

Penerapan Metode *K-Nearest Neighbor* Untuk Sistem Pendukung Keputusan Identifikasi Penyakit *Diabetes Melitus*

Umikulsum Indah Lestari^{*1}, Anis Yusrotun Nadhiroh², Cahyuni Novia³

^{1,2,3} Universitas Nurul Jadid; Paiton-Probolinggo, telp/fax : 0888 30 77077

³Jurusan Teknik Informatika, Fakultas Teknik, UNUJA, Probolinggo

e-mail: ^{*1}umikulsumindah@gmail.com, ²ayusrotun@gmail.com, ³vhie771108@gmail.com

Abstrak

Diabetes melitus merupakan penyakit tidak menular (PTM) yang telah menyebar ke seluruh dunia. Faktor yang dapat menyebabkan seseorang menderita diabetes melitus diantaranya tekanan darah tinggi, obesitas, riwayat penyakit diabetes melitus dalam keluarga, usia, serta pola hidup dan diet tidak sehat. Faktor lain yang menjadi pemicu tingginya tingkat kematian akibat diabetes melitus adalah lambatnya diagnosa karena jumlah tenaga medis yang sedikit terutama di kota kecil. Beberapa uji klinis harus dilakukan untuk mengetahui seseorang terjangkit penyakit diabetes melitus atau tidak, proses uji klinis tersebut memakan waktu yang tidak sebentar. Berdasarkan permasalahan tersebut, maka dibuat sebuah sistem pendukung keputusan untuk identifikasi penyakit diabetes melitus dengan menggunakan metode klasifikasi K-Nearest Neighbor (KNN) dengan 8 variabel yaitu jumlah kehamilan, kadar glukosa, tekanan darah, ketebalan lipatan kulit trisep, insulin, index masa tubuh (IMT), riwayat penyakit diabetes mellitus dalam keluarga, serta usia. Berdasarkan implementasi dan pengujian, dengan nilai $k=23$ didapatkan tingkat akurasi sebesar 0.96 atau 96%, tingkat akurasi tersebut dinilai cukup tinggi sehingga penelitian ini dinilai telah berhasil menerapkan metode KNN untuk sistem pendukung keputusan identifikasi penyakit diabetes melitus sejak dini.

Kata kunci : *K-Nearest Neighbor, identifikasi, diabetes melitus.*

Abstract

Diabetes mellitus is a non-communicable disease that has spread throughout the world. Factors that can cause a person to suffer from diabetes mellitus include high blood pressure, obesity, history of diabetes mellitus in the family, age, as well as unhealthy lifestyle and diet. Another factor that triggers the high rate of death due to diabetes mellitus is the delay in diagnosis due to the small number of medical personnel, especially in small towns. Several clinical trials must be carried out to determine whether someone has diabetes mellitus or not; and the clinical trial process takes a long time. Based on these problems, a decision support system was made for the identification of diabetes mellitus using the K-Nearest Neighbor classification method with 8 variables, namely the number of pregnancies, glucose levels, blood pressure, triceps skinfold thickness, insulin, body mass index (BMI), family history of diabetes mellitus, and age. Based on implementation and testing, with a value of $k = 23$, an accuracy rate of 0.96 or 96% is obtained, the accuracy level is considered high enough so that this study is considered to have succeeded in implementing the KNN method for a decision support system for early identification of diabetes mellitus.

Keywords: *K-Nearest Neighbor, identification, diabetes melitus*



1. PENDAHULUAN

Diabetes melitus merupakan penyakit tidak menular (PTM) yang telah menyebar ke seluruh dunia. Diabetes melitus mempengaruhi lebih dari 415 juta orang pada tahun 2015, dan angka tersebut diperkirakan akan meningkat menjadi 642 juta pada tahun 2040 [1]. Diabetes melitus dapat disebabkan oleh berbagai variabel, antara lain tekanan darah tinggi, obesitas, riwayat keluarga yang terkena diabetes melitus, usia, serta gaya hidup dan makanan yang tidak sehat [2]. Gejala awal diabetes melitus memang seringkali tidak terlihat. Oleh sebab itu, banyak orang baru menyadari adanya diabetes melitus ketika telah mengalami komplikasi [3] diabetes melitus yang tidak terkontrol dapat menyebabkan komplikasi jangka panjang, merusak pembuluh darah dan saraf serta organ penting didalam tubuh bahkan kematian [4] Faktor lain yang menjadi pemicu tingginya tingkat kematian akibat diabetes mellitus adalah lambatnya diagnosa pada pasien karena terbatasnya jumlah tenaga medis, utamanya kota-kota kecil. Saat ini, telah banyak rumah sakit yang menggunakan database untuk mengumpulkan dan menyimpan data terkait diabetes melitus, namun data yang dikumpulkan belum dapat digunakan secara optimal. Beberapa uji klinis harus dilakukan untuk menentukan apakah seseorang menderita diabetes melitus, proses uji klinis tersebut memakan waktu yang tidak sebentar. [5]

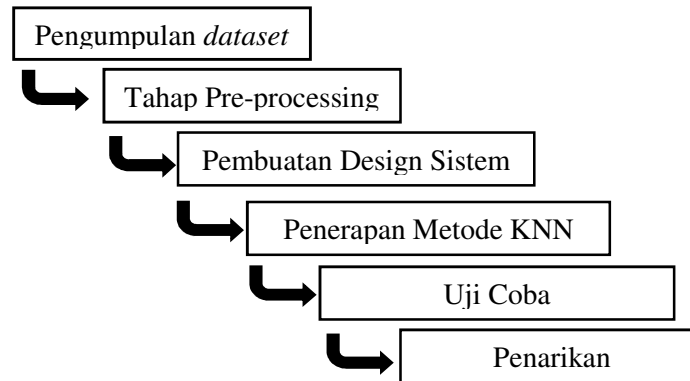
Terdapat beberapa penelitian yang telah dilakukan oleh beberapa peneliti sebelumnya, terkait identifikasi penyakit *diabetes melitus*. Penelitian yang dilakukan oleh Hadli Fadli Santoso, Eko Darwiyanto, dan Untari Novia Wisesty dengan judul “Analisis dan Implementasi Metode Certainty Factor Pada Sistem Pendukung Pengambilan Keputusan untuk Penyakit Diabetes Melitus”. Metode yang dipergunakan dalam penelitian ini ialah metode Certainty Factor. Penelitian ini menghasilkan SPPK diabetes melitus dengan akurasi 80%, SPPK ini dinilai sangat membantu dalam proses diagnosa diabetes melitus [6]. Penelitian kedua dilakukan oleh Risky Aswi Ramadhani dan Ratih Kumalasari Niswatin dengan judul “Sistem Diagnosa Diabetes Menggunakan Metode K-NN”. Data yang dipergunakan pada saat penelitian ini sebanyak 24 data latih dan 7 data uji, mencakup enam variabel: di malam hari sering buang air kecil, luka/goresan sulit sembuh, mudah lelah, gusi berdarah, penurunan cepat dan tiba-tiba dari berat badan, dan keinginan kuat untuk makan dan minum, serta menggunakan nilai $k=3$. Hasil uji akurasi metode KNN mencapai 85,71% [7]. Penelitian ketiga dilakukan oleh Noviandi dengan judul “Implementasi Algoritma Decision Tree C4.5 Untuk Prediksi Penyakit Diabetes”. Penelitian ini bertujuan untuk membangun model prediksi kemungkinan seseorang terjangkit penyakit diabetes melitus menggunakan algoritma C4.5. Database Pima Indians Diabetes (PPID) digunakan untuk membuat prediksi yang berjumlah 768 data dengan delapan variabel yaitu jumlah kehamilan, kadar glukosa, tekanan darah, ketebalan lipatan kulit trisep, kadar insulin, index masa tubuh (IMT), riwayat diabetes melitus dalam keluarga, serta usia. Akurasi hasil prediksi sebesar 70.32% [8].

Berdasarkan latar belakang yang telah diuraikan, maka saat ini dibutuhkan sistem pendukung keputusan (SPK) untuk identifikasi penyakit diabetes melitus. (SPK) adalah sistem yang menghubungkan informasi terkomputerisasi dengan pengguna atau pengambil keputusan[9]. Pada penelitian saat ini mengajukan topik yang serupa dengan metode penelitian yang berbeda. Penelitian ini menggunakan suatu teknik dalam data mining yaitu klasifikasi K-Nearest Neighbor (KNN)dengan 8 variabel yaitu jumlah kehamilan,kadar gula darah (glukosa), tekanan darah, ketebalan lipatan kulit trisep, kadar insulin, index masa tubuh (IMT), riwayat penyakit diabetes melitus dalam keluarga, serta usia. Metode klasifikasi K-Nearest Neighbor (KNN)merupakan salah satu metode yang digunakan untuk mengklasifikasikan objek berdasarkan data latih dengan jarak terdekat[10]. Sistem pendukung keputusan ini diharapkan dapat menghasilkan tingkat akurasi yang jauh lebih tinggi dari penelitian sebelumnya, sehingga

diharapkan dapat mencegah kemungkinan terburuk yang akan terjadi akibat penyakit diabetes melitus.

2. METODE PENELITIAN

Model penelitian yang dirancang pada penelitian ini bertujuan untuk memudahkan proses identifikasi penyakit diabetes melitus.



Gambar 1. Rancangan Penelitian

2.1 Pengumpulan Dataset

Dataset seputar penyakit diabetes melitus berasal dari data publik Pima Indians Diabetes Database. Dataset tersebut terdiri dari 8 variabel yaitu jumlah kehamilan, kadar gula darah (glukosa), tekanan darah, ketebalan lipatan kulit trisep, kadar insulin, index masa tubuh (IMT), riwayat diabetes melitus dalam keluarga, serta usia. Selain itu, peneliti juga melakukan wawancara dengan seorang dokter untuk menggali informasi lebih dalam terkait penyakit diabetes mellitus.

2.2 Tahap Pre-processing

Analisis dataset mengungkapkan bahwa tidak semua variabel memiliki data atribut yang lengkap, dan bahwa kelengkapan atribut memiliki dampak yang signifikan terhadap hasil klasifikasi. Jumlah atribut data tidak lengkap pada masing – masing variable yaitu:

Tabel 1. Data Atribut yang Hilang

No	Variabel	Jumlah Data Hilang
1	Jumlah kehamilan	110
2	Kadar gula darah	5
3	Tekanan darah	35
4	Ketebalan lipatan kulit trisep	227
5	Insulin	374
6	Index Massa Tubuh (IMT)	11
7	Riwayat diabetes dalam keluarga	0
8	Usia	0
9	Hasil prediksi pasien	0

Untuk mengatasi ketidaklengkapan atribut data dari masing-masing variabel tersebut, maka diberlakukan aturan: 1) Nilai nol dalam variabel jumlah kehamilan diasumsikan bahwa pasien belum pernah melahirkan. 2) Nilai nol dalam variabel kadar gula/glukosa, tekanan darah ketebalan lipatan kulit, insulin dan index masa tubuh dihilangkan.

Setelah dilakukan penanganan sesuai dengan kedua aturan yang diatas, didapat 394 data dari 768 data yang siap diolah ketahap selanjutnya. 394 data tersebut kemudian dibagi menjadi data uji dan data latih dengan perbandingan 1:7 atau 50 data uji dan 344 data latih.

2.3 Pembuatan Design Sistem

Pembuatan design sistem dimaksudkan untuk mempermudah pengguna sistem dalam pengoperasian sistem pendukung keputusan yang dibuat. Design sistem pendukung keputusan ini dibuat dengan menggunakan QtDesigner.

2.4 Penggunaan Metode KNN

Penggunaan metode KNN dilakukan pada dua kali percobaan, pertama pada sistem yang dibuat serta pada perhitungan dengan menggunakan excel, perhitungan dengan excel dilakukan untuk melihat hasil perhitungan jarak euclidean yang dihasilkan. Langkah-langkah yang dilakukan dalam algoritma KNN yaitu:

1. Menentukan nilai k.
2. Menghitung kuadrat jarak euclid masing-masing objek terhadap data baru. Dengan rumus :

$$\left(\sum_{k=1}^m (x_{ik} - x_{jk})^2 \right)^{1/2} \quad (1)$$

Dimana: X_{ik} = nilai X pada data latih
 X_{jk} = nilai X pada data uji
m = batas jumlah banyaknya data

3. Mengurutkan objek dari mulai jarak terdekat sampai jarak terjauh
4. Mengumpulkan label class Y (klasifikasi Nearest Neighborhood).
5. Melakukan prediksi berdasarkan label class Y yang paling mayoritas.

2.4.5 Uji Coba

Tahapan uji coba dilakukan untuk mengetahui keberhasilan dari sistem yang dibuat. Proses uji coba dilakukan pada data uji dengan menggunakan nilai k paling optimal. Data uji yang digunakan berjumlah 50 data. Pada proses uji coba, tingkat keberhasilan metode diukur berdasarkan tingkat akurasi yang dihasilkan. Accuracy atau akurasi merupakan rasio prediksi benar positif dan benar negatif (TP + TN) dibandingkan dengan keseluruhan data (TP + FP + TN + FN). Rumus perhitungan akurasi dapat dilihat dibawah ini.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

2.4.6 Penarikan Kesimpulan

Setelah melakukan seluruh rangkaian proses pada model penelitian yang dilakukan, pada tahapan ini akan dilakukan analisis terhadap hasil dari uji coba dari sistem yang digunakan. Sehingga dapat disimpulkan keunggulan dan kelemahan dari sistem yang dibuat. Tahap akhir ini akan menentukan tindakan yang akan dilakukan peneliti selanjutnya untuk mengembangkan penelitian yang telah dilakukan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Setelah peneliti melakukan wawancara dan pengumpulan dataset serta pemrosesan data awal maka diperoleh hasil sebagai berikut :

Tabel 2. Hasil Wawancara Dokter

No	Pertanyaan	Jawaban
1	Apakah penyakit diabetes melitus merupakan salah satu penyakit dengan jumlah pasien tertinggi?	Ya, jumlah pasien penderita diabetes melitus dapat dikatakan cukup tinggi, penyakit ini masuk pada 10 penyakit terbanyak di Puskesmas Paiton.
2	Bagaimana proses identifikasi penyakit diabetes melitus saat ini?	Sampai saat ini proses identifikasi terhadap pasien masih mengandalkan hasil tes dari laboratorium
3	Apakah saja faktor faktor yang menyebabkan seseorang menderita diabetes melitus?	Banyak faktor yang bisa menyebabkan seseorang menderita diabetes melitus seperti hipertensi atau tekanan darah tinggi, kadar gula berlebih, obesitas yang ditandai dengan index massa tubuh $\geq 25\text{kg/m}^2$, riwayat penyakit diabetes melitus dalam keluarga, usia ≥ 45 tahun, jumlah kehamilan (terutama kehamilan bermasalah, seperti bayi lahir meninggal serta bayi lahir besar), ketebalan lipatan kulit trisep, jumlah kadar insulin dalam tubuh, kurangnya aktivitas fisik serta pola hidup dan diet tidak sehat.

Tabel 3. Data Latih

No	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	Pedigree	Age	Outcome
1	0	137	40	35	168	43.1	2.288	33	1
2	3	78	50	32	88	31	0.248	26	1
3	2	197	70	45	543	30.5	0.158	53	1
4	1	189	60	23	846	30.1	0.398	59	1
5	5	166	72	19	175	25.8	0.587	51	1
6	0	118	84	47	230	45.8	0.551	31	1
7	1	103	30	38	83	43.3	0.183	33	0
8	1	115	70	30	96	34.6	0.529	32	1
...
344	1	101	50	15	36	24.2	0.526	26	0

Tabel 4. Data Uji

No	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	sPedigree	Age	Outcome
1	1	99	72	30	18	38.6	0.412	21	0
2	6	92	62	32	126	32	0.085	46	0
3	4	154	72	29	126	31.3	0.338	37	0
4	1	109	60	8	182	25.4	0.947	21	0
5	3	111	58	31	44	29.5	0.43	22	0
6	2	98	60	17	120	34.7	0.198	22	0
7	1	143	86	30	330	30.1	0.892	23	0
8	1	119	44	47	63	35.5	0.28	25	0
...
50	7	97	76	32	91	40.9	0.871	32	1

3.2 Hasil Pembuatan Design Sistem

Hasil pembuatan design sistem tersebut kemudian disimpan dengan nama diabetes.ui.

The screenshot shows a graphical user interface for a diabetes detection application. The window is titled "Deteksi Diabetes". The main title is "APLIKASI DETEKSI DIABETES" and the author is "By : Umikulsun Indah Lestari". The interface includes several input fields for patient data: "Berapa kali pasien melahirkan", "Kadar glukosa / kadar gula pasi", "Tekanan darah pasien", "Ketebalan lipatan kulit pasien", "Kadar insulin pasien", "BMI", "Riwayat diabetes keluarga", and "Umur". Below these fields is a "Prediksi Pasien" button and an asterisk symbol.

Gambar 2. Tampilan Sistem

3.3 Hasil Penggunaan Metode KNN

3.3.1 Penggunaan Metode KNN Pada Sistem

1. Mencari Nilai k Terbaik

Langkah pertama yang perlu dilakukan dalam membuat sistem pendukung keputusan adalah mencari nilai k terbaik dengan tingkat akurasi tertinggi

```

2 knn_clf = KNeighborsClassifier(n_neighbors=3)
3 knn_clf.fit(X_train,y_train)
4
5 y_pred = knn_clf.predict(X_test)
6
7 round(accuracy_score(y_test,y_pred),3)

```

Gambar 3. Proses Mencari Nilai k Paling Optimal

Pada tahap ini dilakukan pengujian terhadap angka ganjil dalam rentan 1-30, sehingga menghasilkan tingkat akurasi sebagai berikut:

Tabel 5. Hasil Akurasi Nilai K

K	Akurasi	K	Akurasi	K	Akurasi
1	80%	11	90%	21	94%
3	82%	13	94%	23	96%
5	90%	15	92%	25	92%
7	88%	17	94%	27	92%
9	90%	19	94%	29	92%

Berdasarkan uji hasil akurasi tersebut, maka didapatkan nilai k terbaik adalah 23 dengan tingkat akurasi mencapai 96%

2. Pembuatan Sistem

Pada tahap ini hasil pembuatan design serta model akurasi paling optimal yang telah dibuat sebelumnya akan disatukan sehingga nantinya akan menghasilkan suatu sistem pendukung keputusan yang siap digunakan kapan saja.

```

def diabetes():
    banyakMelahirkan = float(window.input_melahirkan.text())
    kadarGlukosa = float(window.input_glukosa.text())
    tekananDarah = float(window.input_darah.text())
    ketebalanLipatanKulit = float(window.input_kulit.text())
    kadarInsulin = float(window.input_insulin.text())
    bmi = float(window.input_bmi.text())
    derajatRiwayatDiabetes = float(window.input_riwayat.text())
    umur = float(window.input_umur.text())

    dataPasien = np.array((banyakMelahirkan,kadarGlukosa,tekananDarah, ketebalanLipatanKulit,kadarInsulin, bmi,derajatRiwaya
    data = np.reshape(dataPasien, (1, -1))
    isDiabetes = model.predict(data)

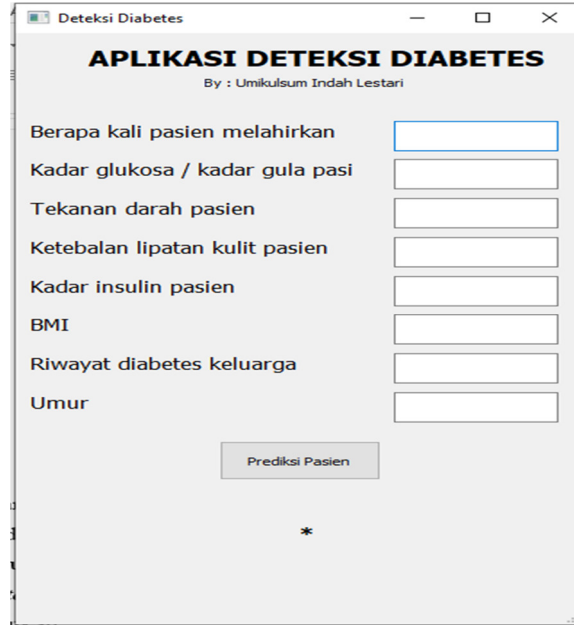
    if(isDiabetes==1):
        window.label_kesimpulan.setText('POSITIF DIABETES')
    elif(isDiabetes==0):
        window.label_kesimpulan.setText('NEGATIF DIABETES')

```

Gambar 4. Proses Prediksi Pada Sistem

Pada tahap ini dilakukan prediksi data baru terhadap data latih yang ada, kemudia hasil dari prediksi tersebut disimpan dalam satu variable yaitu isDiabetes. Apabila hasil prediksi dari isDiabetes menghasilkan angka 1 maka sistem akan menampilkannya sebagai positif diabetes,

sedangkan jika hasil prediksi menghasilkan angka 0, maka sistem akan menampilkannya sebagai negatif diabetes.



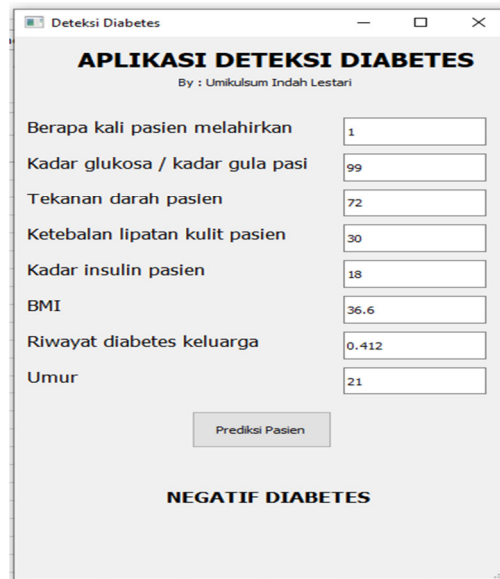
Gambar 5. Hasil Tampilan Awal Sistem

Berdasarkan Gambar diatas dapat dilihat bahwa terdapat 8 variabel inputan yang perlu diisi sebelum melakukan prediksi terhadap pasien, tampilan awal dari hasil prediksi hanya berupa tanda bintang (*) saja, tanda bintang (*) tersebut nantinya akan berubah sesuai menjadi positif diabetes atau negatif diabetes sesuai dengan prediksi sistem terhadap pasien yang datanya sedang diinputkan. Untuk itu dilakukan pengujian awal dengan mengambil dua data uji yaitu satu data positif dan satu data negatif.

Data uji:

Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome
1	99	72	30	18	38.6	0.412	21	0

Prediksi yang dihasilkan sistem terhadap data uji tersebut dapat dilihat dibawah ini:



Gambar 6. Hasil Prediksi Sistem

3.3.2 Perhitungan Metode KNN

Berdasarkan prediksi yang telah dilakukan pada Gambar 6 dapat dilihat bahwa sistem hanya menampilkan hasil akhir atau hasil prediksinya saja tanpa proses perhitungan algoritma jarak Euclidean. Dengan nilai k = 23. Sehingga menghasilkan:

Tabel 6. Urutan Jarak Euclidean

Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome	Jarak	Urutan ke
1	87	78	27	32	34.6	0.101	22	0	20.05234951	1
2	88	58	26	16	28.4	0.766	22	0	21.05149201	2
3	83	58	31	18	34.3	0.336	25	0	22.16970401	3
0	95	85	25	36	37.4	0.247	24	1	23.35523978	4
1	89	76	34	37	31.2	0.192	23	0	23.49060238	5
5	88	66	21	23	24.4	0.342	30	0	23.6990485	6
1	88	62	24	44	29.9	0.422	23	0	31.82279215	7
1	79	80	25	37	25.4	0.583	22	0	32.01982575	8
1	81	72	18	40	26.6	0.283	24	0	33.24179058	9
0	100	70	26	50	30.8	0.597	21	0	33.26971934	10
1	95	66	13	38	19.6	0.334	25	0	33.43659797	11
0	78	88	29	40	36.9	0.434	21	0	34.43676065	12
2	99	70	16	44	20.4	0.235	27	0	35.27423038	13
5	126	78	27	22	29.6	0.439	40	0	35.32705378	14
1	90	62	12	43	27.2	0.58	24	0	35.62286097	15
1	101	50	15	36	24.2	0.526	26	0	35.6282612	16
2	83	66	23	50	32.2	0.497	22	0	37.52288935	17
3	74	68	28	45	29.7	0.293	23	0	38.22596187	18
5	77	82	41	42	35.8	0.156	35	0	38.74152212	19
2	88	74	19	53	29	0.229	22	0	39.56252632	20
4	99	76	15	51	23.2	0.223	21	0	39.70133148	21
4	85	58	22	49	27.8	0.306	28	0	39.89550396	22
2	96	68	13	49	21.1	0.647	26	0	40.09121132	23
...
1	189	60	23	846	30.1	0.398	59	1	833.9024225	344

Berdasarkan hasil tersebut dapat kita hitung jumlah prediksi positif (outcome 1) sebanyak 1 dan prediksi negatif (outcome 0) sebanyak 22 data. jadi dapat disimpulkan bahwa data yang baru saja diinputkan adalah negatif diabetes.

3.4 Hasil Uji Coba

Proses uji coba dilakukan pada data uji dengan menggunakan nilai k paling optimal. Data uji yang digunakan berjumlah 50 data. 50 data tersebut kemudian akan diberikan variabel nilainya untuk menghitung tingkat akurasi dari sistem yang dibuat.

Tabel 7. Variabel Nilai

Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome	Keterangan	Variabel Nilai
1	99	72	30	18	38.6	0.412	21	0	Benar	TN
6	92	62	32	126	32	0.085	46	0	Benar	TN
4	154	72	29	126	31.3	0.338	37	0	Salah	FN
1	109	60	8	182	25.4	0.947	21	0	Benar	TN
3	111	58	31	44	29.5	0.43	22	0	Benar	TN
2	98	60	17	120	34.7	0.198	22	0	Benar	TN
1	143	86	30	330	30.1	0.892	23	0	Benar	TN
1	119	44	47	63	35.5	0.28	25	0	Benar	TN
6	108	44	20	130	24	0.813	35	0	Benar	TN
8	124	76	24	600	28.7	0.687	52	1	Benar	TP
3	176	86	27	156	33.3	1.154	52	1	Benar	TP
2	112	78	50	140	39.4	0.175	24	0	Benar	TN
2	82	52	22	115	28.5	1.699	25	0	Benar	TN
6	123	72	45	230	33.6	0.733	34	0	Benar	TN
0	188	82	14	185	32	0.682	22	1	Benar	TP

1	89	24	19	25	27.8	0.559	21	0	Benar	TN
1	109	38	18	120	23.1	0.407	26	0	Benar	TN
1	109	58	18	116	28.5	0.219	22	0	Benar	TN
1	181	78	42	293	40	1.258	22	1	Benar	TP
1	92	62	25	41	19.5	0.482	25	0	Benar	TN
0	152	82	39	272	41.5	0.27	27	0	Benar	TN
1	111	62	13	182	24	0.138	23	0	Benar	TN
3	106	54	21	158	30.9	0.292	24	0	Benar	TN
7	168	88	42	321	38.2	0.787	40	1	Benar	TP
8	196	76	29	280	37.5	0.605	57	1	Benar	TP
2	68	62	13	15	20.1	0.257	23	0	Benar	TN
2	112	86	42	160	38.4	0.246	28	0	Benar	TN
0	94	70	27	115	43.5	0.347	21	0	Benar	TN
4	90	88	47	54	37.7	0.362	29	0	Benar	TN
0	102	78	40	90	34.5	0.238	24	0	Benar	TN
1	128	82	17	183	27.5	0.115	22	0	Benar	TN
2	94	76	18	66	31.6	0.649	23	0	Benar	TN
2	106	56	27	165	29	0.426	22	0	Benar	TN
1	100	74	12	46	19.5	0.149	28	0	Benar	TN
0	102	86	17	105	29.3	0.695	27	0	Benar	TN
3	103	72	30	152	27.6	0.73	27	0	Benar	TN
0	126	86	27	120	27.4	0.515	21	0	Benar	TN
1	167	74	17	144	23.4	0.447	33	1	Benar	TP
0	179	50	36	159	37.8	0.455	22	1	Benar	TP
2	174	88	37	120	44.5	0.646	24	1	Benar	TP
1	91	54	25	100	25.2	0.234	23	0	Benar	TN
1	117	60	23	106	33.8	0.466	27	0	Benar	TN
5	123	74	40	77	34.1	0.269	28	0	Benar	TN
1	106	70	28	135	34.2	0.142	22	0	Benar	TN
2	155	52	27	540	38.7	0.24	25	1	Benar	TP
2	101	58	35	90	21.8	0.155	22	0	Benar	TN
1	120	80	48	200	38.9	1.162	41	0	Benar	TN
3	187	70	22	200	36.4	0.408	36	1	Benar	TP
8	167	106	46	231	37.6	0.165	43	1	Benar	TP
7	97	76	32	91	40.9	0.871	32	1	Salah	FP

Berdasarkan tabel variabel nilai diatas maka didapat :

True Positive (TP) atau Benar Positif = 12 data

True Negative (TN) atau Benar Negatif = 36 data

False Positive (FP) atau Salah Positif = 1 data

False Negative (FN) atau Salah Negatif = 1 data

Sehingga kemudian dapat dihitung tingkat akurasi dari sistem yang dibuat dengan rumus akurasi

$$: accuracy = \frac{TP+TN}{TP+FP+TN+FN} = \frac{12+36}{12+1+36+1} = 0.96$$

4. KESIMPULAN

Berdasarkan implementasi dan pengujian, dengan nilai k=23 didapatkan tingkat akurasi sebesar 0.96 atau 96%, tingkat akurasi tersebut dinilai cukup tinggi sehingga penelitian ini dinilai telah berhasil menerapkan metode KNN untuk sistem pendukung keputusan identifikasi penyakit diabetes melitus sejak dini.

5. SARAN

Pada penelitian selanjutnya diharapkan dapat mengidentifikasi penyakit diabetes dengan pengembangan metode yang lain dan terbaru dengan tingkat akurasi yang lebih tinggi, salah satunya adalah metode C 4.5 dengan menggunakan cakupan variable yang lebih luas lagi seperti gejala gejala yang dialami penderita penyakit diabetes, beberapa diantaranya adalah gejala sering haus dan sering kencing di malam hari. Aplikasi ini juga diharapkan dapat dikembangkan kedalam versi mobile atau android yang dapat turut serta menampilkan hasil perhitungan jarak euclidean.

UCAPAN TERIMA KASIH

Terimakasih penulis ucapkan kepada Dokter Nina dan Dokter Nindri selaku Kepala serta Dokter di Puskesmas Paiton yang telah sudi meluangkan waktunya untuk wawancara.

DAFTAR PUSTAKA

- [1] International Diabetes Federation. 2015, *IDF Diabetes Atlas, Seventh Edition 2015*. [online], <http://www.idf.org/idf-diabetes-atlas-seventh-edition> (2015).
- [2] American Diabetes Association. 2018, *American Diabetes Association Standards of Medical Care In Diabetes*, ADA, USA.
- [3] American Diabetes Association. 2017, *Standards of Medical Care in Diabetes 2017*, ADA, USA.
- [4] American Diabetes Association. 2020, *Standards of Medical Care in Diabetes*, ADA, USA.

- [5] Santoso, R. R., Megasari, R., & Hambali, Y. A. 2020, *Implementasi Metode Machine Learning Menggunakan Algoritma Evolving Artificial Neural Network pada Kasus Prediksi Diagnosis Diabetes*, JATIKOM: Jurnal Aplikasi dan Teori Ilmu Komputer, 85-97.
- [6] Santoso, H. F., Darwiyanto, E., & Wisesty, U. N. 2016, *Analisis dan Implementasi Metode Certainty Factor pada Sistem Pendukung Pengambilan Keputusan untuk Penyakit Diabetes Melitus*. e-Proceeding of Engineering, 5191.
- [7] Ramadhani, R. A., & Niswatin, R. K. 2018, *Sistem Diagnosa Diabetes Menggunakan Metode KNN*. Jurnal Sains dan Informatika, 98-104.
- [8] Noviani. 2018. *Implementasi Algoritma Decision Tree C4.5 Untuk Prediksi Penyakit Diabetes*. Jurnal INOHIM, Volume 6 Nomor 1, 1-5.
- [9] Sihotang, H. T., & Siboro, M. S. 2016, *Aplikasi Sistem Pendukung Keputusan Penentuan Siswabermasalah Menggunakan Metode SAW pada Sekolah SMP Swasta Mulia Pratama Medan*, JIPN (Journal of Informatics Pelita Nusantara), 1-6.
- [10] Baharuddin, M. M., Hasanuddin, T., & Azis, H. 2019, *Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca*, ILKOM Jurnal Ilmiah Volume 11 Nomor 3, 269-274.