clusters Number Effect Based on K-Means Method for Tourist Attractions Segmentation," vol. 2406, no. 1, p. 012024, January,2022, https://doi.org/10.1088/1742-6596/2406/1/012024.

[16] G. P. I. Rani, A. Aziz, and M. P. T. Sulistyono, "Implementasi Euclidean dan Chebyshev Distance pada K-Medoids Clustering," vol. 6, no. 2, pp. 710–715, October,2022, https://doi.org/10.36040/jati.v6i2.5443.

[17] D. A. Ilfiana, "Pengklasteran Puskesmas di Kabupaten Kudus Menggunakan Metode K-Means dengan Perbandingan Jarak Euclidean dan Chebyshev," in *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 5, 2022, pp. 787–798.

[18] S. Suraya, M. Sholeh, and D. Andayati, "Comparison of distance metric in k-mean algorithm for clustering wheat grain datasheet," vol. 15, no. 2, pp. 73–83, May,2023, https://doi.org/10.35335/cit.Vol15.2023.408.pp73-83.

[19] D. P. Sari, S. Ridmadhanti, R. Erda, N. J. Margiyanti, T. Y. Handayani, and R. A. Tarigan, "Deteksi Dini Anemia pada Remaja di Pulau Nguan Kecamatan Galang Kota Batam Tahun 2020," vol. 4, no. 1, pp. 1–8, April,2020, https://doi.org/10.52643/pamas.v4i1.767.

[This page is intentionally left blank.]