

## BAB V

### PENGUJIAN DAN ANALISIS

Pada bab ini akan dilakukan pengujian dan analisis dari penerapan *Fuzzy k-Nearest Neighbor* (Fk-NN) untuk diagnosa penderita *liver* berdasarkan *Indian Liver Patient Dataset*.

#### 5.1 Pengujian

Pengujian dilakukan untuk mengetahui hasil akurasi dari implementasi yang dilakukan. Pengujian dilakukan dengan 3 macam jenis pengujian yaitu pengujian tingkat akurasi terhadap data latih, pengujian tingkat akurasi terhadap banyaknya nilai  $k$ , dan pengujian tingkat akurasi terhadap sebaran kelas.

##### 5.1.1 Pengujian tingkat akurasi terhadap data latih

Pada pengujian tingkat akurasi terhadap data latih, data latih yang digunakan meliputi 80, 110, 140, 170, 200, dan 230 data. Tabel 5.1 berikut ini menunjukkan pengaruh data latih terhadap tingkat akurasi, dengan data uji yang sama yaitu 50 data uji dan nilai  $k=3$ .

Tabel 5.1 Hasil uji terhadap jumlah data latih

Data Latih	FKNN Akurasi (%)	KNN Akurasi (%)
80	64 %	52%
110	57 %	45%
140	68 %	56%
170	76 %	62%
200	60 %	52%
230	64 %	54%

Dari hasil yang terdapat pada tabel 5.1 dapat disimpulkan bahwa adanya fluktuatif di data latih membuat hasil akurasi tidak baik.

### 5.1.2 Pengujian tingkat akurasi terhadap nilai k

Pada pengujian tingkat akurasi terhadap pengaruh nilai k, data latih yang digunakan adalah 80 data. Sedangkan data uji yang digunakan adalah 50 data. Untuk nilai k yang digunakan adalah  $k = 2$  sampai dengan  $k = n$ . Hasil dari pengujian tingkat akurasi terhadap nilai k dapat dilihat pada tabel 5.2.

Tabel 5.2 Hasil uji terhadap nilai k

k	FKNN Akurasi (%)	KNN Akurasi (%)
2	68 %	66%
3	64 %	52%
4	62 %	62%
5	70 %	60%
6	70 %	68%
7	68 %	66%
8	70 %	70%
9	72 %	68%
10	70 %	70%
14	74 %	70%
15	72 %	70%
16	74 %	70%
19	74 %	72%
20	74 %	74%
25	74 %	74%
30	74 %	74%
35	74 %	74%
40	74 %	74%
45	74 %	74%
50	74 %	74%
80	74%	74%

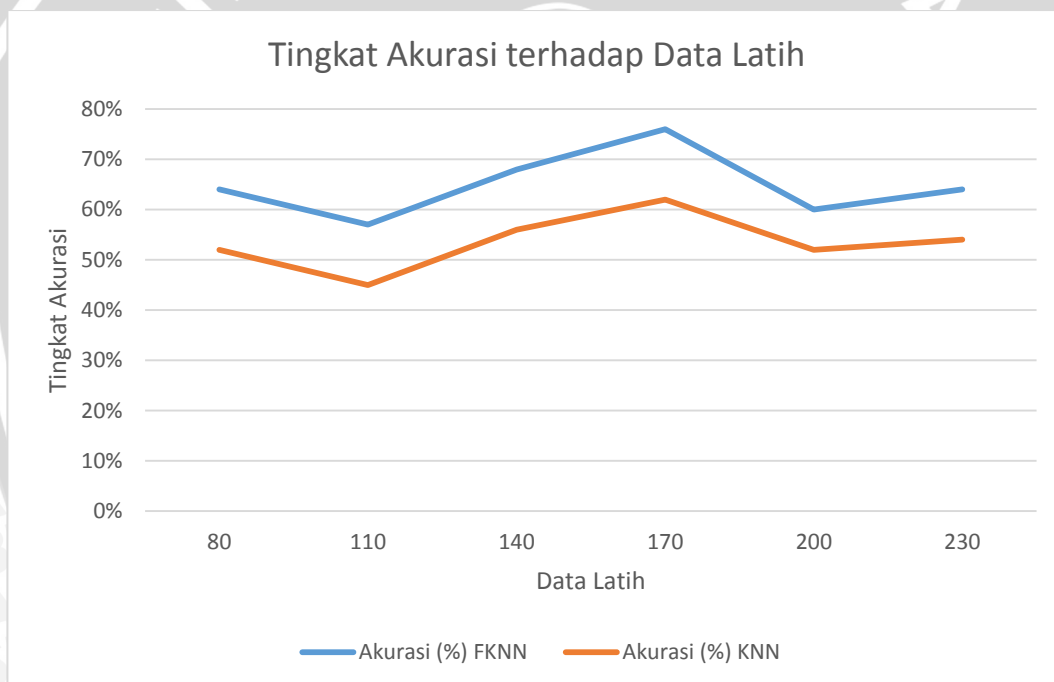
### 5.1.3 Pengujian tingkat akurasi terhadap sebaran data

Pada pengujian tingkat akurasi terhadap sebaran data, data latih yang digunakan meliputi 70, 90, 110, 130, 150, dan 334 data. Sebaran data 1 terdiri dari 70% kelas 1 dan 30% kelas 2, sebaran data 2 terdiri dari 30% kelas 1 dan 70% kelas 2, sedangkan sebaran data 3 terdiri dari 50% kelas 1 dan 50% kelas 2. Tabel 5.3 berikut ini menunjukkan pengaruh sebaran data terhadap tingkat akurasi, dengan data uji yang sama yaitu 50 data uji dimana khusus untuk sebaran data 3 terdiri dari 50% kelas 1 dan 50% kelas 2 serta nilai  $k=3$ .

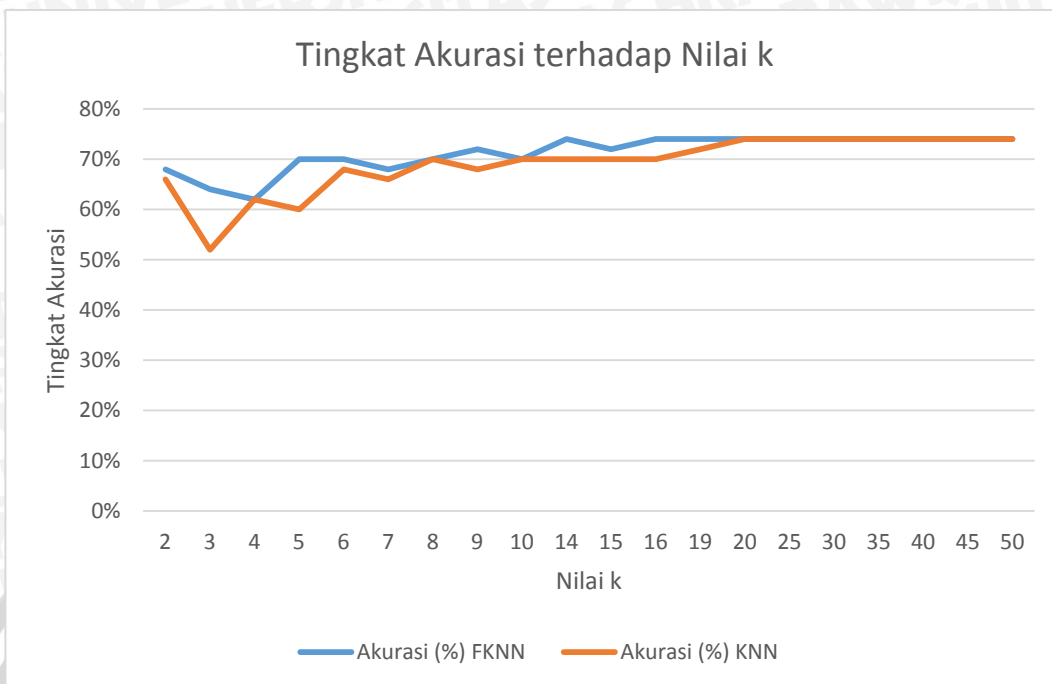
Tabel 5.3 Hasil uji terhadap sebaran data

Data Latih	Sebaran data 1 dengan 70% kelas 1 dan 30% kelas 2	Sebaran data 2 dengan 30% kelas 1 dan 70% kelas 2	Sebaran data 3 dengan 50% kelas 1 dan 50% kelas 2
70	64 %	32%	48%
90	56 %	36%	48%
110	56 %	28%	54%
130	60 %	32%	44%
150	68 %	38%	57%
200	56%	26%	56%

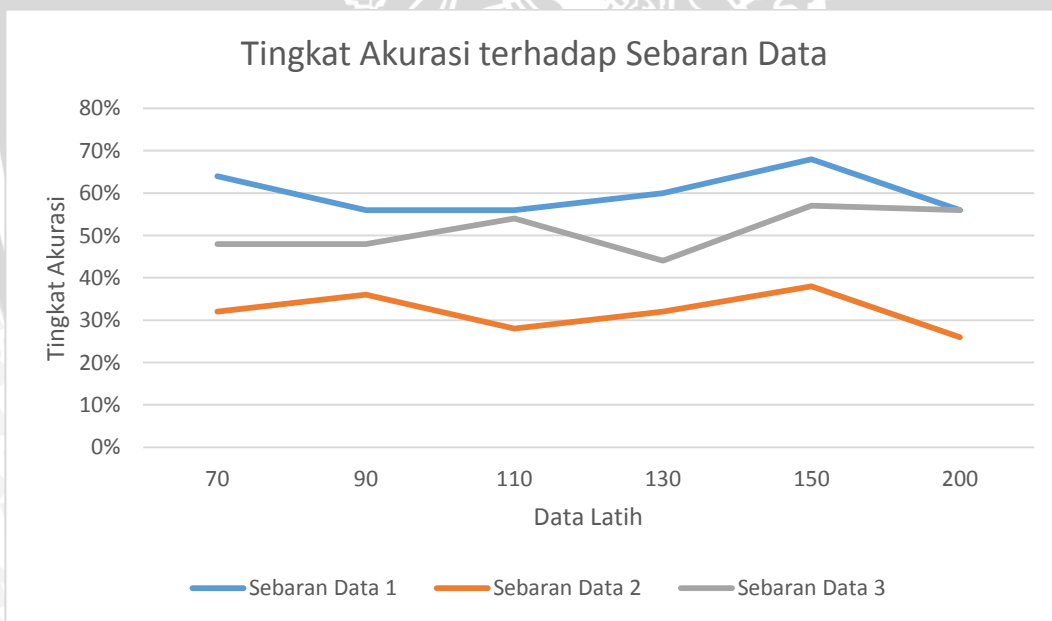
Berdasarkan pengujian tingkat akurasi terhadap data latih, nilai k dan sebaran kelas, maka dibuatlah grafik hubungan tingkat akurasi Fk-NN dengan tingkat akurasi k-NN untuk melihat keakuratan sistem. Grafik ditunjukkan pada grafik 5.1-5.3.



Gambar 5.1 Grafik Tingkat Akurasi terhadap Data Latih



Gambar 5.2 Grafik Tingkat Akurasi terhadap Nilai k



Gambar 5.3 Grafik Tingkat Akurasi terhadap Sebaran Data

## 5.2 Analisis

Pada sub bab ini akan dilakukan analisa dari pengujian dan analisis dari penerapan *Fuzzy k-Nearest Neighbor* (Fk-NN) untuk diagnosa penderita *liver* berdasarkan *Indian Liver Patient Dataset*. Analisis ini dilakukan untuk mengetahui karakteristik data dari pengujian yang telah dilakukan.

### 5.2.1 Analisis hasil pengujian tingkat akurasi terhadap data latih

Pada pengujian tingkat akurasi terhadap data latih dapat disimpulkan bahwa hasil akurasi yang paling bagus sebesar 76% dalam data latih yang sebanyak 170 data. Dari hasil yang ada pada tabel 5.1 dapat disimpulkan bahwa adanya fluktuatif di data latih membuat hasil akurasi tidak baik.

### 5.2.2 Analisis hasil pengujian tingkat akurasi terhadap nilai k

Pada pengujian tingkat akurasi terhadap nilai k dapat disimpulkan bahwa hasil akurasi terbaik sebesar 74%. Akurasi semakin membaik dengan makin banyaknya jumlah k, tetapi fluktuatif data dapat merusak hasil akurasi.

### 5.2.3 Analisis hasil pengujian Akurasi FKNN dengan Akurasi KNN

Rata – rata akurasi untuk pengujian dengan kombinasi jumlah data latih yang berbeda menghasilkan masing – masing paling baik dihasilkan KNN sebesar 62% sedangkan FKNN paling baik sebesar 74%.

### 5.2.4 Analisis hasil pengujian tingkat akurasi terhadap sebaran data

Rata – rata akurasi untuk sebaran data kelas 1 paling besar didapatkan 68%, sebaran data kelas 2 paling besar didapatkan 36%, dan disebaran data kelas 3 paling besar didapatkan 57%. Banyaknya jumlah kelas di suatu data latih sangat mempengaruhi akurasi data.

## BAB VI PENUTUP

### 6.1 Kesimpulan

Berdasarkan hasil penelitian untuk mengetahui tingkat akurasi dari hasil klasifikasi pada *Indian Liver Patient Dataset* (ILPD) dengan menggunakan metode *Fuzzy k-Nearest Neighbor* (FKNN), dapat disimpulkan bahwa:

1. Metode *Fuzzy k-Nearest Neighbor* (FK-NN) bisa diterapkan untuk mendiagnosis penderita *liver* dengan menggunakan 10 parameter yang terdapat pada data *Indian Liver Patient Dataset* (ILPD). Langkah pertama adalah melakukan normalisasi pada data latih dan data uji dengan menggunakan *min-max normalization*. Kemudian data latih dan data uji yang telah dinormalisasi dihitung menggunakan persamaan *euclidean distance*. Output data penderita *liver* yang digunakan ditransformasikan kedalam bentuk *fuzzy*. Pada akhirnya didapatkan hasil diagnose penderita *liver* yang dihitung menggunakan persamaan *Fuzzy k-Nearest Neighbor* (FK-NN).
2. Berdasarkan pengujian tingkat akurasi terhadap data latih dihasilkan akurasi sebesar 76% dalam 170 data latih, terhadap nilai k dihasilkan akurasi sebesar 74%, terhadap sebaran data kelas 1 akurasi sebesar 68%, kelas 2 akurasi sebesar 36%, dan kelas 3 paling besar didapatkan 57%.

### 6.2 Saran

Beberapa saran yang dapat dijadikan bahan pertimbangan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian ini tidak melakukan pengujian terhadap optimasi k dan data latih. Diharapkan pada penelitian selanjutnya untuk meneliti optimasi k dan data latih.
2. Diharapkan pada penelitian berikutnya menggunakan metode lain atau FKNN yang telah dimodifikasi agar mendapat hasil yang lebih maksimal.