

Komparasi Jarak Euclidean dan Manhattan Pada Algoritma K-Nearest Neighbor Dalam Mendeteksi Penyakit Diabetes Mellitus

Agustin Ely Rahayu, Abd. Charis Fauzan, Harliana*

Falkutas Eksakta, Program Studi Ilmu Komputer, Universitas Nahdlatul Ulama Blitar, Blitar, Indonesia

Email: ¹Agustinely25@gmail.com, ²abdcharis@unublitar.ac.id, ³*harliana@unublitar.ac.id

Email Penulis Korespondensi: harliana@unublitar.ac.id

Submitted: 28/10/2022; Accepted: 31/12/2022; Published: 31/12/2022

Abstrak—Diabetes Mellitus merupakan penyakit kronis. Penyakit ini disebabkan oleh peningkatan kadar gula darah dalam tubuh, maka bisa menyebabkan penyakit seperti serangan jantung, obesitas, serta ipenyakit imata, ginjal, dan saraf. Deteksi penyakit Diabetes Mellitus ini biasanya dilakukan dengan uji laboratorium, sehingga pasien harus menjalani beberapa tes medis untuk memberikan nilai input ke sistem diagnosa komputerisasi yang terbukti mahal dan waktu antrian yang panjang. Dari permasalahan tersebut dibutuhkan sistem kecerdasan buatan untuk mendiagnosa penyakit ini secara lebih mudah dan cepat. Maka dari itu, peneliti bertujuan menggunakan sistem cerdas untuk menghasilkan akurasi tertinggi dari hasil uji coba klasifikasi dengan menggunakan metode K-Nearest Neighbor (K-NN) dengan jarak Euclidean dan jarak Manhattan. Klasifikasi kelas yang digunakan yaitu iperhitungan ikehamilan, igula idalam idarah, itekanan idarah, iketebalan lipatan ikulit, iinsulin, iberat ibadan, ifungsi isilsilah iketuranan idiabetes, idan iumur. Data penelitian berupa dataset berjumlah 450 dataset dan data tersebut dibagi menjadi dua untuk mengetahui akurasi tertinggi data uji 80% dan 20% untuk data latih. Akurasi tertinggi menggunakan jarak Euclidean yaitu 84% dengan nilai K=5, dan yang kedua jarak Manhattan memiliki akurasi tertinggi yaitu 82% dengan nilai K=7.

Kata Kunci: Diabetes Mellitus; K-Nearest Neighbor (K-NN); Jarak Euclidean; Jarak Manhattan

Abstract—Diabetes Mellitus is a chronic disease. This disease is caused by an increase in blood sugar levels in the body, it can cause diseases such as heart disease, obesity, and eye, kidney, and nerve diseases. Detection of Diabetes Mellitus is usually carried out by laboratory tests, so that patients have to undergo several medical tests to provide input values to a computerized diagnostic system which has proven to be expensive and has long queue times. From these problems, an artificial intelligence system is needed to diagnose this disease more easily and quickly. Therefore, the researcher aims to use an intelligent system to produce the highest accuracy from the results of the classification test using the K-Nearest Neighbor (K-NN) method with Euclidean distance and Manhattan distance. The class classifications used were pregnancy calculations, blood sugar in blood, blood pressure, skin fold thickness, insulin, body weight, diabetes genealogy dysfunction, and age. The research data in the form of datasets amounted to 450 datasets and the data was divided into two to determine the highest accuracy of 80% test data and 20% for training data. The highest accuracy using Euclidean distance is 84% with a value of K=5, and secondly, the Manhattan distance has the highest accuracy of 82% with a value of K=7.

Keywords: Diabetes Mellitus; K-Nearest Neighbor (K-NN); Euclidean Distance; Manhattan Distance

1. PENDAHULUAN

Diabetes Mellitus adalah salah satu permasalahan yang sudah umum pada dunia kesehatan saat ini [1]. Diabetes Mellitus ini disebabkan oleh tingginya gula darah dan menurunnya insulin dalam tubuh [2]. Biasanya penderita diabetes mellitus akan memiliki keluhan seperti cepat lapar, produksi urine berlebihan, rasa haus yang berlebih, penurunan berat badan, dan kesemutan [3]. Penyakit diabetes ini bisa mengakibatkan berbagai macam penyakit lainnya seperti penyakit gagal ginjal, depresi, obesitas, serta penglihatan menurun, saraf, dan lainnya jika tidak ditangani dengan baik dan tepat [4]. Penyakit diabetes ini tergolong penyakit yang paling susah ditangani dan mematikan, oleh karena itu untuk menghindari resiko terjadinya komplikasi bisa dihindari dan dibutuhkan perawatan medis yang tepat agar resiko terjadinya komplikasi bisa dihindari [5]. Dengan adanya penyakit Covid-19 saat ini yang meninggal sejumlah 3.656 per tanggal 13 Juli 2020 [6]. Dan penyakit diabetes ini adalah salah satu penyakit bawaan yang banyak dialami oleh pasien Covid-19. Oleh karena itu, kadar gula yang tinggi dalam tubuh dapat berdampak signifikan pada virus yang menginfeksi manusia, meningkatkan peradangan, dan merusak sistem kekebalan tubuh [7].

Maka dari itu dalam analisis ini dapat membantu tim medis dan masyarakat untuk mempermudah memprediksi penyakit diabetes mellitus. Penyakit diabetes ini rata-rata masyarakat yang memiliki tingkat kesibukan yang padat sehingga mengganggu metabolisme pada tubuh [8]. Dan perubahan gaya hidup yang tidak teratur berdampak terhadap perubahan pola penyakit yang terjadi di masyarakat, salah satunya adalah diabetes mellitus. Sehingga masyarakat dapat menyadari pentingnya pola hidup sehat dan olahraga untuk menghindari resiko penyakit dan mengurangi makanan cepat saji [9]. Dalam deteksi penyakit diabetes mellitus ini biasanya dilakukan dengan uji laboratorium, sehingga pasien harus menjalani beberapa tes medis untuk memberikan nilai input ke sistem diagnosa komputerisasi yang terbukti mahal dan waktu antrian yang panjang. Dari permasalahan tersebut dibutuhkan sistem kecerdasan buatan untuk mendiagnosa penyakit ini secara lebih efektif dan efisien. Sistem cerdas adalah suatu sistem yang mampu memecahkan suatu permasalahan. Sistem cerdas pada komputer ini di terapkan untuk membantu pengambilan keputusan menggunakan data dan metode tertentu untuk memecahkan suatu permasalahan. Penerapan metode dalam sistem cerdas ini adalah metode K-Nearest Neighbor (K-NN)

dengan jarak Euclidean dan Manhattan[10]. Tujuan dalam penelitian ini menerapkan metode KNN dengan jarak Euclidean dan jarak Manhattan untuk mengklasifikasi akurasi deteksi diabetes[11]. Karena metode KNN adalah model pendukung keputusan yang dapat mengklarifikasi data berdasarkan jarak paling dekat [5].

Terdapat beberapa penelitian sebelumnya tentang deteksi diabetes dengan menggunakan metode yang lainnya. Seperti penelitian deteksi diabetes dengan metode K-Means dan C4.5 dengan akurasi berjumlah 72% [12]. Kemudian penelitian untuk klasifikasi transportasi bus menggunakan metode K-NN dengan jarak euclidean dan jarak manhattan dan menghasilkan akurasi sejumlah 84% dengan K=3[13]. Selanjutnya penelitian menggunakan metode penghitungan jarak antar data pada algoritma iK-Nearest iNeighbor dan lvq untuk klasifikasi data dan menghasilkan akurasi sejumlah 76.76% [14]. Berikutnya penelitian menggunakan metode

K – Nearest Neighbor untuk deteksi dini Diabetes Mellitus menghasilkan akurasi 89.33% [15]. Kemudian penelitian data set pada penderita penyakit diabetes dengan menghasilkan akurasi sebesar 39% pada klasifikasi metode K-Nearest Neighbor [16].

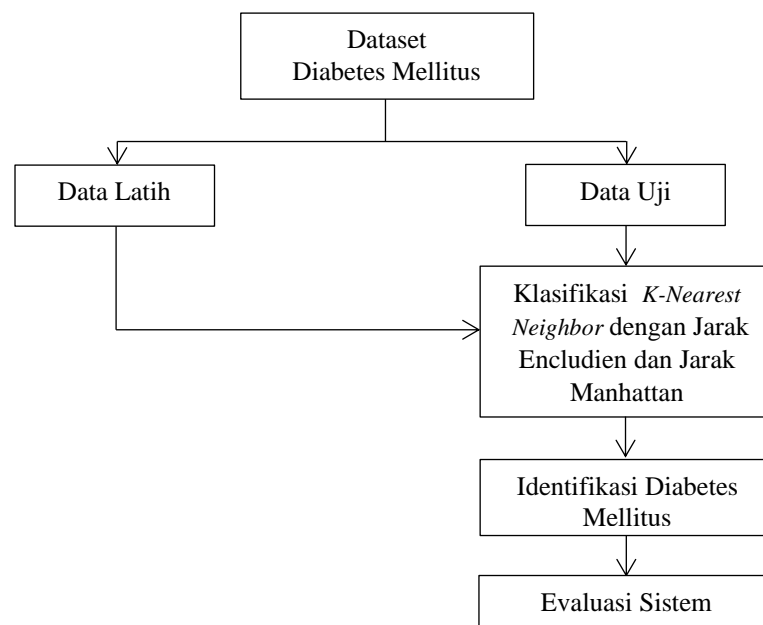
Metode yang digunakan peneliti dalam deteksi diabetes mellitus adalah metode K-Nearest Neighbor. K-Nearest Neighbor merupakan sebuah metode untuk melakukan klasifikasi Diabetes Mellitus berdasarkan dataset yang sebelumnya telah dikenali. Metode iK-Nearest iNeighbor ini termasuk algoritma supervised learning, dimana hasil dari query instance yang baru, dikategorikan berdasarkan sebagian besar pada kategori dari iK-Nearest iNeighbor. Hasil kelas klasifikasi ini diperoleh dari kelas yang paling banyak muncul. Algoritma ini bekerja berdasarkan jarak terpendek dari sampel uji ke sampel pelatihan untuk menentukan K-Nearest Neighbor yang paling dekat. Setelah mengumpulkan K-Nearest Neighbor, mayoritas K-Nearest Neighbor diperoleh dan digunakan sebagai prediksi dari sampel uji. Dekat atau jauhnya biasanya dihitung berdasarkan jarak Euclidean dan jarak Manhattan [16].

Implementasi metode K-Nearest Neighbor pada penelitian ini adalah sistem yang dapat mendeteksi penyakit diabetes dengan melakukan komputasi K-Nearest Neighbor dengan mengacu pada proses komputasi dataset. Kemudian menggunakan jarak Euclidean dan jarak Manhattan untuk menghitung nilai terdekat. Penelitian ini bertujuan untuk mendapatkan hasil akurasi tertinggi dari klasifikasi diabetes mellitus menggunakan metode K-Nearest Neighbor dengan jarak Euclidean dan jarak Manhattan. Klasifikasi yang digunakan adalah positif dan negatif. Tujuan penting dari aplikasi data mining ini dalam mendiagnosa adalah mendeteksi penyakit dan mengidentifikasi faktor resiko. Maka dapat menjadi dasar paling penting untuk pemilihan metode pengobatan yang tepat [17].

2. METODOLOGI PENELITIAN

2.1 Alur Sistem

Gambar 1 menunjukkan alur sistem skrining diabetes. Dimulai, dengan membagi dua bagian data set diabetes: data pelatihan dan data uji. Kemudian hitung jarak euclidean dan manhattan dan klasifikasikan berdasarkan jarak terdekat menggunakan metode K-Nearest Neighbor. Dan evaluasi sistem adalah tahap terakhir.



Gambar 1. Alur Sistem

2.2 Data Set

Penelitian ini diawali dengan langkah pengumpulan data terlebih dahulu. Data yang digunakan sebagai dataset merupakan data diabetes mellitus silsilah riwayat diabetes tipe tiga, sumber data berasal dari Laboratory The Johns Hopkins University Johns Hopkins Road Laurel dengan melalui akses link berikut : <https://datahub.io/machine-learning/diabetes#readme> dataset penelitian berupa dataset berjumlah 450 data et dan data tersebut dibagi menjadi dua untuk mengetahui akurasi tertinggi data uji 80% dan 20% untuk data latih. Tabel 1 adalah data set diabetes mellitus.

Tabel 1. Data set

Kehamilan	Glukosa Plasma	Tekanan Darah	Ketebalan Lipatan	Insulin	Berat Badan	iFungsi Silsilah Diabetes Tipe 3	Umur	Hasil	Kehamilan
6	148	72	35	0	33.6	0.627	50	1	6
1	85	66	29	0	26.6	0.351	31	0	1
8	183	64	0	0	23.3	0.672	32	1	8
1	89	66	23	94	28.1	0.167	21	0	1
0	137	40	35	168	43.1	2.288	33	1	0
5	116	74	0	0	25.6	0.201	30	0	5
3	78	50	32	88	31	0.248	26	1	3
10	115	0	0	0	35.3	0.134	29	0	10
.
.
.
.
2	99	70	16	44	20.4	0.235	27	0	2
6	103	72	32	190	37.7	0.324	55	0	6
5	111	72	28	0	23.9	0.407	27	0	5
8	196	76	29	280	37.5	0.605	57	1	8
5	162	104	0	0	37.7	0.151	52	1	5

2.3 Perhitungan Jarak Euclidean

Jarak Euclidean adalah metode perhitungan jarak yang digunakan untuk mengukur dua titik dalam ruang Euclidean yang meliputi dua, tiga, atau lebih bidang Euclidean [18]. Untuk mengukur kemiripan data menggunakan rumus jarak Euclidean, menggunakan persamaan (1).

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{1}$$

Keterangan:

d = jarak antara x dan y

x = data pusat cluster

y = data untuk atribut i

i = data individu

n = data total

xi = data center cluster ke- i

yi = data individu ke- i.

2.4 Perhitungan Jarak Manhattan

Jarak Manhattan digunakan untuk menghitung selisih mutlak antara koordinat pasangan benda[18]. Menggunakan persamaan (2).

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \tag{2}$$

Keterangan:

d = jarak antara x dan y

x = data pusat cluster

y = data tentang atribut.

2.5 Evaluasi Sistem

Confusion matrix digunakan untuk evaluasi sistem dalam penelitian ini. Karena confusion matrix memungkinkan kita untuk mengukur klasifikasi menggunakan KNN. klasifikasi diabetes ini dapat diketahui berdasarkan sejumlah accuracy, precision, dan recall/sensitivity.

Tabel 2. Confusion Matrix

	TP	FP
	FN	TN

Accuracy digunakan untuk mengetahui jumlah presentase dari data set diabetes mellitus yang diprediksi positif dan negatif dengan persamaan (3).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \tag{3}$$

Menggunakan precision dapat diketahui persentase dari data set diabetes mellitus yang benar saja.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

Sedangkan Recall dapat diketahui persentase diabetes mellitus yang diprediksi negatif dibandingkan dengan keseluruhan dataset diabetes mellitus yang sebenarnya diprediksi positif. Recall digunakan agar hasil klasifikasi ideal[19].

$$\text{Recall} = \frac{TP}{TP+FN} \tag{5}$$

3. HASIL DAN PEMBAHASAN

Seperti yang sudah dijelaskan, langkah yang harus dilakukan dalam penelitian ini adalah memisahkan data latih dan data uji. Peneliti menggunakan 450 data diabetes dan 80-20% dari data uji adalah data latih. Langkah selanjutnya adalah menerapkan metode K-NN menggunakan jarak euclidean dan manhattan untuk memilih nilai K dalam penelitian ini yaitu K = 3, 5, dan 7 [16]. Hasil dari klasifikasi pada Jarak Euclidean pada Tabel 3.

a. Confusion Matrix pada Jarak Euclidean

Tabel 3. Penggunaan Kelas Sesuai K pada Jarak Euclidean

	K=3	K=5	K=7
	0	0	0
	1	1	1
	1	1	1
	0	0	0
	0	0	0
	.	.	.
	1	1	1
	1	1	1
	1	1	1
	1	1	1
	0	0	0

Selanjutnya Tabel 4 akan menunjukkan dalam bentuk confusion matrix pada k=3.

Tabel 4. Confusion Matrix K=3 pada Jarak Euclidean

Jumlah 90	Sesuai	Tidak Sesuai
Positif	18	8
Negatif	53	11

Dari hasil klasifikasi pada Tabel 4 yang merupakan, confusion matrix K=3 pada jarak euclidean di peroleh hasil yang diujikan. Dari kumpulan data tersebut, 18 data positif diprediksi valid, 8 data positif diprediksi negatif, 53 data bernilai negatif diprediksi valid, dan 11 negatif diperkirakan tidak sesuai. Hasil perhitungan dalam jarak euclidean pada metode K-NN menggunakan K=3 yang berarti diranking untuk menentukan jarak terdekat dan terkecil yang akan muncul dalam klasifikasi K-NN, seperti pada Tabel 3.

Tabel 5. Confusion Matrix K=5 pada Jarak Euclidean

Jumlah 90	Sesuai	Tidak Sesuai
Positif	20	6
Negatif	56	8

Tabel 5 menunjukkan hasil dari Confusion Matrix K=5 pada Jarak Euclidean yang diperoleh 20 data positif yang sesuai, 6 data positif yang tidak sesuai, 56 data negatif yang di prediksi valid, dan ada data yang di prediksi tidak sesuai yaitu 8 data bernilai negatif. Hal ini dikarenakan perhitungan jarak euclidean pada metode K=NN menggunakan K=5. Artinya diranking untuk menentukan jarak terdekat dan terkecil yang muncul pada klasifikasi K-NN.

Tabel 6. Confusion Matrix K=7 pada Jarak Euclidean

Jumlah 90	Sesuai	Tidak Sesuai
Positif	17	6
Negatif	56	11

Dari Tabel 6, menunjukkan hasil dari Confusion Matrix K=7 pada Jarak Euclidean yang diprediksi 17 data positif yang sesuai, 6 data positif yang tidak sesuai, 56 data negatif yang diprediksi sesuai, dan 11 data negatif yang tidak sesuai. Hasil perhitungan jarak euclidean pada metode K=NN menggunakan K=7, maka akan diranking untuk mencari jarak terdekat dan terkecil, seperti pada Tabel 3.

b. Confusion Matrix pada Jarak Manhattan

Tabel 7. Penggunaan Kelas Sesuai K pada Jarak Manhattan

	K=3	K=5	K=7
	0	0	0
	1	1	1
	1	1	1
	0	0	0
	0	0	0
	.	.	.
	0	1	1
	0	0	0
	1	1	1
	1	1	1
	0	0	0

Pada Tabel 8 confusion matrix ini menunjukkan pada k=3.

Tabel 8. Confusion Matrix K=3 pada Jarak Manhattan

Jumlah 90	Sesuai	Tidak Sesuai
Positif	16	7
Negatif	56	11

Dari hasil Confusion Matrix K=3 pada Jarak Manhattan diperoleh diprediksi data positif 16 yang sesuai, 7 data positif yang tidak sesuai, 56 data negatif yang diprediksi tidak sesuai, dan 11 data negatif yang diprediksi tidak sesuai. Hasil perhitungan dalam jarak manhattan pada metode K-NN menggunakan K=3 yang berarti diranking untuk menentukan jarak terdekat dan terkecil yang akan muncul dalam klasifikasi K-NN, seperti pada Tabel 7.

Tabel 9. Confusion Matrix K=5 pada Jarak Manhattan

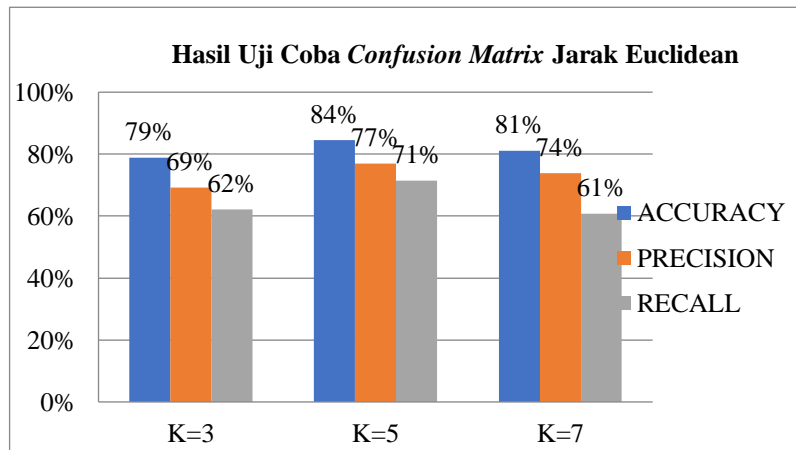
Jumlah 90	Sesuai	Tidak Sesuai
Positif	17	8
Negatif	54	11

Tabel 9 merangkum 17 data positif yang valid, 8 data positif yang diprediksi tidak sesuai, 54 data negatif yang diprediksi valid, dan 11 data negatif yang tidak sesuai. Hal ini dikarenakan perhitungan jarak manhattan pada metode K=NN menggunakan K=5. Artinya diranking untuk menentukan jarak terdekat dan terkecil yang muncul pada klasifikasi K-NN.

Tabel 10. Confusion Matrix K=7 pada Jarak Manhattan

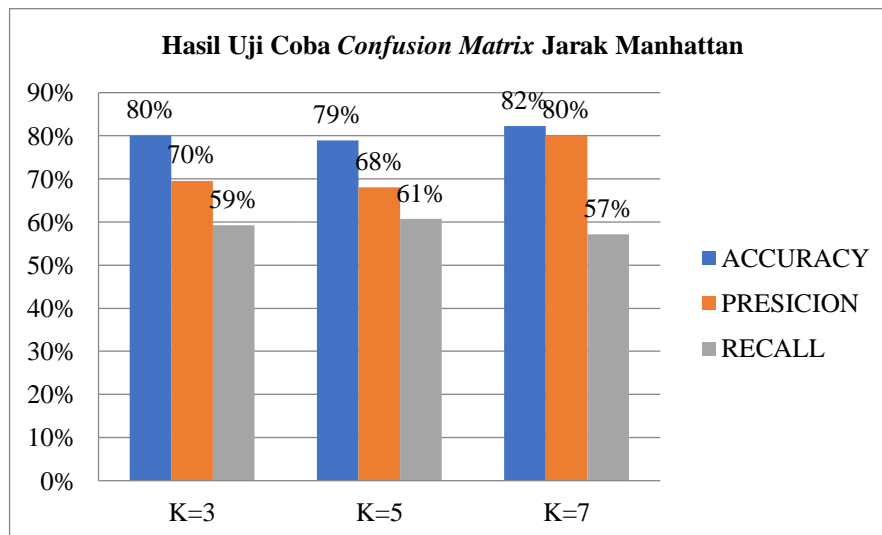
Jumlah 90	Sesuai	Tidak Sesuai
Positif	16	4
Negatif	58	12

Hasil dari Confusion Matrix K=7 pada Jarak Manhattan menunjukkan 16 data positif yang sesuai, 4 data yang positif tidak sesuai, data yang di prediksi sesuai ada 58 data bernilai positif, dan diprediksi 12 data yang bernilai negatif yang tidak sesuai. Hasil perhitungan jarak manhattan pada metode K=NN menggunakan K=7, maka akan diranking untuk mencari jarak terdekat dan terkecil, seperti pada Tabel 7. Selanjutnya adalah uji coba hasil tersebut ke dalam bentuk confusion matrix untuk memperoleh Accuracy, Precision, Recall,



Gambar 2. Hasil Uji Coba Confusion Matrix Jarak Euclidean

Hasil uji coba pada Confusion Matrix dengan jarak Euclidean pada Gambar 2, memperoleh hasil yaitu K=3 Accuracy 79%, Precision 69% ,dan Recall 62%
 K=5 Accuracy 84%, Precision 77% ,dan Recall 71%
 K=7 Accuracy 81%, Precision 74% ,dan Recall 61% .



Gambar 3. Hasil Uji Coba Confusion Matrix Jarak Manhattan

Hasil uji coba pada Confusion Matrix dengan jarak Euclidean pada Gambar 2, memperoleh hasil yaitu K=3 Accuracy 80%, Precision 70% ,dan Recall 59%
 K=5 Accuracy 79%, Precision 68% ,dan Recall 61%
 K=7 Accuracy 82%, Precision 80% ,dan Recall 57%

Hasil uji coba yang mempunyai nilai K yang terbaik pada confusion matrix jarak euclidean adalah K= 5 yang memiliki akurasi yang terbanyak dibandingkan dengan K=3 hanya memiliki akurasi sebanyak 79%, dan K=7 memiliki akurasi 81%. Nilai K terbaik pada confusion matrix jarak euclidean adalah K=7 yang memiliki akurasi 82% dibandingkan dengan K=3 yang hanya memiliki akurasi sebesar 80%, dan K=5 memiliki akurasi 79% saja. Nilai K tertinggi confusion matrix untuk jarak manhattan adalah K=7 dengan akurasi 82% dibandingkan K=3 dengan akurasi hanya 80% dan K=5 dengan akurasi hanya 79%.

4. KESIMPULAN

Berdasarkan hasil uji coba pada metode KNN, dapat diperoleh hasil yaitu pada jarak euclidean memiliki hasil K=3 Accuracy 79%, Precision 69% ,dan Recall 62%, K=5 Accuracy 84%, Precision 77% ,dan Recall 71%, K=7

Accuracy 81%, Precision 74% ,dan Recall 61%. Sedangkan pada jarak manhattan memiliki hasil K=3 Accuracy 80%, Precision 70% ,dan Recall 59%, K=5 Accuracy 79%, Precision 68% ,dan Recall 61%, K=7 Accuracy 82%, Precision 80% ,dan Recall 57%. Dari hasil tersebut uji coba pada jarak euclidean memiliki akurasi yang paling tinggi yaitu 84% pada K=5 dibandingkan dengan jarak manhattan.

REFERENCES

- [1] V. Agustina, M. I. Tekege, F. Carolin, A. D. Wulandari, A. Weya, and O. G. C. Lampongajo, “Deteksi Dini Penyakit Diabetes Melitus,” *J. Magistrorum Sch.*, vol. 02, no. 02, pp. 300–309, 2021.
- [2] K. Saxena, Z. Khan, and S. Singh, “Diagnosis of Diabetes Mellitus using K Nearest Neighbor Algorithm,” *Int. J. Comput. Sci. Trends Technol.*, vol. 2, no. 4, pp. 36–43, 2014.
- [3] H. Bhatt, S. Saklani, and K. Upadhayay, “Anti-oxidant and anti-diabetic activities of ethanolic extract of *Primula Denticulata* Flowers,” *Indones. J. Pharm.*, vol. 27, no. 2, pp. 74–79, 2016, doi: 10.14499/indonesianjpharm27iss2pp74.
- [4] S. Kusumadewi, “Aplikasi Informatika Medis untuk Penatalaksanaan Diabetes Melitus Secara Terpadu,” *Semin. Nas. Apl. Teknol. Inf.*, vol. 2009, no. Snti, pp. C-22-C-27, 2009.
- [5] F. M. Hana, “Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5,” *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 4, no. 1, pp. 32–39, 2020, doi: 10.47970/siskom-kb.v4i1.173.
- [6] Infodatin 2020, “Infodatin-2020-Diabetes-Melitus.pdf.”
- [7] I. Hikmawati and R. Setiyabudi, “Hipertensi Dan Diabetes Militus Sebagai Penyakit Penyerta Utama Covid-19 Di Indonesia Hypertension and Diabetes Mellitus As Covid-19 Comorbidities in Indonesia,” *Pros. Semin. Nas. Lppm Ump*, vol. 0, no. 0, pp. 95–100, 2020, [Online]. Available: <https://semnaslppm.ump.ac.id/index.php/semnaslppm/article/view/224/219%0Ahttps://semnaslppm.ump.ac.id/index.php/semnaslppm/article/view/224>.
- [8] J. I. Kesehatan and S. Husada, “Early Detection of Diabetes Mellitus Risk in Stikes Megarezky Makassar Teaching Staff,” *Juni*, vol. 11, no. 1, pp. 540–547, 2020, doi: 10.35816/jiskh.v10i2.343.
- [9] E. Hasmin et al., “Sistem Pakar Prediksi Penyakit Diabetes Menggunakan Metode K-NN Berbasis Android,” vol. 8, no. 2, pp. 359–370, 2022.
- [10] A. Y. Pratama and S. Yunita, “Komparasi Metode Weighted Product (WP) Dan Simple Additive Weighting (SAW) Pada Sistem Pendukung Keputusan Dalam Menentukan Pemberian Beasiswa,” vol. 4, no. September, pp. 12–24, 2022, doi: 10.30865/json.v4i1.4593.
- [11] N. Hidayati and A. Hermawan, “K-Nearest Neighbor (K-NN) algorithm with Euclidean and Manhattan in classification of student graduation,” vol. 2, no. 2, pp. 86–91, 2021.
- [12] A. Prasatya, R. R. A. Siregar, and R. Arianto, “Penerapan Metode K-Means Dan C4.5 Untuk Prediksi Penderita Diabetes,” *Petir*, vol. 13, no. 1, pp. 86–100, 2020, doi: 10.33322/petir.v13i1.925.
- [13] R. K. Dinata, H. Akbar, and N. Hasdyna, “Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 104–111, 2020, doi: 10.33096/ilkom.v12i2.539.104-111.
- [14] A. K. Dan, L. V. Q. Untuk, and K. Data, “Analisis perbandingan pada metode penghitungan jarak antar data pada algoritma k-nn dan lvq untuk klasifikasi data,” 2020.
- [15] Rzki Mutiara Sari, “Aplikasi Deteksi Dini Diabetes Mellitus Menggunakan Modified K – Nearest Neighbor,” 2018.
- [16] A. Maulida, “Penerapan Metode Klasifikasi K-Nearest Neighbor Pada Dataset Penderita Penyakit Diabetes,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020.
- [17] S. Widodo and H. Brawijaya, “Clustering Kanker Serviks Berdasarkan Perbandingan Euclidean dan Manhattan Menggunakan Metode K-Means,” vol. 5, no. April, pp. 687–694, 2021, doi: 10.30865/mib.v5i2.2947.
- [18] M. Nishom, “Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square,” *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.
- [19] A. P. Agustin, A. C. Fauzan, and Harliana, “Implementasi K-Nearest Neighbor Dengan Jarak Minkowski Untuk Deteksi Dini Covid-19 Pada Citra Ct-Scan Paru - Paru,” *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 4, no. 1, pp. 23–30, 2022.