

BAB III

METODE PENELITIAN

3.1 Analisis Sistem

Analisis sistem merupakan proses awal yang dilakukan sebelum membuat suatu sistem dengan menggunakan suatu metode, dengan tujuan mendapatkan pemahan yang komprehensif tentang sistem yang akan dibuat atau dikembangkan dan memahami permasalahan yang diharapkan dapat diusulkan perbaikannya dalam penelitian tentang sistem klasifikasi barang berdasarkan tingkat penjualannya.

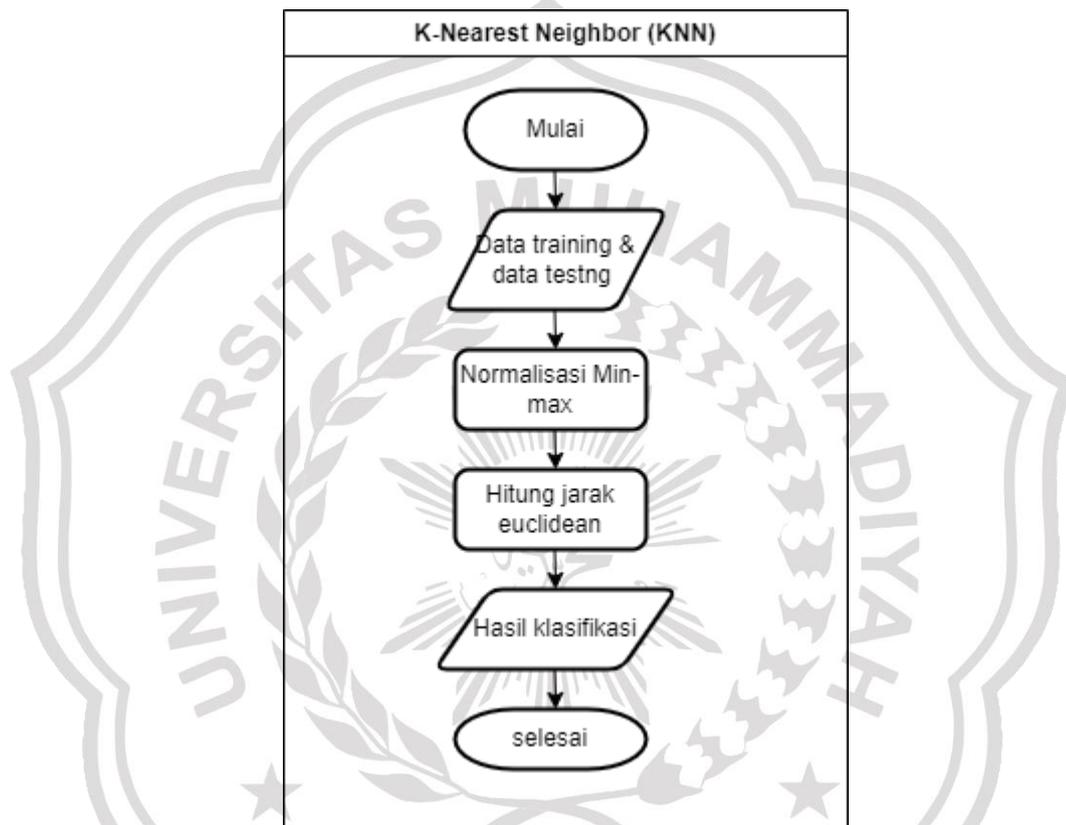
Sesuai dengan hasil penelitian yang dilakukan, didapatkan permasalahan yang dihadapi oleh Toko Solo dalam mengklasifikasikan barang berdasarkan tingkat penjualannya, penelitian yang akan dilakukan ini akan menggunakan metode *Modified K-Nearest Neighbor (MKNN)*. Dengan menggunakan metode ini, penelitian ini bertujuan untuk mengembangkan suatu model klasifikasi yang dapat mengklasifikasikan barang berdasarkan tingkat penjualannya secara akurat. Selain itu, studi kasus Toko Solo akan dilakukan untuk memberikan keakuratan dan kevalidan dari model yang dikembangkan.

Dalam membuat penelitian ini peneliti mengambil Data dari toko Solo untuk pengolahan data dengan tujuan agar toko Solo mengetahui bagaimana cara mengelompokkan barang berdasarkan tingkat penjualannya, diharapkan dapat memberikan toko Solo wawasan yang lebih baik dalam mengklasifikasikan produk mereka, memungkinkan mereka untuk menyesuaikan strategi pemasaran mereka dan meningkatkan daya saing di pasar *fashion* yang kompetitif.

3.2 Perancangan Sistem

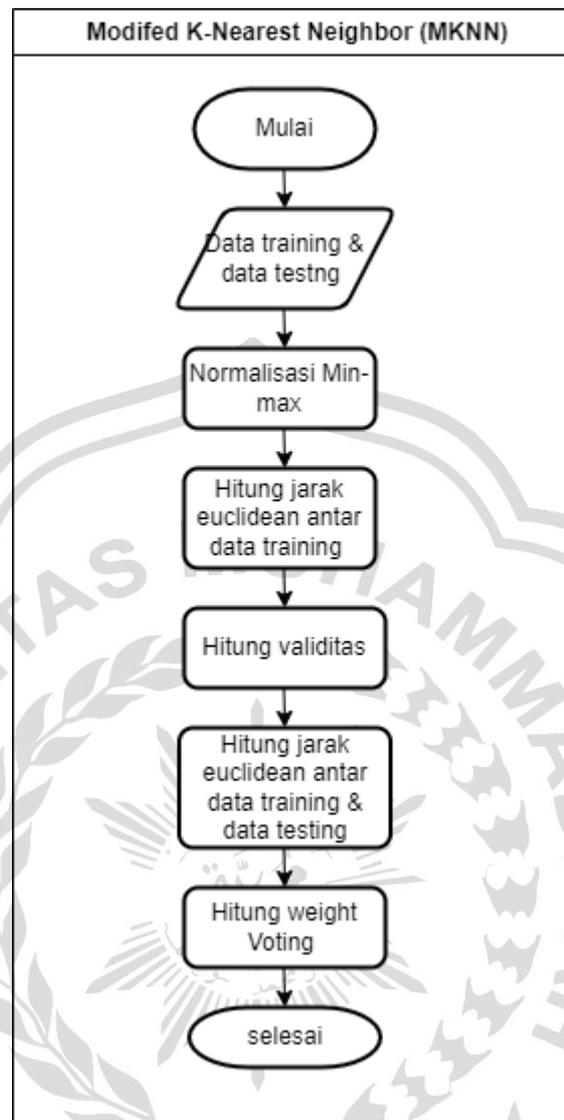
Pada bagian ini akan menjelaskan mengenai diagram alir kerja yaitu *Flowchart*, *Diagram Konteks*, *Diagram Berjenjang*, *Data Flow Diagram* (DFD).

3.2.1 Flowchart *K-Nearest Neighbor* dan *Modifed K-Nearest Neighbor*



Gambar 3. 1 *Flowchart sistem K-Nearest Neighbor*

Pada *flowchart* algoritme Sistem *K-Nearest Neighbor* diawali dengan memasukkan *data training* dan *data testing*, selanjutnya dilakukan normalisasi *min-max*, memasukkan nilai dari *K* tetangga terdekat untuk memberikan titik uji sehingga *K* dari data training yang paling dekat dengan titik uji. Lalu jarak antara data training dan data testing dihitung untuk mencari jarak terdekat berdasarkan nilai dari *K* tetangga terdekat. Ketika kita memasukkan nilai *K* maka jarak yang diambil adalah jarak terdekat. Hasil yang diperoleh dari perhitungan metode *K-Nearest Neighbor* adalah klasifikasi barang berdasarkan tingkat penjualannya.

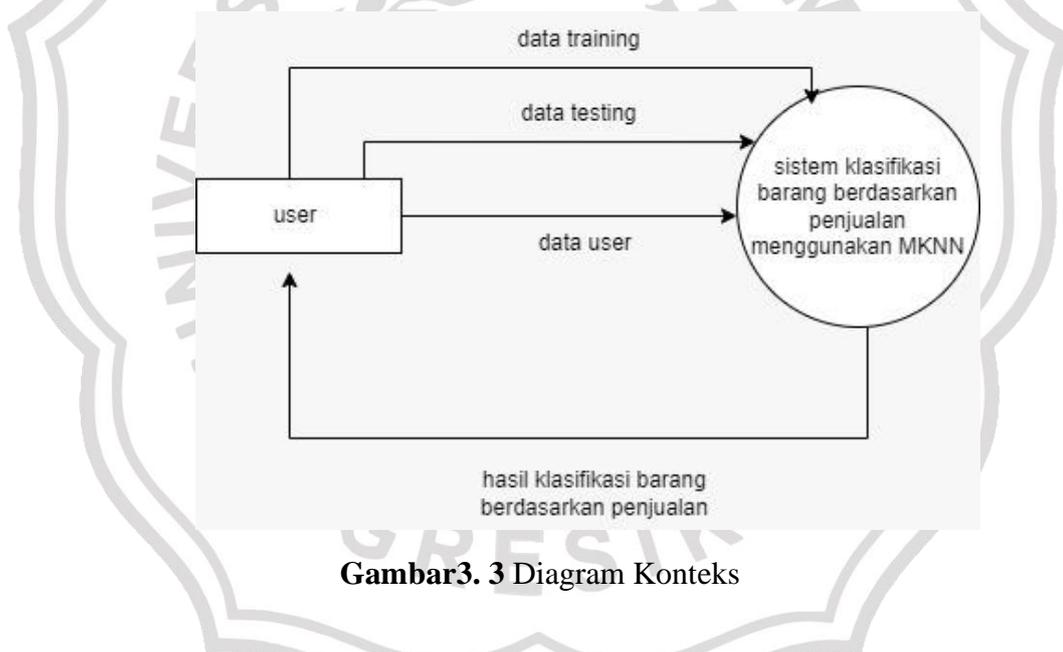


Gambar 3. 2 Flowchart sistem *Modifed K-Nearest Neighbor (MKNN)*

Pada *flowchart* algoritme Sistem *Modifed K-Nearest Neighbor* diawali dengan memasukkan *data training* dan *data testing*, selanjutnya dilakukan normalisasi *min-max*, menghitung jarak dengan euclidean antar data training, menghitung validitas, memasukkan nilai dari K tetangga terdekat untuk memberikan titik uji sehingga K dari data training yang paling dekat dengan titik uji. Lalu jarak antara data training dan data testing dihitung untuk mencari jarak terdekat berdasarkan nilai dari K tetangga terdekat, menghitung nilai *weight voting* . Ketika kita memasukkan nilai K maka jarak yang diambil adalah jarak terdekat. Hasil yang diperoleh dari perhitungan metode *Modifed K-Nearest Neighbor* adalah klasifikasi barang berdasarkan tingkat penjualannya.

3.2.2 Diagram Konteks

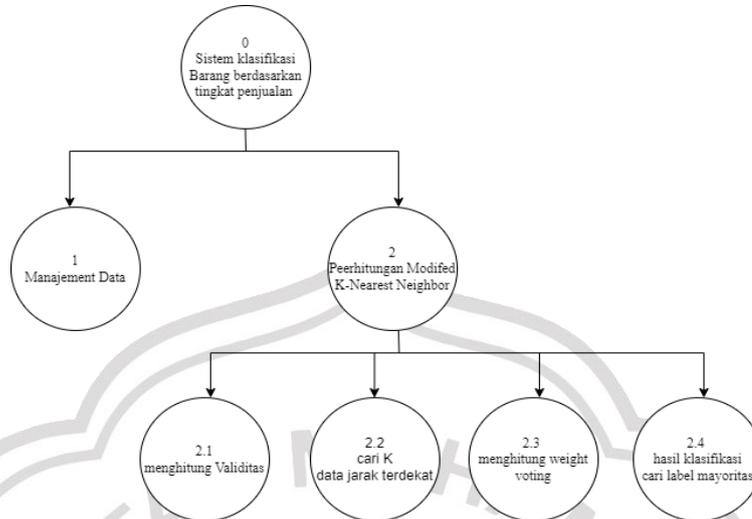
Diagram Konteks adalah sebuah bagian level dari Data Flow Diagram yang digunakan untuk menetapkan konteks serta batasan sistem pada sebuah pemodelan. Hal ini termasuk hubungan dengan entitas-entitas diluar system itu sendiri, seperti sistem, kelompok organisasi, penyimpanan data eksternal lain. Diagram konteks sering disebut juga dengan Level-0 dan menjadi penentu utama pada sebuah sistem yang dimodelkan dalam Data Flow Diagram. Namun untuk membangun suatu sistem DFD utuh masih dibutuhkan 2 level lanjutan, yaitu Level 1 yang bertujuan memecah sistem menjadi lebih kecil serta Level 2 yang bertujuan untuk membuat rincian dari system yang akan dibuat.



Gambar3. 3 Diagram Konteks

User merupakan entity dalam Sistem Klasifikasi barang berdasarkan tingkat penjualan. User memasukkan Data training dan Data testing yang sudah ditentukan. Data training digunakan sebagai data Training yang akan dihitung. Data testing digunakan sebagai data pengujian terhadap hasil perhitungan Data training yang telah didapatkan. Setelah mendapatkan hasil dari pengujian menggunakan metode *Modifed K-Nearest Neighbor*, dilanjutkan dengan melakukan evaluasi dari hasil pengujian.

3.2.3 Diagram Berjenjang



Top Level : 0 Sistem Klasifikasi Barang

Level 0 : 1. Manajemen Data

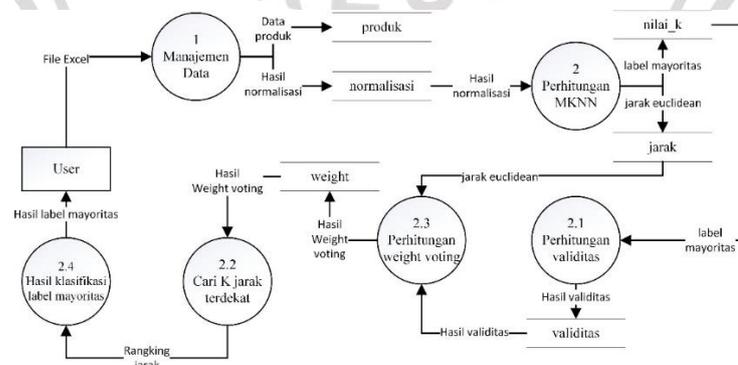
2. Perhitungan *Modified K-Nearest Neighbor*

Level 1 : 2.1 Menghitung Validitas

2.2 Cari K data jarak terdekat

2.3 Menghitung *Weight Voting*

3.2.4 Data Flow Diagram



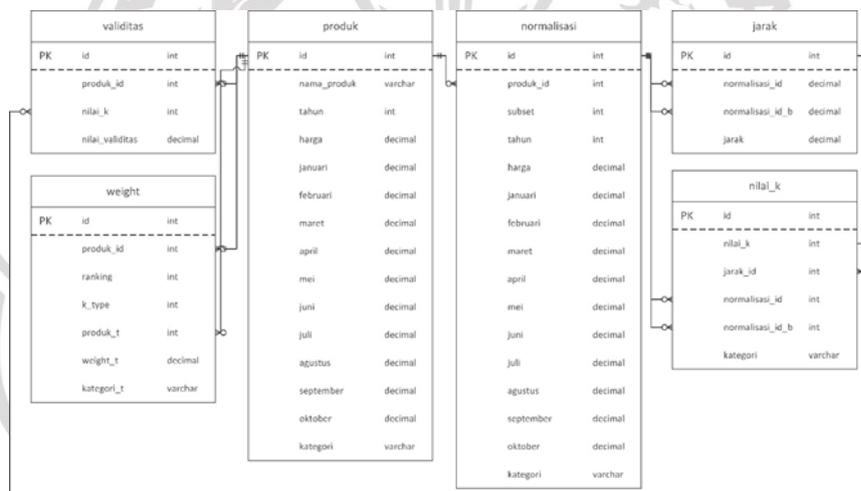
Gambar 3. 5 Data Flow Diagram

Berdasarkan pada data flow diagram terdapat beberapa proses yang terjadi. Antara lain:

1. *User* memasukkan Data training dan Data testing pada sistem dan data tersebut disimpan ke dalam database.
2. Data training dan Data testing dilakukan perhitungan menggunakan metode *Modifed K-Nearest Neighbor*.
3. Hasil perhitungan akan disimpan pada *database* dan data hasil diteruskan kepada *User*.

3.3 Perancangan Basis Data

Sebagai tempat penyimpanan data pada sistem Klasifikasi Barang berdasarkan tingkat penjualannya dibutuhkan sebuah database yang terdiri dari tabel user, data latih, data uji, dan klasifikasi. Model konseptual dalam menggambarkan hubungan antara tabel dan memodelkan struktur data dari tabel termuat pada *Entity Relationship Diagram*. *Entity Relationship Diagram* tersebut sebagai berikut:



Gambar 3. 6 Entity Relationship Diagram System

3.4 PERANCANGAN ANTARMUKA SISTEM

1. Halaman Data Latih halaman data latih digunakan *users* untuk memasukan nilai dari K, menambahkan data latih kedalam *database* sistem dan melihat rincian dari data testing. Tampilan rancangan halaman data latih terdapat pada **Gambar 3.7**

Data Latih

Tambah Data
Import Excel
Hapus

ID	Brand	TES 1	TES 2	TES 3	TES 4	TES 5	Price_Range	Aksi

Data Latih

Data Uji

2. Halaman Data Uji

Halaman data uji digunakan *users* untuk menambahkan data uji kedalam *database* sistem dan melihat rincian dari data uji. Tampilan rancangan halaman data uji terdapat pada **Gambar 3.8**

Data Uji

Tambah Data
Import Excel
Hapus

ID	Brand	TES 1	TES 2	TES 3	TES 4	TES 5	Price_Range	Aksi

Data Latih

Data Uji

Gambar 3. 8 Perancangan Halaman Data uji

3. Halaman Hasil prediksi

Halaman hasil prediksi menjadi halaman untuk menampilkan Semua hasil perhitungan *Modifed K – Nearest Neighbor*. Tampilan rancangan halaman klasifikasi terdapat pada **Gambar 3.9**

Hasil

Data Uji

TES 1	TES 2	TES 3	TES 4	TES 5

Prediksi

Data Latih Nilai K Prediksi

Data Uji

Jarak Euclidean

Data Latih	TES 1	TES 2	TES 3	TES 4	TES 5	Price_Range	Jarak Euclidean

Gambar 3. 9 Perancangan Halaman Hasil

3.5 Analisis Data

Data awal pada penelitian ini memiliki 8 variabel dari total 247 data. Klasifikasi barang berdasarkan tingkat penjualan dibutuhkan 10 variabel yang digunakan untuk penentuan klasifikasi barang yakni sangat laris, cukup laris dan kurang laris yang terdiri dari penjualan bulan Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, oleh karena itu perlu dilakukan preprocessing data untuk melakukan proses seleksi pada data. Data awal dapat dilihat pada tabel 3.1 berikut.

Tabel 3. 1 Data Penjualan Asli

No	Kode	Tgl Penjualan	Nama Barang	Size	Qty	Harga
1	001791	01/01/2023	Bahan hem tulis katun pola 2m	2M	1	800.000
2.	001833	08/04/2023	Hem PA Dh	L	1	237,000
3	001926	12/5/2023	Daster abaya Kancing danar hadi	All Size	1	177.000
4	001934	1/07/2023	Blouse danar hadi	S	1	310.000

5	001942	6/08/2023	Hem lengan pendek danar hadi	M	1	213.000
6	001946	17/08/2023	Sabun/lerak danar hadi	50 ml	2	25.000
7	001970	22/08/2023	Gamis set annisa caltry navy	XL	1	525.000
8	001976	15/09/2023	bahan hem danar hadi	2,5 M	3	188.000
9	001026	21/09/2023	Blouse danar hadi	M	1	310.000
.....						
247	001027	5/10/2023	Koko/takwa samase kurta 3/4 white	M	2	249.000

Pada tabel 3.1 dapat ditemukan data yang masih belum bisa diolah untuk proses MKNN antara lain tanggal penjualan yang masih berbentuk penjualan harian belum global menjadi penjualan tiap bulan, oleh karena itu dibutuhkan proses analisa data mining untuk membuat data bisa diolah kedalam metode MKNN

3.5.1 Analisis *Knowledge Discovery Database*

Analisa tahap *Knowledge Discovery Database* bertujuan untuk mengekstraksi atau mengidentifikasi pola. Analisis data mining pada penelitian ini menggunakan transformasi dan preprocessing.

1. Preprocessing

Preprocessing dilakukan untuk mengelola data menjadi data yang bisa diolah oleh sistem. Berikut analisa preprocessing data.

a. *Data Selection*

Proses selection yang dilakukan adalah menyeleksi variabel yang dibutuhkan berdasarkan kriteria yang sudah ditentukan. Langkah seleksi data yaitu dengan menghapus variabel yang tidak diperlukan. Data yang digunakan pada penelitian ini adalah data penjualan harian Toko Solo pada bulan Januari hingga Oktober tahun 2023 Data digunakan untuk mengetahui klasifikasi barang berdasarkan tingkat penjualannya, data awal berjumlah 8 variabel, pada klasifikasi barang berdasarkan hasil wawancara dari pemilik toko serta karyawan toko dibutuhkan 10 variabel yang digunakan untuk proses pembagian klasifikasi barang berdasarkan tingkat penjualannya dan

2 variabel sebagai variabel pendukung. Terdapat beberapa variabel data penjualan dalam penelitian ini dapat dilihat pada tabel 3.2 berikut

Tabel 3.2 Variabel data penjualan untuk klasifikasi barang

No	Nama fitur	Keterangan
1.	Nama_Barang	Nama Barang yang terjual
2.	Jan	Penjualan Bulan Januari
3.	Feb	Penjualan Bulan Februari
4.	Mar	Penjualan Bulan Maret
5.	Apr	Penjualan Bulan April
6.	Mei	Penjualan Bulan Mei
7.	Jun	Penjualan Bulan Juni
8.	Jul	Penjualan Bulan Juli
9.	Agus	Penjualan Bulan Agustus
10.	Sep	Penjualan Bulan September
11.	Okt	Penjualan Bulan Oktober
12.	Kategori	Status klasifikasi Barang

Berdasarkan Tabel 3.2 beberapa variabel yang digunakan dari data penjualan Toko Solo pada penelitian ini, pada tabel 3.2 nama_barang merupakan primary key, dan kategori merupakan output dari data yaitu barang termasuk kategori sangat laris, cukup laris atau kurang laris.

3.5.2 Analisis Input

Proses penentuan klasifikasi barang berdasarkan tingkat penjualannya pada sistem yang akan dibangun membutuhkan data kriteria-kriteria klasifikasi barang di Toko Solo Gresik. Selanjutnya data akan digunakan sebagai data *training* dan data *testing*.

1. Menentukan Variabel

Variabel yang digunakan pada penelitian tugas akhir ini sebanyak 10 yaitu Nama_Barang, Jan, Feb, Mar, Apr, Mei, Jun, Jul, Agus, Sep, Okt, dan

Kategori. Berikut variabel yang akan menjadi masukan. Berikut data masukan dapat dilihat pada Tabel 3.3 berikut:

Tabel 3.3 Data Masukan (*input*)

Variabel	Nama fitur	Keterangan
X1	Nama _Barang	Nama Barang yang terjual
X2	Jan	Penjualan Bulan Januari
X3.	Feb	Penjualan Bulan Februari
X4	Mar	Penjualan Bulan Maet
X5	Apr	Penjualan Bulan April
X6	Mei	Penjualan Bulan Mei
X7	Jun	Penjualan Bulan Juni
X8	Jul	Penjualan Bulan Juli
X9.	Agus	Penjualan Bulan Agustus
X10.	Sep	Penjualan Bulan September
X11	Okt	Penjualan Bulan Oktober
K	Kategori	Status klasifikasi Barang

Pada tabel 3.3 diatas merupakan variabel yang dibutuhkan untuk klasifikasi Barang berdasarkan tingkat penjualannya. Variabel tersebut merupakan hasil dari wawancara dari Ibu Yanti selaku pemilik Toko Solo.

3.5.3 Analisis *output*

Keluaran (*output*) yang diperoleh dari penelitian ini ada 3 yaitu klasifikasi barang termasuk kategori Sangat laris, Cukup Laris, atau Tidak Laris. Berikut data keluaran dapat dilihat pada Tabel 3.4 berikut.

Tabel 3.4 *Output* Klasifikasi

Keluaran (Output)	Kelas
Sangat Laris	1
Cukup Laris	2
Kurang Laris	3

Tabel 3.4 diatas merupakan keluaran (*output*) dari penelitian, output yang dihasilkan yaitu kategori kelas barang Sangat Laris, Cukup Laris atau

Kurang laris yang dimana untuk rentang nilai keluaran berdasarkan total penjualan bulan Januari hingga Oktober dapat dilihat pada tabel 3.5 Berikut.

Tabel 3.5 Rentang nilai keluaran (*Output*)

Rentang Total Penjualan	Kategori
160-219	Sangat Laris
86-159	Cukup Laris
12-85	Kurang Laris

3.5.4 Analisis Data Training dan Data Testing

Pada tahapan ini dilakukan pembagian data untuk melatih data dengan metode *Modified K-Nearest Neighbor*. Pelatihan *Modified K Nearest Neighbor* yang dilakukan terbagi menjadi 90:10% dengan 111 data, Berikut data latih 90% dapat dilihat pada Tabel 3.6 berikut: Data Training 90 %

Tabel 3.6 Data Training 90 %

NO	Nama_Barang	X 1	X2	X3	X4	X 5	X 6	X 7	X 8	X 9	X10	K
1	Bahan Hem 2m	3	2	6	1	3	2	1	1	2	2	3
2	Hem PA (L)	22	21	23	20	21	20	22	21	19	20	1
3	Gamis Annisa	3	6	8	2	1	3	2	3	5	3	3
4	Hem PA (M)	16	15	30	21	17	17	14	21	11	18	1
5	Bahan Hem dh	10	11	18	11	8	12	11	16	12	12	2
6	Gamis Rumaysha	5	8	12	6	8	5	7	7	4	9	1
7	Hem PE (M)	22	21	26	19	21	20	21	24	19	26	1

NO	Nama_Barang	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	K
8	Hem PA (2L)	7	10	16	8	11	9	14	11	6	13	2
9	Sirwal Samase (S)	5	7	11	5	6	4	5	2	6	5	1
10	Sirwal Samase (M)	19	21	21	18	20	22	17	22	19	21	1
11	Hem PA (3L)	3	4	7	2	1	3	1	4	1	1	1
12	Hem Anak Ulya	13	12	18	10	7	9	11	19	10	6	2
13	Iner 1 Warna	18	20	22	21	17	21	20	21	19	23	1
.....
	
100	Gamis anak Syafira	6	4	9	5	7	8	2	4	1	5	3

Pada Tabel 3.6 di atas merupakan 90% data *training* dengan jumlah data 100. X1 sampai X10 merupakan variabel masukan yang akan digunakan dalam tugas akhir ini kemudian dilanjutkan ke data *testing* penelitian. Tahapan pembagian data uji dilakukan pembagian data untuk menguji data dengan metode *Modified K-Nearest Neighbors*. Pengujian *Modified K-Nearest Neighbors* yang dilakukan terbagi menjadi 10% dengan 11 data,. Berikut 10% data uji yang digunakan pada tabel 4.14:

Tabel 3.7 Data *Testing* 10%

NO	Nama_Bara ng	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	K
101	Gamis Albani Corner	1	3	6	2	1	1	1	1	1	3	3

NO	Nama_Barang	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	K
102	Hem PA Tulis (M)	9	10	17	14	7	8	6	7	9	10	2
103	Hem PA Tulis (L)	3	7	8	4	6	3	1	1	1	1	3
104	Hem PA Tulis (XL)	2	1	4	3	5	2	1	1	1	2	3
105	Peci Rajut Samase	21	20	25	19	18	21	19	21	22	20	1
106	Hem Pa (4L)	5	6	9	7	8	9	7	7	8	5	3
107	Hem PE (4L)	7	9	11	7	6	4	5	1	3	2	3
108	Bahan Sutra 2M	9	11	15	16	15	18	16	14	15	16	2
109	Bahan Sarimbit DH	8	7	12	8	8	7	3	8	7	6	3
110	Jarik Sidomukti	3	1	5	1	1	1	1	1	1	1	3
111	Jarik dh	9	10	14	11	9	8	6	6	4	3	3

Pada Tabel 3.7 di atas merupakan 10% data *testing* dengan jumlah data 11. X1 sampai X10 merupakan variabel masukan yang akan digunakan dalam tugas akhir ini

3.5.5 Normalisasi Data

normalisasi data dengan menggunakan *Min-Max* Normalisasi persamaan (2.1) dimana data yang memiliki sebaran nilai yang jauh agar

dinormalisasikan dalam rentang [0-1]. Berikut adalah perhitungan normalisasi data *training* dan data *testing* sesuai dengan persamaan dapat dilihat pada tabel 3.8 berikut.

Tabel 3.8 Hasil Normalisasi *Min-max*

Data	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	K
1	0,08 69	0,04 34	0,17 24	0	0,08 33	0,04 76	0	0	0,04 76	0	3
2	0,91 30	0,86 95	0,75 86	0,82 60	0,83 33	0,90 47	0,87 5	0,86 95	0,85 71	0,76	1
3	0,08 69	0,21 73	0,24 13	0,04 34	0	0,09 52	0,04 16	0,08 69	0,19 04	0,08	3
4	0,65 217	0,60 86	1	0,86 95	0,66 66	0,76 19	0,54 16	0,86 95	0,47 6	0,68	1
....
100	0,21 73	0,13 04	0,27 58	0,17 39	0,25	0,33 33	0,04 16	0,13 04	0	0,16	1

Perhitungan normalisasi data training dapat dilakukan dengan sebagai berikut :

Tabel 3.9 Nilai Awal Variabel *Max & Min*

NORMALISASI DATA MIN-MAX										
	JAN	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt
MIN(X)	1	1	1	1	1	1	1	1	1	1
MAX	24	24	30	24	25	22	25	24	22	26

Adapun rumus normalisasi sebagai berikut :

Rumus Normalisasi min-max

$$z = \frac{x - \min(x)}{[\max(x) - \min(x)]} \quad (3.1)$$

Dimana:

X : nilai data yang akan di normalisasi

$\min(x)$: nilai minimum

$\max(x)$: nilai maximum

Sehingga didapat perhitungan normalisasi Min-Max terhadap data training sebagai berikut :

Variabel X1

$$\text{Data ke-1} = \frac{3-1}{24-1} = \frac{2}{23} = 0,0869$$

$$\text{Data ke-2} = \frac{22-1}{24-1} = \frac{21}{23} = 0,9130$$

$$\text{Data ke-3} = \frac{3-1}{24-1} = \frac{2}{23} = 0,0869$$

$$\text{Data ke-4} = \frac{16-1}{24-1} = \frac{15}{23} = 0,65217$$

;

$$\text{Data ke-100} = \frac{6-1}{24-1} = \frac{5}{23} = 0,2173$$

Perhitungan dilakukan sampai dengan variabel X10 terhadap data ke 1 hingga 100, begitupun dengan data *Testing* juga dilakukan normalisasi Min-Max.

3.5.5 Analisis metode *Modified K-Nearest Neighbor*

Pada tahapan ini dilakukan proses tahapan metode *Modified K-Nearest Neighbor*, adapun langkah-langkah perhitungan metode *Modified K-Nearest Neighbor* sebagai berikut :

1. Penentuan Nilai K

Nilai K yang digunakan dalam prosedur tahapan metode *Modified K-Nearest Neighbor* (MKNN) ini adalah K=1, K=3, K=5, K=7 dan K=9

2. Perhitungan jarak Euclidean antar data Training

Tahapan Perhitungan jarak Euclidean antar data Training bertujuan untuk mendapatkan tetangga terdekat untuk pencarian validasi data training, adapun rumus euclidean seperti pada persamaan . dari rumus persamaan diatas dilakukan perhitungan terhadap data training berjumlah 100 data sehingga didapat perhitungan sebagai berikut, dapat dilihat pada tabel 3.10 berikut

Tabel 3.10 Hasil perhitungan jarak antar data Training

No Data	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Jarak
2	0,91 3	0,86 96	0,75 86	0,82 61	0,83 33	0,90 48	0,87 5	0,86 96	0,85 71	0,76	2,538 26
3	0,08 69	0,21 74	0,24 14	0,04 35	0 0	0,09 52	0,04 17	0,08 7	0,19 05	0,08	0,286 72
4	0,65 22	0,60 87	0,60 1	0,86 96	0,66 67	0,76 19	0,54 17	0,86 96	0,47 62	0,68	2,150 83
5	0,39 13	0,43 48	0,58 62	0,43 48	0,29 17	0,52 38	0,41 67	0,65 22	0,52 38	0,44	1,376 76
6	0,17 39	0,30 43	0,37 93	0,21 74	0,29 17	0,19 05	0,25 09	0,26 09	0,14 29	0,32	0,686 67
7	0,91 3	0,86 96	0,86 21	0,78 26	0,83 33	0,90 48	0,83 33	0,85 1	0,85 71	1	2,664 57
8	0,26 09	0,39 13	0,51 72	0,30 43	0,41 67	0,38 1	0,54 17	0,43 48	0,23 81	0,48	1,155 03
9	0,17 39	0,26 09	0,34 48	0,17 39	0,20 83	0,14 29	0,16 67	0,04 35	0,23 81	0,16	0,480 66
10	0,78 26	0,86 96	0,68 97	0,73 91	0,79 17	0,66 1	0,66 67	0,91 3	0,85 71	0,8	2,441 8

Perhitungan jarak antar data training ke-1 terhadap data traing 2,3,4 hingga 100 dengan rumus Ecludien sebagai berikut :

$$d_{x_1 d_{x_2}} =$$

$$\sqrt{(0,0869 - 0,9130)^2 + (0,0434 - 0,86951)^2 + (0,1724 - 0,7586)^2 + \dots}$$

$$= 2,53826$$

$$d_{x_1 d_{x_4}} =$$

$$\sqrt{(0,869 - 0,869)^2 + (0,0434 - 0,2173)^2 + (0,1724 - 0,2413)^2 + \dots}$$

$$= 0,28672$$

$$\sqrt{(0,869 - 0,6522)^2 + (0,0434 - 0,6087)^2 + (0,1724 - 1)^2 + \dots}$$

$$= 2,15083$$

Perhitungan jarak euclidean dilakukan hingga selesai dengan data training ke 100 terhadap data lainnya sebagai data testing.

3. Sehingga berdasarkan perhitungan jarak untuk data training dilakukan pengurutan terhadap tertangga terdekat yakni $K=1,3,5,7,9$, sebagai contoh hasil pengurutan data training ke 1 yang dimana berperan sebagai data testing terhadap data training lainnya dapat dilihat pada tabel 3.11 sebagai berikut

Tabel 3.11 hasil pengurutan jarak

No data	<i>distance</i>	Rangking	kategori
39	0,110867	1	3
62	0,140934	2	3
53	0,146571	3	3
84	0,159952	4	3
85	0,160946	5	3
61	0,172685	6	3
82	0,162795	7	3
18	0,177781	8	3
44	0,187	9	3

Pergurutan tertangga terdekat dilakukan hingga selesai dengan data traing ke 100 sebagai data testing terhadap data training lainnya.

4. Perhitungan validitas data training

Dimana dalam perhitungan ini nilai kedekatan data 1 dengan data lainnya adalah berdasarkan kesamaan dari kelas atau klasifikasi data asli, jika kelas data sama, maka nilai validasi=1, jika kelas dari data tidak sama, maka nilai validasi=0, berikut contoh perhitungan validitas data ke 4.

Untuk K=1

$$\begin{aligned} \text{validitas (!)} &= \frac{1}{1} \times s(\text{lbl}(1,4)) \\ &= 1 \end{aligned}$$

Untuk K=3

$$\begin{aligned} \text{validitas (!)} &= \frac{1}{3} \times s(\text{lbl}(1,4)) \\ &= 1 \end{aligned}$$

Untuk K=5

$$\begin{aligned} \text{validitas (!)} &= \frac{1}{5} \times s(\text{lbl}(1,4)) \\ &= \frac{4}{5} = 0,8 \end{aligned}$$

Untuk K=7

$$\begin{aligned} \text{validitas (!)} &= \frac{1}{7} \times s(\text{lbl}(1,4)) \\ &= \frac{5}{7} = 0,714 \end{aligned}$$

Untuk K=9

$$\begin{aligned} \text{validitas (!)} &= \frac{1}{9} \times s(\text{lbl}(1,4)) \\ &= \frac{7}{9} = 0,77 \end{aligned}$$

5. Perhitungan jarak ecluidean antar data training dan data testing

Tahap perhitungan dimulai dengan menghitung jarak datatraining, perhitungan jarak ecluidean antara data traing dan data testing dilakukan untuk semua datatraing dan semua data testing. dari rumus persamaan

diatas dilakukan perhitungan terhadap data testing berjumlah 11 data terhadap data training 100 data sehingga didapat perhitungan sebagai berikut, dapat dilihat pada tabel 3.12 perhituangan jarak antar data testing ke 101 terhadap semua data training

Tabel 3.12 hasil perhitungan jarak data testing

No Data	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Jarak
1	0,087	0,0435	0,1724	0	0,0833	0,0476	0	0	0,0476	0	0,1709
2	0,913	0,8696	0,7586	0,8261	0,8333	0,9048	0,875	0,8696	0,8571	0,7676	2,5757
3	0,087	0,2174	0,2414	0,0435	0	0,0952	0,0417	0,087	0,1905	0,0808	0,2898
4	0,6522	0,6087	0,861	0,8696	0,6667	0,7619	0,5417	0,8696	0,4762	0,6868	2,174
5	0,3913	0,4348	0,5862	0,4348	0,2917	0,5238	0,4167	0,6522	0,5238	0,4444	1,4003
6	0,1739	0,3043	0,3793	0,2174	0,2917	0,1905	0,2509	0,2609	0,1429	0,3232	0,6932
7	0,913	0,8696	0,8621	0,7826	0,8333	0,9048	0,8333	1	0,8571	1	2,6938
.....
100	0,1739	0,2609	0,3448	0,1739	0,2083	0,1429	0,1667	0,0435	0,2381	0,1616	0,5136

Perhitungan jarak antar data training ke-1 terhadap data traing 2,3,4 hingga 100 dengan rumus Ecludien sebagai berikut :

$$d_{x_{101}d_{x_1}} =$$

$$\sqrt{(0,0869 - 0,9130)^2 + (0,0434 - 0,86951)^2 + (0,1724 - 0,7586)^2 + \dots}$$

$$= 0,1709$$

$$d_{x_{101}d_{x_2}} =$$

$$\sqrt{(0,0869 - 0,9130)^2 + (0,0434 - 0,86951)^2 + (0,1724 - 0,7586)^2 + \dots}$$

$$= 2,5757$$

$$d_{x_{101}d_{x_3}} =$$

$$\sqrt{(0,0869 - 0,9130)^2 + (0,0434 - 0,86951)^2 + (0,1724 - 0,7586)^2 + \dots \dots}$$

$$= 0,2898$$

Perhitungan jarak euclidean dilakukan hingga selesai dengan data testing ke-111 terhadap terhadap data training lainnya.

6. Perhitungan weight voting

Perhitungan dilakukan dengan memasukan nilai validasi data training yang telah diperoleh sebelumnya dan data jarak euclidean antara data testing dengan data training, selanjutnya data hasil weight voting diurutkan dari terbesar hingga terkecil .sebagai contoh pada data testing ke-1 pada sampel ke-101 yakni sebagai berikut. Perhitungan weight voting dilakukan hingga data training ke-100 dan data testing ke- 11 Setelah dilakukan perhitungan weight voting, maka dilakukan proses pengurutan nilai *weight voting* dari bobot terbesar sampai dengan bobot terkecil (descending) tiap data testing berdasarkan nilai K optimal yaitu $K = 7$. Kemudian dilakukan voting kelas berdasarkan nilai weight voting yang bertujuan untuk memilih klasifikasi pada data testing yang paling banyak muncul. Berikut nilai weight voting untuk data testing ke-1 pada sampel ke-100 yang dapat dilihat pada Tabel 3.13 sebagai berikut.

Tabel 3.13 Hasil Pengurutan Nilai Weight Voting pada Data Testing ke-1

Data Training data	Data Testing ke-1 $W_{(a,b)}$	Nilai K	Hasil Klasifikasi Data Testing	Klasifikasi Awal Data Testing ke-1
D-85	1,581414	1	3	3
D-39	1,565056		3	
D-80	1,560218	3	3	

D-44	1,530778		3	
D-61	1,51554	5	3	
D-1	1,490438			
D-61	1,471654	7	3	
:		:	:	:
D-59	0,509853		3	

Berdasarkan Tabel 3.13 diketahui bahwa data-85 merupakan data training ke-1 yang terdekat dari data testing pertama ditandai dengan nilai bobot yang paling besar dibandingkan bobot lainnya. Secara keseluruhan dari tetangga pertama sampai dengan tetangga kelima, hasil prediksi klasifikasi pada data testing ke-1 adalah label kelas 3 (kurang laris). Selanjutnya dilakukan hal yang sama tahapan weight voting (bobot) pada data testing ke-2, data testing ke-3, sampai dengan data testing ke-11

3.5.6 Akurasi Prediksi MKNN

Setelah diperoleh hasil prediksi klasifikasi tiap data testing, selanjutnya melakukan perbandingan hasil klasifikasi dari metode MK-NN dengan klasifikasi data aslinya untuk semua percobaan 11 data testing. Berikut tabel hasil perbandingan klasifikasi yang dapat dilihat pada Tabel 3.14.

Tabel 3.14 Hasil Perbandingan Klasifikasi Metode MKNN.

Data Testing	Prediksi MK-NN	Klasifikasi Data Asli
D-101	3	3
D-102	2*	3*
D-103	3	3
D-104	3	3
D-105	1	1
D-106	3	3
D-107	3	3
D-108	2	2
D-109	3	3
D-110	3	3
D-111	3	2

Berdasarkan Tabel 3.15 dapat dilihat bahwa sel yang diberi tanda (*) memiliki perbedaan klasifikasi dengan klasifikasi pada data aslinya.

$$\text{Akurasi} = x = \frac{10}{11} \times 100\%$$

$$= 91\%$$

Terdapat 1 barang tidak tepat diklasifikasikan dan sisanya terdapat 10 barang yang tepat diklasifikasikan pada status klasifikasi barang. Berikut perhitungan akurasi prediksi tepat dari metode MKNN yaitu sebagai berikut. Berdasarkan perhitungan akurasi dari metode MKNN dengan menggunakan K optimal yakni K = 7 diperoleh nilai akurasinya yaitu sebesar 91% yang menggambarkan bahwa ketepatan klasifikasi dengan menggunakan 7 tetangga terdekat (7-NN) pada proporsi 90% : 10% sebagian besar data testing memiliki kesamaan hasil klasifikasi dengan klasifikasi dari data aslinya. Metode MK-NN dapat mengklasifikasikan barang berdasarkan tingkat penjualannya yakni dari 1 barang yang Sangat laris, 2 barang yang cukup laris dan sisanya terdapat 8 barang yang diklasifikasi kurang laris.

3.6 PENGUJIAN *BLACK BOX*

Tabel 3. 15 Pengujian *Black Box*

Test Case	Input Data	Expected Output	Status
Form Input Data Latih	Import Excel	Data barang berdasarkan tingkat penjualannya tersimpan baik dalam database dan tampil di antarmuka	<i>Succes / Failed</i>
Form Input Data Uji	Import Excel	Data barang berdasarkan tingkat penjualannya tersimpan baik dalam database dan tampil di antarmuka	<i>Succes / Failed</i>
Form Hapus Data	Klik tombol hapus All	Semua Data Terhapus	<i>Succes / Failed</i>
Hasil	Klik tombol prediksi dan Nilai K	Menampilkan semua hasil perhitungan <i>Modified K-Nearest Neighbor</i>	<i>Succes / Failed</i>

Berdasarkan tabel 3.15 dapat diketahui bahwa ada tahapan dalam melakukan pengujian Black Box, berikut merupakan tahapan untuk melakukan pengujian Black Box. Langkah pertama ada memasukan data latih ke Excel dan dalam penelitian kali ini peneliti berhasil melakukan tahapan pertama. Tahapan kedua adalah memasukan data uji ke excel, pada tahap ini peneliti juga berhasil menyelesaikan tahapan kedua. Tahapan ketiga, peneliti menghapus dat agar semua data terhapus , peneliti juga berhasil melakukan tahapan ketiga ini. Tahapan terakhir

yaitu mendapatkan hasil, pada tahapan ini peneliti mendapatkan hasil perhitungan ModifiedK- Nearest Neighbor dan peneliti berhasil melakukan tahapan tersebut.

3.7 Instrument Pengujian

Pada tahap ini terdapat dua proses pengujian untuk mendapatkan hasil prediksi barang berdasarkan tingkat penjualannya. Pengujian pertama menggunakan sistem yang dibuat menggunakan algoritme *Modified K -Nearest Neighbor* menggunakan jarak Euclidean. Kemudian melakukan pengujian *Modified K- Nearest Neighbor*.

Sedangkan Langkah – Langkah dari pengujian kedua menggunakan sistem yang dibuat menggunakan algoritme *Modified K – Nearest Neighbor* adalah sebagai berikut:

1. Masukkan data training dan data testing yang akan diuji.
2. Sistem akan memproses menggunakan algoritme *Modified K – Nearest Neighbor*.
3. akan diuji dengan menggunakan proporsi data 90:10 %

3.8 Evaluasi Confusion Matrix

Tabel 3.16 Evaluasi Performa *Confusion Matrix*

Three – Class Prediction		Actual Class		
		1	2	3
Predicted Class	1	1	1	8
	2	0	2	0
	3	0	0	8

$$\text{True Positive 1 (TP1)} = P11 = 1$$

$$\text{True Positive 2 (TP2)} = P22 = 2$$

$$\text{True Positive 3 (TP3)} = P33 = 8$$

$$\text{False Positive 1 (FP1)} = P21 + P31 = 0 + 0 = 0$$

$$\text{False Positive 2 (FP2)} = P12 + P32 = 0 + 0 = 0$$

$$\text{False Positive 3 (FP3)} = P13 + P23 = 8 + 0 = 8$$

$$\text{False Negative 1 (FN1)} = P12 + P13 = 2 + 0 = 2$$

$$\text{False Negative 2 (FN2)} = P21 + P23 = 1 + 0 = 0$$

$$\text{False Negative 3 (FN3)} = P31 + P32 = 0 + 0 = 0$$

$$\text{Accuracy} = \frac{p11 + p22 + p33}{p11+p12+p13+p21+p22+p23+p31+p32+p33} \quad (3.2)$$

$$\text{Precision 1} = \frac{TP1}{TP1+FP1} \quad (3.3)$$

$$\text{Precision 2} = \frac{TP2}{TP2+FP2} \quad (3.4)$$

$$\text{Precision 3} = \frac{TP3}{TP3+FP3} \quad (3.5)$$

$$\text{Recall 1} = \frac{TP1}{TP1+FN1} \quad (3.6)$$

$$\text{Recall 2} = \frac{TP2}{TP2+FN2} \quad (3.7)$$

$$\text{Recall 3} = \frac{TP3}{TP3+FN3} \quad (3.8)$$

3.9 SPESIFIKASI KEBUTUHAN

Kebutuhan perangkat keras dan perangkat lunak dalam pembuatan sistem Klasifikasi barang berdasarkan tingkat penjualannya di Toko Solo sebagai berikut:

3.9.1 PERANGKAT LUNAK

Perangkat lunak adalah program yang dibutuhkan untuk membuat suatu sistem dan menjalankan perangkat keras. Dalam penelitian ini perangkat lunak yang digunakan :

- a. Sistem Operasi Windows (7,8,10,11) 64 Bit.
- b. XAMPP version 7.3.7-1
- c. MySQL database server

- d. Google chrome

3.9.2 PERANGKAT KERAS

Perangkat keras adalah komponen yang membentuk suatu laptop/komputer. Dalam penelitian ini perangkat keras yang digunakan:

- a. Processor AMD Ryzen 3 5300U
- b. RAM 8 GB
- c. SSD 256 GB
- d. Monitor
- e. Keyboard
- f. Mouse
- g. Printer.

