

Contents list available at [www.jurnal.unimed.ac.id](http://www.jurnal.unimed.ac.id)

**CESS**  
**(Journal of Computing Engineering, System and Science)**

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



**Klasifikasi Demam Menggunakan Metode *Neighbor Weighted K-Nearest Neighbor***

***Fever Classification Using the Neighbor Weighted K-Nearest Neighbor Method***

Surayya Safira Milania<sup>1\*</sup>, Cucu Suhery<sup>2</sup>, Tedy Rismawan<sup>3</sup>

<sup>1,2,3</sup> Program Studi Rekayasa Sistem Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Tanjungpura

Jalan Prof. Dr. H. Hadari Nawawi Pontianak

email: <sup>1</sup>[surayyasafira@student.untan.ac.id](mailto:surayyasafira@student.untan.ac.id), <sup>2</sup>[csuhery@siskom.untan.ac.id](mailto:csuhery@siskom.untan.ac.id), <sup>3</sup>[tedyrismawan@siskom.untan.ac.id](mailto:tedyrismawan@siskom.untan.ac.id)

**ABSTRAK**

Demam merupakan gejala atau reaksi tubuh terhadap suatu infeksi atau penyakit. Demam dapat disebabkan karena adanya infeksi virus, bakteri, dan parasit. Serta demam akibat gigitan nyamuk. Beberapa penyakit penyebab demam yang perlu diwaspadai antara lain Demam Berdarah *Dengue* (DBD), Demam *Tifoid*, dan Malaria dikarenakan gejala klinis dari ketiga penyakit tersebut sangat mirip dan sulit untuk dibedakan. Akibat dari gejala yang mirip, seringkali menyebabkan kesulitan dalam mendapatkan diagnosis awal sehingga kurang tepat dalam penanganan. Oleh karena itu, pada penelitian ini dibangun sebuah sistem yang dapat mengklasifikasikan demam menggunakan metode *Neighbor Weighted K-Nearest Neighbor*. Data yang digunakan berjumlah 300 data dengan komposisi rasio data latih dan data uji sebesar 70%:30% sehingga data latih yang digunakan berjumlah 210 data dan data uji berjumlah 90 data. Penelitian ini dilakukan dengan mengamati variasi nilai ketetanggaan (K) dan nilai *exp* (E) terhadap akurasi sistem klasifikasi demam. Hasil pelatihan menunjukkan bahwa nilai K dan E yang bervariasi tidak mempunyai pengaruh terhadap akurasi tersebut. Hasil pengujian yang dilakukan mendapatkan akurasi sebesar 100% pada setiap variasi nilai K dan E.

**Kata Kunci:** Demam, DBD, Malaria, Demam *Tifoid*, *Neighbor Weighted K-Nearest Neighbor*

**ABSTRACT**

Fever is a symptom of the body's reaction to an infection or disease. Fever can be caused by viral, bacterial, or parasitic infections, as well as fever due to mosquito bites. Several diseases that cause fever that need to be watched out for include dengue hemorrhagic fever (DHF), typhoid fever, and malaria because the clinical symptoms of these three diseases are very

\*Penulis Korespondensi:

email: [surayyasafira@student.untan.ac.id](mailto:surayyasafira@student.untan.ac.id)

similar and difficult to distinguish. As a result of similar symptoms, it often causes difficulties in getting an early diagnosis, so treatment is not appropriate. Therefore, in this study, a system was developed that could classify fever using the neighbor weighted K-nearest neighbor method. The data used totaled 300, with a composition ratio of 70% training data to 30% test data, for a total of 210 training data and 90 test data. This research was conducted by observing the variation in the value of neighborliness (K) and the value of exp (E) on the accuracy of the fever classification system. The results of the training show that the varying K and E values have no effect on accuracy. The results of the tests carried out obtained an accuracy of 100% for each variation in the values of K and E.

**Keywords:** *Fever, DHF, Malaria, Typhoid Fever, Neighbor Weighted K-Nearest Neighbor*

---

## 1. PENDAHULUAN

Demam merupakan gejala atau reaksi tubuh terhadap suatu infeksi atau penyakit. Penyebab demam juga sangat bervariasi antara lain demam setelah mendapatkan imunisasi, demam dikarenakan adanya infeksi virus dan bakteri seperti Covid-19, Meningitis, Tifus, Disentri, Cacar Air, dan Infeksi Saluran Kemih. Serta demam akibat gigitan nyamuk seperti DBD, Malaria, Chikungunya [1]. Beberapa penyakit penyebab demam yang perlu diwaspadai antara lain Demam Berdarah *Dengue* (DBD), Demam *Tifoid*, dan Malaria, dikarenakan semuanya termasuk penyakit endemik yang banyak dijumpai di Indonesia dan seringkali menyebabkan kejadian luar biasa (KLB), serta dapat menyebabkan kematian.

Demam *Tifoid* merupakan penyakit infeksi akut sistem pencernaan yang disebabkan oleh bakteri *Salmonella Typhi* atau *Salmonella Paratyphi* [2]. Di Indonesia, demam *Tifoid* Dan *Paratifoid* termasuk urutan ke 3 dari 10 penyakit terbanyak pada pasien rawat inap di rumah sakit dengan jumlah 55.098 kasus dan *Case Fatality Rate* (CFR) sebesar 2,06% [3].

Selain Demam Tifoid, DBD juga merupakan salah satu masalah kesehatan masyarakat yang utama di Indonesia. Penyakit DBD disebabkan oleh virus *Dengue* dari genus *Flavivirus*, famili *Flaviviridae* yang ditularkan ke manusia melalui gigitan nyamuk *aedes* [4]. Kasus DBD di Indonesia pada tahun 2020 tercatat sebanyak 108.303 kasus dengan jumlah kasus kematian sebanyak 747 kasus [5].

Masalah kesehatan masyarakat di Indonesia lainnya adalah penyakit Malaria yang dapat menyebabkan kematian terutama pada kelompok risiko tinggi seperti bayi, balita, dan ibu hamil. Penyakit ini disebabkan oleh parasit *plasmodium* yang ditularkan melalui gigitan nyamuk *anopheles* betina [6]. Berdasarkan data profil kesehatan Indonesia tahun 2020, kasus positif Malaria di Indonesia tercatat sebanyak 254.050 kasus [5].

Untuk mengetahui jenis penyakit demam yang diderita oleh seseorang maka harus diketahui gejala klinis yang dialami oleh orang tersebut. Namun, gejala klinis pada penyakit Demam *Tifoid*, DBD, dan Malaria sangat sulit dibedakan karena terdapat beberapa gejala klinis yang sama seperti sakit kepala, nyeri perut, mual dan muntah, letargi, anoreksia, delirium, serta malaise. Akibat dari gejala yang mirip dari ketiga penyakit tersebut seringkali menyebabkan kesulitan dalam mendapatkan diagnosis awal sehingga kurang tepat dalam penanganan. Oleh karena itu, diperlukan sebuah sistem yang dapat mengklasifikasikan demam menggunakan data mining guna mencegah terjadinya *medical error*. Salah satu metode yang dapat digunakan untuk melakukan klasifikasi pada data mining adalah metode *Neighbor Weighted K-Nearest Neighbor* (NWKNN).

Metode NWKNN pernah diterapkan pada kasus klasifikasi penyakit *Diabetes Mellitus* Tipe 1 dan Tipe 2 berdasarkan gejala dengan rata-rata akurasi yang dihasilkan sebesar 92.3% dengan menggunakan nilai ketetanggaan (K) = 15 dan exp (E) = 2 [7]. Penelitian lain yang mengimplementasikan metode NWKNN pernah diterapkan pada kasus diagnosis hama dan penyakit tanaman Bawang Merah berdasarkan gejala. Akurasi yang dihasilkan pada penelitian ini sebesar 100% dengan menggunakan nilai K = 2 dan nilai E = 4 [8]. Selain itu, penelitian terkait kasus demam berdasarkan gejala pernah diterapkan dengan menggunakan metode *K-Nearest Neighbor – Certainty Factor*. Penelitian ini bertujuan untuk mengetahui sejak dini jenis penyakit demam yang diderita oleh pasien. Hasil pengujian variasi nilai ketetanggaan (K) didapatkan akurasi tertinggi sebesar 88.37% dengan menggunakan nilai k=3, k=5, dan k=8 [9].

Berdasarkan permasalahan dan penelitian terkait yang telah dijabarkan, serta mengingat gejala klinis pada Demam *Tifoid*, DBD, dan Malaria sangat sulit untuk dibedakan. Maka dilakukan suatu penelitian klasifikasi demam menggunakan NWKNN guna mencegah terjadinya kesalahan diagnosis dan penanganan yang kurang tepat.

## 2. DASAR/TINJAUAN TEORI

### 2.1. Klasifikasi

Klasifikasi merupakan sebuah metode untuk mengelompokkan sesuatu berdasarkan kategori atau kelas. Pada metode klasifikasi mempunyai variabel target berbentuk kategori. Contoh algoritma pada metode klasifikasi adalah *K-Nearest Neighbor*, *Neural Network*, dan *Decision Tree* [10].

### 2.2. Metode *Neighbor Weighted K-Nearest Neighbor*

Metode *Neighbor Weighted K-Nearest Neighbor* (NWKNN) merupakan pengembangan dari metode KNN untuk menyelesaikan masalah data yang tidak seimbang [11]. Langkah-langkah pada algoritma NWKNN adalah sebagai berikut [12].

- 1) Menentukan nilai ketetanggaan (K) dan *exp* (E).
- 2) Menghitung nilai kedekatan ketetanggaan antara data latih dan data uji dengan menggunakan rumus *Cosine Similarity* (CosSim). Rumus CosSim dapat dilihat pada Persamaan 1.

$$CosSim(q, d_j) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^m (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^m (w_{ij})^2 \cdot \sum_{i=1}^m (w_{iq})^2}} \quad (1)$$

Keterangan:

- q : Data Uji
- d<sub>j</sub> : Data Latih
- W<sub>ij</sub> : Bobot nilai i pada data latih j
- W<sub>iq</sub> : Bobot nilai i pada data uji q
- m : Banyak jumlah data

- 3) Mengurutkan hasil perhitungan nilai kedekatan ketetanggaan dengan cara mengurutkan nilai dari yang tertinggi ke nilai yang terendah.
- 4) Menghitung jumlah data pada masing-masing kategori kelas klasifikasi.
- 5) Melakukan perhitungan bobot dengan menggunakan Persamaan 2.

$$Weight_i = \frac{1}{\left( \frac{Num(C_i^d)}{\min\{Num(C_j^d) | j=1, \dots, k^*\}} \right)^{1/E}} \quad (2)$$

Keterangan:

- $Weight_i$  : Bobot kelas i
- $Num(C_i^d)$  : Banyaknya data latih d pada kelas i
- $Num(C_j^d)$  : Banyaknya data latih d pada kelas j, dimana j termasuk dalam himpunan k tetangga terdekat
- $Exp$  : Eksponen (nilai E lebih dari 1)

6) Setelah mendapatkan nilai bobot dilanjutkan dengan menghitung nilai skor pada masing-masing kategori. Perhitungan nilai skor dapat dilihat pada Persamaan 3.

$$Score(X, C_i) = Weight_i * \left( \sum_{d_j \in N_{WKNN}(x)} \left( Sim(d, d_j) * \delta(d_j, C_i) \right) \right) \quad (3)$$

Keterangan :

- $Weight_i$  : Bobot kelas i
- $d_j \in N_{WKNN}(x)$  : Data latih  $d_j$  pada kumpulan tetangga terdekat data uji X
- $\delta(d_j, C_i)$  : Akan bernilai 1 jika nilai jarak  $\in C_i$  dan bernilai 0 jika nilai jarak  $\notin C_i$
- $Sim(d, d_j)$  : Nilai CosSim antara data uji dan data latih
- $C_i$  : Jenis atau kelas i

### 2.3. K-Fold Cross Validation

Metode *K-Fold Cross Validation* merupakan salah satu metode *Cross Validation* untuk melakukan pengujian. Konsep *K-Fold Cross Validation* yaitu dengan cara membagi data ke dalam subset terpisah dengan ukuran yang sama sebanyak k atau *fold* [13]. Jika data dalam ukuran besar, maka metode ini sangat efisien untuk digunakan karena setiap data telah digunakan sebagai pelatihan dan pengujian [14].

### 2.4. Confusion Matrix

*Confusion matrix* merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu model klasifikasi dalam memperoleh *accuracy*, *precision*, *recall*, dan *F1-Score*. Untuk mengukur kinerja suatu model klasifikasi dengan dimensi 3x3 dapat menggunakan tabel *confusion matrix multiclass* yang dapat dilihat pada Tabel 2.

**Tabel 1. Confusion Matrix Multiclass 3x3**

<b>Confusion Matrix</b>		<b>Predicted</b>		
		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>
<b>Actual</b>	<b>Class 1</b>	TP	C <sub>12</sub>	C <sub>13</sub>
	<b>Class 2</b>	C <sub>21</sub>	TP	C <sub>23</sub>
	<b>Class 3</b>	C <sub>31</sub>	C <sub>32</sub>	TP

Keterangan:

- 1) TP (*True Positive*) menunjukkan jumlah data hasil klasifikasi yang bernilai positif dan data aktualnya benar sebagai positif.
- 2) FP (*False Positive*) menunjukkan jumlah data hasil klasifikasi yang bernilai positif tetapi data aktualnya sebagai negatif.
- 3) TN (*True Negative*) menunjukkan jumlah data hasil klasifikasi yang bernilai negatif dan data aktualnya benar sebagai negatif.
- 4) FN (*False Negative*) menunjukkan jumlah data hasil klasifikasi yang bernilai negatif tetapi data aktualnya sebagai positif.

Berdasarkan perhitungan *confusion matrix*, maka dapat diukur kinerja suatu model klasifikasi sebagai berikut [15].

a. *Accuracy*

*Accuracy* merupakan suatu pengukuran jumlah yang terkoreksi dengan benar dibagi dengan seluruh data. Rumus perhitungan *accuracy* dapat dilihat pada Persamaan 4.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

b. *Precision*

*Precision* merupakan suatu pengukuran hasil klasifikasi oleh sistem yang sesuai dengan kategori atau data aktualnya. Rumus perhitungan *precision* dapat dilihat pada Persamaan 5.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

c. *Recall*

*Recall* merupakan suatu pengukuran tingkat keberhasilan sistem dalam mengenali suatu kategori. Rumus perhitungan *recall* dapat dilihat pada Persamaan 6.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

d. *F1-Score*

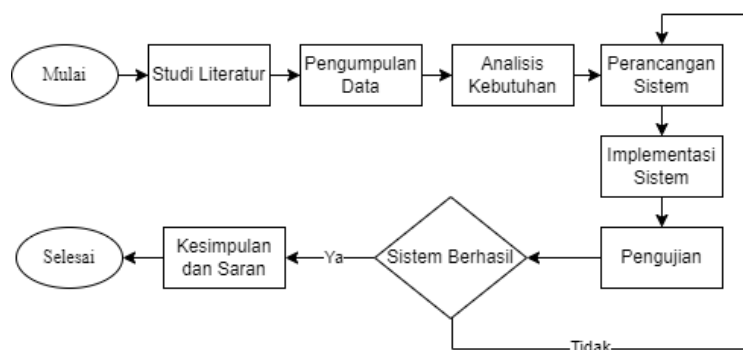
*F1-Score* merupakan kombinasi antara *precision* dan *recall*. Rumus perhitungan *F1-Score* dapat dilihat pada Persamaan 7.

$$F1 - score = 2 * \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

### 3. METODE

#### 3.1 Metode Penelitian

Metode penelitian merupakan sebuah proses atau langkah-langkah yang dilakukan dalam penelitian. Terdapat 7 proses dalam penelitian ini, yaitu studi literatur, pengumpulan data, analisis kebutuhan, perancangan sistem, implementasi sistem, dan pengujian serta kesimpulan dan saran. Metode penelitian yang digunakan pada penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Diagram Alir Penelitian

a. Studi Literatur

Pada tahap studi literatur dilakukan pengumpulan jurnal, buku maupun artikel yang berkaitan dengan penelitian ini seperti metode *NWKNN*, *K-Fold Cross Validation*, *Confusion Matrix*, *Flowchart*, *ERD*, *MySQL*, *Python*, *Laravel*, dan pengertian demam beserta jenisnya.

b. Pengumpulan Data

Pada tahap pengumpulan data dilakukan melalui observasi langsung ke RSUD Pemangkat, Kabupaten Sambas dengan menggunakan 29 daftar gejala penyakit yang telah dikonsultasikan dengan Dr. Esti Nur Ekasari sebagai salah satu dokter umum di RSUD Pemangkat. Daftar gejala penyakit DBD, Malaria, dan Demam Tifoid dapat dilihat pada Tabel 1.

**Tabel 2.** Daftar Gejala Penyakit DBD, Malaria, dan Demam Tifoid

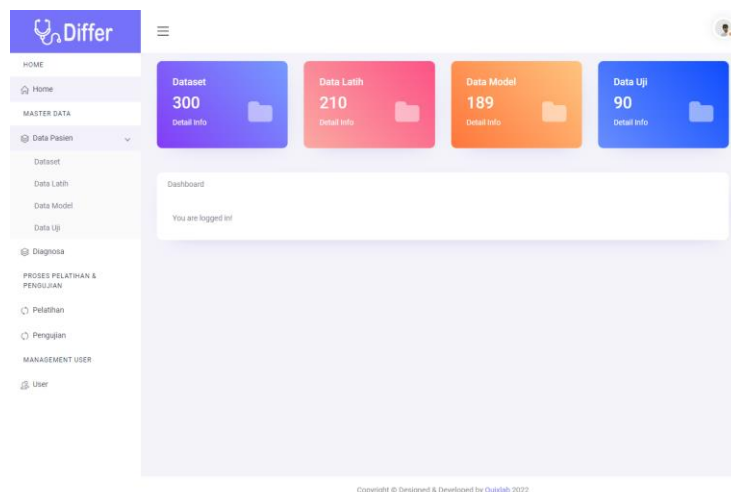
No.	Kode gejala	Nama Gejala
1	G1	Demam intermittent (demam putus-putus)
2	G2	Demam terutama malam hari
3	G3	Demam sepanjang hari
4	G4	Demam menggigil (15menit-1jam)
5	G5	Berkeringat (2-4jam)
6	G6	Lama Demam 2-7 hari
7	G7	Lama Demam 10-14 hari
8	G8	Lama Demam > 14 Hari
9	G9	Sakit Kepala
10	G10	Nyeri Perut
11	G11	Diare atau Susah BAB (Konstipasi)
12	G12	Muntah dan mual
13	G13	Anoreksia (Kehilangan Nafsu Makan)
14	G14	Malaise (lemah & lesu)
15	G15	Nyeri otot dan sendi
16	G16	Petechie (pendarahan kulit / mimisan)
17	G17	Syok
18	G18	Melena (Tinja Berwarna hitam atau gelap)
19	G19	Hematuria (Kencing Berdarah)
20	G20	Nyeri Punggung
21	G21	Riwayat berpergian ke daerah endemik malaria
22	G22	Pucat
23	G23	Lidah kotor
24	G24	Perut Kembung (Meteorismus)
25	G25	Letargi (Penurunan Kesadaran)
26	G26	Delirium (Kebingungan Parah, Tidak merespons sekitar)
27	G27	Pembesaran hati
28	G28	Pembesaran Limpa
29	G29	Ikterus (Sakit Kuning)

- c. Analisis Kebutuhan  
Pada tahap analisis kebutuhan dilakukan analisis kebutuhan perangkat keras dan perangkat lunak yang dibutuhkan oleh sistem klasifikasi demam.
- d. Perancangan Sistem  
Pada tahap perancangan sistem dilakukan perancangan *Database*, ERD, DFD, *Flowchart* dan *Mock-up* antarmuka.
- e. Implementasi  
Pada tahap implementasi dilakukan pembuatan kode program metode NWKNN dengan menggunakan bahasa Python. Selanjutnya diimplementasikan rancangan DFD dan ERD ke dalam sistem klasifikasi demam berbasis *website* dengan menggunakan *framework* Laravel.
- f. Pengujian  
Pada tahap pengujian dilakukan pengujian terhadap sistem yang telah dibuat dengan menggunakan metode *blackbox* dan mengukur akurasi metode dengan menggunakan *confusion matrix*.
- g. Kesimpulan dan Saran  
Pada tahap kesimpulan dilakukan penarikan kesimpulan yang akan menjawab pertanyaan pada rumusan masalah. Setelah dilakukan penarikan kesimpulan, diberikan saran untuk penelitian selanjutnya untuk mendapatkan hasil yang lebih baik.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Implementasi Sistem

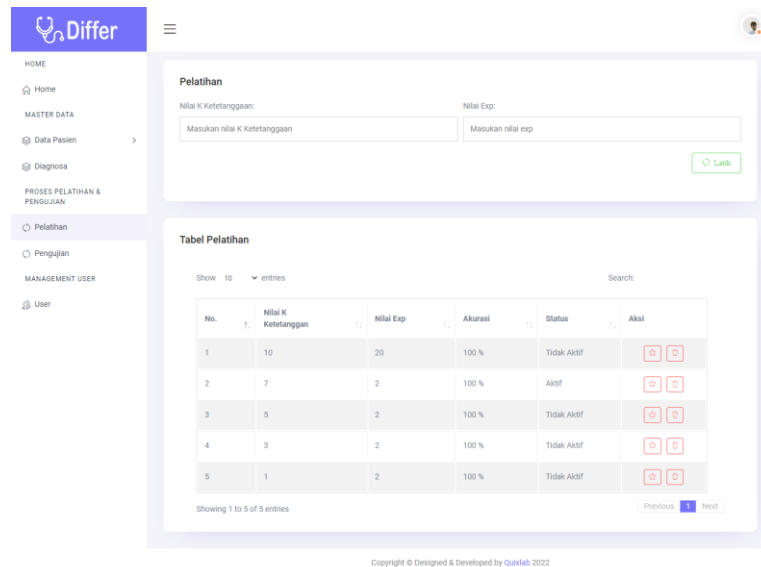
Implementasi sistem pada penelitian ini berupa sistem klasifikasi demam berbasis *website* menggunakan *framework* Laravel. Berikut tampilan halaman *home* admin yang dapat dilihat pada Gambar 1.



**Gambar 2.** Halaman *Home* Admin

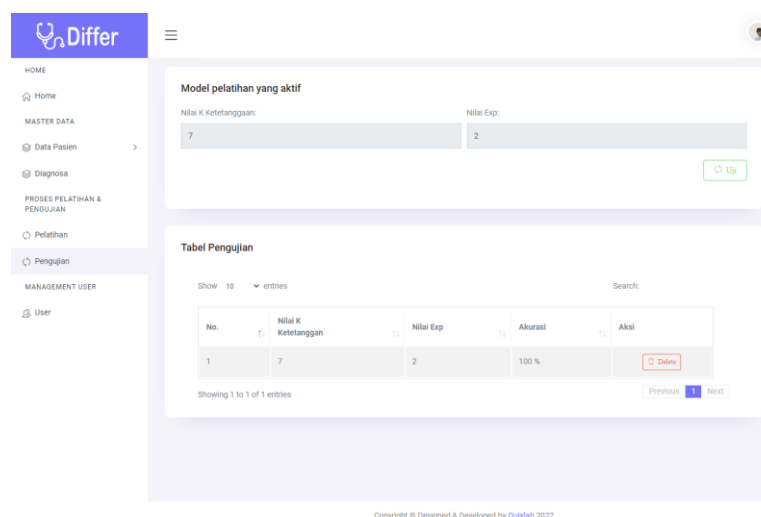
Pada halaman *home* admin terdapat informasi berhasil *login*, jumlah dataset pasien demam, jumlah data latih, jumlah data model, dan jumlah data uji. Selain itu, fitur menu yang terdapat pada halaman admin terdiri dari menu *home*, data pasien, diagnosa, pelatihan, pengujian, dan manajemen *users*. Fitur menu data pasien terdapat informasi terkait data pasien, data latih, data model, dan data uji yang terdiri dari nama pasien, jenis kelamin, umur, diagnosa, dan gejala yang diderita. Fitur menu diagnosa terdapat informasi diagnosa penyakit yang terdiri dari nama diagnosa, deskripsi, langkah-langkah pengobatan dan pencegahan yang harus

dilakukan. Fitur menu pelatihan terdapat informasi terkait hasil proses pelatihan yang terdiri dari nilai K, nilai *exp*, akurasi, dan status. Berikut tampilan halaman pelatihan dapat dilihat pada Gambar 2.



**Gambar 3.** Halaman Pelatihan

Pada halaman pelatihan terdapat form yang digunakan untuk melakukan proses pelatihan metode NWKNN dengan memasukkan nilai ketetangaan (K) dan nilai *exp* (E). Tombol latih berfungsi untuk melakukan proses pelatihan. Setelah berhasil melakukan proses pelatihan, hasil pelatihan akan ditampilkan pada tabel pelatihan yang terdiri dari nilai ketetangaan, nilai *exp*, akurasi, dan status. Admin juga dapat menghapus hasil pelatihan dengan menekan tombol hapus dan mengubah status menjadi aktif atau tidak aktif dengan menekan tombol status. Status pelatihan yang aktif akan digunakan untuk proses pengujian. Berikut tampilan halaman pengujian dapat dilihat pada Gambar 3.



**Gambar 4.** Halaman Pengujian

Pada halaman pengujian terdapat *form* yang berisi model pelatihan yang aktif. Proses pengujian dapat dilakukan dengan menekan tombol uji. Kemudian, hasil pengujian akan



ditampilkan pada tabel pengujian yang terdiri dari nilai ketetangaan, nilai *exp*, dan akurasi. Admin dapat menghapus hasil pengujian dengan menekan tombol hapus pada tabel pengujian.

#### 4.2. Pengujian *Confusion Matrix*

Pengujian metode *confusion matrix* bertujuan untuk menguji akurasi dari nilai ketetangaan (K) dan nilai *exp* (E) yang dihasilkan dari proses pelatihan. Nilai K dan nilai E yang bervariasi dalam proses pelatihan bertujuan untuk mendapatkan nilai K dan E yang optimal. Pelatihan dilakukan menggunakan 210 data yang terbagi menjadi 10-*fold* dengan menggunakan metode *K-Fold Cross Validation* sehingga diperoleh pembagian data latih sebanyak 189 data dan data uji sebanyak 21 data pada setiap *fold*. Variasi nilai K yang digunakan adalah 2, 4, 8, 16, 32, 64, 128 dan nilai E yang digunakan adalah 2, 4, 8, 16. Hasil pelatihan variasi nilai K dengan menggunakan metode *K-Fold Cross Validation* dapat dilihat pada Tabel 2.

**Tabel 3.** Hasil Pelatihan Variasi Nilai K dengan menggunakan Metode *K-Fold Cross Validation*

Nilai K	Akurasi (%)									
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
2	100	100	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100	100	100	100
32	100	100	100	100	100	100	100	100	100	100
64	100	100	100	100	100	100	100	100	100	100
128	100	100	100	100	100	100	100	100	100	100

Pada Tabel 2. menunjukkan hasil dari proses pelatihan setiap variasi nilai K untuk *fold* 1 hingga *fold* 10 menghasilkan akurasi sebesar 100%. Sedangkan, hasil pelatihan variasi nilai E dengan menggunakan metode *K-Fold Cross Validation* dapat dilihat pada Tabel 3.

**Tabel 4.** Hasil Pelatihan Variasi Nilai K dengan menggunakan Metode *K-Fold Cross Validation*

Nilai E	Akurasi (%)									
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
2	100	100	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100	100	100	100

Pada Tabel 4. menunjukkan akurasi pelatihan setiap variasi nilai E untuk *fold* 1 hingga *fold* 10 mencapai 100%. Dalam hal ini, nilai K dan E yang bervariasi tidak mempunyai pengaruh terhadap akurasi sistem klasifikasi demam menggunakan metode NWKNN. Nilai K dan E yang optimal dari hasil proses pelatihan akan digunakan pada proses pengujian. Oleh karena itu, dilakukan pengujian untuk setiap variasi nilai K dan E yang optimal dengan menggunakan 90 data uji dan data latih sebanyak 189 data dari proses pelatihan. Pengujian nilai K dan E hasil

proses pelatihan dihitung menggunakan metode *confusion matrix* digunakan untuk mencari nilai *accuracy*, *recall*, *precision*, dan *F1-score*.

**Tabel 5.** Hasil Pengujian Nilai K = 2 dan E = 2

<i>Confusion Matrix</i>		<i>Predicted</i>		
		<b>DBD (P)</b>	<b>Malaria (N)</b>	<b>Tifoid (N)</b>
<b>Actual</b>	<b>DBD (P)</b>	<i>True Positive (TP)</i> 64	<i>False Negative (FN)</i> 0	<i>False Negative (FN)</i> 0
	<b>Malaria (N)</b>	<i>False Positive (FP)</i> 0	<i>True Negative (TN)</i> 16	<i>True Negative (TN)</i> 0
	<b>Tifoid (N)</b>	<i>False Positive (FP)</i> 0	<i>True Negative (TN)</i> 0	<i>True Negative (TN)</i> 10

Tabel 4. menunjukkan hasil pengujian dengan menggunakan nilai K = 2 dan E = 2. Berdasarkan tabel tersebut dapat dihitung *accuracy*, *recall*, *precision*, dan *F1-score* sebagai berikut.

$$Accuracy = \frac{(TP + TN)}{(\text{Total data})} \times 100\% = \frac{64 + 16 + 10}{90} \times 100\% = 100\%$$

$$Recall = \frac{TP}{TP + FN} \times 100\% = \frac{64}{64 + 0 + 0} \times 100\% = 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\% = \frac{64}{64 + 0 + 0} \times 100\% = 100\%$$

$$F1 - score = \left( 2 * \frac{Precision \times Recall}{Precision + Recall} \right) \times 100\% = \left( 2 * \frac{1 \times 1}{1 + 1} \right) \times 100\% = 100\%$$

Berdasarkan perhitungan tersebut, hasil pengujian nilai K = 2 dan E = 2 didapatkan nilai *accuracy* sebesar 100%, *recall* sebesar 100%, *precision* sebesar 100%, dan *F1-score* sebesar 100%. Hasil dari keseluruhan pengujian setiap variasi nilai K dan E menggunakan cara yang sama yaitu dengan metode *confusion matrix* untuk mencari nilai *accuracy*, *recall*, *precision*, dan *F1-score* dapat dilihat pada Tabel 6.

**Tabel 6.** Hasil Pengujian Setiap Variasi Nilai K dan E

No	Nilai K	Nilai E	<i>Accuracy</i> (%)	<i>Precision</i> (%)	<i>Recall</i> (%)	<i>F1-Score</i> (%)
1	2	2	100	100	100	100
2		4	100	100	100	100
3		8	100	100	100	100
4		16	100	100	100	100
5	4	2	100	100	100	100
6		4	100	100	100	100
7		8	100	100	100	100
8		16	100	100	100	100
9	8	2	100	100	100	100
10		4	100	100	100	100
11		8	100	100	100	100
12		16	100	100	100	100

13	16	2	100	100	100	100
14		4	100	100	100	100
15		8	100	100	100	100
16		16	100	100	100	100
17	32	2	100	100	100	100
18		4	100	100	100	100
19		8	100	100	100	100
20		16	100	100	100	100
21	64	2	100	100	100	100
22		4	100	100	100	100
23		8	100	100	100	100
24		16	100	100	100	100
25	128	2	100	100	100	100
26		4	100	100	100	100
27		8	100	100	100	100
28		16	100	100	100	100
<b>Rata-rata</b>			100	100	100	100

Pada Tabel 6. dapat dilihat bahwa setiap variasi nilai K dan E yang digunakan untuk pengujian menghasilkan nilai *accuracy*, *recall*, *precision*, dan *F1-score* sebesar 100% dari 90 data uji.

## 5. KESIMPULAN

Berdasarkan penelitian yang dilakukan maka dapat disimpulkan bahwa hasil dari proses pelatihan dengan menggunakan 210 data latih menunjukkan setiap variasi nilai K dan E mendapatkan akurasi sebesar 100% pada setiap fold. Dalam hal ini nilai K dan E yang bervariasi tidak mempunyai pengaruh terhadap akurasi sistem klasifikasi demam menggunakan metode *Neighbor Weighted K-Nearest Neighbor* (NWKNN). Sedangkan, hasil dari proses pengujian klasifikasi demam menggunakan metode NWKNN mendapatkan akurasi sebesar 100% pada setiap variasi nilai K dan E dengan menggunakan data uji sebanyak 90 data. Oleh karena itu, penelitian dengan menggunakan metode NWKNN berhasil diimplementasikan pada kasus klasifikasi demam.

## REFERENSI

- [1] Halodoc, "Demam - Gejala, Penyebab, dan Pengobatan," *Halodoc*, 2022. <https://www.halodoc.com/kesehatan/demam> (diakses 27 Juli 2022).
- [2] Y. Levani dan A. D. Prasetya, "Demam tifoid : manifestasi klinis, pilihan terapi dan pandangan dalam islam," *AL-IQRA Med. J. J. Berk. Ilm. Kedokt.*, vol. 3, no. 1, hal. 10–16, 2020.
- [3] Kementerian Kesehatan RI, *Profil Kesehatan Indonesia Tahun 2011*. Jakarta: Kementerian Kesehatan RI, 2012.
- [4] Kementerian Kesehatan RI, "Buletin Jendela Epidemiologi Demam Berdarah Dengue," *Buletin Jendela Epidemiologi*, vol. 2, Jakarta, 2010.
- [5] Kementerian Kesehatan RI, *Profil Kesehatan Indonesia 2020*. Jakarta: Kementerian Kesehatan RI, 2021.

- [6] Kementerian Kesehatan RI, *Buku Saku Tatalaksana Kasus Malaria*. Jakarta: Direktorat Jenderal Pencegahan dan Pengendalian Penyakit Kementerian Kesehatan Republik Indonesia, 2020.
- [7] D. Z. Maliha, E. Santoso, dan M. T. Furqon, "Penerapan Metode Neighbor Weighted K-Nearest Neighbor Dalam Klasifikasi Diabetes Mellitus," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, hal. 2910–2915, 2019, [Daring]. Tersedia pada: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4817>
- [8] M. V. Rosyidah, B. D. Setiawan, dan M. T. Furqon, "Diagnosis Hama Penyakit Tanaman Bawang Merah Menggunakan Algoritma Neighbors Weighted K-Nearest Neighbors (NWKNN)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. Vol 3 No 3 (2019), hal. 404–415, 2019, [Daring]. Tersedia pada: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4864/2271>
- [9] E. N. Shofia, R. R. M. Putri, dan A. Arwan, "Sistem Pakar Diagnosis Penyakit Demam : DBD , Malaria dan Tifoid Menggunakan Metode K-Nearest Neighbor – Certainty Factor," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 5, hal. 426–435, 2017.
- [10] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*, vol. 9780470908. Canada: John Wiley & Sons, Inc., Hoboken, New Jersey., 2005. doi: 10.1002/9781118874059.
- [11] S. Tan, "Neighbor-weighted K-nearest neighbor for unbalanced text corpus," *Expert Syst. Appl.*, vol. 28, hal. 667–671, 2005, doi: 10.1016/j.eswa.2004.12.023.
- [12] P. N. Fadila, Indriati, dan D. E. Ratnawati, "IDENTIFIKASI JENIS ATTENTION DEFICIT HYPERACTIVITY DISORDER ( ADHD ) PADA ANAK USIA DINI MENGGUNAKAN METODE NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR ( NWKNN )," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, hal. 194–200, 2016.
- [13] Mustika dkk., *DATA MINING DAN APLIKASINYA*. Bandung: WIDINA BHAKTI PERSADA BANDUNG (GRUP CV. WIDINA MEDIA UTAMA), 2021.
- [14] B. Santoso, A. I. S. Azis, dan Zohrahayaty, *Machine Learning & Reasoning Fuzzy Logic Algoritma, Manual, Matlab, & Rapid Miner*. Yogyakarta: Penerbit DEEPUBLISH, 2020.
- [15] W. Setiawan, *Topik Khusus Kecerdasan Komputasional: Deep Learning untuk Image dan Speech Recognition*. Malang: Media Nusa Creative (MNC Publishing), 2020. [Daring]. Tersedia pada: [https://www.google.co.id/books/edition/Topik\\_Khusus\\_Kecerdasan\\_Komputasional/A4ZOEAAAQBAJ?hl=id&gbpv=1&kptab=getbook](https://www.google.co.id/books/edition/Topik_Khusus_Kecerdasan_Komputasional/A4ZOEAAAQBAJ?hl=id&gbpv=1&kptab=getbook)