


Artikel - Kapti

 Universitas Muhammadiyah Surakarta

Document Details

Submission ID

trn:oid::6284:105158863

Submission Date

Jul 20, 2025, 11:38 PM GMT+7

Download Date

Jul 20, 2025, 11:45 PM GMT+7

File Name

Artikel - Kapti

File Size

41.0 KB

82 Pages




5,785 Words

35,891 Characters

17% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Top Sources

- 19%  Internet sources
- 1%  Publications
- 8%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

- 19% Internet sources
- 1% Publications
- 8% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	sistemasi.ftik.unisi.ac.id	11%
2	Internet	www.mdpi.com	4%
3	Internet	ijcis.net	1%
4	Internet	ejournal.poltekharber.ac.id	1%
5	Internet	research.tue.nl	<1%
6	Internet	www.researchgate.net	<1%
7	Internet	pmc.ncbi.nlm.nih.gov	<1%
8	Internet	docobook.com	<1%
9	Internet	journal2.uad.ac.id	<1%
10	Internet	repositorio.usp.br	<1%
11	Submitted works	Athens University of Economics and Business on 2024-09-20	<1%

12	Internet	bright-journal.org	<1%
13	Submitted works	Universitas Amikom on 2019-01-04	<1%
14	Internet	p3m.sinus.ac.id	<1%
15	Internet	jutif.if.unsoed.ac.id	<1%
16	Internet	link.springer.com	<1%
17	Submitted works	Universitas Brawijaya on 2017-06-21	<1%
18	Submitted works	BPP College of Professional Studies Limited on 2023-10-29	<1%
19	Submitted works	University of Sussex on 2024-01-26	<1%
20	Internet	repository.unri.ac.id	<1%
21	Internet	www.djes.info	<1%
22	Internet	repositorio.unifesp.br	<1%
23	Submitted works	Universitas Kristen Satya Wacana on 2025-07-06	<1%
24	Internet	www.sbu.se	<1%
25	Submitted works	Liberty University on 2025-04-29	<1%

26	Internet	idss.iocspublisher.org	<1%
27	Internet	repository.uinsu.ac.id	<1%
28	Submitted works	President University on 2017-01-19	<1%
29	Internet	core.ac.uk	<1%
30	Internet	jonedu.org	<1%
31	Submitted works	IUBH - Internationale Hochschule Bad Honnef-Bonn on 2024-08-16	<1%
32	Submitted works	Universidad Pontificia Bolivariana on 2024-10-12	<1%
33	Internet	ebin.pub	<1%
34	Internet	eltikom.poliban.ac.id	<1%
35	Internet	kc.umn.ac.id	<1%
36	Submitted works	STIKOM Surabaya on 2013-05-27	<1%
37	Submitted works	Management & Science University on 2021-10-01	<1%
38	Submitted works	Nottingham Trent University on 2024-09-03	<1%
39	Submitted works	Universitas Putera Batam on 2021-03-16	<1%

40	Submitted works	University of Technology, Sydney on 2024-06-14	<1%
41	Internet	doaj.org	<1%
42	Internet	ejournal.undiksha.ac.id	<1%
43	Internet	ojs.trigunadharma.ac.id	<1%
44	Internet	pe.nchu.edu.tw	<1%
45	Internet	techniumscience.com	<1%
46	Submitted works	University of Greenwich on 2025-06-20	<1%
47	Internet	dergipark.org.tr	<1%
48	Internet	docplayer.info	<1%
49	Internet	repository.its.ac.id	<1%
50	Submitted works	University of Nottingham on 2024-09-05	<1%
51	Internet	idtesis.com	<1%
52	Internet	specrada.bsmu.edu.ua	<1%
53	Internet	vestnik.rsmu.press	<1%

54	Submitted works	RDI Distance Learning on 2025-01-13	<1%
55	Submitted works	STT PLN on 2020-09-23	<1%
56	Submitted works	Universitas Dian Nuswantoro on 2023-07-19	<1%
57	Internet	adoc.pub	<1%
58	Internet	colegiodemedicos.es	<1%
59	Internet	es.scribd.com	<1%
60	Internet	ouci.dntb.gov.ua	<1%
61	Internet	painmedicine.org.ua	<1%
62	Internet	polodelconocimiento.com	<1%
63	Submitted works	Universitas Brawijaya on 2018-07-13	<1%
64	Submitted works	Universitas Muria Kudus on 2016-08-10	<1%
65	Submitted works	University of Northumbria at Newcastle on 2025-05-12	<1%

38 IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK SEGMENTASI
PELANGGAN

38 Implementation of K-Means Clustering Algorithm for
Customer Segmentation

1Kapti*, 2Dwi Astuti

1Teknik Informatika, STMIK Bina Patria

30 2Sistem Informasi, STMIK Bina Patria

1,2Jl. Raden Saleh No.7, Potrobangsari, Kec. Magelang Utara, Kota Magelang,

37 Jawa Tengah, Indonesia

*e-mail: tensmart18@stmikbinapatria.ac.id

(received: ?, revised: ?, accepted: ? diisi oleh editor)

Abstrak

39 Seiring dengan berkembangnya teknologi digital, perusahaan dihadapkan pada volume data pelanggan yang semakin besar dan kompleks. Untuk itu, diperlukan pendekatan yang lebih efektif dalam segmentasi pelanggan agar dapat memahami pola perilaku dan preferensi konsumen dengan lebih baik. Salah satu metode yang umum digunakan dalam segmentasi pelanggan adalah algoritma K-Means Clustering. Meskipun K-Means banyak diaplikasikan, algoritma ini memiliki keterbatasan, seperti ketergantungan terhadap inisialisasi centroid awal dan kerentanannya terhadap outlier,

20 yang dapat mempengaruhi akurasi dan stabilitas hasil segmentasi. Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means Clustering dalam segmentasi pelanggan dengan membagi dalam tiga kluster yaitu pelanggan sangat rajin, cukup rajin dan kurang rajin belanja. Tujuan pengembangan sistem ini adalah untuk meningkatkan efisiensi dalam pengelompokan data pelanggan berdasarkan data transaksi dan preferensi pembelian pelanggan yang relevan. Parameter yang digunakan dalam klusterisasi ini berdasarkan jumlah pembelian, jumlah transaksi dan jumlah barang. Hasil pengujian fungsionalitas menunjukkan sistem dinyatakan baik karena seluruh skenario uji berjalan sukses. Hasil Pengujian dengan metode Silhouette Coefficient (SC) menghasilkan status strong struktur dengan nilai rata SC sebesar 0,97

Kata kunci: K-Means Clustering, Segmentasi Pelanggan, Pengelompokan, Pemasaran Digital

Abstract

50 As digital technology advances, companies are faced with increasingly large and complex volumes of customer data. Therefore, a more effective approach to customer segmentation is needed to better understand consumer behavior patterns and preferences. One method commonly used in customer segmentation is the K-Means Clustering algorithm. Although K-Means is widely applied, this algorithm has limitations, such as dependence on initial centroid initialization and its susceptibility to outliers, which can affect the accuracy and stability of segmentation results. This study aims to implement the K-Means Clustering algorithm in customer segmentation by

44 dividing customers into three clusters: very diligent customers, moderately diligent customers, and less diligent customers. The purpose of developing this system is to improve the efficiency of grouping customer data based on transaction data and relevant customer purchasing preferences. The parameters used in this clustering are based on the number of purchases, the number of transactions, and the number of items. 47 The results of functional testing indicate that the system is declared good because all test scenarios run successfully. The test results using the Silhouette Coefficient (SC) method produced a strong structure status with an average SC value of 0.97.

1 Keywords: Clustering; Grouping; Drugs; K-Means Clustering; Health Center.

Pendahuluan

35 Perkembangan teknologi informasi yang pesat telah membawa perubahan signifikan dalam berbagai sektor bisnis, termasuk dalam pengelolaan dan analisis data pelanggan. Data pelanggan yang dihasilkan dari transaksi e-commerce, aplikasi, dan platform digital lainnya semakin besar dan kompleks, menuntut perusahaan untuk mengadopsi teknik analisis data yang lebih canggih untuk memahami perilaku pelanggan. 8 Segmentasi pelanggan menjadi salah satu metode penting dalam pemasaran yang memungkinkan perusahaan untuk mengelompokkan pelanggan berdasarkan karakteristik atau perilaku tertentu sehingga dapat memberikan layanan yang lebih personal dan efektif.

Segmentasi pelanggan tradisional sering kali mengandalkan data demografis atau geografis yang terbatas. Namun, dalam era digital ini, segmentasi yang lebih tepat

8 harus didasarkan pada data transaksi dan preferensi pembelian pelanggan yang lebih dinamis. Salah satu teknik yang populer digunakan untuk segmentasi pelanggan adalah algoritma K-Means Clustering. Algoritma ini memungkinkan pengelompokan pelanggan berdasarkan kemiripan data mereka dalam ruang fitur yang multidimensional. Namun, meskipun K-Means banyak digunakan dalam berbagai aplikasi, algoritma ini memiliki kelemahan yang signifikan, seperti ketergantungan pada inisialisasi centroid awal dan kerentanannya terhadap outlier[1].

56 Keterbatasan ini dapat menyebabkan hasil segmentasi yang tidak stabil dan kurang akurat. Untuk itu, diperlukan pengembangan algoritma K-Means yang lebih adaptif, seperti penerapan metode K-Means++ untuk inisialisasi centroid yang lebih cerdas dan teknik robust untuk menangani outlier. Dengan mengintegrasikan teknik-teknik ini, diharapkan dapat diperoleh hasil segmentasi yang lebih stabil dan akurat, yang pada gilirannya dapat meningkatkan efektivitas strategi pemasaran dan pengalaman pelanggan.

8 Algoritma K-Means Clustering telah menjadi salah satu metode yang banyak digunakan untuk segmentasi pelanggan[2]. Namun, algoritma ini menghadapi beberapa keterbatasan, di antaranya adalah ketergantungannya terhadap inisialisasi centroid awal dan kerentanan terhadap outlier, yang dapat menyebabkan hasil segmentasi yang tidak stabil dan tidak akurat [3]-[4]. Hal ini dapat membahayakan keputusan strategis perusahaan dan berdampak negatif pada pengalaman pelanggan. Dengan mengalami salah pengelompokan, perusahaan dapat membuang sumber daya dalam upaya pemasaran yang tidak efektif, dan berpotensi kehilangan loyalitas pelanggan dan pendapatan [5]-[6].

Untuk itu riset ini dirancang untuk menjawab tantangan tersebut dengan mengembangkan algoritma K-Means yang lebih adaptif yang mengintegrasikan teknik inisialisasi centroid yang lebih cerdas (K-Means++) dan metode yang lebih robust untuk menangani outlier [7]. Penelitian ini bertujuan tidak hanya untuk meningkatkan kualitas segmentasi tetapi juga untuk memberikan panduan praktik bagi perusahaan dalam menerapkan strategi pemasaran yang lebih efektif dan personal[8]-[9].

Urgensi penelitian ini dapat dilihat dari beberapa dimensi kritis. Secara ekonomi, penelitian ini berpotensi mencegah pemborosan sumber daya pemasaran dan mengoptimalkan strategi target pasar[10]-[11]. Dari perspektif teknologi, penelitian ini memberikan kontribusi metodologis dalam bidang data mining dan menghadirkan solusi inovatif bagi keterbatasan algoritma eksisting[12]. Secara strategis, penelitian ini sejalan dengan Prioritas Riset Nasional dan mendukung transformasi digital yang sedang berlangsung [13].

Penelitian ini mengusulkan pendekatan komprehensif untuk mengatasi keterbatasan algoritma K-Means clustering dalam segmentasi pelanggan [14]. Strategi utama difokuskan pada perbaikan inisialisasi centroid melalui adopsi teknik K-Means++, yang memungkinkan pemilihan titik pusat awal yang lebih cerdas dan tersebar, secara signifikan mengurangi bias inisialisasi tradisional[15].

Pendekatan inovatif ini akan dikombinasikan dengan metode penanganan outlier yang robust, menggunakan teknik deteksi berbasis jarak untuk mengelola data ekstrim tanpa menghilangkan informasi penting [16].

1 Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan algoritma K-Means Clustering yang lebih adaptif untuk segmentasi pelanggan, dengan fokus

49 pada penggunaan data transaksi dan preferensi pembelian. Penelitian ini juga akan melakukan evaluasi komparatif terhadap algoritma K-Means yang telah dimodifikasi dan K-Means konvensional untuk mengukur perbaikan dalam akurasi dan stabilitas segmentasi. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam penerapan algoritma K-Means untuk strategi pemasaran yang lebih personal dan berbasis data [17].

Tinjauan Literatur

2 Penelitian ini memiliki keterkaitan dengan penelitian terdahulu seperti penelitian dengan judul, "Applying Machine Learning in Marketing: An Analysis Using the NMF and k-Means Algorithms", Penelitian ini menggunakan NMF untuk topic modeling dan k-means untuk analisis komparatif dalam pemasaran. Fokusnya adalah meningkatkan segmentasi pelanggan dan kampanye periklanan. Penelitian ini menggunakan k-means namun lebih menekankan pada penerapan machine learning secara umum, sedangkan penelitian pengusul fokus pada pengembangan k-means untuk segmentasi berdasarkan data transaksi dan preferensi pembelian [18].

32 Penelitian dengan judul, "Customer Segmentation in E-commerce: A Context-Aware Quality Framework for Comparing Clustering Algorithms", Penelitian ini 23 membandingkan beberapa algoritma clustering, termasuk k-means, dalam konteks e-commerce. Meskipun menggunakan k-means, penelitian ini juga memperkenalkan kerangka kerja kualitas dan membandingkan dengan algoritma lain. Penelitian 55 pengusul fokus pada pengembangan algoritma k-means untuk segmentasi pelanggan berdasarkan data yang lebih spesifik [19].

Penelitian dengan judul, "Enhancing K-Means Clustering with Post-Redistribution", Penelitian ini mengusulkan penyempurnaan algoritma k-means dengan redistribusi pasca-clustering untuk meningkatkan kualitas cluster. Fokusnya pada perbaikan algoritma k-means berbeda dengan penelitian pengusul yang tidak mengusulkan modifikasi serupa tetapi tetap berfokus pada pengembangan k-means untuk segmentasi pelanggan [20].

12 Penelitian dengan judul, "Applied Density-Based Clustering Techniques for Classifying High-Risk Customers: A Case Study of Commercial Banks in Vietnam", Penelitian ini membandingkan k-means dengan algoritma berbasis kepadatan (DBSCAN, HDBSCAN) dalam mengklasifikasikan pelanggan berisiko tinggi. Penelitian pengusul lebih terfokus pada k-means untuk segmentasi pelanggan berdasarkan kriteria jumlah pembelian, jumlah transaksi dan jumlah barang[21].

Penelitian ini menghadirkan lima kebaruan utama dalam pengembangan algoritma K-Means untuk segmentasi pelanggan. Pertama, algoritma dibuat lebih adaptif dengan mengintegrasikan K-Means++ untuk inisialisasi centroid, sehingga meningkatkan kualitas cluster. Kedua, diterapkan metode robust untuk mendeteksi dan menangani outlier secara efektif guna menjaga integritas data. Ketiga, penelitian ini memfokuskan pada data transaksi dan preferensi pembelian pelanggan, bukan hanya data demografis atau perilaku online seperti pada penelitian sebelumnya, sehingga memberikan segmentasi yang lebih relevan untuk strategi pemasaran. Keempat, algoritma dikembangkan secara spesifik untuk konteks industri ritel atau e-commerce agar lebih aplikatif dalam dunia bisnis nyata. Kelima, digunakan metrik evaluasi yang

disesuaikan dengan tujuan bisnis, bukan hanya metrik standar, agar penilaian kinerja algoritma lebih tepat dan kontekstual.

Metode Penelitian

Metodologi penelitian ini dirancang untuk mengembangkan algoritma K-Means Clustering yang lebih adaptif untuk segmentasi pelanggan berdasarkan data transaksi dan preferensi pembelian. Metode penelitian ini mengikuti empat tahapan utama seperti yang disebutkan dalam ringkasan, yaitu:

1. Pengumpulan Data Transaksi dan Preferensi Pelanggan

Sumber Data

Data transaksional diambil dari perusahaan ritel/e-commerce di Kota Magelang, data ini mencakup data pembelian pelanggan minimal 1 tahun terakhir. Data diambil di salah satu cabang Perusahaan PT Indomart di Kota Magelang melalui kerjasama mitra secara resmi.

Data preferensi pembelian dari survei pelanggan dan rekam jejak digital

Data demografis pelanggan sebagai data pendukung

1.2 Strategi Pengumpulan Data

Pengumpulan data dilakukan melalui kegiatan FGD dengan survei untuk identifikasi data

Pengambilan data menggunakan teknik API integration dan database querying

Pengumpulan data preferensi melalui survei online terhadap minimal 500 pelanggan

Pengumpulan data juga dilakukan secara langsung melalui wawancara

2. Pra Pemrosesan Data

Pembersihan data untuk menangani missing values dan data yang tidak konsisten

Transformasi data melalui normalisasi dan standardisasi

Feature engineering untuk menghasilkan variabel RFM (Recency, Frequency, Monetary)

Identifikasi awal outlier menggunakan metode statistik standar

Pengolahan Data dilakukan dengan Software Lisensi Analisis Data SPSS

3. Pengembangan Algoritma K-Means Adaptif

3.1 Integrasi K-Means++ untuk Inisialisasi Centroid

Implementasi teknik K-Means++ untuk pemilihan centroid awal yang optimal

Pengembangan formula probabilistik untuk pemilihan centroid yang meminimalisir overlapping [22].

Pengujian stabilitas inisialisasi pada berbagai skenario dataset

Pengujian Testing Model dilakukan dengan tools Software Lisensi Rapidminer

3.2 Pengembangan Metode Robust untuk Penanganan Outlier

Implementasi teknik Local Outlier Factor (LOF) untuk deteksi outlier

Pengembangan mekanisme pembobotan untuk mengurangi pengaruh outlier pada perhitungan centroid

Integrasi metode DBSCAN untuk identifikasi noise points sebelum diproses K-Means [23].

Pengembangan Metode dilakukan dengan tools Software Lisensi Matlab

3.3 Integrasi Feature Weighting

Pengembangan mekanisme pembobotan fitur berdasarkan signifikansi dalam pembentukan cluster

Implementasi distance measure yang dapat mengakomodasi bobot fitur berbeda

Pengembangan mekanisme adaptasi bobot berdasarkan karakteristik data

4. Evaluasi Komparatif

4.1 Metrik Evaluasi Internal

Silhouette Index untuk mengukur kohesi dan separasi cluster dengan batas nilai SC

$\geq 0,71$ (struktur kuat)

Davies-Bouldin Index untuk mengukur rasio scatter within-cluster terhadap between-cluster [24].

Calinski-Harabasz Index untuk mengevaluasi variance ratio

Dunn Index untuk mengukur kompaksi dan separasi cluster

4.2 Optimasi Parameter Algoritma

Pengembangan teknik penentuan jumlah cluster optimal menggunakan metode Elbow dan Silhouette

Implementasi adaptive learning rate untuk mempercepat konvergensi

Pengembangan teknik early stopping untuk efisiensi komputasi

4.3 Benchmarking Algoritma

Perbandingan dengan K-Means konvensional

Perbandingan dengan algoritma clustering lain (DBSCAN, Hierarchical Clustering, Fuzzy C-Means)[25].

Pengujian pada dataset sintetis dengan outlier yang dikontrol

Pengujian pada dataset riil dari perusahaan mitra

4.4 Analisis Stabilitas

Pengujian multiple runs dengan inisialisasi berbeda

Analisis standar deviasi hasil clustering

Analisis sensitivitas terhadap parameter input

Pengujian cross-validation untuk memastikan generalisasi algoritma

5. Validasi melalui Studi Kasus Riil

5.1 Implementasi Strategi Pemasaran

Pengembangan strategi pemasaran yang dipersonalisasi berdasarkan segmen pelanggan hasil clustering

Implementasi strategi pada kelompok eksperimen di perusahaan mitra

Pelaksanaan melakukan testing untuk mengukur efektivitas strategi pemasaran berbasis segmentasi baru [26].

Model dikembangkan dalam suatu web yang dibuat dengan tools Software Software

Lisensi Dreamweaver

5.2 Pengukuran Performa Bisnis

Evaluasi tingkat respons pelanggan terhadap kampanye pemasaran

Pengukuran Customer Retention Rate antar segmen

Analisis perubahan Customer Lifetime Value (CLV)[27].

Evaluasi ROI kampanye pemasaran berdasarkan segmentasi

5.3 Validasi dengan Domain Expert

Evaluasi interpretabilitas segmen oleh pakar pemasaran

Penilaian kegunaan segmentasi dalam pengambilan keputusan strategis

Analisis kualitatif terhadap manfaat segmentasi dalam personalisasi layanan

Implementasi dan pembuktian hasil dilakukan melalui uji validasi dengan membandingkan antara hasil perhitungan manual dengan hasil prototipe.

Validasi eksternal dilakukan melalui penilaian langsung oleh stakeholder dan konsumen menggunakan kuesioner.

Adapun Alir Alur Penelitian ini dapat dilihat pada gambar 1.

Gambar 1. Diagram Alur Penelitian

Hasil dan Pembahasan

Implementasi Algoritma K-Means Clustering

Pada penelitian ini peneliti membatasi dan fokus pada implementasi dari Algoritma K-Means Clustering serta dikembangkan pada suatu sistem untuk mengelompokan segmentasi pelanggan. Data yang digunakan dalam penelitian ini mengambil sebanyak 20 data sampel secara acak dari PT Indomart Magelang. Data yang diambil berdasarkan parameter seperti jumlah pembelian, jumlah transaksi dan jumlah barang.

Adapun tahapan dari Implementasi Algoritma K-Means Clustering sebagai berikut:

Menentukan Dataset Pelanggan

Pada penelitian ini dari data pelanggan yang telah dikumpulkan kemudian dibuat dalam bentuk dataset seperti disajikan pada tabel 1.

Tabel 1. Dataset Pelanggan

No

Nama Pelanggan

Alamat

Jumlah Pembelian

Jumlah Transaksi

Jumlah Barang

1

Ahmad

Dukun

Rp 87.625

4

60

2

Ardian

Kaliangkrik

Rp 34.250

3

15

3

Bagas

Kajoran

Rp 40.000

1

5,5

4

Budianto

Dukun

Rp 37.000

2

22

5

Cerly

Kajoran

Rp 256.750

5

36

6

Farid

Bandongan

Rp 139.000

3

40

7

Fitriani

Dukun

Rp 271.750

5

142

8

Galih

Kaliangkrik

Rp 41.750

4

30

9

Gading

Ngablak

Rp 49.250

5

20

10

Herianto

Bandongan

Rp 145.750

5

27

11

Hermawan

Secang

Rp 312.500

7

129

12

Ikah

Salam

Rp 36.500

3

64

13

Joko

Bandongan

Rp 49.250

4

39

14

Kamal

Tegalrejo

Rp 49.250

3

41

15

Kurnia

Ngablak

Rp 40.750

2

23

16

Lutfy

Pakis

Rp 49.250

4

45

17

Mona

Sawangan

Rp 149.250

6

65

18

Nasir

Ngluwar

Rp 254.250

6

45

19

Opta

Pakis

Rp 45.250

4

32

20

Prabu

Ngluwar

Rp 17.500

2

9

MIN :

Rp 17.500

1

5,5

MAX :

Rp 312.500

7

142

1 Menghitung Normalisasi

Berikut ini contoh perhitungan Normalisasi pada baris pertama.

17 Jumlah Pembelian = $(\text{Nilai} - \text{Min}(\text{nilai})) / (\text{Max}(\text{nilai}) - (\text{Min}(\text{nilai})))$

$$= (87625-17500)/(312500-17500)$$
$$= 0,24$$

17 Jumlah Transaksi = $(\text{Nilai}-\text{Min}(\text{nilai})) / (\text{Max}(\text{nilai}) - (\text{Min}(\text{nilai})))$

$$= ((4-1)/(7-1))$$
$$= 0,5$$

17 Jumlah Barang = $(\text{Nilai}-\text{Min}(\text{nilai})) / (\text{Max}(\text{nilai}) - (\text{Min}(\text{nilai})))$

$$= ((60-5,5)/(142-5,5))$$
$$= 0,40$$

1 Adapun rekap proses perhitungan normalisasi ditunjukkan pada tabel 2.

Tabel 2. Normalisasi

No

Jumlah Pembelian

Jumlah Transaksi

Jumlah Barang

1

0,24

0,50

0,40

2

0,06

0,33

0,07

3

0,08

0,00

0,00

4

0,07

0,17

0,12

5

0,81

0,67

0,22

6

0,41

0,33

0,25

7

0,86

0,67

1,00

8

0,08

0,50

0,18

9

0,11

0,67

0,11

10

0,43

0,67

0,16

11

1,00

1,00

0,90

12

0,06

0,33

0,43

13

0,11

0,50

0,25

14

0,11

0,33

0,26

15

0,08

0,17

0,13

16

0,11

0,50

0,29

17

0,45

0,83

0,44

18

0,80

0,83

0,29

19

0,09

0,50

0,19

20

0,00

0,17

0,03

Menghitung Iterasi 1

Pada perhitungan iterasi ke 1 dilakukan dengan menghitung nilai parameter berdasarkan dibandingkan dengan centroid awal. Berikut ini penulis sajikan pada contoh baris pertama yaitu

$$\text{Cluster 1} = \sqrt{(0,24-0,06)^2 + ((0,50-0,17)^2) + ((0,40-0,03)^2)} = 0,53$$

$$\text{Cluster 2} = \sqrt{(0,24-0,11)^2 + ((0,50-0,33)^2) + ((0,40-0,25)^2)} = 0,26$$

$$\text{Cluster 3} = \sqrt{(0,24-0,86)^2 + ((0,50-0,83)^2) + ((0,40-1)^2)} = 0,93$$

Berdasarkan proses perhitungan normalisasi yang telah dilakukan sehingga menghasilkan iterasi ke 1 seperti disajikan pada tabel 3.

Tabel 3. Hasil Perhitungan Iterasi ke 1

No

Cluster 1

Cluster 2

Cluster 3

Posisi Cluster

Jarak Terdekat

Pelanggan

1

0,53

0,26

0,93

2

0,26

Ahmad

2

0,17

0,19

1,33

1

0,19

Ardian

3

0,17

0,42

1,52

1

0,17

Bagas

4

0,09

0,21

1,36

1

0,09

Budianto

5

0,92

0,78

0,80

2

0,78

Cerly

6

0,45

0,30

1,00

2

0,30

Farid

7

1,35

1,11

0,16

3

0,16

Fitriani

8

0,36

0,19

1,18

2

0,19

Galih

9

0,50

0,37

1,18

2

0,37

Gading

10

0,64

0,48

0,96

2

0,48

Herianto

11

1,53

1,29

0,24

3

0,24

Hermawan

12

0,43

0,18

1,10

2

0,18

lkah

13

0,40

0,17

1,12

2

0,17

Joko

14

0,29

0,01

1,17

2

0,01

Kamal

15

0,10

0,21

1,35

1

0,10

Kurnia

16

0,42

0,17

1,09

2

0,17

Lutfy

17

0,87

0,63

0,70

2

0,63

Mona

18

1,03

0,86

0,71

3

0,71

Nasir

19

0,37

0,18

1,16

2

0,18

Opta

20

0,06

0,30

1,46

1

0,06

Prabu

Berdasarkan hasil iterasi ke 1 maka diperoleh jumlah Cluster $C1 = 5$, $C2 = 12$ dan $C3 = 3$

Menentukan Centroid Baru

Nilai Centroid baru ini digunakan sebagai salah satu tahapan dalam perhitungan iterasi dari masing-masing cluster. Adapun hasil centroid baru seperti disajikan gambar 2.

Gambar 2. Proses Perhitungan Centroid Baru – Iterasi ke 2

Pada perhitungan iterasi ke 2 dilakukan dengan menghitung nilai parameter berdasarkan dibandingkan dengan centroid baru. Berikut ini penulis sajikan pada contoh baris pertama yaitu

$$\text{Cluster 1} = \sqrt{(0,24-0,056)^2 + ((0,50-0,167)^2) + ((0,40-0,0689)^2)} = 0,50$$

$$\text{Cluster 2} = \sqrt{(0,24-0,251)^2 + ((0,50-0,528)^2) + ((0,40-0,2643)^2)} = 0,14$$

$$\text{Cluster 3} = \sqrt{(0,24-0,888)^2 + ((0,50-0,833)^2) + ((0,40-0,7314)^2)} = 0,80$$

Setelah perhitungan centroid baru selesai, maka hasil dari iterasi ke 2 seperti tampil pada tabel 4.

Tabel 4. Hasil Perhitungan Iterasi ke 2

1

No

Cluster 1

Cluster 2

Cluster 3

Posisi Cluster

Jarak Terdekat

Pelanggan

1

0,50

0,14

0,80

2

0,14

Ahmad

2

0,17

0,34

1,17

1

0,17

Ardian

3

0,18

0,62

1,37

1

0,18

Bagas

4

0,05

0,43

1,22

1

0,05

Budianto

5

0,92

0,58

0,54

3

0,58

Cerly

6

0,43

0,25

0,84

2

0,25

Farid

7

1,33

0,97

0,32

3

0,32

Fitriani

8

0,35

0,19

1,03

2

0,19

Galih

9

0,50

0,25

1,01

2

0,25

Gading

10

0,63

0,25

0,75

2

0,25

Herianto

11

1,51

1,09

0,27

3

0,27

Hermawan

12

0,40

0,32

1,01

2

0,32

Ikah

13

0,38

0,15

0,98

2

0,15

Joko

14

0,26

0,24

1,04

2

0,24

Kamal

15

0,06

0,42

1,21

1

0,06

Kurnia

16

0,40

0,15

0,96

2

0,15

Lutfy

17

0,86

0,40

0,53

2

0,40

Mona

18

1,03

0,63

0,45

3

0,45

Nasir

19

0,36

0,17

1,02

2

0,17

Opta

20

0,07

0,50

1,32

1

0,07

Prabu

Berdasarkan hasil iterasi ke 1 maka diperoleh jumlah Cluster C1 = 5, C2 = 11 dan C3 = 4 (Iterasi ke 2 ini belum valid dengan dibandingkan hasil iterasi ke 1, untuk itu dilanjutkan perhitungan iterasi ke 3 dengan cara menentukan nilai centroid baru seperti disajikan pada gambar 3.

Gambar 3. Proses Perhitungan Centroid Baru – Iterasi ke 3

Pada perhitungan iterasi ke 2 dilakukan dengan menghitung nilai parameter berdasarkan dibandingkan dengan centroid baru. Berikut ini penulis sajikan pada contoh baris pertama yaitu

$$\text{Cluster 1} = \sqrt{(0,24-0,056)^2 + ((0,50-0,167)^2) + ((0,40-0,0689)^2)} = 0,50$$

$$\text{Cluster 2} = \sqrt{(0,24-0,200)^2 + ((0,50-0,515)^2) + ((0,40-0,2681)^2)} = 0,14$$

$$\text{Cluster 3} = \sqrt{(0,24-0,869)^2 + ((0,50-0,792)^2) + ((0,40-0,6044)^2)} = 0,72$$

43 Setelah perhitungan centroid baru selesai, maka hasil dari iterasi ke 3 seperti tampil pada tabel 4.

1 Tabel 4. Hasil Perhitungan Iterasi 3

No

Cluster 1

Cluster 2

Cluster 3

Posisi Cluster

Jarak Terdekat

Pelanggan

1

0,50

0,14

0,72

2

0,14

Ahmad

2

0,17

0,30

1,07

1

0,30

Ardian

3

0,18

0,59

1,27

1

0,18

Bagas

4

0,05

0,40

1,13

1

0,05

Budianto

5

0,92

0,63

0,41

3

0,63

Cerly

6

0,43

0,28

0,74

2

0,28

Farid

7

1,33

1,00

0,41

3

0,41

Fitriani

8

0,35

0,15

0,94

2

0,15

Galih

9

0,50

0,24

0,92

2

0,24

Gading

10

0,63

0,30

0,64

2

0,30

Herianto

11

1,51

1,13

0,39

3

0,39

Hermawan

12

0,40

0,28

0,94

2

0,28

lkah

13

0,38

0,10

0,89

2

0,10

Joko

14

0,26

0,20

0,95

2

0,20

Kamal

15

0,06

0,39

1,11

1

0,06

Kurnia

16

0,40

0,10

0,87

2

0,10

Lutfy

17

0,86

0,44

0,46

2

0,44

Mona

18

1,03

0,68

0,32

3

0,32

Nasir

19

0,36

0,13

0,92

2

0,13

Opta

20

0,07

0,47

1,22

1

0,07

Prabu

Berdasarkan hasil iterasi ke 1 maka diperoleh jumlah Cluster $C1 = 5$, $C2 = 11$ dan $C3 = 4$ (Iterasi ke 3 ini telah sama atau sesuai dengan iterasi ke 2) sehingga dengan demikian dataset ini telah valid.

Tabel 5. Hasil Akhir Klasterisasi Sagmentasi Pelanggan

No

No. ID

Nama Pelanggan

Alamat

Klaster

Keterangan

1

2

Ardian

Kaliangkrik

C1

Kurang Rajin Belanja

2

3

Bagas

Kajoran

C1

Kurang Rajin Belanja

3

4

Budianto

Dukun

C1

Kurang Rajin Belanja

4

15

Kurnia

Ngablak

C1

Kurang Rajin Belanja

5

20

Prabu

Ngluwar

C1

Kurang Rajin Belanja

No

No. ID

Nama Pelanggan

Alamat

Klaster

Keterangan

1

1

Ahmad

Dukun

C2

Cukup Rajin Belanja

2

6

Farid

Bandongan

C2

Cukup Rajin Belanja

3

8

Galih

Kaliangkrik

C2

Cukup Rajin Belanja

4

9

Gading

Ngablak

C2

Cukup Rajin Belanja

5

10

Herianto

Bandongan

C2

Cukup Rajin Belanja

6

12

Ikah

Salam

C2

Cukup Rajin Belanja

7

13

Joko

Bandongan

C2

Cukup Rajin Belanja

8

14

Kamal

Tegalrejo

C2

Cukup Rajin Belanja

9

16

Lutfy

Pakis

C2

Cukup Rajin Belanja

10

17

Mona

Sawangan

C2

Cukup Rajin Belanja

11

19

Opta

Pakis

C2

Cukup Rajin Belanja

No

No. ID

Nama Pelanggan

Alamat

Klaster

Keterangan

1

5

Cerly

Kajoran

C3

Sangat Rajin Belanja

2

7

Fitriani

Dukun

C3

Sangat Rajin Belanja

3

11

Hermawan

Secang

C3

Sangat Rajin Belanja

4

18

Nasir

Ngluwar

C3

Sangat Rajin Belanja

Berdasarkan hasil klasterisasi dari tabel 5 maka dapat diringkas bahwa C1 (Kurang Rajin Belanja) yaitu 4 orang sedangkan C2 (Cukup Rajin Belanja) ada 11 orang dan C3 Sangat Rajin Belanja yaitu 4 orang.

Implementasi Sistem

Pada Implementasi sistem merupakan bagian dari pengembangan sistem kedalam suatu program aplikasi. Sistem dibuat dengan menggunakan bahasa Pemrograman PHP dan Database MySQL. Adapun tampilan program Aplikasi dari Sistem Klasterisasi Sagmentasi Pelanggan dengan menggunakan Metode K-Means Clustering sebagai berikut:

Tampilan Halaman Utama

Program Aplikasi Klasterisasi Sagmentasi Pelanggan ini ketika dijalankan maka akan tampil halaman awal sistem seperti disajikan pada gambar 4.

Gambar 4. Tampilan Halaman Utama Aplikasi

Tampilan Data Pelanggan

Pada aplikasi ini memiliki menu untuk mengelola data pelanggan yang digunakan sebagai objek dalam klasterisasi sagmentasi pelanggan. Adapun tampilan data pelanggan di program aplikasi seperti disajikan pada gambar 5.

Gambar 5. Tampilan Data Pelanggan

Tampilan Halaman Login Sistem

36 Pada Aplikasi ini telah dilengkapi halaman login aplikasi yang digunakan untuk masuk kedalam sistem. Tampilan halaman login aplikasi seperti disajikan pada gambar 6.

Gambar 6. Tampilan Halaman Login Aplikasi

Tampilan Halaman Administrator

1 Apabila telah berhasil login, maka akan tampil langsung dihalaman dashbord sistem yang berfungsi untuk mengelola sistem secara keseluruhan, seperti disajikan pada gambar 7.

Gambar 7. Tampilan Halaman Administrator

Tampilan Halaman Input Data Pelanggan

1 Pada Aplikasi ini telah dilengkapi fitur untuk menginputkan data pelanggan kedalam sistem. Input data pelanggan ini berguna sebagai bagian data dasar untuk melakukan proses klasterisasi pada segmentasi pelanggan, adapun tampilan dari halaman input data pelanggan disajikan di gambar 8.

Gambar 8. Tampilan Halaman Input Data Pelanggan

1 Setelah data input data pelanggan selesai diinputkan kedalam aplikasi maka data pelanggan tersebut akan tersimpan kedalam sistem seperti daftar pelanggan pada gambar 9.

Gambar 9. Tampilan Daftar Data Pelanggan

Tampilan Halaman Setting Centroid

Pada sistem ini terdapat menu untuk memberikan nilai centroid awal yang digunakan sebagai pendukung dalam proses perhitungan algoritma k-means clustering. Tampilan setting nilai centroid awal seperti disajikan pada gambar 10.

Gambar 10. Tampilan Daftar Data Pelanggan

Proses Kluster Pelanggan dengan Algoritma K-Means Clustering

Pada menu inti dari aplikasi ini adalah proses clusterisasi dengan menggunakan metode Algoritma K-Means Clustering, caranya klik menu K-Means Clustering sehingga tampil digambar 11.

Gambar 11. Tampilan Halaman Proses Cluster Data Pelanggan

Setelah mengklik Proses maka secara otomatis aplikasi akan melakukan proses clusterisasi dengan perhitungan algoritma k-means cluster melalui beberapa iterasi seperti disajikan pada gambar 11.

Gambar 11. Tampilan Halaman Proses Iterasi Clustering

Hasil Clustering

Pada proses akhir dari sistem ini terbentuknya 3 kluster yaitu kategori obat ringan, kategori obat sedang dan kategori obat berat. Tampilan hasil kluster ini dibentuk berdasarkan proses perhitungan dengan algoritma k-means clustering seperti disajikan pada gambar 12.

Gambar 12. Tampilan Hasil Klusterisasi Data Pelanggan

Pengujian Sistem

1 Pengujian sistem merupakan salah satu tahapan dalam proses pengembangan perangkat lunak yang bertujuan untuk memastikan bahwa sistem telah berjalan sesuai dengan desain yang direncanakan. Tahapan ini difokuskan pada evaluasi menyeluruh terhadap sistem setelah seluruh komponen terintegrasi, guna menilai kelayakan sistem secara keseluruhan. Dalam penelitian ini, dilakukan tiga jenis pengujian, yaitu uji fungsionalitas sistem, uji validitas, dan uji silhouette coefficient.

Uji Fungsionalitas Sistem

1 Uji fungsionalitas sistem merupakan jenis pengujian yang bertujuan untuk memverifikasi bahwa setiap fitur dalam sistem beroperasi sesuai dengan spesifikasi yang telah ditentukan. Pengujian ini dilakukan untuk memastikan bahwa sistem dapat berjalan dengan baik tanpa mengalami kegagalan saat diimplementasikan, sebagaimana ditunjukkan pada tabel 6.

Tabel 6. Rekapitulasi Hasil Uji Fungsionalitas

3 No

Jenis Uji

Komponen Sistem yang diuji

Skenario Uji

Hasil yang diharapkan

Hasil yang dihasilkan

Satus Uji

3 Hasil Pengujian

1

Uji Normal

Form Login Admin

Masukan username dan password dan dengan benar

Tampil halaman menu utama admin

1 Muncul pesan "Login Sukses, Selamat Datang admin"

Tampil halaman admin

3 Normal

Diterima

Uji Salah

Form Login Admin

Masukan username dan password dan dengan salah

Muncul pesan kesalahan

Muncul pesan "Code Salah!"

Tidak masuk admin

Normal

Diterima

2

Uji Normal

4 Form Input Data Pelanggan

Masukan data obat secara lengkap dan benar

Data tersimpan dengan baik dan benar

Muncul pesan "Penyimpanan Berhasil"

Tersimpan dengan baik

Normal

Diterima

Uji Salah

Form Input Data Pelanggan

Masukan data obat secara tidak lengkap

Tidak bisa menyimpan

Tidak bisa disimpan

Normal

Diterima

3

Uji Normal

4 Form Input Data Pelanggan

Masukan data pelanggan secara lengkap dan benar

Data siswa tersimpan dengan baik dan benar

Muncul pesan "Penyimpanan Berhasil"

4 Tersimpan dengan baik

Normal

Diterima

Uji Salah

Form Input Data Pelanggan

Masukan data pelanggan secara tidak lengkap

Tidak bisa menyimpan

Tidak bisa disimpan

Normal

Diterima

4

Uji Normal

Form Input Centroid Awal

3 Masukan data Centroid Awal secara lengkap dan benar

Dataperiode tersimpan dengan baik dan benar

Muncul pesan "Penyimpanan Berhasil"

Tersimpan dengan baik

Normal

Diterima

Uji Salah

Form Input Centroid Awal

Masukan data Centroid Awal secara tidak lengkap

Tidak bisa menyimpan

Tidak bisa disimpan

Normal

Diterima

5

Uji Normal

Proses Cluster

3 Masukan data cluster secara lengkap dan benar

Data kuota tersimpan dengan baik dan benar

Muncul pesan "Penyimpanan Berhasil"

Tersimpan dengan baik

Normal

Diterima

Uji Salah

Proses Cluster

Masukan data cluster secara tidak lengtakap

Tidak bisa menyimpan

1 Tidak bisa disimpan

Normal

Diterima

Berdasarkan hasil pengujian uji fungsionalitas yang telah dilakukan maka dapat diambil kesimpulan bahwa seluruh scenario uji menghasilkan hasil "Diterima" artinya sistem telah diterima berfungsi dengan baik.

Uji Validitas

48 Pengujian Validitas digunakan untuk menguji sistem dengan cara membandingkan

1 antara hasil perhitungan manual dengan hasil dari program aplikasi. Tahap

Perhitungan Algoritma K-Means Clustering sebagai berikut:

Menentukan Dataset

Menentukan Nilai MIN dan MAX pada masing-masing Variabel berikut:

MIN => Jumlah Pembelian : 17500 || Jumlah Transaksi : 1 || Jumlah Barang : 5,5

MIN => Jumlah Pembelian : 312500 || Jumlah Transaksi : 7 || Jumlah Barang : 142

1 Menghitung Nilai Normalisasi

Berikut ini contoh perhitungan Normalisasi pada baris pertama.

Jumlah Pembelian = $(87,625-17500)/(7-1) = 0,24$

$$\text{Jumlah Transaksi} = (4-1)/(7-1) = 0,50$$

$$\text{Jumlah Barang} = (60-5,5)/(142-5,5) = 0,40$$

1 Adapun rekap proses perhitungan normalisasi ditunjukkan pada tabel 7.

Tabel 7. Data Normalisasi

No

Jumlah Pembelian

Jumlah Transaksi

Jumlah Barang

1

0,24

0,50

0,40

2

0,06

0,33

0,07

3

0,08

0,00

0,00

4

0,07

0,17

0,12

5

0,81

0,67

0,22

6

0,41

0,33

0,25

7

0,86

0,67

1,00

8

0,08

0,50

0,18

9

0,11

0,67

0,11

10

0,43

0,67

0,16

11

1,00

1,00

0,90

12

0,06

0,33

0,43

13

0,11

0,50

0,25

14

0,11

0,33

0,26

15

0,08

0,17

0,13

16

0,11

0,50

0,29

17

0,45

0,83

0,44

18

0,80

0,83

0,29

19

0,09

0,50

0,19

20

0,00

0,17

0,03

1 Pada pembahasan ini peneliti langsung berfokus pada hasil klasterisasi agar dapat lebih fokus pada hasil penelitian.

Hasil Klastering

1 Berdasarkan proses perhitungan algoritma k-means clustering untuk mengelompokan data pelanggan maka dapat direkap hasil clustering dari data pelanggan seperti disajikan di tabel 8.

Tabel 8. Hasil Klaster Data Pelanggan

No

Cluster 1

Cluster 2

Cluster 3

Posisi Cluster

Jarak Terdekat

Nama Pelanggan

1

0,50

0,14

0,72

2

0,14

Ahmad

2

0,17

0,30

1,07

1

0,30

Ardian

3

0,18

0,59

1,27

1

0,18

Bagas

4

0,05

0,40

1,13

1

0,05

Budianto

5

0,92

0,63

0,41

3

0,63

Cerly

6

0,43

0,28

0,74

2

0,28

Farid

7

1,33

1,00

0,41

3

0,41

Fitriani

8

0,35

0,15

0,94

2

0,15

Galih

9

0,50

0,24

0,92

2

0,24

Gading

10

0,63

0,30

0,64

2

0,30

Herianto

11

1,51

1,13

0,39

3

0,39

Hermawan

12

0,40

0,28

0,94

2

0,28

lkah

13

0,38

0,10

0,89

2

0,10

Joko

14

0,26

0,20

0,95

2

0,20

Kamal

15

0,06

0,39

1,11

1

0,06

Kurnia

16

0,40

0,10

0,87

2

0,10

Lutfy

17

0,86

0,44

0,46

2

0,44

Mona

18

1,03

0,68

0,32

3

0,32

Nasir

19

0,36

0,13

0,92

2

0,13

Opta

20

0,07

0,47

1,22

1

0,07

Prabu

Berdasarkan hasil klasterisasi dari implementasi algoritma k-means clustering maka dapat diambil kesimpulan bahwa C1 : Pelanggan Kurang Rajin Belanja = 5 data, C2 Pelanggan Cukup Rajin Belanja = 11 data dan Pelanggan Rajin Belanja = 4 data. Hasil cluster ini bila dibandingkan dengan hasil diprogram aplikasi dari gambar 12 diperoleh hasil yang sama, maka dengan demikian ini sistem ini dinyatakan valid.

1 Uji Silhouette Coefficient

Silhouette coefficient merupakan metrik evaluasi yang digunakan dalam analisis klaster untuk menentukan jumlah klaster yang optimal dalam proses data mining.

23 Metrik ini menilai seberapa baik suatu objek berada dalam klaster yang sesuai

dibandingkan dengan kedekatannya terhadap kluster lainnya. Rekapitulasi hasil perhitungan uji Silhouette coefficient disajikan pada tabel 9.

Tabel 9. Rekapitulasi Uji Silhouette Coefficient

No

Jumlah Pembelian

Jumlah Transaksi

Jumlah Barang

ai

bi

si

1

0,5034

0,1373

0,7249

0,4815

0,4552

0,9737

2

0,1667

0,3050

1,0750

0,5593

0,5155

0,9562

3

0,1815

0,5938

1,2729

0,7238

0,6827

0,9589

4

0,0531

0,4014

1,1264

0,5684

0,5269

0,9585

5

0,9190

0,6309

0,4051

0,6792

0,6517

0,9725

6

0,4342

0,2795

0,7366

0,5090

0,4834

0,9744

7

1,3293

0,9982

0,4149

0,8709

0,9142

1,0433

8

0,3522

0,1483

0,9404

0,5251

0,4803

0,9552

9

0,5041

0,2402

0,9183

0,5909

0,5542

0,9633

10

0,6337

0,3003

0,6354

0,5459

0,5232

0,9773

11

1,5116

1,1314

0,3884

0,9565

1,0105

1,0540

12

0,3965

0,2780

0,9424

0,5741

0,5390

0,9649

13

0,3808

0,0965

0,8907

0,5020

0,4560

0,9540

14

0,2589

0,2042

0,9529

0,5135

0,4720

0,9585

15

0,0637

0,3946

1,1142

0,5653

0,5242

0,9589

16

0,4030

0,0962

0,8739

0,5027

0,4577

0,9550

17

0,8556

0,4360

0,4565

0,5960

0,5827

0,9867

18

1,0252

0,6816

0,3246

0,7021

0,6771

0,9750

19

0,3582

0,1302

0,9239

0,5158

0,4708

0,9550

20

0,0704

0,4693

1,2168

0,6289

0,5855

0,9566

SC =>

0,9726

1 Dari hasil uji Silhouette Coefficient (SC) menghasilkan nilai SC rata-rata sebesar 0,97

1 (strong structure) artinya model struktur tersebut memiliki tingkat struktur yang kuat.

Kesimpulan

1 Berdasarkan hasil pembahasan yang telah yang dilakukan maka dapat ditarik kesimpulan Bahwa Hasil Penelitian ini adalah Aplikasi Pengelompokan Data Pelanggan dengan dengan Algoritma K-Means Clustering yang mampu mengelompokan tiga cluster yaitu pada cluster yaitu C1 : Pelanggan Kurang Rajin Belanja = 5 data, C2 Pelanggan Cukup Rajin Belanja = 11 data dan Pelanggan Rajin Belanja = 4 data. Hasil uji validitas telah dinyatakan valid karena berdasarkan perbandingan antara perhitungan manual dengan program aplikasi telah diperoleh hasil yang sama. Hasil pengujian sistem yang telah dilakukan pada uji fungsionalitas diperoleh diperoleh hasil "Diterima" pada seluruh skenario uji hal ini memberikan tanda bahwa sistem telah berfungsi dengan baik dan terakhir pada uji model Silhoutte Coeficient (SC) maka diperoleh hasil rata-rata yang baik yakni 0,97 hampir mendekati angka 1 sehingga sistem ini dinyatakan strong structure.

Ucapan Terima Kasih

1 Pada kesempatan ini Tim Peneliti mengucapkan terimakasih banyak pada Direktorat Penelitian dan Pengabdian Masyarakat (DPPM) di Lingkungan Kemendikti Saintek yang telah memberikan pendaan penelitian dari program HIBAH Dikti Tahun 2025, Pada Skema Penelitian Dosen Pemula (PDP), semoga hasil penelitian ini dapat memberikan manfaat untuk masyarakat.

Referensi

60 Oti EU, Olusola MO, Eze FC, Enogwe SU, "Comprehensive Review of K-Means Clustering Algorithms," *Int. J. Adv. Sci. Res. Eng.*, vol. 7, no. 8, pp. 64–69, 2021. [Online]. Available: <https://doi.org/10.31695/IJASRE.2021.34050>.

6 Oti EU, Olusola MO, Eze FC, Enogwe SU. Comprehensive Review of K-Means Clustering Algorithms. *Int J Adv Sci Res Eng.* 2021;07(08):64–9 Available from: <https://doi.org/10.31695/IJASRE.2021.34050>

2 Zhou Y, Wang R, Ding R, Shi D, Ye Q. Investigation on hierarchical control for driving stability and safety of intelligent HEV during car-following and lane-change process.

40 *Sci China Technol Sci* [Internet]. 2022;65(1):53–76. Available from: <https://doi.org/10.1007/s11431-021-1891-8>

11 Duarte V, Zuniga-Jara S, Contreras S. Machine Learning and Marketing: A Systematic Literature Review. *IEEE Access.* 2022;10(August):93273–88. Available from: <http://doi.org/10.1109/ACCESS.2022.3202896>

19 Vázquez XH. Alexandra França DOCTORAL DISSERTATION Adapting for Sustainable Success: Navigating Change in Strategic Management. 2023;. Available from: <https://www.investigacion.biblioteca.uvigo.es/xmlui/handle/11093/5523>

2 Zhang L, Lu Z. Advances in InSAR Imaging and Data Processing—A Review. *Remote Sens.* 2022;14(17):1–7. Available from: <https://doi.org/10.3390/rs14174307>

7 Sáez-de-Cámara X, Flores JL, Arellano C, Urbieta A, Zurutuza U. Clustered federated learning architecture for network anomaly detection in large scale heterogeneous IoT networks. *Comput Secur.* 2023;131. Available from: <https://doi.org/10.1016/j.cose.2023.103299>

2 Singh P, Khoshaim L, Nuwisser B, Alhassan I. How Information Technology (IT) Is Shaping Consumer Behavior in the Digital Age: A Systematic Review and Future Research Directions. *Sustain.* 2024;16(4). Available from:

<https://doi.org/10.3390/su16041556>

21 Hassan DO, Hassan BA. A comprehensive systematic review of machine learning in the retail industry: classifications, limitations, opportunities, and challenges. *Neural Comput Appl [Internet]*. 2025;37(4):2035–70. Available from:

<https://doi.org/10.1007/s00521-024-10869-w>

2 Dritsas E, Trigka M. Exploring the Intersection of Machine Learning and Big Data : A Survey. 2025. Available from: <https://doi.org/10.3390/make7010013>

25 Jabeen S. The Landscape of CRM Research: A Bibliometric Analysis of Key Trends and Future Directions. *J Relatsh Mark [Internet]*. 2025 :1–41. Available from:

<https://doi.org/10.1080/15332667.2025.2462884>

41 Zhou S, Bi S, Qi G. Rule Mining Trends from 1987 to 2022: A Bibliometric Analysis and Visualization. *Data Intell.* 2023;1–44. Available from:

https://doi.org/10.1162/dint_a_00239

2 Liu H, Luo Y, Geng J, Yao P. Research hotspots and frontiers of product R&D management under the background of the digital intelligence era-bibliometrics based on citespace and histcite. *Appl Sci.* 2021;11(15). Available from:

<https://doi.org/10.3390/app11156759>

5 Yuan J, Chen C, Yang W, Liu M, Xia J, Liu S. A survey of visual analytics techniques for machine learning. *Comput Vis Media.* 2021;7(1):3–36. Available from :

<http://doi.org/10.1007/s41095-020-0191-7>

2 Yang J, Wang YK, Yao X, Lin CT. Adaptive Initialization Method for K-Means

52 Algorithm. *Front Artif Intell.* 2021;4(November):1–13. Available from:

<https://doi.org/10.3389/frai.2021.740817>

Yang J, Rahardja S, Rahardja S, member student. Dimensional Outlier Detection.

10 2024; Available from: <https://doi.org/10.36227/techrxiv.171172948.85055456/v1>

2 Kumar A, Saini R, Kumar R. A Comparative Analysis of Machine Learning Algorithms for Breast Cancer Detection and Identification of Key Predictive Features. *Trait du*

58 *Signal.* 2024;41(1):127–40. Available from: <https://doi.org/10.18280/ts.410110>

2 Gallego V, Freixes A, Ligan J. Applying Machine Learning in Marketing: An Analysis

33 Using the NMF and K-Means Algorithms. *Lect Notes Comput Sci (including Subser*

62 *Lect Notes Artif Intell Lect Notes Bioinformatics)* . 2025;14778 LNCS:14–26. 2024.

Available from : <https://doi.org/10.3390/info15070368>

2 Wasilewski A. Customer segmentation in e-commerce: a context-aware quality

framework for comparing clustering algorithms. *J Internet Serv Appl.* 2024;15(1):160–

10 78. Available from: <https://doi.org/10.5753/jisa.2024.3851>

Takie A, Selmi E, Zerarka MF, Cheriet A. Ingénierie des Systèmes d ' Information

31 Enhancing K-Means Clustering with Post-Redistribution. 2024;29(2):429–36. Available

31 from: <https://doi.org/10.18280/isi.290204>

15 Nhat NM. Applied Density-Based Clustering Techniques for Classifying High-Risk

Customers: A Case Study of Commercial Banks in Vietnam. *J Appl Data Sci.*

12 2024;5(4):1639–53. Available from: <https://doi.org/10.47738/jads.v5i4.344>

54 Nugroho BI, Rafhina A, Ananda PS, Gunawan G. Journal of Intelligent Decision

Support System (IDSS) Customer segmentation in sales transaction data using k-

26 means clustering algorithm. J Intell Decis Support Syst. 2024;7(2):130–6. Available from: <https://doi.org/10.35335/idss.v7i2.236>

18 Grünewälder S. M Achine L Earning : W Hat Is M Achine L Earning ? 2019 Third Int
61 Conf I-SMAC (IoT Soc Mobile, Anal Cloud). 2015;11(113):13–21. Available from:
<https://doi.org/10.28945/5086>

16 Saxena A, Agarwal A, Pandey BK, Pandey D. Examination of the Criticality of
Customer Segmentation Using Unsupervised Learning Methods. Circ Econ Sustain
22 [Internet]. 2024;4(2):1447–60. Available from: <https://doi.org/10.1007/s43615-023-00336-4>

2 Hidayati S, Darmaliana AT, Riski R. Comparison of K-Means, Fuzzy C-Means, Fuzzy
Gustafson Kessel, and DBSCAN for Village Grouping in Surabaya Based on Poverty
22 Indicators. J Pendidik Mat. 2022;5(2):185, Available from:
<https://doi.org/10.1007/s43615-023-00336-4>

34 Ardana CH, Khoyum AAAAA, Faisal M. Segmentasi Pelanggan Penjualan Online
Menggunakan Metode K-means Clustering. JISKA (Jurnal Inform Sunan Kalijaga).
42 2024;9(1):1–9. Available from: <https://doi.org/10.14421/jiska.2024.9.1.1-9>

2 Firmansyah EB, Machado MR, Moreira JLR. How can Artificial Intelligence (AI) be
used to manage Customer Lifetime Value (CLV)—A systematic literature review. Int J
45 Inf Manag Data Insights [Internet]. 2024;4(2):100279. Available from:
<https://doi.org/10.1016/j.jjime.2024.100279>