

BAB III

ANALISA DAN HASIL

3.1. Arsitektur Sistem

Arsitektur sistem dalam *Data mining* merupakan kerangka kerja yang menjelaskan bagaimana data dikelola, diproses, dianalisis, dan disajikan untuk mengekstraksi informasi yang berguna dari dataset besar. Data mining sendiri adalah proses analisis data untuk menemukan pola, hubungan, atau wawasan tersembunyi yang mendukung pengambilan keputusan. Dalam implementasinya, arsitektur sistem data mining harus dirancang dengan baik agar setiap proses berjalan efisien, mulai dari pengumpulan data hingga penyajian hasil analisis. Sistem data mining biasanya digunakan dalam berbagai bidang, termasuk perdagangan, kesehatan, pendidikan, dan keuangan, untuk membantu organisasi atau bisnis memahami tren, memprediksi kebutuhan, atau membuat keputusan strategis berbasis data. Dengan arsitektur sistem yang baik, proses data mining dapat dioptimalkan untuk menghasilkan hasil yang akurat dan relevan.

Dalam penelitian ini, arsitektur sistem data mining dirancang untuk memanfaatkan metode *Naïve Bayes* dan metode *K-means clustering* dalam menganalisis kepuasan pelanggan pada penjualan mochi terlaris. Sistem ini bertujuan untuk membantu mengklasifikasikan tingkat kepuasan pelanggan, serta mengelompokkan pelanggan berdasarkan karakteristik dan preferensi. Proses analisis menggunakan data mining memungkinkan pengambilan keputusan berbasis data yang lebih akurat dan efisien.

3.2. Desain Aktivitas Sistem

Desain sistem dalam data mining mencakup perencanaan dan implementasi struktur yang memungkinkan penggalian informasi berharga dari data. Hal ini melibatkan pemilihan metode dan algoritma yang sesuai, dan penyesuaian parameter untuk memaksimalkan akurasi dan efisiensi proses penambangan data. Desain aktivitas sistem menjadi kunci dalam pengembangan sistem informasi, perangkat lunak, atau bahkan dalam konteks bisnis.

Selain itu, desain aktivitas sistem mencakup pemikiran tentang manajemen data dan informasi yang terlibat dalam setiap aktivitas. Ini melibatkan identifikasi data yang dibutuhkan, data tersebut akan dikumpulkan atau dihasilkan, serta data akan disimpan, diproses, dan dipertukarkan di antara aktivitas-aktivitas tersebut. Keberhasilan desain aktivitas sistem tergantung pada pemahaman mendalam terhadap kebutuhan pengguna, tujuan sistem, dan konteks di mana sistem akan beroperasi. Dengan merinci setiap aktivitas, para pengembang dapat membangun sistem yang efisien, mudah dipelajari, dan sesuai dengan kebutuhan bisnis atau fungsional yang diinginkan. Desain aktivitas sistem, ketika dilakukan dengan baik, dapat menjadi landasan yang kuat untuk menghasilkan sistem yang dapat meningkatkan produktivitas, efisiensi, dalam konteks yang diinginkan.

3.3. Analisa Data

Analisa data merupakan suatu proses dalam menyederhanakan data agar mudah di baca dan didefinisikan. Penelitian ini menggunakan data selama 3 bulan pada bulan Mei 2024 sampai bulan Juli 2024 untuk diterapkan pada tahapan langkah data mining. Data tersebut diperoleh dari data manual pada saat observasi. Data kepuasan pelanggan akan di ubah ke dalam format data. Data kemudian digunakan untuk membangun model klasifikasi berbasis dua algoritma utama diantaranya adalah *Naïve Bayes* , algoritma ini dipilih karena kemampuannya dalam mengolah data dengan pendekatan probabilitas. Sementara algoritma *K-means clustering* dipilih karena memiliki kompleksitas waktu yang lebih rendah sehingga dapat menangani dataset besar dengan lebih baik.

Setiap algoritma di uji dengan menggunakan 2 dataset yang telah dibagi menjadi data latih dan data uji dengan evaluasi menggunakan matric seperti akurasi, recall, precision dan F1-Score.

$$Precision = \frac{TP}{FP+TP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

3.4. Pengelolaan Data Menggunakan Metode *Naïve Bayes*

Untuk mempermudah pengolahan data pada data mining, maka data akan diubah ke dalam bentuk table seperti pada table dibawah ini menggunakan metode *Naïve Bayes*.

Tabel 3.1. Data Kepuasan Pelanggan

Responden	Jenis Kelamin	Rasa Mochi	Kemasan	Pelayanan	Tingkat Kepuasan
R1	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R2	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R3	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas
R4	Perempuan	Mochi Stroberi	Tidak Menarik	Cukup Baik	Puas
R5	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R6	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
R7	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R8	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Tidak Puas
R9	Perempuan	Mochi	Menarik	Baik	Puas

		Anggur			
R10	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R11	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas
R12	Perempuan	Mochi Stroberi	Menarik	Baik	Puas
R13	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R14	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R15	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R16	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R17	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
R18	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R19	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R20	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
...	...	...	...	...	...
R80	Perempuan	Mochi Stroberi	Menarik	Baik	Puas

Data diatas merupakan data yang dikumpulkan dan akan digunakan untuk membantu proses klasifikasi ataupun perhitungan pada metode *Naïve Bayes*. Tetapi tidak hanya itu saja, data diatas akan dibagi dan dipecah menjadi beberapa table.

Sebelum melakukan pengolahan data, data set dibagi menjadi 2 bagian diantaranya Data Training dan Data Testing. Dengan ketentuan Data Training berjumlah 80% dari jumlah seluruh data set, Data Training merupakan data yang digunakan sebagai data latih pada data set. Sedangkan Data Testing merupakan data yang digunakan sebagai data uji yang berjumlah 20% dari jumlah seluruh data set yang ada.

Tabel 3.2. Data Training

Responden	Jenis Kelamin	Rasa Mochi	Kemasan	Pelayanan	Tingkat Kepuasan
R1	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R2	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R3	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas
R4	Perempuan	Mochi Stroberi	Tidak Menarik	Cukup Baik	Puas
R5	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R6	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
R7	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R8	Perempuan	Mochi	Tidak	Baik	Tidak Puas

		Stroberi	Menarik		
R9	Perempuan	Mochi Anggur	Menarik	Baik	Puas
R10	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R11	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas
R12	Perempuan	Mochi Stroberi	Menarik	Baik	Puas
R13	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R14	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R15	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R16	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R17	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
R18	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R19	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R20	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R21	Perempuan	Mochi Stroberi	Tidak Menarik	Cukup Baik	Tidak Puas

R22	Perempuan	Mochi Stroberi	Menarik	Baik	Puas
R23	Perempuan	Mochi Anggur	Tidak Menarik	Cukup Baik	Tidak Puas
R24	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R25	Perempuan	Mochi Stroberi	Menarik	Baik	Puas
R26	Perempuan	Mochi Coklat	Tidak Menarik	Cukup Baik	Tidak Puas
R27	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R28	Perempuan	Mochi Stroberi	Menarik	Sangat Baik	Puas
R29	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R30	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
...	...	...	...	...	...
R64	Perempuan	Mochi Anggur	Menarik	Baik	Puas

Data diatas merupakan data training atau data latih yang digunakan untuk membantu proses klasifikasi dengan menggunakan metode *Naïve Bayes* . Data diatas juga nantinya akan dibagi menjadi beberapa bagian per variabel guna untuk mencari probabilitas pada setiap variabel.

Tabel 3.3. Data Testing

Responden	Jenis Kelamin	Rasa Mochi	Kemasan	Pelayanan	Tingkat Kepuasan
R1	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R2	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R3	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas
R4	Perempuan	Mochi Stroberi	Tidak Menarik	Cukup Baik	Puas
R5	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R6	Laki-laki	Mochi Stroberi	Menarik	Sangat Baik	Puas
R7	Perempuan	Mochi Anggur	Menarik	Sangat Baik	Puas
R8	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Tidak Puas
R9	Perempuan	Mochi Anggur	Menarik	Baik	Puas
R10	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R11	Perempuan	Mochi Stroberi	Tidak Menarik	Baik	Puas

R12	Perempuan	Mochi Stroberi	Menarik	Baik	Puas
R13	Perempuan	Mochi Coklat	Menarik	Baik	Puas
R14	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R15	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas
R16	Perempuan	Mochi Coklat	Menarik	Sangat Baik	Puas

Tabel diatas merupakan data uji yang akan digunakan untuk menguji kinerja model setelah dilatih. Data uji merupakan data yang labelnya belum diketahui oleh model algoritma. Model algoritma akan memprediksi berdasarkan nilai yang telah dilatih. Hasil prediksi tersebut nantinya akan dibandingkan dengan label yang sebenarnya untuk mengevaluasi akurasi model. Data tersebut akan dihitung probabilitasnya untuk menentukan kepuasan pelanggan pada mochi terlaris.

Tabel 3.4. Variabel Jenis Kelamin

Variabel	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Jenis Kelamin	Laki-Laki	6	3	$6/50 = 0,12$	$3/14 = 0,23$
	Perempuan	44	11	$44/50 = 0,88$	$11/14 = 0,77$

Tabel 3.5. Variabel Rasa Mochi

Variabel	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
----------	---------	------	------------	----------	----------------

Rasa Mochi	Mochi Stroberi	30	3	$30/50 = 0,59$	$3/13 = 0,23$
	Mochi Anggur	7	5	$7/50 = 0,14$	$5/13 = 0,38$
	Mochi Coklat	14	5	$14/50 = 0,27$	$5/13 = 0,38$

Tabel 3.6. Variabel Kemasan

Variabel	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Kemasan	Menarik	48	2	$48/51 = 0,94$	$2/13 = 0,15$
	Tidak Menarik	3	11	$3/51 = 0,06$	$11/13 = 0,85$

Tabel 3.7. Variabel Pelayanan

Variabel	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Pelayanan	Sangat Baik	33	0	$33/51 = 0,65$	$0/13 = 0$
	Baik	17	4	$17/51 = 0,33$	$4/13 = 0,31$
	Cukup Baik	1	9	$1/51 = 0,02$	$9/13 = 0,69$

Tabel 3.8. Variabel Status

Variabel		P (Puas & Tidak Puas)
Tingkat Kepuasan	Puas	$51/64 = 0,80$
	Tidak Puas	$13/64 = 0,20$

Setelah data selesai dikelompokkan berdasarkan atributnya, maka dari itu data bisa diolah dengan menggunakan metode *Naïve Bayes*. Dibawah ini merupakan pengolahan data dengan menggunakan data pertama diantaranya sebagai berikut :

$$P(\text{Kategori}) = P(\text{Jenis kelamin}|\text{Perempuan}) \times P(\text{Rasa mochi}|\text{Mochi coklat}) \\ P(\text{Kemasan}|\text{Menarik}) \times P(\text{Pelayanan}|\text{Baik})$$

$$P(\text{Puas}) = P(\text{Perempuan}|\text{Puas}) \times P(\text{Mochi Coklat}|\text{Puas}) \times P(\text{Menarik}|\text{Puas}) \times \\ P(\text{Baik}|\text{Puas}) \times P(\text{Status}|\text{Puas}) \\ = 0,88 \times 0,27 \times 0,94 \times 0,33 \times 0,80 \\ = 0,05896 \text{ (Puas)}$$

$$P(\text{Tidak Puas}) = P(\text{Perempuan}|\text{Tidak Puas}) \times P(\text{Mochi Coklat}|\text{Tidak Puas}) \times \\ P(\text{Menarik}|\text{Tidak Puas}) \times P(\text{Baik}|\text{Tidak Puas}) \times P(\text{Status}|\text{Tidak Puas}) \\ = 0,77 \times 0,38 \times 0,15 \times 0,31 \times 0,20 \\ = 0,00272 \text{ (Tidak Puas)}$$

Hasil dari pengolahan data Responden 1 yaitu nilai Puas lebih tinggi dari nilai Tidak Puas. Maka dapat disimpulkan untuk data Responden 1 Puas dengan penjualan mochi terlaris. Untuk perhitungan data selanjutnya masih dengan cara yang sama. Adapun hasil yang telah didapatkan dari perhitungan yang telah dilakukan adalah sebagai berikut :

Tabel 3.9. Hasil Klasifikasi *Naïve Bayes*

Responden	Jenis Kelamin	Tingkat Kepuasan		Tingkat Kepuasan	Puas	Tidak Puas
R1	Perempuan	Puas	P R E D I K S I	Puas	0,06	0,00
R2	Perempuan	Puas		Puas	0,06	0,00
R3	Perempuan	Puas		Puas	0,01	0,01
R4	Perempuan	Puas		Puas	0,00	0,00
R5	Perempuan	Puas		Puas	0,06	0,00
R6	Laki-laki	Puas		Puas	0,03	0,00
R7	Perempuan	Puas		Puas	0,06	0,00
R8	Perempuan	Tidak Puas		Tidak Puas	0,00	0,01
R9	Perempuan	Puas		Puas	0,03	0,00
R10	Perempuan	Puas		Puas	0,12	0,00
R11	Perempuan	Puas		Tidak Puas	0,01	0,01
R12	Perempuan	Puas		Puas	0,13	0,00
R13	Perempuan	Puas		Puas	0,06	0,00
R14	Perempuan	Puas		Puas	0,12	0,00
R15	Perempuan	Puas		Puas	0,12	0,00
R16	Perempuan	Puas		Puas	0,12	0,00

Hasil pada tabel diatas merupakan hasil perhitungan manual yang dilakukan dengan menggunakan Microsoft Excel dengan mencantumkan hasil prediksi klasifikasi disetiap data uji.

3.3.1. Evaluasi Model *Naïve Bayes*

Tabel 3.10. *Naïve Bayes*

Confusion Table	Prediction	Class	
		Puas	Tidak Puas
	PUAS	14 (TP)	0 (FP)
	TIDAK PUAS	1 (FN)	1(TN)

Tabel diatas merupakan hasil dari uji performa dengan metode naïve bayes, dimana hasil dari True Positive (TP) yaitu 14, True Negative (TN) yaitu 1, False Positive (FP) yaitu 0 dan False Negative (FN) yaitu 1.

Untuk mengetahui hasil *Accuracy* yaitu dengan menggunakan persamaan sebagai berikut:

$$\begin{aligned}
 Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \\
 &= \frac{14+1}{14+1+0+1} = \frac{15}{16} = 0,94 * 100\% = 94\%
 \end{aligned}$$

Untuk mengetahui hasil *Precision* yaitu dengan menggunakan persamaan sebagai berikut:

$$\begin{aligned}
 Precision &= \frac{TP}{TP + FP} \\
 &= \frac{14}{14+0} = \frac{14}{14} = 1 * 100\% = 100\%
 \end{aligned}$$

Untuk mengetahui hasil *Recall* yaitu dengan menggunakan persamaan sebagai berikut:

$$\begin{aligned}
 Recall &= \frac{TP}{TP + FN} \\
 &= \frac{14}{14+1} = \frac{14}{15} = 0,93 * 100\% = 93\%
 \end{aligned}$$

Untuk mengetahui hasil *F1-Score* yaitu dengan menggunakan persamaan sebagai berikut:

$$\begin{aligned}
 F1 - Score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \\
 &= 2 \times \frac{1 \times 0,93}{1+0,93} = \frac{93}{94} = 0,98 * 100\% = 98\%
 \end{aligned}$$

Setelah mengetahui hasil dari uji performa diatas, maka klasifikasi kepuasan pelanggan pada mochi terlaris dengan menggunakan metode *Naïve Bayes* dinyatakan Puas. Hal ini dikarenakan memiliki nilai akurasi yang tinggi yaitu mencapai 94%.

3.5. Pengelolaan Data Menggunakan Metode *K-means clustering*

K-means clustering merupakan algoritma *clustering* dengan mengelompokkan data yang memiliki kriteria yang mirip dengan cara menentukan titik tengah *cluster(centroid)*, mengelompokkan data yang memiliki jarak/karakteristik yang dekat dengan *centroid* yang telah ditentukan, semua data harus masuk kedalam *cluster*, setelah setiap *centroid* memiliki anggota, kemudian anggota setiap *cluster* diolah kembali untuk membuat *cluster* baru. Pengelompokkan pada K-Means dilakukan secara berulang-ulang sampai *centroid* tidak berubah/tetap. Adapun langkah-langkah algoritma K-Means sebagai berikut :

1. Menentukan jumlah *cluster*.
2. Memasukkan data pada setiap *cluster* secara acak.
3. Menentukan *centroid* dengan menghitung rata-rata data pada setiap *cluster*.
4. Menghitung karakteristik setiap data dengan membandingkan dengan *centroid* awal.
5. Mengelompokkan hasil karakteristik setiap data pada *cluster* yang memiliki jarak terdekat.
6. Ulangi langkah ke-3 hingga tidak ada data yang berpindah *cluster* dan nilai *centroid* tidak berubah.

Menghitung jarak antara data dan *centroid* yaitu menggunakan persamaan

1 *Eucliden* :

$$dik = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \dots \dots \dots \text{Persamaan 1.} \quad \text{Persamaan } Euclidean$$

Pengelompokan data pada setiap *cluster* dengan jarak yang paling dekat, dapat menggunakan persamaan 2 seperti berikut:

$\text{Min} \sum_k^k \text{dik} = \sqrt{\sum_j^m (\text{Cij} - \text{Ckj})^2}$ Persamaan 2. Mengelompokan data pada *cluster* yang dekat.

Menghitung *centroid* baru dapat menggunakan persamaan 3 seperti berikut:

$$\text{Ckj} = \frac{\sum_i^p x_{ij}}{p}$$
 Persamaan 3. Menghitung *centroid* baru.

Keterangan:

dik = Jarak antara *i* (data) dengan *k* (*centroid*)

m = Banyak kriteria data

Cij = Koordinat data

Ckj = Koordinat *centroid*

P = Banyak anggota *cluster*

Xij = Anggota dari *centroid*

Untuk mempermudah pengolahan data pada data mining, maka data akan di Transformasikan ke dalam bentuk numerik dengan skala seperti pada table dibawah ini menggunakan metode *K-means clustering*.

Tabel 3.11. Transformasi Data

Responden	Rasa Mochi	Kemasan	Pelayanan
R1	1	1	2
R2	2	1	2
R3	3	2	2
R4	1	2	3
R5	2	1	1
R6	3	1	1
R7	1	1	1
R8	2	2	3
R9	3	1	2
R10	1	1	1
R11	2	2	2

R12	3	1	2
R13	1	1	2
R14	2	1	1
R15	3	1	1
R16	1	1	1
R17	2	1	1
R18	3	1	1
R19	1	1	1
R20	2	1	1
R21	3	2	3
R22	1	1	1
R23	2	2	3
R24	3	1	1
R25	1	1	2
R26	2	2	3
R27	3	1	1
R28	1	1	1
R29	2	1	1
R30	3	1	1
R31	1	1	1
R32	2	2	3
R33	3	1	1
R34	1	1	1
R35	2	2	3
R36	3	1	2
R37	1	1	1
R38	2	1	1
R39	3	1	1
R40	1	1	1

Keterangan :

Rasa Mochi		Pelayanan		Kemasan	
Mochi Coklat	= 1	Sangat Baik	= 1	Menarik	= 1

Mochi Stroberi = 2 Baik = 2 Tidak Menarik = 2
Mochi Anggur = 3 Cukup Baik = 3

3.5.1. Langkah-langkah algoritma *K-means clustering*

Keberhasilan algoritma sangat tergantung pada pemilihan nilai *k* (jumlah klaster) yang tepat dan inisialisasi awal *centroid* yang baik. Berikut langkah-langkahnya :

1. Inisialisasi: Memilih (*k*) *centroid* awal secara acak.

Tabel 3.12. *Centroid* Awal Acak

<i>Centroid</i>	Rasa Mochi	Kemasan	Pelayanan
<i>Cluster 1</i>	2	2	1
<i>Cluster 2</i>	1	2	3
<i>Cluster 3</i>	3	1	2

2. Penugasan: Menghitung jarak setiap *dataset* yang ada terhadap setiap pusat *cluster* berdasarkan jarak terdekat ke *centroid*. Dengan menggunakan persamaan *Euclidean Distance Space*, seperti berikut :
 - a. Jarak antara data pertama dengan *centroid* pertama :

$$C0 = \sqrt{(1 - 2)^2 + (1 - 2)^2 + (2 - 1)^2} = 1,000$$

- b. Jarak antara data pertama dengan *centroid* kedua :

$$C1 = \sqrt{(1 - 1)^2 + (2 - 2)^2 + (2 - 2)^2} = 2,449$$

- c. Jarak antara data pertama dengan *centroid* ketiga :

$$C2 = \sqrt{(3 - 3)^2 + (1 - 1)^2 + (2 - 2)^2} = 1,414$$

Hasil dari perhitungan dari keseluruhan data terhadap tiap pusat *cluster* awal. Dan menentukan anggota setiap *cluster* dengan jarak terdekat pada masing-masing data, disajikan pada table 3.13 berikut :

Tabel 3.13. Hasil Iterasi 1

Responde n	Rasa Mochi	Kemasa n	Pelayana n	C0	C1	C2	JT	C
R5	2	1	1	1.000	2.449	1.414	1.000	0
R7	1	1	1	1.414	2.236	2.236	1.414	0
R10	1	1	1	1.414	2.236	2.236	1.414	0
R11	2	2	2	1.000	1.414	1.414	1.000	0
R14	2	1	1	1.000	2.449	1.414	1.000	0
R16	1	1	1	1.414	2.236	2.236	1.414	0
R17	2	1	1	1.000	2.449	1.414	1.000	0
R19	1	1	1	1.414	2.236	2.236	1.414	0
R20	2	1	1	1.000	2.449	1.414	1.000	0
R22	1	1	1	1.414	2.236	2.236	1.414	0
R28	1	1	1	1.414	2.236	2.236	1.414	0
R29	2	1	1	1.000	2.449	1.414	1.000	0
R31	1	1	1	1.414	2.236	2.236	1.414	0
R34	1	1	1	1.414	2.236	2.236	1.414	0
R37	1	1	1	1.414	2.236	2.236	1.414	0
R38	2	1	1	1.000	2.449	1.414	1.000	0
R40	1	1	1	1.414	2.236	2.236	1.414	0
R41	2	1	1	1.000	2.449	1.414	1.000	0
R46	1	1	1	1.414	2.236	2.236	1.414	0
R47	2	1	1	1.000	2.449	1.414	1.000	0
R55	1	1	1	1.414	2.236	2.236	1.414	0
R56	2	2	2	1.000	1.414	1.414	1.000	0
R58	1	1	1	1.414	2.236	2.236	1.414	0
R61	1	1	1	1.414	2.236	2.236	1.414	0
R62	2	1	1	1.000	2.449	1.414	1.000	0
R74	2	1	1	1.000	2.449	1.414	1.000	0
R76	1	1	1	1.414	2.236	2.236	1.414	0

R1	1	1	2	1.732	1.414	2.000	1.414	1
R4	1	2	3	2.236	0.000	2.449	0.000	1
R8	2	2	3	2.000	1.000	1.732	1.000	1
R13	1	1	2	1.732	1.414	2.000	1.414	1
R23	2	2	3	2.000	1.000	1.732	1.000	1
R25	1	1	2	1.732	1.414	2.000	1.414	1
R26	2	2	3	2.000	1.000	1.732	1.000	1
R32	2	2	3	2.000	1.000	1.732	1.000	1
R35	2	2	3	2.000	1.000	1.732	1.000	1
R43	1	2	3	2.236	0.000	2.449	0.000	1
R49	1	2	3	2.236	0.000	2.449	0.000	1
R50	2	2	3	2.000	1.000	1.732	1.000	1
R52	1	1	2	1.732	1.414	2.000	1.414	1
R64	1	1	2	1.732	1.414	2.000	1.414	1
R67	1	2	3	2.236	0.000	2.449	0.000	1
R70	1	2	2	1.414	1.000	2.236	1.000	1
R73	1	2	3	2.236	0.000	2.449	0.000	1
R77	2	2	3	2.000	1.000	1.732	1.000	1
R79	1	1	2	1.732	1.414	2.000	1.414	1
R80	1	2	2	1.414	1.000	2.236	1.000	1
R2	2	1	2	1.414	1.732	1.000	1.000	2
R3	3	2	2	1.414	2.236	1.000	1.000	2
R6	3	1	1	1.414	3.000	1.000	1.000	2
R9	3	1	2	1.732	2.449	0.000	0.000	2
R12	3	1	2	1.732	2.449	0.000	0.000	2
R15	3	1	1	1.414	3.000	1.000	1.000	2
R18	3	1	1	1.414	3.000	1.000	1.000	2
R21	3	2	3	2.236	2.000	1.414	1.414	2
R24	3	1	1	1.414	3.000	1.000	1.000	2
R27	3	1	1	1.414	3.000	1.000	1.000	2
R30	3	1	1	1.414	3.000	1.000	1.000	2

R33	3	1	1	1.414	3.000	1.000	1.000	2
R36	3	1	2	1.732	2.449	0.000	0.000	2
R39	3	1	1	1.414	3.000	1.000	1.000	2
R42	3	1	2	1.732	2.449	0.000	0.000	2
R44	2	1	2	1.414	1.732	1.000	1.000	2
R45	3	1	2	1.732	2.449	0.000	0.000	2
R48	3	1	1	1.414	3.000	1.000	1.000	2
R51	3	1	2	1.732	2.449	0.000	0.000	2
R53	2	1	2	1.414	1.732	1.000	1.000	2
R54	3	2	3	2.236	2.000	1.414	1.414	2
R57	3	1	1	1.414	3.000	1.000	1.000	2
R59	2	1	2	1.414	1.732	1.000	1.000	2
R60	3	1	2	1.732	2.449	0.000	0.000	2
R63	3	1	1	1.414	3.000	1.000	1.000	2
R65	2	1	2	1.414	1.732	1.000	1.000	2
R66	3	2	2	1.414	2.236	1.000	1.000	2
R68	2	1	2	1.414	1.732	1.000	1.000	2
R69	3	1	2	1.732	2.449	0.000	0.000	2
R71	2	1	2	1.414	1.732	1.000	1.000	2
R72	3	1	1	1.414	3.000	1.000	1.000	2
R75	3	2	3	2.236	2.000	1.414	1.414	2
R78	3	1	1	1.414	3.000	1.000	1.000	2

Keterangan :

R : Responden **C2** : *Cluster 2*
C0 : *Cluster 0* **JT** : Jarak Terdekat
C1 : *Cluster 1* **C** : *Cluster*

3. Pembaruan *Centroid*: Menentukan *centroid* baru dengan cara menghitung banyaknya anggota setiap *cluster*, kemudian menjumlahkan

nilai dari semua anggota dibagi dengan banyaknya anggota setiap *cluster* :

- a. *Cluster* pertama mempunyai 27 anggota yaitu R5, R7, R10, R11, R14, R16, R17, R19, R20, R22, R28, R29, R31, R34, R37, R38, R40, R41, R46, R47, R55, R56, R58, R61, R62, R74, dan R76 maka :

$$C0,1 = \frac{(2+1+1+2+2+1+2+1+2+1+1+2+1+1+1+2+1+2+1+2+1+2+1+1+2+2+1)}{27} = 1,44$$

$$C0,2 = \frac{(1+1+1+2+1+1+1+1+1+1+1+1+1+1+1+1+1+1+2+1+1+1+1+1)}{27} = 1,07$$

$$C0,3 = \frac{(1+1+1+2+1+1+1+1+1+1+1+1+1+1+1+1+1+1+2+1+1+1+1+1)}{27} = 1,07$$

- b. *Cluster* kedua mempunyai 20 anggota yaitu R1, R4, R8, R13, R23, R25, R26, R32, R35, R43, R49, R37, R50, R52, R64, R67, R70, R73, R77, R79, dan R80 maka :

$$C1,1 = \frac{(1+1+2+1+2+1+2+2+2+1+1+2+2+1+1+1+1+1+2+1+1+2)}{20} = 1,35$$

$$C1,2 = \frac{(1+2+2+1+2+1+2+2+2+2+2+2+2+1+1+2+2+2+2+1+2+1)}{20} = 1,70$$

$$C1,3 = \frac{(2+3+3+2+3+2+3+3+3+3+3+3+3+2+2+3+2+3+3+2+2+2)}{20} = 2,60$$

- c. *Cluster* ketiga mempunyai 33 anggota yaitu R2, R3, R6, R9, R12, R15, R18, R21, R24, R27, R30, R33, R36, R39, R42, R44, R45, R48, R51, R53, R54, R57, R59, R60, R63, R65, R66, R68, R69, R71, R72, R75, dan R78 maka :

$$C2,1=(2+3+3+3+3+3+3+3+3+3+3+3+3+3+3+2+3+3+3+2+3+3+2+3+3+2+3+3+2+3+3+2+3+3+2+3+3+3)/33= 2.73$$

$$C2,2=(1+2+1+1+1+1+1+2+1+1+1+1+1+1+1+1+1+1+1+2+1+1+1+1+1+2+1+1+1+1+1+2+1+1+1+1+2+1)/33= 1.12$$

$$C2,3=(2+2+1+2+2+1+1+3+1+1+1+1+2+1+2+2+2+1+2+2+3+1+2+2+1+2+2+2+2+2+1+3+1)/33= 1.64$$

Tabel 3.14. Hasil *Centroid* Baru

<i>Centroid</i>	Rasa Mochi	Kemasan	Pelayanan
<i>Cluster 1</i>	1,44	1,07	1,07
<i>Cluster 2</i>	1.35	1.70	2,60
<i>Cluster 3</i>	2,73	1.12	1.64

- Iterasi: Ulangi langkah kedua yaitu menghitung jarak setiap data terhadap setiap *centroid*. Kemudian langkah ketiga yaitu menentukan anggota setiap *cluster* dengan data yang jaraknya paling dekat, selanjutnya langkah keempat menentukan *centroid* yang baru, Ulangi terus langkah kedua sampai keempat (iterasi) hingga posisi data pada *cluster* tidak berubah.

Tabel 3.15. Hasil Iterasi 2

Responden	Rasa Mochi	Kemasan	Pelayanan	C0	C1	C2	JT	C
R5	2	1	1	0.565	1.863	0.974	0.565	0
R7	1	1	1	0.457	1.781	1.845	0.457	0
R10	1	1	1	0.457	1.781	1.845	0.457	0
R14	2	1	1	0.565	1.863	0.974	0.565	0
R16	1	1	1	0.457	1.781	1.845	0.457	0
R17	2	1	1	0.565	1.863	0.974	0.565	0

R19	1	1	1	0.457	1.781	1.845	0.457	0
R20	2	1	1	0.565	1.863	0.974	0.565	0
R22	1	1	1	0.457	1.781	1.845	0.457	0
R28	1	1	1	0.457	1.781	1.845	0.457	0
R29	2	1	1	0.565	1.863	0.974	0.565	0
R31	1	1	1	0.457	1.781	1.845	0.457	0
R34	1	1	1	0.457	1.781	1.845	0.457	0
R37	1	1	1	0.457	1.781	1.845	0.457	0
R38	2	1	1	0.565	1.863	0.974	0.565	0
R40	1	1	1	0.457	1.781	1.845	0.457	0
R41	2	1	1	0.565	1.863	0.974	0.565	0
R46	1	1	1	0.457	1.781	1.845	0.457	0
R47	2	1	1	0.565	1.863	0.974	0.565	0
R55	1	1	1	0.457	1.781	1.845	0.457	0
R58	1	1	1	0.457	1.781	1.845	0.457	0
R61	1	1	1	0.457	1.781	1.845	0.457	0
R62	2	1	1	0.565	1.863	0.974	0.565	0
R74	2	1	1	0.565	1.863	0.974	0.565	0
R76	1	1	1	0.457	1.781	1.845	0.457	0
R11	2	2	2	1.422	0.934	1.197	0.934	1
R56	2	2	2	1.422	0.934	1.197	0.934	1
R1	1	1	2	1.030	0.986	1.769	0.986	1
R4	1	2	3	2.183	0.610	2.370	0.610	1
R8	2	2	3	2.208	0.820	1.778	0.820	1
R13	1	1	2	1.030	0.986	1.769	0.986	1
R23	2	2	3	2.208	0.820	1.778	0.820	1
R25	1	1	2	1.030	0.986	1.769	0.986	1
R26	2	2	3	2.208	0.820	1.778	0.820	1
R32	2	2	3	2.208	0.820	1.778	0.820	1
R35	2	2	3	2.208	0.820	1.778	0.820	1
R43	1	2	3	2.183	0.610	2.370	0.610	1

R49	1	2	3	2.183	0.610	2.370	0.610	1
R50	2	2	3	2.208	0.820	1.778	0.820	1
R52	1	1	2	1.030	0.986	1.769	0.986	1
R64	1	1	2	1.030	0.986	1.769	0.986	1
R67	1	2	3	2.183	0.610	2.370	0.610	1
R70	1	2	2	1.383	0.757	1.972	0.757	1
R73	1	2	3	2.183	0.610	2.370	0.610	1
R77	2	2	3	2.208	0.820	1.778	0.820	1
R79	1	1	2	1.030	0.986	1.769	0.986	1
R80	1	2	2	1.383	0.757	1.972	0.757	1
R2	2	1	2	1.082	1.128	0.822	0.822	2
R3	3	2	2	2.033	1.781	0.989	0.989	2
R6	3	1	1	1.559	2.403	0.703	0.703	2
R9	3	1	2	1.812	1.890	0.470	0.470	2
R12	3	1	2	1.812	1.890	0.470	0.470	2
R15	3	1	1	1.559	2.403	0.703	0.703	2
R18	3	1	1	1.559	2.403	0.703	0.703	2
R21	3	2	3	2.643	1.724	1.645	1.645	2
R24	3	1	1	1.559	2.403	0.703	0.703	2
R27	3	1	1	1.559	2.403	0.703	0.703	2
R30	3	1	1	1.559	2.403	0.703	0.703	2
R33	3	1	1	1.559	2.403	0.703	0.703	2
R36	3	1	2	1.812	1.890	0.470	0.470	2
R39	3	1	1	1.559	2.403	0.703	0.703	2
R42	3	1	2	1.812	1.890	0.470	0.470	2
R44	2	1	2	1.082	1.128	0.822	0.822	2
R45	3	1	2	1.812	1.890	0.470	0.470	2
R48	3	1	1	1.559	2.403	0.703	0.703	2
R51	3	1	2	1.812	1.890	0.470	0.470	2
R53	2	1	2	1.082	1.128	0.822	0.822	2
R54	3	2	3	2.643	1.724	1.645	1.645	2

R57	3	1	1	1.559	2.403	0.703	0.703	2
R59	2	1	2	1.082	1.128	0.822	0.822	2
R60	3	1	2	1.812	1.890	0.470	0.470	2
R63	3	1	1	1.559	2.403	0.703	0.703	2
R65	2	1	2	1.082	1.128	0.822	0.822	2
R66	3	2	2	2.033	1.781	0.989	0.989	2
R68	2	1	2	1.082	1.128	0.822	0.822	2
R69	3	1	2	1.812	1.890	0.470	0.470	2
R71	2	1	2	1.082	1.128	0.822	0.822	2
R72	3	1	1	1.559	2.403	0.703	0.703	2
R75	3	2	3	2.643	1.724	1.645	1.645	2
R78	3	1	1	1.559	2.403	0.703	0.703	2

Tabel 3.16. Hasil *Centroid* Baru

<i>Centroid</i>	Rasa Mochi	Kemasan	Pelayanan
<i>Cluster 1</i>	1,40	1,00	1,00
<i>Cluster 2</i>	1.41	1.73	2,55
<i>Cluster 3</i>	2,79	1.15	1.70

Tabel 3.17. Hasil Iterasi 3

Responden	Rasa Mochi	Kemasan	Pelayanan	C0	C1	C2	JT	C
R5	2	1	1	0.600	1.807	1.063	0.600	0
R7	1	1	1	0.400	1.756	1.925	0.400	0
R10	1	1	1	0.400	1.756	1.925	0.400	0
R14	2	1	1	0.600	1.807	1.063	0.600	0
R16	1	1	1	0.400	1.756	1.925	0.400	0
R17	2	1	1	0.600	1.807	1.063	0.600	0
R19	1	1	1	0.400	1.756	1.925	0.400	0

R20	2	1	1	0.600	1.807	1.063	0.600	0
R22	1	1	1	0.400	1.756	1.925	0.400	0
R28	1	1	1	0.400	1.756	1.925	0.400	0
R29	2	1	1	0.600	1.807	1.063	0.600	0
R31	1	1	1	0.400	1.756	1.925	0.400	0
R34	1	1	1	0.400	1.756	1.925	0.400	0
R37	1	1	1	0.400	1.756	1.925	0.400	0
R38	2	1	1	0.600	1.807	1.063	0.600	0
R40	1	1	1	0.400	1.756	1.925	0.400	0
R41	2	1	1	0.600	1.807	1.063	0.600	0
R46	1	1	1	0.400	1.756	1.925	0.400	0
R47	2	1	1	0.600	1.807	1.063	0.600	0
R55	1	1	1	0.400	1.756	1.925	0.400	0
R58	1	1	1	0.400	1.756	1.925	0.400	0
R61	1	1	1	0.400	1.756	1.925	0.400	0
R62	2	1	1	0.600	1.807	1.063	0.600	0
R74	2	1	1	0.600	1.807	1.063	0.600	0
R76	1	1	1	0.400	1.756	1.925	0.400	0
R11	2	2	2	1.536	0.849	1.197	0.849	1
R56	2	2	2	1.536	0.849	1.197	0.849	1
R1	1	1	2	1.077	0.997	1.820	0.997	1
R4	1	2	3	2.272	0.670	2.369	0.670	1
R8	2	2	3	2.315	0.794	1.743	0.794	1
R13	1	1	2	1.077	0.997	1.820	0.997	1
R23	2	2	3	2.315	0.794	1.743	0.794	1
R25	1	1	2	1.077	0.997	1.820	0.997	1
R26	2	2	3	2.315	0.794	1.743	0.794	1
R32	2	2	3	2.315	0.794	1.743	0.794	1
R35	2	2	3	2.315	0.794	1.743	0.794	1
R43	1	2	3	2.272	0.670	2.369	0.670	1
R49	1	2	3	2.272	0.670	2.369	0.670	1

R50	2	2	3	2.315	0.794	1.743	0.794	1
R52	1	1	2	1.077	0.997	1.820	0.997	1
R64	1	1	2	1.077	0.997	1.820	0.997	1
R67	1	2	3	2.272	0.670	2.369	0.670	1
R70	1	2	2	1.470	0.734	2.002	0.734	1
R73	1	2	3	2.272	0.670	2.369	0.670	1
R77	2	2	3	2.315	0.794	1.743	0.794	1
R79	1	1	2	1.077	0.997	1.820	0.997	1
R80	1	2	2	1.470	0.734	2.002	0.734	1
R2	2	1	2	1.166	1.084	0.858	0.858	2
R3	3	2	2	2.135	1.704	0.926	0.926	2
R6	3	1	1	1.600	2.334	0.744	0.744	2
R9	3	1	2	1.887	1.832	0.400	0.400	2
R12	3	1	2	1.887	1.832	0.400	0.400	2
R15	3	1	1	1.600	2.334	0.744	0.744	2
R18	3	1	1	1.600	2.334	0.744	0.744	2
R21	3	2	3	2.750	1.677	1.569	1.569	2
R24	3	1	1	1.600	2.334	0.744	0.744	2
R27	3	1	1	1.600	2.334	0.744	0.744	2
R30	3	1	1	1.600	2.334	0.744	0.744	2
R33	3	1	1	1.600	2.334	0.744	0.744	2
R36	3	1	2	1.887	1.832	0.400	0.400	2
R39	3	1	1	1.600	2.334	0.744	0.744	2
R42	3	1	2	1.887	1.832	0.400	0.400	2
R44	2	1	2	1.166	1.084	0.858	0.858	2
R45	3	1	2	1.887	1.832	0.400	0.400	2
R48	3	1	1	1.600	2.334	0.744	0.744	2
R51	3	1	2	1.887	1.832	0.400	0.400	2
R53	2	1	2	1.166	1.084	0.858	0.858	2
R54	3	2	3	2.750	1.677	1.569	1.569	2
R57	3	1	1	1.600	2.334	0.744	0.744	2

R59	2	1	2	1.166	1.084	0.858	0.858	2
R60	3	1	2	1.887	1.832	0.400	0.400	2
R63	3	1	1	1.600	2.334	0.744	0.744	2
R65	2	1	2	1.166	1.084	0.858	0.858	2
R66	3	2	2	2.135	1.704	0.926	0.926	2
R68	2	1	2	1.166	1.084	0.858	0.858	2
R69	3	1	2	1.887	1.832	0.400	0.400	2
R71	2	1	2	1.166	1.084	0.858	0.858	2
R72	3	1	1	1.600	2.334	0.744	0.744	2
R75	3	2	3	2.750	1.677	1.569	1.569	2
R78	3	1	1	1.600	2.334	0.744	0.744	2

Berdasarkan hasil *clustering*, dengan melakukan 3 iterasi maka diperoleh tiga kelompok pelanggan dengan jumlah anggota sebagai berikut:

1. Klaster 0 : 25 pelanggan (R5, R7, R10, R14, R16, R17, R19, R20, R22, R28, R29, R31, R34, R37, R38, R40, R41, R46, R47, R55, R58, R61, R62, R74, dan R76).
2. Klaster 1 : 22 pelanggan (R1, R4, R8, R11, R13, R23, R25, R26, R32, R35, R43, R49, R50, R52, R56, R64, R67, R70, R73, R77, R79, dan R80).
3. Klaster 2 : 33 pelanggan (R2, R3, R6, R9, R12, R15, R18, R21, R24, R27, R30, R33, R36, R39, R42, R44, R45, R48, R51, R53, R54, R57, R59, R60, R63, R65, R66, R68, R69, R71, R72, R75, dan R78).

Berdasarkan karakteristik hasil *clustering* ini, kita dapat melakukan interpretasi berdasarkan pola data:

1. Klaster 0 (25 pelanggan)

Kemungkinan kelompok ini memiliki tingkat kepuasan rendah. Pelanggan dalam klaster ini mungkin mengalami keluhan terhadap kualitas layanan, atau aspek lain yang tidak memenuhi harapan mereka. Pengusaha perlu

memperhatikan pelanggan dalam kelompok ini karena mereka memiliki potensi untuk berhenti menggunakan layanan jika tidak ada perbaikan.

2. Klaster 1 (22 pelanggan)

Pelanggan dalam klaster ini kemungkinan memiliki kepuasan sedang. Mereka mungkin memiliki pengalaman yang cukup baik, tetapi ada beberapa aspek yang bisa ditingkatkan untuk memastikan kepuasan lebih tinggi. Fokus peningkatan layanan pada kelompok ini dapat membantu mengubah mereka menjadi pelanggan yang lebih loyal.

3. Klaster 2 (33 pelanggan)

Ini adalah kelompok pelanggan dengan kepuasan tertinggi. Mereka kemungkinan besar merasa puas dengan layanan, kualitas produk, harga, atau faktor lain yang memengaruhi pengalaman mereka. Pelanggan dalam klaster ini berpotensi menjadi pelanggan tetap dan bahkan bisa memberikan rekomendasi positif kepada orang lain.

Sebagian besar pelanggan masuk dalam Klaster 2 (paling puas), yang menunjukkan bahwa layanan atau produk saat ini secara umum mendapat tanggapan positif. Masih ada pelanggan dalam Klaster 0 dan 1, yang berarti ada peluang untuk meningkatkan kepuasan mereka melalui berbagai strategi seperti peningkatan layanan pelanggan, perbaikan kualitas produk, atau promosi yang lebih menarik. Perlu dilakukan analisis lebih lanjut untuk mengetahui faktor spesifik yang menyebabkan pelanggan masuk ke klaster tertentu.

Analisis *K-means clustering* ini memberikan gambaran yang jelas mengenai pola kepuasan pelanggan. Dengan memahami karakteristik setiap klaster, perusahaan dapat menyusun strategi yang lebih efektif untuk meningkatkan kepuasan pelanggan secara keseluruhan.