

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Sistem

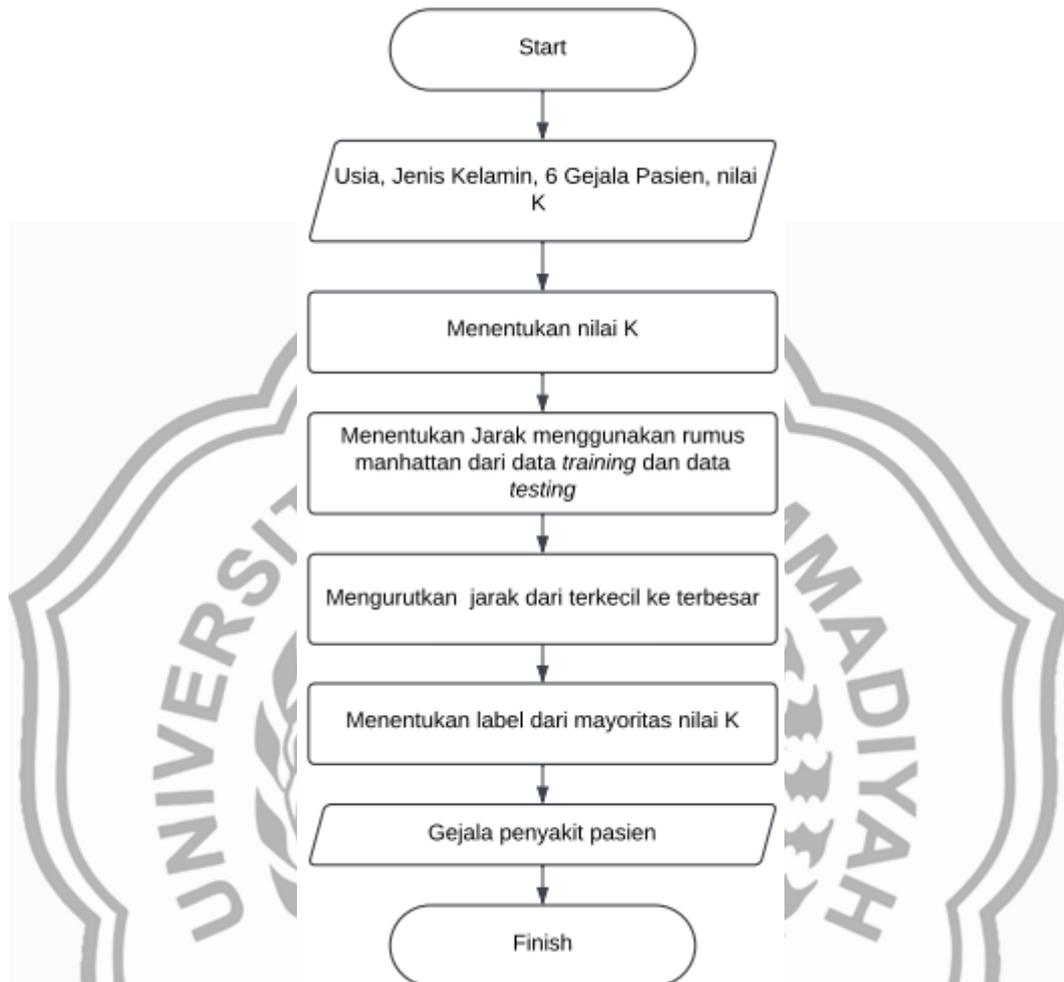
Penyakit demam tifoid merupakan penyakit yang disebabkan oleh bakteri dengan penularan melalui *fecal* dan *oral* dari konsumsi makanan dan minuman yang terkontaminasi atau kebersihan dalam pengolahan makanan. Pada data pasien yang telah dikumpulkan oleh peneliti saat melakukan pemeriksaan oleh dokter beberapa diantaranya memiliki gejala seperti demam tak kunjung turun 3-7 hari atau lebih dari 7 hari, demam disertai menggigil, sakit kepala mual dan munta, nyeri perut, dan nyeri persendian. Pasien yang mengalami gejala tersebut dilakukan pemeriksaan fisik oleh dokter dalam analisis penyakit demam tifoid dan dapat memberikan hasil diagnosa pada pasien. Pada contoh pasien datang untuk melakukan pemeriksaan terkait gejala yang timbul seperti keluhan demam selama empat hari, mengalami nyeri perut, kepala sering sakit, dan mengalami nyeri persendian saat demam. Gejala pasien tersebut dilakukan pemeriksaan pada ruangan dokter untuk mendapatkan hasil pemeriksaan fisik pada pasien berkaitan dengan gejala yang dialami dan dapat memberikan hasil diagnosis pasien terkait demam tifoid.

Analisa terhadap pasien untuk mendapatkan gejala yang menghasilkan diagnosa pasien diperlukan dalam pembuatan sistem prediksi penyakit demam tifoid. Sistem tersebut digunakan untuk pengujian data pasien yang telah dikumpulkan oleh peneliti. Gejala yang telah dikumpulkan dalam sebuah data diuji untuk dapat menentukan hasil diagnosa pasien penyakit demam tifoid dengan penerapan teknologi dalam proses diagnosa.

3.2 Hasil Analisis Sistem

Penelitian diagnosa demam tifoid pada pasien dari data yang telah dikumpulkan dipilih menggunakan algoritme *K-Nearest Neighbor*. Penerapan *K-Nearest Neighbor* diharapkan dapat memberikan kontribusi bagi dokter dalam penanganan pasien demam tifoid. Berikut ini merupakan alur dari algoritme *K-*

Nearest Neighbor dalam prediksi pasien demam tifoid dapat dilihat pada Gambar 3.1 :



Gambar 3. 1 Diagram Alir Algoritme *K-Nearest Neighbor*

Pada gambar 3.1 merupakan gambaran tahapan dari Algoritme *K-Nearest Neighbor* pada sistem uji gejala demam tifoid. Tahap pertama dokter memasukkan data testing beserta nilai K , tahap kedua menentukan nilai K (jumlah jarak terdekat) pengambilan tersebut berdasarkan nilai K yang di input oleh dokter, tahap ketiga proses perhitungan data *training* dan data *testing* dengan rumus jarak *Manhattan*, tahap keempat pengurutan nilai jarak antar data tahap ini membantu proses penentuan label dengan melihat mayoritas K , tahap kelima menentukan kelas hasil dari nilai K terbanyak, setelah itu sistem akan mengeluarkan output berupa hasil perhitungan beserta prediksi dari pengujian data yang dimasukkan oleh dokter. Dalam proses menentukan nilai kelas yang dapat menghasilkan klasifikasi penyakit

demam tifoid membutuhkan data untuk perhitungan algoritme *K-Nearest Neighbor*. Data pasien demam tifoid berdasarkan kunjungan pasien selama bulan Januari – April 2024 dalam proses pengumpulan data , peneliti mendapatkan 6 gejala pada pasien yang diduga terdiagnosis demam tifoid, dapat dilihat pada tabel berikut ini :

Tabel 3.1 Atribut Data (Input)

Atribut	Kode	Type Data	Keterangan
Usia	usia	Numerik	Usia pasien
Jenis Kelamin	jk	Numerik	Jenis kelamin pasien laki-laki bernilai 1 dan jenis kelamin perempuan bernilai 0
Demam	G01	Numerik	Demam Akut (<7 Hari) : 0 Demam Sub-akut (>7 Hari) : 1
Menggigil	G02	Numerik	Demam disertai menggigil di tuliskan dengan angka 1 = ya dan 0 = tidak
Sakit Kepala	G03	Numerik	Sakit kepala adalah salah satu gejala pasien demam tifoid, dituliskan menjadi 2 kategori Ya = 1 dan Tidak = 0
Mual dan Muntah	G04	Numerik	Mual dan muntah gejala gangguan pencernaan yang dapat menjadi pemicu demam tifoid dituliskan pada sebuah data ya bernilai 1 dan tidak bernilai 0
Nyeri Perut	G05	Numerik	Nyeri pada perut gejala gangguan pencernaan dituliskan pada sebuah data ya bernilai 1 dan tidak bernilai 0
Nyeri Persendian	G06	Numerik	Nyeri persendian pada pasien diduga terdiagnosa demam tifoid dapat disebabkan oleh demam tinggi pada pasien, dapat dituliskan dalam sebuah data dengan ya = 1 dan tidak = 0

Pada Tabel 3.1 atribut dalam penelitian ini berupa data gejala pasien dari hasil pengumpulan data pasien pada klinik kesehatan, terdiri dari usia, jk (jenis kelamin), demam, menggigil, sakit kepala, mual dan muntah, nyeri perut, nyeri persendian.

Perolehan data pasien pada penelitian demam tifoid menghasilkan output berupa hasil klasifikasi pasien yang dinyatakan terdiagnosis demam tifoid dapat dilihat pada Tabel 3.2 :

Tabel 3.2 Atribut data (output)

Atribut	Tipe Data	Keterangan
Kelas	Kategori	Positive (+) dan Negative (-)

3.3 Representasi Model

Data pasien dikumpulkan dari sebuah klinik kesehatan dengan indikasi penyakit demam tifoid. Proses pengujian terhadap sebuah algoritme dari pasien terkumpul sejumlah 127 data pasien. Pengumpulan data meliputi gejala penyakit demam tifoid yang akan di klasifikasi menggunakan algoritme *K-Nearest Neighbor*. Data pasien dapat dilihat pada Tabel 3.3 :

Tabel 3.3 Data Pasien

No	usia	jk	G01	G02	G03	G04	G05	G06	Kelas
1	20	L	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
2	15	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Tidak	Tidak	+
3	19	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Ya	Tidak	+
4	13	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
5	30	L	3 Hari - 7 Hari	Ya	Ya	Tidak	Tidak	Ya	+
6	56	L	3 Hari - 7 Hari	Ya	Ya	Ya	Tidak	Ya	+
7	69	L	3 Hari - 7 Hari	Tidak	Tidak	Ya	Tidak	Tidak	+
8	12	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
9	14	L	>7 Hari	Tidak	Ya	Ya	Tidak	Tidak	-
10	6	L	3 Hari - 7 Hari	Tidak	Ya	Ya	Ya	Tidak	+
11	14	L	< 3 Hari	Tidak	Ya	Ya	Ya	Tidak	-
12	15	L	3 Hari - 7 Hari	Tidak	Tidak	Tidak	Tidak	Tidak	+
13	32	P	3 Hari - 7 Hari	Ya	Tidak	Tidak	Tidak	Tidak	+
14	50	P	Tidak Demam	Tidak	Ya	Ya	Ya	Tidak	-
15	48	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Ya	Tidak	+
16	21	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Tidak	Tidak	+
17	53	P	3 Hari - 7 Hari	Ya	Ya	Ya	Ya	Tidak	+
18	10	L	3 Hari - 7 Hari	Tidak	Ya	Tidak	Ya	Tidak	+
...
127	11	P	>7 Hari	Tidak	Ya	Ya	Ya	Tidak	+

Data pasien yang telah dikumpulkan akan melalui *preprocessing*. Pada beberapa data tersebut, dilakukan pembersihan data untuk memperbaiki data agar sesuai dengan gejala penyakit demam tifoid. Dari total 127 data pasien setelah melalui proses pembersihan data tersisa 93 data yang sesuai dengan gejala demam tifoid. Hasil pembersihan data secara lengkap pada Tabel 3.4 :

Tabel 3.4 Hasil Cleaning Data

No	usia	jk	G01	G02	G03	G04	G05	G06	Kelas
1	20	L	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
2	15	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Tidak	Tidak	+
3	19	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Ya	Tidak	+
4	13	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
5	30	L	3 Hari - 7 Hari	Ya	Ya	Tidak	Tidak	Ya	+
6	56	L	3 Hari - 7 Hari	Ya	Ya	Ya	Tidak	Ya	+
7	69	L	3 Hari - 7 Hari	Tidak	Tidak	Ya	Tidak	Tidak	+
8	12	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Tidak	Tidak	+
9	14	L	>7 Hari	Tidak	Ya	Ya	Tidak	Tidak	-
10	6	L	3 Hari - 7 Hari	Tidak	Ya	Ya	Ya	Tidak	+
11	15	L	3 Hari - 7 Hari	Tidak	Tidak	Tidak	Tidak	Tidak	+
12	32	P	3 Hari - 7 Hari	Ya	Tidak	Tidak	Tidak	Tidak	+
13	48	P	3 Hari - 7 Hari	Tidak	Ya	Tidak	Ya	Tidak	+
14	21	P	3 Hari - 7 Hari	Tidak	Ya	Ya	Tidak	Tidak	+
15	53	P	3 Hari - 7 Hari	Ya	Ya	Ya	Ya	Tidak	+
16	10	L	3 Hari - 7 Hari	Tidak	Ya	Tidak	Ya	Tidak	+
17	5	P	3 Hari - 7 Hari	Tidak	Tidak	Tidak	Tidak	Tidak	-
18	72	P	3 Hari - 7 Hari	Tidak	Tidak	Ya	Ya	Tidak	+
...
93	11	P	>7 Hari	Tidak	Ya	Ya	Ya	Tidak	+

Sebelum memasuki tahap perhitungan menggunakan algoritme *K-Nearest Neighbor*, untuk mempersiapkan data yang berbeda type diperlukan proses normalisasi data dan pengubahan data berupa teks menjadi data numerik untuk mendukung kinerja dalam pengolahan data algoritme KNN. Penggunaan normalisasi data MINMAX untuk mengatasi data yang tidak sesuai dengan rentang tertentu pada *preprocessing* penelitian. Berikut contoh pada data nomor 1 dilakukan proses normalisasi dan perubahan data teks menjadi data numerik :

- a. usia : Pada atribut usia data dikelola menggunakan rumus MINMAX, contoh perhitungan nilai usia menggunakan rumus MINMAX dengan data pasien nomor 1 berusia 20 tahun , nilai terkecil (X_{max}) pada usia pasien adalah 1 tahun

dan nilai terbesar (X_{max}) usia pasien adalah 72 pada Tabel 3.4. Berikut adalah perhitungan data nomor 1 dengan Persamaan 2.1:

$$X_{new} = \frac{(20 - 1)}{(72 - 1)}$$

$$X_{new} = \frac{19}{71}$$

$$X_{new} = 0,268$$

- b. jk (Jenis Kelamin) : Pada atribut jk atau jenis kelamin memiliki format awal berupa sebuah teks yang perlu dirubah menjadi numerik untuk proses perhitungan algoritme *K-Nearest Neighbor*. Dalam proses perubahan data pada atribut jenis kelamin tidak memiliki urutan atau peringkat antar kategori seperti pada tabel berikut ini :

Tabel 3.5 Kategorial Jenis Kelamin

Jenis Kelamin	Nilai Atribut
Laki – Laki	0
Perempuan	1

Pada Tabel 3.5 diterapkan pada sebuah data nomor 1 pada tabel 3.4 pasien memiliki jenis kelamin laki-laki dan data tersebut dirubah menjadi numerik bernilai 0.

- c. G01 (Gejala Demam) : pada atribut G01 yaitu demam pada pasien dua kategori yaitu demam akut 3-7 hari dan demam sub-akut >7 hari, perubahan data teks ke numerik dalam gejala demam memberikan nilai tertinggi terhadap data pasien yang bergejala demam >7 hari, pemberian nilai pada gejala demam pada Tabel 3.6 :

Tabel 3.6 Kategorial G01 (Gejala Demam)

G01 (Gejala Demam)	Keterangan	Nilai kategori
3 – 7 Hari	Demam Akut	0
>7 Hari	Demam Sub-akut	1

Pada Tabel 3.6 diterapkan pada sebuah data nomor 1 pada tabel 3.4 pasien memiliki gejala demam 3-7 hari dan data tersebut dirubah menjadi numerik bernilai 0.

- d. G02 (Gejala Demam Disertai Menggigil) : pada atribut G02 terdapat pasien yang terdiagnosa demam tifoid mengalami gejala demam disertai menggigil. Perubahan data dari teks ke numerik dengan pemberian teks Ya pada pasien mengalami demam disertai menggigil atau pasien Tidak mengalami gejala tersebut digunakan untuk pemberian nilai numerik pasien agar dapat dikelola pada algoritme KNN berikut tabel pemberian nilai numerik terhadap data gejala demam disertai menggigil :

Tabel 3.7 Kategorial G02 (Gejala Demam disertai Menggigil)

G01 (Gejala Demam disertai Menggigil)	Nilai Kategori
Ya	1
Tidak	0

Pada Tabel 3.7 diterapkan pada sebuah data nomor 1 pada tabel 3.4 pasien Tidak memiliki gejala demam disertai menggigil dan data tersebut dirubah menjadi numerik bernilai 0.

- e. G03 (Gejala Sakit Kepala) : pada atribut G03 terdapat pasien yang terdiagnosa demam tifoid mengalami gejala sakit kepala. Pemberian teks Ya dan Tidak untuk menentukan pasien mengalami gejala tersebut dilakukan proses perubahan ke numerik bernilai 1 dan 0 seperti pada tabel 3.7 terkait dengan gejala pasien. Pada data pasien nomor 1 Tabel 3.4 teks menunjukkan Ya, maka pasien tersebut mengalami gejala sakit kepala dan bernilai numerik 1.
- f. G04 (Gejala Mual dan Muntah) : Pemberian teks Ya dan Tidak untuk menentukan pasien mengalami gejala mual dan muntah tersebut dilakukan proses perubahan ke numerik bernilai 1 dan 0 seperti pada tabel 3.7 sesuai dengan gejala pasien. Pada data pasien nomor 1 Tabel 3.4 teks menunjukkan Tidak, maka pasien tersebut tidak mengalami gejala sakit kepala dan bernilai numerik 0.
- g. G05 (Gejala Nyeri perut) : Pemberian teks Ya dan Tidak untuk menentukan pasien mengalami gejala nyeri perut pada kolom tabel G05 untuk proses perubahan teks ke numerik bernilai 1 dan 0 seperti pada tabel 3.7 sesuai dengan gejala pasien. Pada data pasien nomor 1 Tabel 3.4 teks menunjukkan Tidak, maka pasien tersebut tidak mengalami gejala tersebut dan bernilai numerik 0.

- h. G06 (Gejala Nyeri Persendian) : Pemberian teks Ya dan Tidak untuk menentukan pasien mengalami gejala nyeri persendian pada kolom tabel G06 untuk proses perubahan teks ke numerik bernilai 1 dan 0 seperti pada tabel 3.7 sesuai data pasien. Pada data pasien nomor 1 Tabel 3.4 teks menunjukkan Tidak, maka pasien tersebut tidak mengalami gejala tersebut dan bernilai numerik 0.

Data yang telah melalui *preprocessing* disajikan dalam Tabel 3.8. Isi dari tabel tersebut merupakan perubahan dari gejala pada pasien yang berawal dari sebuah teks dirubah menjadi numerik dengan nilai kategori tertentu yang berada pada Tabel 3.4 :

Tabel 3.8 Hasil *Preprocessing*

No	usia	jk	G01	G02	G03	G04	G05	G06	Kelas
1	0,268	0	0	0	1	0	0	0	+
2	0,197	1	0	0	1	1	0	0	+
3	0,254	1	0	0	1	1	1	0	+
4	0,169	1	0	0	1	0	0	0	+
5	0,408	0	0	1	1	0	0	1	+
6	0,775	0	0	1	1	1	0	1	+
7	0,958	0	0	0	0	1	0	0	+
8	0,155	1	0	0	1	0	0	0	+
9	0,183	0	1	0	1	1	0	0	-
10	0,07	0	0	0	1	1	1	0	+
11	0,197	0	0	0	0	0	0	0	+
12	0,437	1	0	1	0	0	0	0	+
13	0,662	1	0	0	1	0	1	0	+
14	0,282	1	0	0	1	1	0	0	+
15	0,732	1	0	1	1	1	1	0	+
16	0,127	0	0	0	1	0	1	0	+
17	0,056	1	0	0	0	0	0	0	-
18	1	1	0	0	0	1	1	0	+
...
93	0,141	1	1	0	1	1	1	0	+

3.3.1 Perhitungan Algoritme *K-Nearest Neighbor* $K = 1$

Algoritme *K-Nearest Neighbor* pada Gambar 3.1 memiliki empat tahap dalam proses perhitungan klasifikasi data yang di uji. Proses pertama adalah menentukan nilai K agar dapat menetapkan kelas dari jumlah mayoritas K terbanyak, kedua

menghitung jarak menggunakan manhattan dari data *training* dan data *testing*, tahap ketiga yaitu mengurutkan hasil jarak yang telah dihitung dari nilai terkecil ke terbesar, tahap keempat menetapkan kelas dari jumlah mayoritas nilai K terbanyak. Berikut merupakan implementasi perhitungan Algoritme *K-Nearest Neighbor* K = 1 :

a. Menentukan Nilai K

Penentuan nilai *K* pada perhitungan klasifikasi demam tifoid menggunakan nilai $K = 1$.

b. Perhitungan Jarak

Pada tahap perhitungan jarak menggunakan rumus *Manhattan*, perhitungan tersebut membutuhkan data *training* dan data *testing* dalam pengujian. Berikut merupakan data *training* berjumlah 10 data yang diambil secara acak pada data yang telah melalui *preprocessing* di Tabel 3.8 :

Tabel 3.9 Data *training* perhitungan jarak

No	Usia	jk	G01	G02	G03	G04	G05	G06	Kelas
1	0,775	0	0	1	1	1	0	1	+
2	0,732	1	0	1	1	1	1	0	+
3	0,056	1	0	0	0	0	0	0	-
4	0,634	0	0	1	0	1	1	1	+
5	0,085	0	0	0	1	0	0	0	-
6	0,324	0	0	0	1	0	0	0	-
7	0,225	0	1	0	1	1	1	0	+
8	0,606	0	0	0	0	0	0	1	-
9	0,239	1	0	0	1	0	0	0	-
10	0,465	1	1	1	0	1	1	0	+

Perhitungan jarak *Manhattan* memerlukan data *testing* sebagai pengujian dalam penentuan nilai K pada algoritme *K-Nearest Neighbor*. Berikut merupakan data *testing* sebanyak 20 data yang diambil dari hasil *preprocessing* data pada Tabel 3.10 :

Tabel 3.10 Data *testing* perhitungan jarak

No	Usia	jk	G01	G02	G03	G04	G05	G06	Kelas
1	0,254	1	0	0	1	1	1	0	+
2	0,183	0	1	0	1	1	0	0	-

3	0,268	0	0	0	0	1	1	0	-
4	0,676	0	0	0	1	1	1	0	-
5	0,338	1	0	1	1	1	1	0	+
6	0,38	0	0	1	1	1	0	0	+
7	0,366	1	1	0	1	1	0	1	-
8	0,282	1	0	0	1	0	0	1	-
9	0,113	1	0	0	0	0	1	0	-
10	0,535	1	0	0	0	1	1	1	+
11	0,366	1	0	0	1	0	0	0	+
12	0,042	1	0	0	0	1	1	0	+
13	0,845	0	0	1	0	1	1	0	-
14	0,845	0	0	0	0	0	0	1	-
15	0,211	0	0	0	0	0	0	1	-
16	0,535	1	0	0	0	0	0	1	-
17	0,099	1	0	0	1	0	0	0	-
18	0,366	1	0	1	1	0	0	1	+
19	0,239	1	0	0	1	1	1	1	+
20	0,141	1	1	0	1	1	1	0	+

Pengolahan data *training* dan data *testing* menggunakan rumus jarak *Manhattan* agar dapat menentukan kelas data baru dengan hasil perhitungan jarak terdekat (*K* tetangga terdekat). Berikut contoh perhitungan dari data *testing* 1 dan data *training* menggunakan Persamaan 2.2 :

$$d = |0,775 - 0,254| + |0 - 1| + |0 - 0| + |1 - 0| + |1 - 1| + |1 - 1| \\ + |0 - 1| + |1 - 0| \\ d = 4,521$$

Perhitungan jarak *Manhattan* dari data *training* ke-1 dan data *testing* ke-1 memperoleh hasil 4,521. Untuk perhitungan data *testing* dan data *training* lainnya dapat menggunakan rumus yang sama agar dapat memperoleh hasil jarak yang akan digunakan untuk mencari *K* terdekat.

c. Mengurutkan Hasil Jarak

Dari proses perhitungan jarak, untuk menemukan tetangga terdekat hasil perhitungan diurutkan dari nilai terkecil ke yang terbesar. Fungsi dari pengurutan tersebut untuk memastikan nilai terdekat pada jarak yang

digunakan untuk proses klasifikasi data. Berikut merupakan hasil pengurutan nilai terkecil hasil perhitungan jarak pada Tabel 3.11 :

Tabel 3.11 Hasil Pengurutan Jarak *Manhattan* $K=1$

No	Kelas Asli	Jarak data <i>testing</i> ke data <i>training</i>	Urutan nilai yang terkecil
1	+	4,521	9
2	+	1,478	1
3	+	3,198	6
4	+	4,380	8
5	+	3,169	5
6	-	3,070	4
7	-	2,029	3
8	-	5,352	10
9	-	2,015	2
10	-	3,211	7

Pada Tabel 3.11 hasil perhitungan jarak menggunakan rumus *Manhattan* diperoleh hasil jarak dengan nilai terkecil yaitu pada nomor 2.

d. Penentuan Kelas

Tahap penentuan kelas dari hasil pengurutan nilai terkecil diperoleh dari Tabel 3.8, ditunjukkan pada nomor 2 kelas asli pada data *training* merupakan Kelas (+) positif terdiagnosa demam tifoid pada tetangga terdekat yang ditentukan dengan $K=1$ menunjukkan hasil dari data uji nomor 1 terhadap data latih dengan jarak 1,478 memiliki label baru berkelas “**Positif**”. Hasil keseluruhan penentuan kelas dari 20 data *testing* yang di uji di paparkan dalam Tabel 3.12 :

Tabel 3.12 Hasil Perhitungan *K-Nearest Neighbor* $K = 1$

No	Usia	jk	G01	G02	G03	G04	G05	G06	Kelas Asli	Prediksi
1	0,254	1	0	0	1	1	1	0	+	+
2	0,183	0	1	0	1	1	0	0	-	+
3	0,268	0	0	0	0	1	1	0	-	+
4	0,676	0	0	0	1	1	1	0	-	+
5	0,338	1	0	1	1	1	1	0	+	+
6	0,38	0	0	1	1	1	0	0	+	+

7	0,366	1	1	0	1	1	0	1	-	+
8	0,282	1	0	0	1	0	0	1	-	-
9	0,113	1	0	0	0	0	1	0	-	-
10	0,535	1	0	0	0	1	1	1	+	+
11	0,366	1	0	0	1	0	0	0	+	-
12	0,042	1	0	0	0	1	1	0	+	-
13	0,845	0	0	1	0	1	1	0	-	+
14	0,845	0	0	0	0	0	0	1	-	-
15	0,211	0	0	0	0	0	0	1	-	-
16	0,535	1	0	0	0	0	0	1	-	-
17	0,099	1	0	0	1	0	0	0	-	-
18	0,366	1	0	1	1	0	0	1	+	+
19	0,239	1	0	0	1	1	1	1	+	+
20	0,141	1	1	0	1	1	1	0	+	+

Hasil perhitungan data *testing* keseluruhan terhadap data *training* menggunakan algoritme *K-Nearest Neighbor* $K=1$ pada tabel 3.12 terdapat 7 pasien terdiagnosis positif penyakit demam tifoid dan 13 pasien terdiagnosis negatif penyakit demam tifoid. Hasil dari proses prediksi tersebut menunjukkan 13 data uji sesuai dengan hasil prediksi dan 7 data uji tidak sesuai dengan hasil prediksi (kolom berwarna merah).

e. Pengujian *Confusion Matrix*

Proses selanjutnya yaitu mengevaluasi kinerja dari algoritme *K-Nearest Neighbor* menggunakan *Confusion Matrix*. Pengujian ini berfungsi untuk mengevaluasi hasil klasifikasi dengan cara membandingkan hasil prediksi model dengan data aktual. Berikut contoh evaluasi pada hasil klasifikasi dari Algoritme *K-Nearest Neighbor* $K = 1$ pada Tabel 3.13 :

Tabel 3.13 *Confusion Matrix* Algoritme *K-Nearest Neighbor* $K = 1$

<i>True Label</i>	<i>Negative</i>	6	5
	<i>Positive</i>	2	7
		<i>Negative</i>	<i>Positive</i>
		<i>Predict Label</i>	

Pada Tabel 3.13 hasil dari kinerja algoritme *K-Nearest Neighbor* $K=1$ menghasilkan 7 (*True Positif*) yang berarti model tersebut diprediksi positif

dan benar, 6 (*True Negative*) model yang diuji diprediksi negatif dan benar, 5 (*False Positive*) yang berarti label asli adalah negatif dan diprediksi positif, dan 2 (*False Negative*) yang berarti kelas asli positif tetapi dari 2 sample tersebut diprediksi negatif. Setelah mengisi dari tabel confusion matrix pengukuran kinerja model dapat menggunakan akurasi, *recall*, dan *precision* menggunakan rumus Persamaan 2.3 , Persamaan 2.4, dan Persamaan 2.5 :

$$\text{Akurasi} = \frac{7+6}{7+6+5+2} \times 100\% = 65\%$$

$$\text{Recall} = \frac{7}{7+2} \times 100\% = 77,78\%$$

$$\text{Precision} = \frac{7}{7+5} \times 100\% = 58,3\%$$

Dari hasil evaluasi kinerja 20 data *testing* terhadap 10 data *training* menggunakan algoritme *K-Nearest Neighbor* mendapatkan hasil akurasi 65%, *recall* 77,78%, dan *precision* 58,3% .

1.3.2 Perhitungan Algoritme *K-Nearest Neighbor* $K = 5$

a. Menentukan Nilai K

Penentuan nilai K pada perhitungan klasifikasi demam tifoid menggunakan nilai $K = 5$.

b. Perhitungan Jarak

Perhitungan jarak menggunakan rumus *Manhattan* dan menggunakan data *training* dan data *testing* pada Tabel 3.6 dan Tabel 3.7. Berikut contoh perhitungan data ke 1 dari data *training* dan data *testing* menggunakan rumus Persamaan 2.2 :

$$d = |0,775 - 0,254| + |0 - 1| + |0 - 0| + |1 - 0| + |1 - 1| + |1 - 1| \\ + |0 - 1| + |1 - 0|$$

$$d = 4,521$$

Perhitungan jarak *Manhattan* dari data *training* ke-1 dan data *testing* ke-1 memperoleh hasil 4,521. Proses perhitungan dilanjutkan sampai data *testing* terhitung keseluruhan dengan data *training* menggunakan rumus jarak *Manhattan*.

c. Mengurutkan Hasil Jarak

Dari proses perhitungan jarak, untuk menemukan tetangga terdekat hasil perhitungan diurutkan dari nilai terkecil ke yang terbesar. Berikut merupakan hasil pengurutan nilai terkecil hasil perhitungan jarak pada Tabel 3.11 :

Tabel 3.14 Hasil Pengurutan Jarak *Manhattan* $K=5$

No	Kelas Asli	Jarak data <i>testing</i> ke data <i>training</i>	Urutan nilai yang terkecil
1	+	4,521	9
2	+	1,478	1
3	+	3,198	6
4	+	4,380	8
5	+	3,169	5
6	-	3,070	4
7	-	2,029	3
8	-	5,352	10
9	-	2,015	2
10	-	3,211	7

d. Penentuan Kelas

Tahap penentuan kelas dari hasil pengurutan nilai terkecil diperoleh dari Tabel 3.14 ditunjukkan pada nomor 2,9,7,6,dan 5. Pada hasil $K=5$ tersebut kelas asli yang menunjukkan negatif berada pada nomor 6,7, dan 9, sedangkan kelas asli yang menunjukkan positif berada pada nomor 2 dan 5. Maka nilai mayoritas kelas terbanyak yang mendekati tetangga adalah “Negatif” Hasil keseluruhan penentuan kelas dari 20 data *testing* yang di uji di paparkan dalam Tabel 3.15 :

Tabel 3.15 Hasil Perhitungan *K-Nearest Neighbor* $K = 5$

No	Usia	jk	G01	G02	G03	G04	G05	G06	Kelas Asli	Prediksi
1	0,254	1	0	0	1	1	1	0	+	-
2	0,183	0	1	0	1	1	0	0	-	-
3	0,268	0	0	0	0	1	1	0	-	+
4	0,676	0	0	0	1	1	1	0	-	-
5	0,338	1	0	1	1	1	1	0	+	+
6	0,38	0	0	1	1	1	0	0	+	-

7	0,366	1	1	0	1	1	0	1	-	+
8	0,282	1	0	0	1	0	0	1	-	-
9	0,113	1	0	0	0	0	1	0	-	-
10	0,535	1	0	0	0	1	1	1	+	+
11	0,366	1	0	0	1	0	0	0	+	-
12	0,042	1	0	0	0	1	1	0	+	+
13	0,845	0	0	1	0	1	1	0	-	+
14	0,845	0	0	0	0	0	0	1	-	-
15	0,211	0	0	0	0	0	0	1	-	-
16	0,535	1	0	0	0	0	0	1	-	-
17	0,099	1	0	0	1	0	0	0	-	-
18	0,366	1	0	1	1	0	0	1	+	-
19	0,239	1	0	0	1	1	1	1	+	+
20	0,141	1	1	0	1	1	1	0	+	+

Hasil perhitungan data *testing* keseluruhan terhadap data *training* menggunakan algoritme *K-Nearest Neighbor K=5* pada tabel 3.15 terdapat 8 gejala pasien terdiagnosis positif demam tifoid dan 12 pasien terdiagnosis negatif penyakit demam tifoid. Hasil dari proses prediksi tersebut menunjukkan 13 data uji sesuai dengan hasil prediksi dan 7 data uji tidak sesuai dengan hasil prediksi (kolom berwarna merah).

e. Pengujian Confusion Matriks

Proses selanjutnya yaitu mengevaluasi kinerja dari algoritme *K-Nearest Neighbor* menggunakan *Confusion Matrix*. Berikut contoh evaluasi pada hasil klasifikasi dari Algoritme *K-Nearest Neighbor K = 5* pada Tabel 3.13 :

Tabel 3.16 *Confusion Matrix* Algoritme *K-Nearest Neighbor K = 5*

<i>True Label</i>	<i>Negative</i>	8	3
	<i>Positive</i>	4	5
		<i>Negative</i>	<i>Positive</i>
		<i>Predict Label</i>	

Pada Tabel 3.16 hasil dari kinerja algoritme *K-Nearest Neighbor K=5* menghasilkan 5 (*True Positif*) yang berarti model tersebut diprediksi positif dan benar, 8 (*True Negative*) model yang diuji diprediksi negatif dan benar, 4 (*False Positive*) yang berarti label asli positif tetapi dari 4 sample tersebut

dinyatakan negatif, dan 3 (*False Negative*) yang berarti label asli negatif tetapi dari 3 sample tersebut dinyatakan positif. Setelah mengisi dari tabel confusion matrix pengukuran kinerja model dapat menggunakan akurasi, *recall*, dan *precision* menggunakan rumus Persamaan 2.3 , Persamaan 2.4, dan Persamaan 2.5 :

$$\text{Akurasi} = \frac{8+5}{8+5+4+3} \times 100\% = 65\%$$

$$\text{Recall} = \frac{8}{8+3} \times 100\% = 72,73\%$$

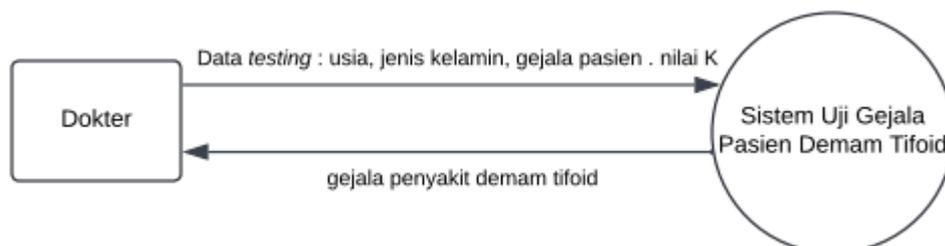
$$\text{Precision} = \frac{8}{8+4} \times 100\% = 66,67\%$$

Hasil evaluasi kinerja algoritme K-Nearest Neighbor K=5 mendapatkan hasil akurasi sebesar 65% dengan recall 72,73% precision sebesar 66,67%.

3.4 Perancangan Sistem

3.4.1 Diagram Konteks

Diagram konteks pada sistem untuk menggambarkan entitas luar atau yang terlibat dalam sistem.

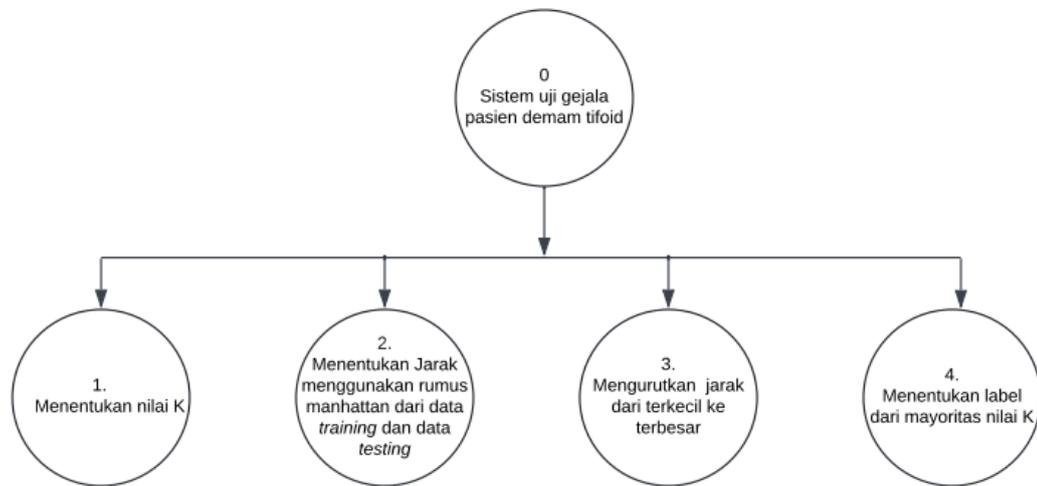


Gambar 3. 2 Diagram Konteks Sistem Uji Gejala Pasien Demam Tifoid

Pada Gambar 3.2 menggambarkan sistem klasifikasi demam tifoid bekerja, dokter merupakan *entity* dalam sistem klasifikasi demam tifoid. dokter menginput data gejala pasien yang diduga terdiagnosis demam tifoid pada sistem klasifikasi. Sistem klasifikasi memberikan *ouput* berupa hasil klasifikasi kepada dokter.

3.4.2 Diagram Jenjang

Diagram jenjang merupakan gambaran hubungan dan tingkatan elemen yang terdapat pada sistem berdasarkan hirarki atau urutan dari sistem tersebut. Berikut merupakan gambaran dari Sistem klasifikasi demam tifoid pada Gambar 3.3 :



Gambar 3. 3 Diagram Jenjang Sistem Uji Gejala Pasien

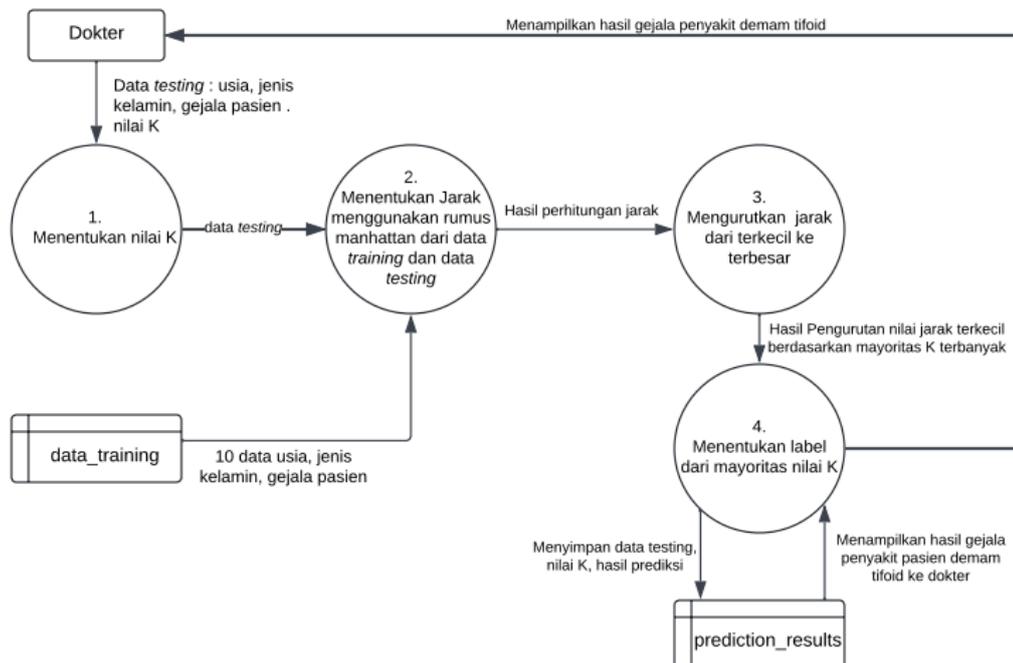
Pada Gambar 3.3 menggambarkan diagram berjenjang dari sistem klasifikasi dengan urutan yang terdiri dari :

2. Top level : Sistem uji gejala penyakit demam tifoid menggunakan algoritme *K-Nearest Neighbor*.
3. Level 1 : Bagian dari proses sistem uji gejala penyakit demam tifoid meliputi proses dalam algoritme *K-Nearest Neighbor*.

3.4.3 Data Flow Diagram (DFD)

3.4.3.1 Data Flow Diagram Level 1

Data Flow Diagram Level 1 merupakan proses utama yang telah dipecah menjadi subproses lebih rinci dan memberikan pandangan kepada *dokter* terkait alur dari sistem. Berikut adalah gambar dari DFD Level 1 pada Gambar 3.4 :



Gambar 3. 4 *Data Flow Diagram* Level 1 Sistem Uji Gejala Pasien

Pada Gambar 3.4 merupakan aktivitas dari DFD level 1 dengan tahapan proses dalam algoritme KNN yang terdapat pada sistem. Berawal dari dokter menginputkan gejala pasien berupa data *testing* beserta nilai K dan 10 data *training*, kemudian ke proses menentukan nilai K, dan menghitung jarak dengan rumus *Manhattan*, proses selanjutnya mengurutkan jarak dari nilai terkecil, dalam proses ini sebagai pengumpulan nilai mayoritas K terdekat dari nilai K yang ditentukan, proses tersebut disimpan dalam *database* *prediction_results* dan ditampilkan ke dokter dalam berupa gejala penyakit demam tifoid pada pasien

3.5 Perancangan Basis Data

Perancangan basis data merupakan proses perancangan struktur penyimpanan data dari sistem yang akan digunakan dalam proses pengujian data pada algoritme *K-Nearest Neighbor*. Pada Sistem Klasifikasi Penyakit Demam Tifoid penggunaan *database* terbagi dari empat tabel yaitu, tabel *users*, tabel *data_training*, dan tabel *prediction-results*. Berikut merupakan rincian dari tiga tabel sistem :

3.5.1 Tabel *users*

Tabel *users* pada *database* sistem klasifikasi digunakan untuk menyimpan data *users* untuk membuka laman klasifikasi sistem dengan memasukkan data pada halaman *login*. Berikut merupakan tabel dari *user* pada Tabel 3.17 :

Tabel 3.17 Tabel *Dokter* Sistem Klasifikasi Penyakit Demam Tifoid

Name	Type	Deskripsi
<i>_id</i>	String	PK, Unique
<i>username</i>	String	
<i>password</i>	String	

3.5.2 Tabel *data_training*

Tabel *data_training* pada *database* sistem uji gejala penyakit demam tifoid digunakan untuk menyimpan data training yang digunakan perhitungan jarak menggunakan rumus Manhattan. Berikut merupakan tabel dari *training* pada Tabel 3.18 :

Tabel 3.18 Tabel *data_training* Sistem Uji Gejala Penyakit Demam Tifoid

Name	Type	Deskripsi
<i>_id</i>	String	PK, Unique
<i>feature1</i>	INT	
<i>feature2</i>	INT	
<i>feature3</i>	INT	
<i>feature4</i>	INT	
<i>feature5</i>	INT	
<i>feature6</i>	INT	
<i>feature7</i>	INT	
<i>feature8</i>	INT	
<i>label</i>	String	

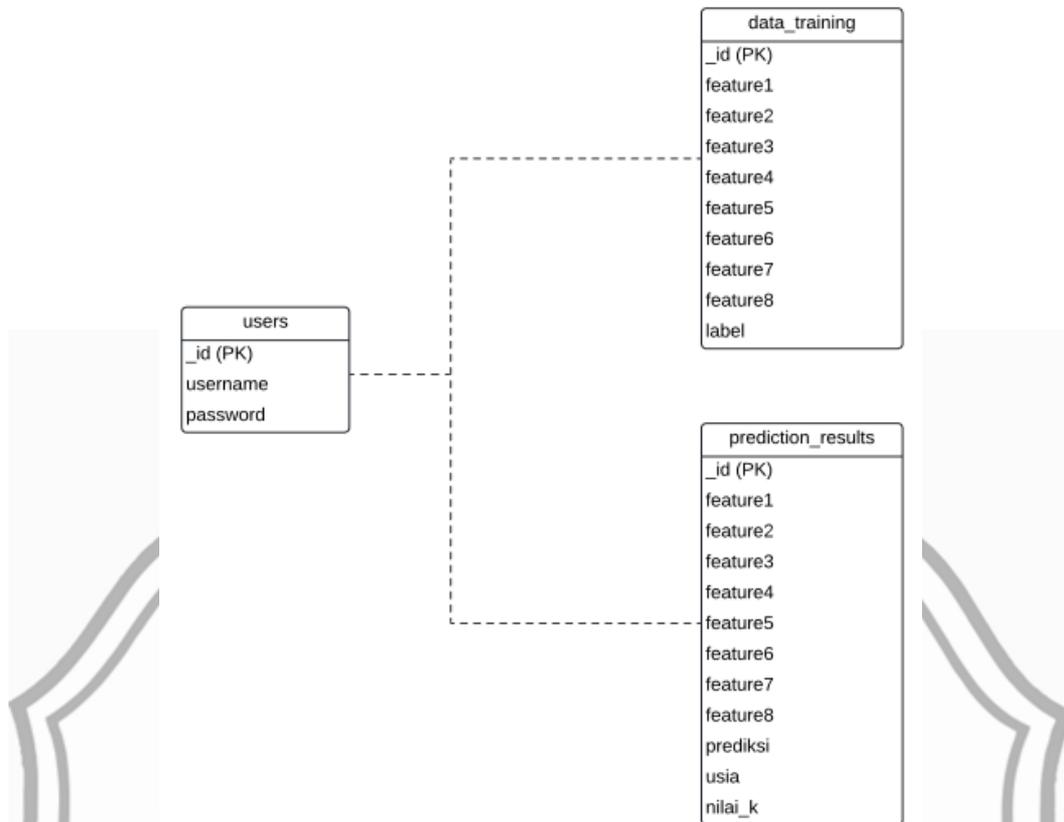
3.5.3 Tabel prediction_results

Tabel prediction_results pada *database* sistem klasifikasi digunakan untuk menyimpan data *testing*, hasil prediksi, usia, dan nilai K pada pengujian KNN dengan rumus jarak Manhattan. Berikut merupakan tabel dari prediction_results pada Tabel 3.19 :

Tabel 3.19 Tabel data_testing Sistem Klasifikasi Penyakit Demam Tifoid

Name	Type	Deskripsi
_id	String	PK, Unique
feature1	INT	
feature2	INT	
feature3	INT	
feature4	INT	
feature5	INT	
feature6	INT	
feature7	INT	
feature8	INT	
prediksi	String	
usia	INT	
nilai_k	INT	

Sistem ini memiliki database yang menyimpan data-data yang diperlukan saat pengujian data menggunakan algoritme KNN. Tabel dokter berhubungan dengan tabel data_training dan tabel prediction_results, tipe relasi dari tabel dokter ke tabel data_training dan tabel prediction_results adalah *one to many*. Berikut merupakan gambaran hubungan setiap tabel yang tesaji dalam ERD pada Gambar 3.6 :



Gambar 3. 5 ERD (*Entity Relationship Diagram*)

3.6 Perancangan Antarmuka

a. Halaman Login

Halaman *login* merupakan tampilan awal dari sebuah sistem. Dokter pada sistem harus memasukkan dokter *username* dan *password* yang sesuai untuk dapat masuk ke dalam sistem dengan klik “button” *login*. Apabila dokter tidak memiliki akun, dokter dapat melakukan pendaftaran akut dengan klik “button” daftar akun. Tampilan rancangan halaman *login* terdapat pada Gambar 3.7 :

Sistem Uji Gejala Demam Tifoid Menggunakan Algoritme KNN

Login

Username

Password

[Don't have an account? Register here](#)

Gambar 3. 6 Perancangan Antarmuka Halaman *Login*

b. Halaman Register User

Dalam halaman *register user* menampilkan form input untuk mendaftarkan sebuah username dan password agar dapat login dan menampilkan halaman Klasifikasi. Berikut rancangan tampilan halaman Register pada Gambar 3.7 :

REGISTER

Username

Password

[Don't have an account? Register here](#)

Gambar 3. 7 Perancangan Antarmuka Halaman Register

Logo	Klasifikasi	Data Klasifikasi	Logout
UJI GEJALA DEMAM TIFOID MENGGUNAKAN ALGORITME K-NEAREST NEIGHBOR			
Usia	:	<input type="text"/>	
Jenis Kelamin	:	<input type="radio"/> Perempuan	<input type="radio"/> Laki-Laki
Demam	:	<input type="radio"/> 3 - 7 Hari	<input type="radio"/> Lebih dari 7 hari
Disertai Mengigil	:	<input type="radio"/> Ya	<input type="radio"/> Tidak
Sakit Kepala	:	<input type="radio"/> Ya	<input type="radio"/> Tidak
Mual dan Muntah	:	<input type="radio"/> Ya	<input type="radio"/> Tidak
Nyeri Perut	:	<input type="radio"/> Ya	<input type="radio"/> Tidak
Nyeri Persendian	:	<input type="radio"/> Ya	<input type="radio"/> Tidak
Nilai K	:	<input type="radio"/> 3	<input type="radio"/> 5 <input type="radio"/> 7 <input type="radio"/> 9
<input type="button" value="Classify"/>			
"HASIL KLASIFIKASI"			

Gambar 3. 9 Perancangan Antarmuka Halaman Klasifikasi

e. Halaman Edit Data Klasifikasi

Halaman edit data klasifikasi merupakan rancangan tampilan untuk merubah sebuah gejala yang telah di input oleh dokter apabila ada kesalahan, sistem menampilkan kembali ke halaman data klasifikasi beserta hasil yang telah dirubah dan prediksi yang di perbarui.

3.7 Perancangan Pengujian Sistem

No	Test	Input	Expected Result	Status
1.	Dokter Login	<i>Username</i> dan <i>password</i>	Sistem dapat menampilkan halaman klasifikasi	
2.	Dokter memasuki halaman klasifikasi	Klik navbar "Klasifikasi"	Sistem dapat menampilkan Halaman Klasifikasi	

3.	Dokter dapat menguji klasifikasi	Usia, jk, G01, G02, G03, G04, G05, G06,	Sistem menampilkan hasil proses klasifikasi	
4.	Dokter dapat melihat data pasien yang telah dilakukan klasifikasi	Klik navbar “Data Klasifikasi”	Sistem dapat menampilkan Halaman Klasifikasi dan Hasil data pasien yang telah diklasifikasi	
5.	Dokter dapat logout dari sistem	Klik Logout	Dokter kembali ke halaman login	

