

**ANALISIS PENGELOMPOKAN PERILAKU PENGGUNAAN
TRANSAKSI *CASHLESS* DI KALANGAN MAHASISWA
KOTA MEDAN MENGGUNAKAN ALGORITMA
*K-MEANS CLUSTERING***

SKRIPSI

DISUSUN OLEH

DESPITA SARAGIH

NPM. 2209010191



PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

2026

**ANALISIS PENGELOMPOKAN PERILAKU PENGGUNAAN
TRANSAKSI *CASHLESS* DI KALANGAN MAHASISWA
KOTA MEDAN, MENGGUNAKAN ALGORITMA
*K-MEANS CLUSTERING***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom) dalam Program Studi Sistem Informasi pada Fakultas
Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah
Sumatera Utara**

DESPITA SARAGIH

NPM. 2209010191

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

2026

LEMBAR PENGESAHAN

Judul Skripsi : Analisis Pengelompokan Perilaku Penggunaan
Transaksi Cashless di Kalangan Mahasiswa Kota
Medan Menggunakan Algoritma *K-Means*
Clustering

Nama Mahasiswa : Despita Saragih

NPM : 2209010191

Program Studi : Sistem Informasi

Menyetujui
Komisi Pembimbing

(Assoc. Prof. Dr. Al-Khowarizmi, M.Kom)
NIDN. 0127099201

Ketua Program Studi

(Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom)
NIDN. 0117088902



Dekan

(Assoc. Prof. Dr. Al-Khowarizmi, M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

**ANALISIS PENGELOMPOKAN PERILAKU PENGGUNAAN
TRANSAKSI CASHLESS DI KALANGAN MAHASISWA KOTA MEDAN
MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, Mei 2026

Yang membuat pernyataan



Despita Saragih

NPM. 2209010191

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Despita Saragih
NPM : 220910191
Program Studi : Sistem Informasi
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**ANALISIS PENGELOMPOKAN PERILAKU PENGGUNAAN
TRANSAKSI CASHLESS DI KALANGAN MAHASISWA KOTA MEDAN
MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, Mei 2026

Yang membuat pernyataan



Despita Saragih

NPM. 2209010191

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Despita Saragih
Tempat dan Tanggal Lahir : Pematang Siantar, 15 Desember 2003
Alamat Rumah : Buah Bolon, Merek Raya, Kec. Raya
Telepon/Faks/HP : 082363586125
E-mail : desvitasaragih15@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SD NEGERI 019321 MEREK RAYA TAMAT: 2016
SMP : SMP NEGERI 3 RAYA TAMAT: 2019
SMA : SMA NEGERI 1 RAYA TAMAT: 2022

KATA PENGANTAR



Puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah memberikan rahmat serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam semoga senantiasa tercurah kepada junjungan kita Nabi Besar Muhammad SAW yang telah membawa umat manusia dari zaman kegelapan menuju zaman yang penuh dengan ilmu pengetahuan seperti sekarang ini. Dalam penyusunan skripsi ini, penulis menyadari bahwa masih terdapat berbagai kekurangan. Namun demikian, skripsi ini dapat diselesaikan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S1) pada Program Studi Sistem Informasi di Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) Universitas Muhammadiyah Sumatera Utara. Skripsi ini berjudul —Analisis Pengelompokan Perilaku Penggunaan Transaksi Cashless di Kalangan Mahasiswa Kota Medan Menggunakan Algoritma K-Means Clustering dapat diselesaikan sesuai dengan harapan penulis.

Selama proses penyusunan skripsi ini, penulis menghadapi berbagai kesulitan dan tantangan. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan dukungan, bantuan, serta doa sehingga skripsi ini dapat diselesaikan dengan baik. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Akrim, M.Pd, selaku Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).

2. Bapak Assoc. Prof. Dr. Al-Khowarizmi, S.Kom., M.Kom, selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU sekaligus Dosen Pembimbing yang telah memberikan bimbingan, arahan serta masukan yang sangat berarti dalam penyusunan skripsi ini.
3. Ibu Dr. Firahmi Rizky, S.Kom., M.Kom, selaku Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
4. Bapak Mhd. Basri, S.Si., M.Kom, selaku Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Bapak Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom, selaku Ketua Program Studi Sistem Informasi.
6. Bapak Mulkan Azhari, S.Kom., M.Kom, selaku Sekretaris Program Studi Sistem Informasi.
7. Penulis menyampaikan terima kasih kepada seluruh dosen FIKTI yang telah memberikan ilmu, bimbingan, dan arahan selama masa perkuliahan. Setiap pengajaran, perhatian, dan nasihat yang diberikan sangat berarti dalam membentuk penulis hingga dapat menyelesaikan skripsi ini. Semoga segala kebaikan para dosen mendapat balasan yang terbaik dari Allah SWT.
8. Penulis menyampaikan rasa terima kasih yang sedalam-dalamnya kepada Ibu dan Ayah tercinta, yang telah membesarkan dan mendidik penulis dengan sepenuh hati sejak kecil hingga saat ini. Setiap langkah yang penulis tempuh adalah jejak dari kasih sayang dan ketabahan yang Ibu dan Ayah tanamkan sejak awal kehidupan. Terima kasih karena selalu mendukung segala kegiatan dan pilihan penulis. Dukungan yang Ibu dan

Ayah berikan baik berupa materi, tenaga, maupun perjuangan yang seringkali tak terlihat menjadi kekuatan yang tidak tergantikan dalam perjalanan ini. Ibu dan Ayah tidak pernah lelah memastikan setiap kebutuhan penulis terpenuhi, tanpa mengeluh dan tanpa meminta balasan. Doa yang Ibu dan Ayah panjatkan setiap hari menjadi pelindung dalam setiap langkah dan menjadi alasan penulis bisa tetap bertahan, berjuang, dan terus melangkah maju. Untuk itu, penulis memohon maaf apabila masih banyak harapan Ibu dan Ayah yang belum mampu penulis wujudkan sampai saat ini. Penulis sadar bahwa kebaikan dan pengorbanan Ibu dan Ayah tidak pernah bisa dibalas dengan apa pun, namun penulis berjanji akan terus berusaha menjadi anak yang dapat membanggakan. Semoga Allah SWT senantiasa melindungi Ibu dan Ayah, melimpahkan kesehatan, kebahagiaan, dan umur yang panjang. Semoga suatu hari nanti penulis diberikan kesempatan untuk membalas, meski hanya sebagian kecil, dari segala kebaikan, cinta, dan ketulusan yang telah Ibu dan Ayah berikan sepanjang hidup ini.

9. Penulis juga menyampaikan terima kasih yang sebesar-besarnya kepada kakak dan abang tercinta, yang selalu hadir sebagai tempat bercerita, berbagi pikiran, dan meminta saran di setiap masa sulit maupun bahagia. Terima kasih atas arahan, dukungan, dan perhatian yang tulus, yang seringkali menjadi penopang ketika penulis merasa ragu atau kehilangan arah. Semoga kasih sayang dan kebaikan kakak serta abang dibalas dengan keberkahan dan kebahagiaan oleh Allah SWT.

10. Penulis juga menyampaikan terima kasih kepada seluruh teman-teman seperjuangan, baik yang hadir dalam perjalanan perkuliahan, masa kecil, maupun dalam berbagai fase kehidupan lainnya. Terima kasih atas kebersamaan yang telah terukir menjadi kenangan berharga, atas tawa yang menguatkan, dan atas ruang cerita yang selalu terbuka kapan pun penulis membutuhkan tempat untuk berbagi. Kehadiran teman-teman sebagai pendengar, penyemangat, dan penopang dalam berbagai kesulitan menjadi bagian penting yang mengiringi proses penyelesaian skripsi ini. Terima kasih karena telah menemani dalam suka maupun duka, serta menjadi bagian dari perjalanan yang penuh warna. Semoga hubungan baik ini tetap terjaga, meski suatu saat masing-masing akan melangkah ke kesibukan dan jalan hidup yang berbeda. Semoga pertemanan ini terus menjadi pengingat bahwa penulis tidak pernah berjalan sendirian.
11. Penulis juga mengucapkan terima kasih kepada semua pihak yang tidak dapat disebutkan satu per satu, atas segala bantuan, dukungan, dan kebaikan yang telah diberikan selama proses penyusunan skripsi ini. Setiap kontribusi yang penulis terima sangat berarti dan menjadi bagian penting dalam tercapainya karya ini. Semoga seluruh kebaikan tersebut mendapatkan balasan terbaik dari Allah SWT.

**“ANALISIS PENGELOMPOKAN PERILAKU PENGGUNAAN
TRANSAKSI *CASHLESS* DI KALANGAN MAHASISWA
KOTA MEDAN, MENGGUNAKAN ALGORITMA
K-MEANS CLUSTERING”**

ABSTRAK

Penelitian ini bertujuan menganalisis pengelompokan perilaku penggunaan transaksi *cashless* khususnya *e-wallet* pada mahasiswa di Kota Medan menggunakan algoritma *K-Means Clustering*. Empat variabel numerik digunakan sebagai dasar pengelompokan, yaitu frekuensi penggunaan, nominal transaksi, kemudahan, dan kepercayaan terhadap keamanan *e-wallet*. Dua variabel nominal jenis aplikasi dan tujuan transaksi dimanfaatkan sebagai informasi pendukung dalam interpretasi hasil. Data diolah melalui sistem aplikasi berbasis *Python* yang mencakup tahap input data, pembersihan, normalisasi, proses clustering, visualisasi, serta penyimpanan hasil ke dalam database *MySQL*. Hasil penelitian menghasilkan tiga cluster utama: pengguna aktif dengan variasi aplikasi dan tujuan transaksi yang luas; pengguna frekuensi rendah yang cenderung memakai satu aplikasi dan bertransaksi untuk kebutuhan dasar; serta pengguna moderat dengan pola aplikasi dan transaksi yang bervariasi namun tidak dominan. Temuan ini memberikan gambaran mengenai pola penggunaan *e-wallet* di kalangan mahasiswa dan dapat menjadi referensi bagi institusi maupun penyedia layanan digital dalam merancang strategi edukasi keuangan yang lebih tepat.

Kata Kunci : *e-wallet*, transaksi *cashless*, *K-Means Clustering* , cluster, data mining.

**“ANALYSIS OF THE CLUSTERING OF CASHLESS TRANSACTION
USAGE BEHAVIOR AMONG UNIVERSITY STUDENTS IN MEDAN,
USING THE K-MEANS CLUSTERING ALGORITHM”**

ABSTRACT

This study aims to analyze the clustering of e-wallet usage behavior among college students in Medan using the K-Means Clustering algorithm. Four numerical variables were used as the basis for clustering: frequency of use, transaction amount, ease of use, and trust in e-wallet security. Two nominal variables—type of application and transaction purpose—were utilized as supporting information in interpreting the results. The data was processed through a Python-based application system that included data input, cleaning, normalization, clustering, visualization, and storage of results in a MySQL database. The research results yielded three main clusters: active users with a wide variety of applications and transaction purposes; low-frequency users who tend to use a single application and transact for basic needs; and moderate users with varied but not dominant application and transaction patterns. These findings provide an overview of e-wallet usage patterns among college students and can serve as a reference for institutions and digital service providers in designing more appropriate financial education strategies.

Keywords : *e-wallet, cashless transactions, K-Means Clustering, cluster, data mining.*

DAFTAR ISI

LEMBAR PENGESAHAN.....	i
PERNYATAAN ORISINALITAS.....	ii
KATA PENGANTAR.....	v
ABSTRAK.....	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
BAB II LANDASAN TEORI.....	7
2.1 Kajian Teori.....	7
2.1.1 Konsep E-Wallet.....	7
2.1.2 Perilaku Pengguna E-Wallet.....	8
2.1.3 Model Teori Perilaku Penggunaan Teknologi.....	8
2.1.4 Indikator Perilaku Penggunaan E-Wallet.....	11
2.2 Konsep Data Mining dan Analisis Clustering.....	12
2.2.1 Definisi Data Mining.....	12
2.2.2 Tahapan Data Mining (CRISP-DM Framework).....	12
2.2.3 Clustering dalam Data Mining.....	13
2.3 Algoritma K-Means Clustering.....	14
2.3.1 Evaluasi Algoritma K-Means Menggunakan WCV dan BCV.....	15

2.4 Unified Modeling Language (UML).....	16
2.5 Skala Pengukuran dan Kuesioner.....	18
2.6 Penelitian Terdahulu.....	19
2.7 Kerangka Berpikir Konseptual.....	21
BAB III ANALISA DAN PERANCANGAN SISTEM.....	24
3.1 Jenis Penelitian.....	24
3.2 Teknik Pengumpulan Data.....	24
3.2.1 Instrumen Pengumpulan Data.....	25
3.2.2 Populasi dan Sampel.....	25
3.2.3 Prosedur Pengumpulan Data.....	26
3.3 Lingkungan Pengembangan.....	27
3.3.1 Perangkat Lunak yang Digunakan.....	27
3.3.2 Jenis Laptop yang Digunakan.....	28
3.4 Strategi Pemecahan Masalah.....	28
3.5 Model Pengembangan Sistem.....	30
3.5.1 Class Diagram.....	30
3.5.2 Use Case Diagram.....	34
3.5.3 Activity Diagram.....	37
3.5.4 Sequence Diagram.....	38
3.5.5 Gambaran Umum dan Alur Kerja Sistem.....	39
3.6 Penerapan Algoritma K-Means Clustering.....	40
3.6.1 Data Penelitian.....	40
3.6.2 Tahapan Perhitungan K-Means Clustering.....	43
3.6.3 Iterasi dan Pembentukan Cluster.....	49
3.6.4 Evaluasi Cluster Menggunakan WCV dan BCV.....	50
3.7 Flowchart Penelitian.....	55

3.8 Rancangan Antarmuka Sistem.....	57
3.8.1 Prinsip Desain Antarmuka.....	57
3.8.2 Struktur Umum Antarmuka Sistem.....	58
3.9 Pengujian Sistem.....	61
3.9.1 Skenario dan Tahapan Pengujian Sistem.....	62
BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM.....	65
4.1 Implementasi Sistem.....	65
4.2 Tampilan Antarmuka Sistem.....	65
4.2.1 Halaman Utama Sistem.....	66
4.2.2 Halaman Home.....	67
4.2.3 Halaman Input Data.....	68
4.2.4 Halaman Pemrosesan Data.....	68
4.2.5 Halaman Proses Clustering.....	69
4.2.6 Halaman Hasil Clustering.....	70
4.2.7 Halaman Visualisasi Cluster.....	71
4.2.8 Halaman Cek Database.....	73
4.3 Hasil Proses Clustering.....	75
4.4 Analisis Karakteristik Cluster.....	77
4.4.1 Karakteristik Cluster 1.....	77
4.4.2 Karakteristik Cluster 2.....	78
4.4.3 Karakteristik Cluster 3.....	78
4.4.4 Informasi Pendukung Berdasarkan Indikator Nominal.....	79
4.5 Analisis Hasil Sistem.....	81
4.6 Kelebihan dan Keterbatasan Sistem.....	82
4.6.1 Kelebihan Sistem.....	82
4.6.2 Keterbatasan Sistem.....	83

BAB V PENUTUP.....	84
5.1 Kesimpulan.....	84
5.2 Saran.....	86
DAFTAR PUSTAKA.....	88
LAMPIRAN.....	92

DAFTAR TABEL

Tabel 2.1 Hubungan Model Teori dengan Penelitian.....	10
Tabel 2.2 Bentuk Skala Likert.....	18
Tabel 2.3 Tabel Penelitian Terdahulu.....	19
Tabel 3.1 Tabel Data Kuesioner.....	30
Tabel 3.2 Tabel Class User.....	31
Tabel 3.3 Tabel Class DataHandler.....	31
Tabel 3.4 Tabel Class KMeansProcessor.....	31
Tabel 3.5 Tabel Class DatabaseManager.....	32
Tabel 3.6 Tabel Class Visualizer.....	32
Tabel 3.7 Tabel Data Responden dalam Bentuk Numerik.....	42
Tabel 3.8 Hasil Perhitungan Awal (iterasi 1).....	45
Tabel 3.9 Tabel Data Responden Data 1.....	46
Tabel 3.10 Tabel Data Responden Data 2, 3, 4, 5, 6, 9, 10.....	47
Tabel 3.11 Tabel Data Responden Data 7 dan 8.....	47
Tabel 3.12 Centroid Baru.....	48
Tabel 3.13 Hasil Perhitungan iterasi 2.....	49
Tabel 3.14 Centroid Cluster Final.....	51
Tabel 3.15 WCV Per Cluster.....	52
Tabel 3.16 BCV Per Cluster.....	54
Tabel 3.17 Jarak Antar Centroid.....	54
Tabel 3.18 Tabel Pengujian Sistem (Black Box System).....	63
Tabel 4.1 Hasil Pengelompokan Data Responden.....	75
Tabel 4.2 Jumlah Anggota pada Setiap Cluster.....	76

DAFTAR GAMBAR

Gambar 2.1 Kerangka Berpikir Konseptual.....	22
Gambar 3.1 Class Diagram.....	33
Gambar 3.2 Use case Diagram.....	34
Gambar 3.3 Activity Diagram.....	37
Gambar 3.4 Sequence Diagram.....	38
Gambar 3.5 Alur Kerja Sistem.....	39
Gambar 3.6 Flowchart Tahapan Penelitian.....	55
Gambar 3.7 Rancangan Layar 1.....	59
Gambar 3.8 Rancangan Layar 2.....	59
Gambar 3.9 Rancangan Layar 3.....	60
Gambar 3.10 Rancangan Layar 4.....	60
Gambar 3.11 Rancangan Layar 5.....	61
Gambar 4.1 Halaman Utama Sistem.....	66
Gambar 4.2 Tampilan Halaman Home.....	67
Gambar 4.3 Tampilan Input Data.....	68
Gambar 4.4 Tampilan Prapemrosesan Data.....	69
Gambar 4.5 Tampilan Proses Clustering.....	70
Gambar 4.6 Tampilan Hasil Clustering.....	71
Gambar 4.7 Visualisasi Grafik Scatter Plot.....	72
Gambar 4.8 Visualisasi Bar Chart dan Interpretasi.....	72
Gambar 4.9 Visualisasi Profiling Cluster.....	73
Gambar 4.10 Tampilan Data yang Telah di Clustering pada Sistem.....	74
Gambar 4.11 Data Clustering yang Tersimpan pada phpMyAdmin.....	74

BAB I

PENDAHULUAN

1.1 Latar Belakang

Revolusi digital dalam bidang keuangan telah mendorong pergeseran besar dari sistem pembayaran konvensional menuju sistem pembayaran non-tunai (*cashless*). Secara global penggunaan digital *payments* terus meningkat seiring luasnya pemanfaatan *smartphone*, internet, dan aplikasi keuangan digital. Transaksi non-tunai mengalami pertumbuhan signifikan didorong oleh adopsi cepat berbagai aspek seperti pembayaran berbasis kartu, *e-wallet*, *mobile banking*, hingga transaksi berbasis kode QR yang kini menjadi bagian dari aktivitas ekonomi sehari-hari masyarakat (Ramli & Hamzah, 2021).

Di Indonesia, penerapan teknologi pembayaran non-tunai didorong oleh regulasi pemerintah seperti Gerakan Nasional Non-Tunai (GNNT) dan kemajuan infrastruktur digital. Salah satu implementasi penting adalah QRIS (*Quick Response Code Indonesian Standard*) yang menjadi instrumen pembayaran digital utama yang digunakan masyarakat dengan berbagai skala transaksi, dengan jumlah merchant mencapai 29,6 juta, dimana 92% di antaranya adalah UMKM (Bank Indonesia, 2023). Fenomena ini sejalan dengan “Blueprint Sistem Pembayaran Indonesia 2030” yang menunjukkan bahwa volume transaksi ritel digital (melalui kanal mobile dan internet) pada 2023 mencapai 25,3 miliar transaksi. Hal ini menegaskan bahwa preferensi masyarakat telah bergeser ke arah pembayaran elektronik (Bank Indonesia, 2019).

Di kalangan mahasiswa, khususnya generasi Z, adopsi pembayaran digital berlangsung lebih cepat dibandingkan kelompok usia lainnya. Mahasiswa dikenal

sebagai pengguna aktif teknologi digital dan cenderung memilih metode pembayaran non-tunai, terutama *e-wallet*, karena dinilai lebih praktis, cepat, dan sesuai dengan mobilitas serta gaya hidup mereka. Perkembangan ini sangat relevan di Kota Medan, sebuah kota besar dengan populasi mahasiswa yang tinggi dari berbagai perguruan tinggi. Fenomena penggunaan *e-wallet* di kalangan mahasiswa mulai menarik perhatian peneliti dan praktisi karena karakteristik unik generasi ini cenderung cepat mengadopsi teknologi digital. Peneliti lokal Medan menunjukkan adanya dampak penggunaan *e-wallet* mendorong perubahan pola konsumsi mahasiswa dan berpotensi meningkatkan perilaku konsumtif karena kemudahan transaksi dan promosi digital (Nababan et al., 2025).

Namun demikian, perilaku penggunaan *e-wallet* di kalangan mahasiswa tidak bersifat homogen. Terdapat perbedaan dalam frekuensi penggunaan, jenis aplikasi *e-wallet* yang digunakan, nilai transaksi, serta tujuan transaksi, baik untuk konsumsi harian, belanja daring, maupun kebutuhan akademik. Perbedaan pola penggunaan ini menunjukkan bahwa setiap mahasiswa memiliki karakteristik perilaku penggunaan *e-wallet* yang berbeda-beda.

Beberapa penelitian internasional menunjukkan bahwa faktor-faktor seperti persepsi kemudahan, manfaat, keamanan dan pengaruh sosial memengaruhi adopsi dompet digital di kalangan pengguna muda. Hal ini dijelaskan melalui kerangka *Technology Acceptance Model* (TAM) dan *Unified Theory of Acceptance and Use of Technology* (UTAUT) yang menekankan peran persepsi kemudahan dan kegunaan dalam membentuk perilaku pengguna teknologi (Lee & Thoo, 2022) ; (Goh et al., 2025)

Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada faktor-faktor yang memengaruhi niat atau adopsi penggunaan seperti literasi keuangan, persepsi risiko, serta kemudahan teknologi. Pendekatan penelitian tersebut umumnya menggunakan metode analisis pengaruh atau deskriptif, sehingga belum banyak peneliti yang melakukan pemetaan atau pengelompokan perilaku pengguna berdasarkan karakteristik penggunaan *e-wallet* itu sendiri. Dengan kata lain, masih terbatas penelitian yang mengelompokkan mahasiswa ke dalam segmen perilaku penggunaan *e-wallet*, seperti pengguna intensif, pengguna moderat, atau pengguna dengan tingkat pemanfaatan rendah.

Kesenjangan penelitian tersebut menunjukkan perlunya pendekatan analisis yang mampu membentuk segmentasi perilaku pengguna *e-wallet* secara lebih spesifik. Pengelompokan perilaku ini penting untuk memahami variasi pola penggunaan *e-wallet* di kalangan mahasiswa secara lebih mendalam. Oleh karena itu, penelitian ini menggunakan algoritma *K-Means Clustering* yang merupakan teknik dalam data mining, yang efektif dalam melakukan pengelompokan data berdasarkan tingkat kemiripan atribut numerik untuk mengelompokkan mahasiswa berdasarkan indikator perilaku penggunaan *e-wallet*.

Algoritma *K-Means* banyak digunakan dalam analisis segmentasi karena memiliki proses perhitungan yang relatif sederhana, efisien dalam mengolah data, serta mampu menghasilkan kelompok data yang memiliki karakteristik serupa dalam satu cluster. Dengan metode ini, perilaku penggunaan *e-wallet* mahasiswa dapat dianalisis dan dikelompokkan sehingga menghasilkan segmentasi pengguna berdasarkan pola transaksi yang dimiliki.

Hasil pengelompokan ini diharapkan dapat menghasilkan informasi mengenai karakteristik setiap kelompok pengguna, sehingga dapat menjadi dasar pemahaman perilaku finansial mahasiswa serta memberikan kontribusi informasi bagi pengembangan edukasi keuangan dan layanan pembayaran digital yang lebih tepat sasaran.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, permasalahan utama dalam penelitian ini adalah bagaimana mengelompokkan perilaku pengguna *e-wallet* pada mahasiswa Kota Medan.

1.3 Batasan Masalah

Untuk menjaga fokus penelitian, maka batasan masalah dari penelitian ini adalah sebagai berikut :

1. Penelitian difokuskan pada mahasiswa yang berdomisili atau sedang menempuh pendidikan di Kota Medan.
2. Objek penelitian dibatasi pada penggunaan *e-wallet* sebagai alat transaksi non-tunai, dan tidak mencakup mobile banking, kartu debit, atau kartu kredit.
3. Data diperoleh melalui kuesioner online berbasis skala Likert yang menggambarkan perilaku dan pola penggunaan *e-wallet*.
4. Metode analisis yang digunakan terbatas pada algoritma *K-Means Clustering*, tanpa membandingkan dengan metode *clustering* lainnya.
5. Variabel yang dianalisis meliputi frekuensi penggunaan, jenis *e-wallet* yang digunakan, tujuan transaksi, nominal transaksi, dan tingkat kepercayaan terhadap keamanan.

6. Penelitian bersifat kuantitatif deskriptif, sehingga tidak menelaah hubungan kausal atau pengaruh antar variabel.
7. Penelitian dibatasi pada periode pengambilan data tertentu, sehingga hasil analisis mencerminkan kondisi pada waktu tersebut

1.4 Tujuan Penelitian

Didasarkan pada rumusan masalah diatas, maka tujuan dari penelitian ini adalah untuk mengelompokkan perilaku penggunaan e-wallet pada mahasiswa Kota Medan menggunakan algoritma *K-Means Clustering* serta mendeskripsikan karakteristik dari setiap cluster yang terbentuk berdasarkan variabel frekuensi penggunaan, nominal transaksi, kemudahan penggunaan, dan kepercayaan terhadap keamanan.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian yang dilakukan saat ini diantaranya ialah:

1. Bagi Institusi Perguruan Tinggi

Penelitian ini dapat menjadi sumber informasi mengenai pola dan karakteristik perilaku penggunaan *e-wallet* di kalangan mahasiswa. Hasil penelitian ini berpotensi dimanfaatkan sebagai dasar dalam penyusunan program literasi keuangan digital, perencanaan kegiatan akademik berbasis teknologi finansial, serta pengembangan kurikulum terkait *data mining* dan analisis perilaku pengguna.

2. Bagi Keilmuan

Penelitian ini memberikan kontribusi bagi pengembangan ilmu Sistem Informasi, khususnya pada bidang data mining dan analisis perilaku pengguna. Penerapan algoritma *K-Means Clustering* dalam penelitian ini

dapat menjadi referensi metodologis bagi penelitian berikutnya yang mengkaji segmentasi pengguna layanan digital berbasis data numerik.

3. Bagi Penulis

Penelitian ini memberikan pengalaman langsung dalam pengembangan sistem analisis berbasis *K-Means Clustering*, mulai dari pengolahan data, implementasi algoritma, visualisasi hasil, hingga interpretasi cluster. Proses ini meningkatkan kemampuan penulis dalam analisis data, pemodelan algoritma, serta perancangan aplikasi berbasis web untuk mendukung penelitian ilmiah.

4. Bagi Masyarakat Umum

Hasil penelitian ini memberikan gambaran mengenai karakteristik perilaku penggunaan *e-wallet* di kalangan mahasiswa Kota Medan. Informasi ini dapat menjadi bahan pertimbangan bagi penyedia layanan keuangan digital, pengembang teknologi finansial, maupun lembaga terkait untuk merancang program edukasi, keamanan, dan fitur layanan yang lebih sesuai dengan kebutuhan pengguna muda.

BAB II

LANDASAN TEORI

2.1 Kajian Teori

2.1.1 Konsep *E-Wallet*

E-wallet atau dompet digital merupakan instrumen pembayaran berbasis elektronik yang digunakan untuk menyimpan data keuangan dan melakukan transaksi secara daring. Menurut (Bank Indonesia, 2024a), *e-wallet* adalah layanan elektronik yang berfungsi untuk menyimpan data instrumen pembayaran, antara lain uang elektronik dan/atau kartu pembayaran, yang dapat digunakan untuk melakukan transaksi pembayaran.

Secara umum, *e-wallet* memiliki beberapa komponen utama seperti saldo digital, QR code untuk pembayaran, fitur *top-up*, transfer antar pengguna, dan riwayat transaksi. Jenis-jenis *e-wallet* sendiri yang paling populer di Indonesia antara lain Dana, GoPay, OVO, ShopeePay, dan LinkAja.

Menurut Ramli & Hamzah (2021), perkembangan *e-wallet* di negara berkembang seperti Indonesia sangat dipengaruhi oleh penetrasi *smartphone*, kemudahan akses internet, dan strategi promosi agresif. Data (Bank Indonesia, 2024b) menunjukkan bahwa volume transaksi digital melalui *e-wallet* meningkat signifikan, mencerminkan perubahan preferensi masyarakat menuju sistem pembayaran non-tunai.

Selain manfaat seperti kemudahan, kecepatan, dan efisiensi, penggunaan *e-wallet* juga menghadapi tantangan seperti risiko keamanan data dan rendahnya literasi keuangan sebagian pengguna (Suryani et al., 2024). Oleh karena itu,

pemahaman terhadap perilaku pengguna menjadi penting untuk memastikan adopsi dan penggunaan yang bijak serta berkelanjutan.

2.1.2 Perilaku Pengguna *E-Wallet*

Perilaku pengguna *e-wallet* mengacu pada pola kebiasaan, motivasi, dan preferensi individu dalam menggunakan layanan keuangan digital. Faktor-faktor yang memengaruhi perilaku tersebut antara lain:

1. Kemudahan penggunaan (*Ease of Use*)
2. Kemanfaatan (*Usefulness*)
3. Kepercayaan (*Trust/Security*)
4. Promosi dan intensif (*Reward, Cashback*)
5. Pengaruh sosial (*Social Influence*)

Lee & Thoo (2022) menjelaskan bahwa kemudahan dan kemanfaatan merupakan faktor paling dominan dalam menentukan niat penggunaan *e-wallet*. Sedangkan Goh et al. (2025) menemukan bahwa faktor sosial dan pengalaman positif memperkuat niat berkelanjutan untuk terus menggunakan *e-wallet*, terutama di kalangan mahasiswa.

Generasi Z, yang merupakan pengguna utama *e-wallet*, menunjukkan perilaku konsumtif yang tinggi terhadap layanan digital karena nilai praktis dan kemudahan yang ditawarkan. Namun demikian, tingkat literasi keuangan yang rendah dapat berimplikasi pada kurangnya kontrol dalam penggunaan *e-wallet* (Ramadanti et al., 2021).

2.1.3 Model Teori Perilaku Penggunaan Teknologi

Teori perilaku penggunaan teknologi menjadi dasar untuk memahami adopsi dan penggunaan *e-wallet*. Beberapa model yang relevan antara lain :

1. *Theory of Planned Behavior* (TPB)

Diperkenalkan oleh Ajzen (1991), teori ini menyatakan bahwa niat berperilaku dipengaruhi oleh sikap, norma subjektif, dan persepsi kontrol perilaku. Dalam konteks ini, teori TPB menjelaskan mengapa mahasiswa niat untuk menggunakan *e-wallet*, baik karena sikap positif terhadap teknologi, pengaruh sosial, maupun faktor kontrol perilaku seperti akses internet dan keamanan.

2. *Technology Acceptance Model* (TAM)

Diperkenalkan oleh Davis (1989), model ini menjelaskan bahwa keputusan seseorang untuk menggunakan teknologi dipengaruhi oleh dua faktor utama, yaitu *Perceived Usefulness* (persepsi kegunaan) dan *Perceived Ease of Use* (persepsi kemudahan penggunaan). Dalam konteks *e-wallet*, mahasiswa akan lebih cenderung menggunakannya apabila aplikasi memberikan manfaat nyata (hemat waktu, mudah diakses) dan tidak memerlukan upaya besar untuk digunakan.

3. *Unified Theory of Acceptance and Use of Technology* (UTAUT)

Model UTAUT dikembangkan oleh Venkatesh et al. (2003), menggabungkan berbagai pendekatan sebelumnya untuk menjelaskan bagaimana faktor sosial dan kondisi fasilitasi memengaruhi niat serta kenyataan penggunaan teknologi. Model ini menekankan empat faktor utama, yaitu *Performance Expectancy*, *Effort Expectancy*, *Social Influence*, dan *Facilitating Conditions*. (Goh et al., 2025) menunjukkan bahwa UTAUT efektif untuk menjelaskan perilaku pengguna muda terhadap *e-wallet* karena mempertimbangkan aspek sosial dan dukungan teknologi.

Penelitian terbaru oleh (Paramita & Cahyadi, 2024) mengonfirmasi bahwa ekspektasi performa, kemudahan pengguna, dan dukungan sosial menjadi determinan utama dalam adopsi QRIS sebagai bagian dari sistem *e-wallet*. Artinya, mahasiswa yang menganggap *e-wallet* lebih praktis dan efisien cenderung lebih aktif menggunakannya dalam transaksi digital.

Tabel 2.1 Hubungan Model Teori dengan Penelitian

Teori	Faktor Utama	Fokus Penjelasan	Keterkaitan dengan <i>E-Wallet</i>
TPB	<i>Attitude, Norm, Control</i>	Sikap dan norma sosial	Sikap mahasiswa terhadap transaksi digital
TAM	<i>PU, PEOU</i>	Persepsi manfaat & kemudahan	Niat menggunakan <i>e-wallet</i>
UTAUT	<i>Expectancy, Influence, Conditions</i>	Harapan dan pengaruh sosial	Kinerja & akses terhadap <i>e-wallet</i>

Theory of Planned Behavior memberikan fondasi psikologis yang luas tentang perilaku manusia, sedangkan TAM dan UTAUT mengoperasionalkan konsep tersebut dalam konteks teknologi digital. Dalam penelitian ini, TAM dan UTAUT berperan sebagai jembatan yang menghubungkan teori perilaku umum (TPB) dengan konteks perilaku digital mahasiswa pengguna *e-wallet*.

Teori *Theory of Planned Behavior* (TPB), *Technology Acceptance Model* (TAM), dan *Unified Theory of Acceptance and Use of Technology* (UTAUT) dalam penelitian ini tidak digunakan untuk membangun model pengujian hubungan antar variabel, melainkan sebagai landasan konseptual dalam pemilihan indikator perilaku penggunaan *e-wallet*. Indikator-indikator yang digunakan dalam penelitian ini disarikan dari konstruk utama pada teori-teori tersebut.

2.1.4 Indikator Perilaku Penggunaan *E-Wallet*

Indikator perilaku penggunaan e-wallet dalam penelitian ini disusun berdasarkan sintesis dari teori TPB, TAM, dan UTAUT, serta disesuaikan dengan karakteristik penggunaan *e-wallet* pada mahasiswa. Indikator ini kemudian digunakan sebagai atribut dalam proses *clustering* menggunakan algoritma *K-Means*.

Perilaku pengguna *e-wallet* dapat diukur melalui enam dimensi utama, yaitu :

1. Frekuensi penggunaan : seberapa sering *e-wallet* digunakan dalam aktivitas transaksi.
2. Jenis aplikasi *e-wallet* : aplikasi yang paling sering digunakan (GoPay, OVO, Dana, ShopeePay, LinkAja, dll).
3. Tujuan transaksi : misalnya untuk konsumsi harian, belanja daring atau kebutuhan pendidikan.
4. Nominal transaksi : rata-rata jumlah transaksi yang digunakan melalui *e-wallet*.
5. Kemudahan penggunaan : persepsi kemudahan dalam memahami dan mengoperasikan aplikasi.
6. Kepercayaan terhadap keamanan : sejauh mana pengguna yakin terhadap keamanan data dan sistem aplikasi.

Indikator kemudahan dan kemanfaatan berakar dari konstruk TAM, sedangkan faktor kepercayaan dan pengaruh sosial diadaptasi dari model UTAUT. Indikator-indikator ini sejalan dengan variabel yang digunakan dalam berbagai penelitian perilaku pengguna *e-wallet* di Asia Tenggara.

2.2 Konsep *Data Mining* dan Analisis *Clustering*

2.2.1 Definisi *Data Mining*

Data Mining adalah proses menemukan pola atau pengetahuan dari kumpulan data besar menggunakan teknik statistik dan algoritma. Menurut Han et al. (2022), tujuan utama *data mining* adalah untuk mengekstraksi informasi tersembunyi dari data yang kompleks sehingga dapat digunakan untuk pengambilan keputusan. Dalam konteks penelitian sosial, *data mining* digunakan untuk mengidentifikasi pola perilaku pengguna berdasarkan kesamaan karakteristik.

2.2.2 Tahapan *Data Mining* (*CRISP-DM Framework*)

Metodologi *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) adalah metodologi standar dalam proyek *data mining* yang digunakan secara luas oleh para praktisi analisis data untuk menyelesaikan permasalahan dan menemukan pengetahuan dari data besar secara sistematis dan terstruktur. *CRISP-DM* merupakan kerangka kerja standar yang terorganisasi untuk pelaksanaan proses ekstraksi pengetahuan dari data (Muis et al., 2025).

CRISP-DM terdiri dari enam fase utama yang secara berulang dapat memberikan kerangka kerja untuk masalah *data mining* (Mauritius & Binsar, 2020), yaitu :

1. *Business Understanding*, tahap awal yang berfokus pada pemahaman tujuan penelitian dan perumusan masalah dari sudut pandang bisnis/tujuan penelitian data.

2. *Data Understanding*, fase ini mencakup pengumpulan data awal, deskripsi data, pemeriksaan kualitas data, dan eksplorasi data untuk memahami karakteristik awal yang akan dianalisis.
3. *Data Preparation*, tahapan pembersihan data (*data cleaning*), transformasi, dan pemilihan fitur yang akan digunakan dalam pemodelan sehingga siap diproses.
4. *Modeling*, yaitu tahapan pemilihan teknik atau pemodelan data yang tepat sesuai kebutuhan, seperti *K-Means Clustering* dalam penelitian ini, serta penyesuaian parameter model.
5. *Evaluation*, tahapan menilai apakah model yang dibangun mampu memenuhi tujuan awal atau perlu perbaikan lebih lanjut.
6. *Deployment*, tahapan penerapan model terhadap lingkungan nyata atau penyajian hasil sehingga dapat digunakan dalam pengambilan keputusan.

Beberapa keunggulan CRISP-DM, diantaranya :

- a. Bersifat netral alat, tidak bergantung pada *software* tertentu.
- b. Struktur prosesnya terbukti banyak digunakan dalam riset dan industri *data mining*.
- c. Dapat diterapkan pada berbagai jenis masalah (*Clustering*, Klasifikasi, segmentasi data, dan sebagainya).

Kerangka ini memastikan penelitian berjalan sistematis dan terarah hingga hasilnya dapat diinterpretasikan dengan baik.

2.2.3 Clustering dalam Data Mining

Clustering merupakan metode *unsupervised learning* yang berfungsi untuk mengelompokkan objek data ke dalam kelompok (*cluster*) berdasarkan kemiripan

antar data. Penelitian sebelumnya menunjukkan bahwa algoritma *K-Means Clustering* efektif digunakan untuk mengelompokkan perilaku penggunaan *e-wallet* berdasarkan karakteristik tertentu sehingga dapat memberikan gambaran segmentasi pengguna yang jelas (Amin et al., 2025). Jenis-jenis *clustering* antara lain :

- a. *Partitioning Method (K-Means)*
- b. *Hierarchical Clustering*
- c. *Density-Based Clustering (DBSCAN)*

2.3 Algoritma *K-Means Clustering*

Algoritma *K-Means* adalah metode *clustering* yang mengelompokkan data ke dalam K kelompok berdasarkan jarak terdekat dengan titik pusat (*centroid*). Proses ini berulang hingga posisi *centroid* stabil atau tidak berubah secara signifikan. Menurut Ikotun et al. (2023), *K-Means* unggul karena kesederhanaannya, kecepatan proses, dan kemampuan menangani data jumlah besar. Namun, algoritma ini sensitif terhadap pemilihan jumlah *cluster* (K) dan keberadaan outlier.

Langkah-langkah utama Algoritma *K-Means* :

1. Tentukan jumlah *cluster* (K).
2. Pilih titik *centroid* awal secara acak.
3. Hitung jarak setiap data ke *centroid* menggunakan *Euclidean Distance*.
4. Kelompokkan data *centroid* terdekat.
5. Perbaharui *centroid* hingga tidak berubah.

Rumus jarak *Euclidean* yang digunakan adalah :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

Keterangan :

$d(x, y)$ = jarak euclidean antara titik x dan y

x_i = nilai atribut data

y_i = nilai atribut centroid

n = jumlah atribut yang digunakan

Setelah jarak dihitung, data dikelompokkan ke *centroid* terdekat. Nilai *centroid* kemudian diperbarui menggunakan persamaan :

$$C_{baru} = \frac{\sum X}{n} \quad (2.2)$$

di mana :

$\sum X$ = jumlah nilai variabel dalam satu *cluster*

n = jumlah data dalam *cluster* tersebut

Proses ini dilakukan berulang kali (iterasi) hingga posisi *centroid* tidak lagi berubah, menandakan bahwa hasil *cluster* telah stabil.

2.3.1 Evaluasi Algoritma K-Means Menggunakan WCV dan BCV

Dalam analisis *clustering*, evaluasi kualitas cluster merupakan aspek penting untuk memastikan bahwa kelompok yang terbentuk benar-benar merepresentasikan pola data secara akurat. Pada algoritma *K-Means*, terdapat dua ukuran utama yang digunakan dalam menilai kualitas cluster, yaitu *Within-Cluster Variation* (WCV) dan *Between-Cluster Variation* (BCV).

Within-Cluster Variation (WCV) merupakan ukuran yang menggambarkan seberapa rapat anggota suatu cluster terhadap centroid-nya. Semakin kecil nilai WCV, semakin homogen anggota cluster, yang berarti pembentukan cluster memiliki kualitas yang baik. WCV dihitung menggunakan rumus berikut :

$$WCV = \sum_{k=1}^K \sum_{i \in C_k} \left\| x_i - \mu_k \right\|^2 \quad (2.3)$$

dengan C_k adalah himpunan anggota cluster ke-k, x_i adalah data ke-i, dan μ_k adalah centroid cluster ke-k.

Sementara itu, *Between-Cluster Variation* (BCV) mengukur variasi antar centroid cluster dengan membandingkannya terhadap rata-rata global seluruh data. BCV yang tinggi menunjukkan bahwa masing-masing cluster berada pada posisi yang cukup terpisah, sehingga perbedaan karakteristik antar cluster semakin jelas. Rumus BCV dinyatakan sebagai:

$$BCV = \sum_{k=1}^K n_k \cdot \left\| \mu_k - \mu \right\|^2 \quad (2.4)$$

dengan n_k adalah jumlah anggota cluster ke-k dan μ merupakan centroid global seluruh dataset.

WCV dan BCV digunakan secara bersamaan untuk menilai kualitas model clustering. Model dikatakan baik apabila memiliki nilai WCV yang rendah dan nilai BCV yang tinggi, yang menunjukkan bahwa anggota cluster kompak dan antar cluster terpisah dengan jelas.

2.4 Unified Modeling Language (UML)

Unified Modeling Language merupakan bahasa pemodelan standar yang digunakan untuk memvisualisasikan, merancang, dan mendokumentasikan sistem

perangkat lunak secara terstruktur. UML menyediakan notasi grafis yang memungkinkan pengembang sistem untuk menggambarkan struktur sistem, interaksi antara pengguna dan sistem, serta alur proses yang terjadi di dalam sistem secara jelas dan sistematis (Maylawati et al., 2018).

Penggunaan UML dalam perancangan sistem informasi bertujuan untuk membantu proses analisis dan desain sistem sebelum tahap implementasi dilakukan. Dengan adanya pemodelan menggunakan UML, sistem yang akan dibangun dapat dipahami secara konseptual sehingga meminimalkan kesalahan pada tahap pengembangan dan mempermudah proses dokumentasi sistem (Narulita et al., 2024).

Beberapa diagram UML yang umum digunakan antara lain *use case diagram*, *class diagram*, *activity diagram*, dan *sequence diagram*. *Use case diagram* digunakan untuk menggambarkan interaksi antara aktor dan sistem, *class diagram* digunakan untuk memodelkan struktur kelas dan hubungan antar kelas dalam sistem, *activity diagram* digunakan untuk menggambarkan alur aktivitas atau proses bisnis yang terjadi di dalam sistem, sedangkan *sequence diagram* digunakan untuk menggambarkan interaksi antar objek dalam suatu sistem secara berurutan.

Dalam penelitian ini, UML digunakan sebagai alat bantu pemodelan sistem pengelompokan perilaku pengguna *e-wallet* berbasis algoritma *K-Means Clustering*. Pemodelan dilakukan untuk memberikan gambaran umum mengenai fungsi sistem, interaksi pengguna dengan sistem, serta alur kerja sistem secara keseluruhan sebelum sistem diimplementasikan.

2.5 Skala Pengukuran dan Kuesioner

Skala Likert merupakan alat ukur psikometrik yang digunakan untuk mengukur sikap, pendapat, atau persepsi seseorang terhadap suatu fenomena sosial berdasarkan pernyataan yang diberikan. Konsep ini pertama kali dikembangkan oleh Rensis Likert (1932) untuk memudahkan kuantifikasi pendapat responden dalam bentuk angka yang dapat dianalisis secara statistik.

Penelitian ini menggunakan Skala Likert (1-5) untuk mengukur persepsi responden terhadap indikator perilaku pengguna *e-wallet*. Skala ini umum digunakan dalam penelitian perilaku karena mampu menggambarkan tingkat persetujuan secara numerik. Huh & Gim (2025) menegaskan bahwa data Likert dapat diperlakukan sebagai variabel kontinu dalam analisis parametrik, termasuk untuk proses pengelompokan berbasis algoritma seperti *K-Means*.

Setiap pertanyaan diukur menggunakan skala 5 tingkat penilaian, yaitu :

Tabel 2.2 Bentuk Skala Likert

Skor	Kategori Jawaban
1	Sangat Tidak Setuju (STS)
2	Tidak Setuju (TS)
3	Netral (N)
4	Setuju (S)
5	Sangat Setuju (SS)

Skala likert dalam penelitian ini digunakan untuk memperoleh data kuantitatif terkait perilaku dan persepsi mahasiswa terhadap penggunaan *e-wallet*, khususnya pada indikator kemudahan penggunaan dan kepercayaan terhadap

keamanan transaksi. Kedua indikator tersebut diukur menggunakan pernyataan dengan lima alternatif jawaban; yaitu Sangat Tidak Setuju hingga Sangat Setuju.

Sementara itu, indikator perilaku penggunaan *e-wallet* lainnya, seperti frekuensi penggunaan, tujuan transaksi, jenis aplikasi *e-wallet*, dan nominal transaksi, diukur menggunakan pertanyaan berbentuk pilihan kategori yang mempresentasikan kondisi aktual responden. Data dari indikator tersebut selanjutnya dikonversi ke dalam bentuk numerik agar dapat diolah menggunakan algoritma *K-Means Clustering*.

Dengan pendekatan tersebut, data yang diperoleh mencerminkan baik aspek persepsi maupun perilaku nyata pengguna *e-wallet*, sehingga memungkinkan proses pengelompokan pengguna berdasarkan kemiripan karakteristik penggunaan secara lebih akurat.

2.6 Penelitian Terdahulu

Penelitian terdahulu berfungsi sebagai acuan dan pembanding dalam penelitian ini.

Tabel 2.3 Tabel Penelitian Terdahulu

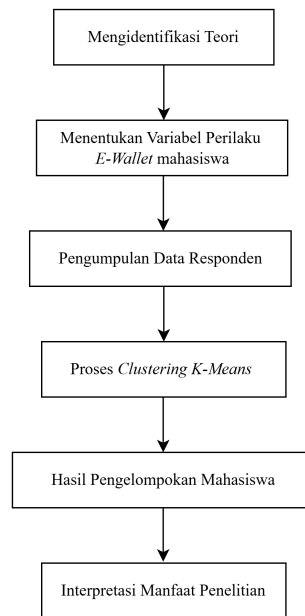
No	Peneliti & Tahun	Judul Penelitian	Tujuan & Metode	Hasil Utama
1	B. Bongado, M. Bernadeth, & A.R. Magallanes (2025)	<i>Influence of Digital Wallets on the Financial Behavior of HEI's Students</i>	Meneliti pengaruh penggunaan <i>e-wallet</i> terhadap perilaku keuangan mahasiswa perguruan tinggi (HEI) menggunakan analisis <i>cluster</i> dan regresi.	Penggunaan <i>e-wallet</i> berpengaruh positif terhadap pengelolaan keuangan mahasiswa.
2	Fatkhul Amin, Fithri	<i>Cluster Analysis of DANA and</i>	Menganalisis segmentasi	<i>K-Means</i> mampu

	Widyanita Yarisma, Ira Setiawati, & Ratih Milati Ilham (2025)	<i>GoPay E-Wallet Service Reviews using K-Means Algorithm</i>	pengguna <i>e-wallet</i> menggunakan <i>K-Means Clustering</i>	membentuk <i>cluster</i> pengguna <i>e-wallet</i> berdasarkan karakteristik layanan.
3	SN. Tazkia, R. Nurcahyo, & M. Habiburrahman (2024)	Segmentasi QRIS Users Using Cluster Analysis in Indonesia	Menggunakan <i>cluster analysis (K-Means)</i> untuk mengelompokkan pengguna QRIS berdasarkan perilaku transaksi dan karakteristik pengguna.	Terbentuk tiga kelompok pengguna yaitu intensif, moderat dan ringan.
4	Sukaris Sukaris, Wiki Renedi, Maulidyah Amalina Rizqi, & Budiyono Pristyadi (2021)	<i>Usage Behavior on Digital Wallet: Perspective of the Theory of UTAUT Models</i>	Analisis perilaku penggunaan <i>e-wallet</i> berbasis UTAUT	faktor kepercayaan dan kemudahan memengaruhi penggunaan.
5	G. Ilieva, T. Yankova, Y. Dzhabarova, & M. Ruseva (2023)	<i>Customer Attitude Toward Digital Wallet Services</i>	Menganalisis sikap pengguna terhadap layanan <i>e-wallet</i> menggunakan SEM dan analisis cluster.	Ditemukan empat kelompok pengguna berdasarkan sikap dan kepuasan layanan digital wallet.
6	A. Suyanto & D.W. Pawestri (2025)	<i>Positioning Analysis of Electronic Wallet Services in Indonesia Based on E-Service Quality</i>	Mengukur persepsi pengguna terhadap kualitas layanan <i>e-wallet</i> dengan pendekatan kuantitatif berbasis survei Likert.	DANA, OVO, dan GoPay menjadi <i>e-wallet</i> paling dominan di Indonesia.
7	M. Abdurahman	<i>An Empirical Analysis of</i>	Menggunakan <i>cluster analysis</i>	Terbentuk tiga <i>cluster</i>

	(2023)	<i>Electronic Payment System (EPS) Among Banking Customers in Kerala</i>	untuk mengidentifikasi kelompok pengguna sistem pembayaran elektronik.	berdasarkan intensitas penggunaan dan persepsi keamanan.
8	SE Maya Sari (2025)	<i>Global Trends in Mobile Payment Adoption: A Systematic Literature Review with Insights for Indonesia</i>	Menyusun tinjauan pustaka global mengenai adopsi mobile payment dan e-wallet di berbagai negara.	Menunjukkan bahwa faktor kemudahan dan literasi digital menjadi kunci adopsi.
9	R. Singhi, V. Sancheti & V. Chaturvedi (2024)	<i>Segmentation of M-Wallet Users Using Cluster Analysis: An Empirical Study in Delhi NCR</i>	Mengelompokkan pengguna m-wallet berdasarkan perilaku dan persepsi mereka menggunakan K-Means Clustering	Tiga segmen ditemukan: <i>heavy users</i> , <i>convenience seekers</i> , dan <i>occasional users</i> .

2.7 Kerangka Berpikir Konseptual

Penelitian ini didasarkan pada teori TAM dan UTAUT yang menjelaskan faktor-faktor perilaku pengguna teknologi. Indikator perilaku mahasiswa digunakan sebagai dasar dalam proses pengelompokan (*clustering*) menggunakan algoritma *K-Means*.



Gambar 2.1 Kerangka Berpikir Konseptual

Keterangan :

1. Identifikasi Teori Perilaku Penggunaan Teknologi

Penelitian ini berlandaskan teori TAM (Davis, 1989), TPB (Ajzen, 1991), dan UTAUT (Venkatesh et al., 2003) yang menjelaskan faktor-faktor yang memengaruhi perilaku individu dalam menggunakan teknologi, seperti kemanfaatan, kemudahan, dan pengaruh sosial.

2. Penentuan Variabel Perilaku Pengguna *E-Wallet*

Berdasarkan teori dan penelitian terdahulu, ditetapkan enam indikator perilaku mahasiswa, yaitu frekuensi penggunaan, jenis aplikasi yang sering digunakan, nominal transaksi, tujuan transaksi, kepercayaan terhadap keamanan, dan kemudahan penggunaan, yang akan menjadi variabel input dalam analisis.

3. Pengumpulan dan Pengolahan Data

Data dikumpulkan melalui kuesioner skala likert (1-5) melalui tahap pembersihan, normalisasi, dan input ke perangkat analisis untuk memastikan data siap digunakan dalam proses *clustering*.

4. Analisis Data (Algoritma *K-Means Clustering*)

Data diolah menggunakan algoritma *K-Means Clustering*, yang mengelompokkan mahasiswa ke dalam beberapa *cluster* berdasarkan tingkat kemiripan perilaku penggunaan *e-wallet* menggunakan jarak Euclidean Distance.

5. Hasil *Clustering*

Analisis menghasilkan beberapa kelompok pengguna, yaitu pengguna intensif, pengguna moderat, dan pengguna ringan yang mencerminkan variasi perilaku penggunaan *e-wallet* di kalangan mahasiswa.

6. Interpretasi dan Manfaat Penelitian

Hasil pengelompokan diinterpretasikan untuk memahami karakteristik tiap segmen dan digunakan sebagai dasar dalam edukasi keuangan digital, pengembangan strategi layanan *e-wallet*, serta kontribusi akademik dalam penelitian perilaku teknologi finansial.

BAB III

ANALISA DAN PERANCANGAN SISTEM

3.1 Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif deskriptif dengan pendekatan *data mining*, yang bertujuan untuk melakukan pengelompokan (*clustering*) terhadap perilaku mahasiswa dalam menggunakan transaksi *cashless*. Pendekatan *data mining* dipilih karena memungkinkan peneliti untuk menemukan pola tersembunyi di dalam data, khususnya pola perilaku penggunaan *e-wallet* yang tidak dapat diamati secara langsung.

Metode utama yang digunakan adalah algoritma *K-Means Clustering*, yaitu algoritma *unsupervised learning* yang mampu mengelompokkan data berdasarkan tingkat kesamaan perilaku transaksi. Hasil dari pengelompokan ini akan menunjukkan segmen mahasiswa berdasarkan intensitas, nominal, dan pola penggunaan transaksi non-tunai, termasuk *e-wallet* dan QRIS.

3.2 Teknik Pengumpulan Data

Pengumpulan data merupakan tahapan krusial dalam penelitian karena tujuan utamanya adalah memperoleh informasi yang diperlukan. Proses ini dapat dilakukan di berbagai lingkungan, menggunakan berbagai sumber, dan beragam metode pengumpulan.

Pengumpulan data dalam penelitian ini dilakukan dengan pendekatan kuantitatif survei menggunakan kuesioner daring yang disebarakan kepada mahasiswa di beberapa perguruan tinggi di Kota Medan. Teknik ini dipilih karena efisien, mampu menjangkau banyak responden secara cepat, dan relevan dengan karakter digital native generasi mahasiswa.

Adapun jenis data yang digunakan dalam penelitian ini adalah data primer, yaitu data yang diperoleh langsung dari hasil penyebaran kuesioner yang diisi oleh mahasiswa Kota Medan. Kuesioner ini disusun menggunakan platform Google Form, berisi pertanyaan tertutup dengan skala Likert lima poin (1-5).

3.2.1 Instrumen Pengumpulan Data

Instrumen utama dalam penelitian ini adalah kuesioner terstruktur, yang disusun berdasarkan teori perilaku penggunaan teknologi (TAM dan UTAUT). Setiap indikator variabel diukur menggunakan skala Likert lima poin, yang dinilai efektif dalam menggambarkan persepsi dan perilaku responden secara kuantitatif (Rokeman, 2024 ; Huh & Gim, 2025). Kuesioner terdiri dari beberapa bagian berdasarkan indikator perilaku pengguna *e-wallet*.

3.2.2 Populasi dan Sampel

Populasi penelitian ini adalah mahasiswa aktif di Kota Medan. Karena jumlah populasi tidak diketahui secara pasti, teknik pengambilan sampel menggunakan non-probability sampling dengan metode *purposive sampling*, yaitu pemilihan responden berdasarkan kriteria :

- a. Mahasiswa aktif di perguruan tinggi negeri maupun swasta di Kota Medan
- b. Pernah menggunakan transaksi *cashless* minimal 3 bulan terakhir.
- c. Bersedia mengisi kuesioner secara lengkap.

Ukuran sampel minimal dihitung menggunakan rumus Slovin :

$$n = \frac{N}{1+N(e)^2} \quad (3.1)$$

dengan :

n = jumlah sampel,

N = ukuran populasi

e = tingkat kesalahan (ditetapkan 10%)

Jika jumlah mahasiswa Kota Medan diasumsikan sekitar 100.000 orang, maka :

$$n = \frac{100.000}{1+100.000(0,1)^2} = 100$$

Sehingga jumlah responden minimal yang dibutuhkan adalah 100 mahasiswa, namun untuk meningkatkan validitas, peneliti menargetkan 150-200 responden.

3.2.3 Prosedur Pengumpulan Data

Terdapat lima poin dalam prosedur pengumpulan data, yaitu :

1. Penyusunan Instrumen

Kuesioner disusun berdasarkan teori TAM dan UTAUT agar indikator pengukuran sesuai dengan konsep perilaku pengguna teknologi.

2. Validasi Ahli (*Expert Judgment*)

Instrumen kuesioner divalidasi melalui *expert judgement* oleh dosen atau pakar terkait sebelum disebarkan ke responden, untuk menilai kelayakan, relevansi, dan kejelasan item pernyataan (*content validity*).

3. Penyebaran Kuesioner

Dilakukan secara online melalui media sosial atau grup mahasiswa agar dapat menjangkau responden dari berbagai universitas di Medan.

4. Rekapitulasi dan Pembersihan Data

Hasil kuesioner dikumpulkan dalam file *Excel*, lalu dilakukan *data cleaning* untuk menghapus data ganda atau kosong sebelum dianalisis.

3.3 Lingkungan Pengembangan

Lingkungan yang dipakai untuk mengembangkan program meliputi jenis komputer dan perangkat lunak apa saja yang dipakai serta sistem operasi apa yang digunakan.

3.3.1 Perangkat Lunak yang Digunakan

Perangkat lunak yang dipakai untuk menjalankan algoritma *K-Means Clustering* :

a. *Python* versi 3.12.10

Digunakan untuk mengimplementasikan sistem *clustering* berbasis algoritma *K-Means* secara mandiri sebagai bahasa pemrograman utama. *Streamlit* sebagai *framework* untuk membangun antarmuka sistem, sedangkan beberapa pustaka pendukung seperti *Pandas* dan *Library scikit-learn* digunakan untuk pengolahan data dan penerapan algoritma *K-Means Clustering*.

b. *Microsoft Excel 2021*

Dimanfaatkan untuk melakukan pembersihan data (*data cleaning*) dan konversi format sebelum data dianalisis. *Excel* juga digunakan untuk tabulasi dan penyajian statistik deskriptif awal.

c. *Google Form*

Digunakan untuk pengumpulan data kuesioner secara daring.

d. XAMPP dan phpMyAdmin

XAMPP digunakan sebagai server lokal yang menyediakan layanan database MySQL. phpMyAdmin digunakan sebagai antarmuka untuk mengelola database yang menyimpan data hasil clustering. Hasil

pengelompokan yang dihasilkan oleh sistem akan secara otomatis disimpan ke dalam database sehingga dapat diakses kembali untuk keperluan analisis maupun dokumentasi data.

3.3.2 Jenis Laptop yang Digunakan

Perangkat keras yang dipakai untuk menjalankan algoritma *K-Means Clustering* memiliki spesifikasi yaitu :

- a. Processor AMD Ryzen 5 5600H with Radeon Graphics (3.30 GHz)
- b. RAM 8 GB
- c. Storage 512GB SSD
- d. Sistem operasi Windows 11 Home 64-bit

3.4 Strategi Pemecahan Masalah

Strategi pemecahan masalah dalam penelitian ini disusun untuk menjawab tujuan penelitian, yaitu mengelompokkan perilaku penggunaan *e-wallet* pada mahasiswa menggunakan algoritma *K-Means Clustering*. Strategi yang digunakan mengacu pada konsep *data mining* dengan pendekatan *CRISP-DM (Cross Industry Standard Process for Data Mining)*.

Tahapan strategi pemecahan masalah dalam penelitian ini meliputi beberapa langkah sebagai berikut :

1. Pemahaman Masalah (*Business Understanding*)

Tahap awal dilakukan dengan mengidentifikasi permasalahan penelitian, yaitu belum adanya pengelompokan yang jelas terkait perilaku penggunaan *e-wallet* pada mahasiswa. Pada tahap ini juga ditentukan tujuan penelitian yaitu membentuk *cluster* pengguna *e-wallet* berdasarkan karakteristik perilaku penggunaan.

2. Pemahaman Data (*Data Understanding*)

Pada tahap dilakukan pengumpulan data melalui kuesioner yang disebarakan kepada mahasiswa.

3. Persiapan Data (*Data Preparation*)

Data hasil kuesioner selanjutnya melalui proses pembersihan data (*data cleaning*) untuk menghilangkan data tidak lengkap atau tidak valid. Selanjutnya dilakukan transformasi data ke dalam bentuk numerik agar dapat diproses menggunakan algoritma *K-Means*.

4. Pemodelan (*Modeling*)

Tahap pemodelan dilakukan dengan menerapkan algoritma *K-Means Clustering* untuk membentuk kelompok perilaku pengguna *e-wallet*. Proses *clustering* dilakukan secara manual pada sebagian data sebagai ilustrasi untuk tahapan algoritma, sedangkan proses pengelompokan terhadap keseluruhan data dilakukan menggunakan sistem clustering yang dikembangkan dalam penelitian ini.

5. Evaluasi (*Evaluation*)

Hasil pengelompokan yang diperoleh selanjutnya dievaluasi untuk memastikan bahwa *cluster* yang terbentuk memiliki perbedaan karakteristik yang jelas dan sesuai dengan tujuan penelitian.

6. Penyajian Hasil (*Deployment*)

Tahap akhir berupa penyajian hasil pengelompokan dalam bentuk visualisasi dan interpretasi perilaku.

3.5 Model Pengembangan Sistem

Model pengembangan sistem pada penelitian ini digunakan untuk menggambarkan bagaimana sistem analisis pengelompokan penggunaan *e-wallet* dibangun secara konseptual dan terstruktur. Pendekatan digunakan adalah *Unified Modeling Language* (UML), yang dipilih karena mampu merepresentasikan struktur data, fungsi sistem, serta alur proses secara jelas dan sistematis.

Penggunaan UML juga sejalan dengan tujuan penelitian yang tidak hanya berfokus pada analisis *data mining*, tetapi juga pada perancangan sistem pendukung analisis tersebut. Pemodelan sistem ini menggunakan *Class Diagram*, *Use Case Diagram* dan *Activity Diagram*, *Sequence Diagram* serta penjelasan desain sistem secara konseptual untuk memberikan gambaran menyeluruh mengenai cara kerja sistem.

Tabel 3.1 Tabel Data Kuesioner

Variabel	Tipe Data	Contoh Nilai
Frekuensi Penggunaan	Numerik	4
Nominal Transaksi	Numerik	3
Jenis <i>E-Wallet</i>	Kategorikal	GoPay
Tujuan Transaksi	Kategorikal	Transportasi
Kepercayaan Keamanan	Numerik	5
Kemudahan Penggunaan	Numerik	4

3.5.1 *Class Diagram*

Class Diagram digunakan untuk menggambarkan struktur statis data dan hubungan antar kelas dalam sistem analisis pengelompokan penggunaan *e-wallet*. Diagram ini merepresentasikan entitas utama yang terlibat dalam sistem beserta

atribut-atribut yang dimilikinya. *Class Diagram* juga menjadi dasar dalam perancangan basis data dan implementasi sistem menggunakan bahasa pemrograman *Python*.

Pada sistem ini, terdapat beberapa kelas utama, yaitu :

a. *Class User*

Memiliki tanggung jawab untuk mengelola proses autentikasi pengguna.

Tabel 3.2 Tabel *Class User*

Method	Keterangan
login()	Melakukan verifikasi username dan password sebelum mengakses sistem

b. *Class DataHandler*

Mengelola data kuesioner yang diunggah dan melakukan prapemrosesan.

Tabel 3.3 Tabel *Class DataHandler*

Method	Keterangan
uploadData()	Membaca file CSV/XLSX dari pengguna
cleanData()	Menghapus duplikasi dan merapikan data
normalizeData()	Melakukan normalisasi menggunakan Min-Max Scaler

c. *Class KMeansProcessor*

Melakukan seluruh perhitungan algoritma K-Means.

Tabel 3.4 Tabel *Class KMeansProcessor*

Method	Keterangan
initializeCentroids()	Menentukan centroid awal
computeDistances()	Menghitung jarak Euclidean setiap data ke centroid

assignClusters()	Menentukan cluster berdasarkan jarak terdekat
updateCentroids()	Memperbarui centroid hingga konvergensi

d. *Class DatabaseManager*

Mengelola penyimpanan hasil cluster ke dalam database (MySQL/phpMyAdmin).

Tabel 3.5 Tabel *Class DatabaseManager*

Method	Keterangan
saveResult()	Menyimpan hasil clustering
loadResult()	membaca data hasil clustering sebelumnya

e. *Class Visualizer*

Menampilkan hasil visualisasi cluster.

Tabel 3.6 Tabel *Class Visualizer*

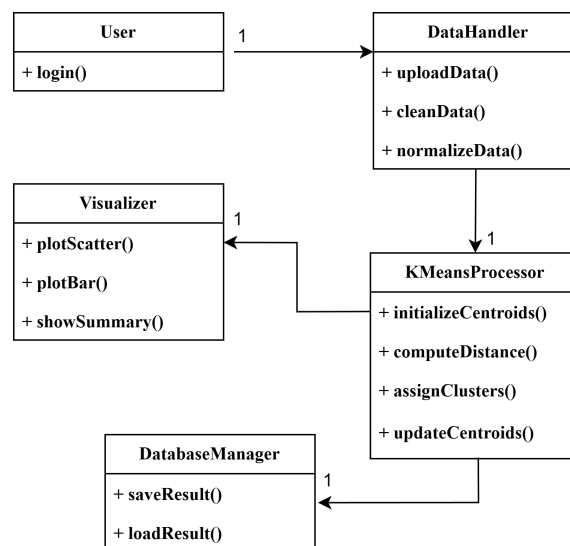
Method	Keterangan
plotScatter()	Membuat visualisasi scatter plot
plotBar()	Menampilkan jumlah anggota per cluster
showSummary()	Menampilkan ringkasan cluster (tabel rata-rata)

Relasi antarclass pada sistem ini menunjukkan alur kerja logis dalam proses pengolahan data, penerapan algoritma *K-Means*, hingga penyimpanan dan visualisasi hasil. *Class User* berperan sebagai pemicu utama proses, di mana pengguna melakukan tindakan seperti mengunggah data kuesioner dan memulai analisis. Ketika pengguna mengunggah data, *class User* berasosiasi dengan *class*

DataHandler sebagai komponen yang bertugas membaca, membersihkan, dan melakukan normalisasi terhadap data mentah.

Setelah proses prapemrosesan selesai, *DataHandler* meneruskan data yang telah siap analisis ke *class KMeansProcessor* melalui relasi asosiasi. *Class* ini memiliki tanggung jawab besar dalam proses inti sistem, yaitu melakukan inisialisasi centroid, menghitung jarak antar data, menentukan keanggotaan cluster, memperbarui centroid, hingga menentukan kondisi konvergensi.

Hasil pengelompokan yang dihasilkan oleh *KMeansProcessor* kemudian dikirimkan ke *class DatabaseManager* melalui relasi asosiasi. *Class* ini bertugas menyimpan hasil *clustering* ke dalam basis data *MySQL* untuk kebutuhan pencatatan dan pelacakan riwayat analisis. Selain itu, *KMeansProcessor* juga memiliki relasi dependency terhadap *class Visualizer*, karena nilai cluster yang dihasilkan menjadi input bagi class tersebut untuk menampilkan grafik sebaran, diagram batang jumlah anggota cluster, serta tabel karakteristik cluster.

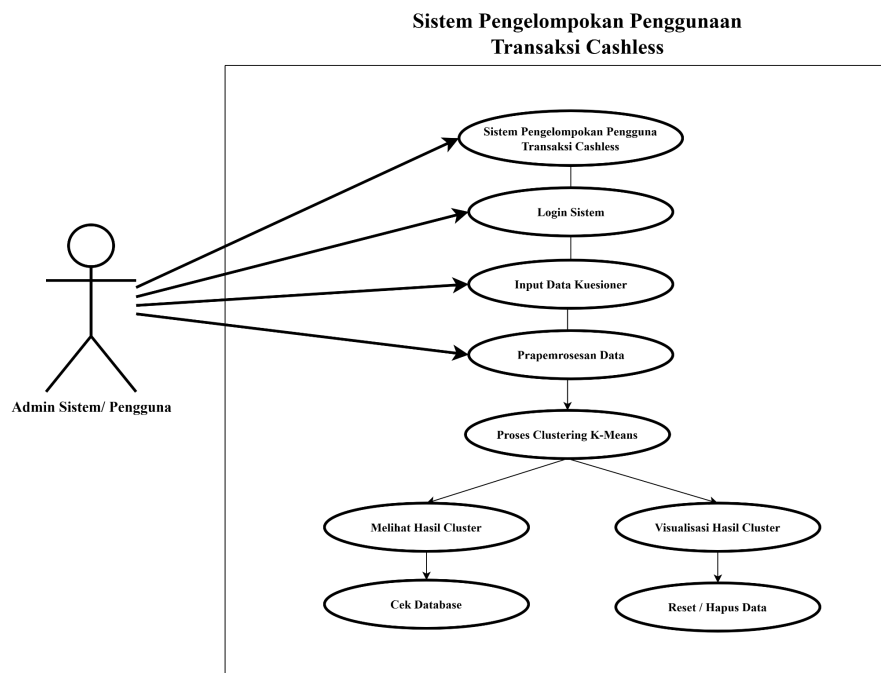


Gambar 3.1 *Class Diagram*

Dengan demikian, keseluruhan relasi antarclass pada sistem ini membentuk alur kerja terstruktur mulai dari input data, prapemrosesan, proses *K-Means*, penyimpanan hasil, hingga visualisasi, yang menunjukkan keterhubungan fungsional setiap komponen sesuai perannya dalam sistem.

3.5.2 Use Case Diagram

Use Case Diagram digunakan untuk menggambarkan interaksi antara pengguna sistem (Admin/Peneliti) dengan seluruh fungsi yang tersedia dalam aplikasi sistem analisis pengelompokan perilaku penggunaan *e-wallet*. Diagram ini menunjukkan fungsi-fungsi utama yang dapat dilakukan oleh pengguna serta batasan sistem secara umum.



Gambar 3.2 Use case Diagram

Aktor utama dalam sistem ini adalah Admin/Pengguna Sistem, yaitu pihak yang memiliki hak akses penuh untuk mengunggah data, menjalankan proses

analisis, serta melihat hasil pengelompokan. *Use Case* utama yang terdapat dalam sistem meliputi :

a. Login sistem

Pengguna harus melakukan autentikasi terlebih dahulu untuk dapat mengakses seluruh fitur aplikasi. Proses login memastikan bahwa hanya pihak yang berwenang yang dapat mengolah dan melihat data penelitian.

b. Input data kuesioner

Use case ini memungkinkan pengguna untuk mengunggah data responden dalam format *CSV* atau *Excel*. Data yang diunggah akan ditampilkan dalam bentuk tabel sebagai langkah awal sebelum dilakukan pemrosesan lebih lanjut.

c. Pra pemrosesan data

Tahap ini berfungsi untuk menyiapkan data agar siap digunakan dalam proses *clustering*. Prapemrosesan terdiri dari dua fungsi inti:

1. *Cleaning data* : menghapus data duplikat, menormalkan format jawaban, serta memastikan validitas data responden.
2. Normalisasi data : melakukan transformasi *range* nilai menggunakan *Min-Max Normalization* agar seluruh variabel memiliki skala yang seragam

d. Proses *Clustering K-Means*

Use case ini melakukan perhitungan algoritma K-Means mulai dari pembentukan centroid awal, perhitungan jarak Euclidean, penentuan cluster, hingga pembaruan centroid pada iterasi berikutnya.

Proses clustering juga mencakup penyimpanan hasil akhir ke database untuk memastikan data tersimpan sebagai riwayat analisis.

e. Melihat hasil cluster

Pengguna dapat melihat hasil akhir pengelompokan dalam bentuk tabel berisi nilai variabel (X1–X4) serta label cluster. Fitur ini membantu analisis perilaku mahasiswa berdasarkan hasil segmentasi yang terbentuk.

f. Visualisasi hasil cluster

Use case ini menampilkan visualisasi berupa *Scatter Plot*, *Bar Chart* distribusi anggota cluster dan Grafik karakteristik cluster. Visualisasi ini membantu peneliti dalam memahami pola data dan karakteristik antar cluster secara lebih intuitif.

g. Cek Database

Fitur ini memungkinkan pengguna melihat seluruh data hasil clustering yang tersimpan di *MySQL (phpMyAdmin)*. Riwayat analisis dapat ditelusuri kembali tanpa harus melakukan proses clustering ulang.

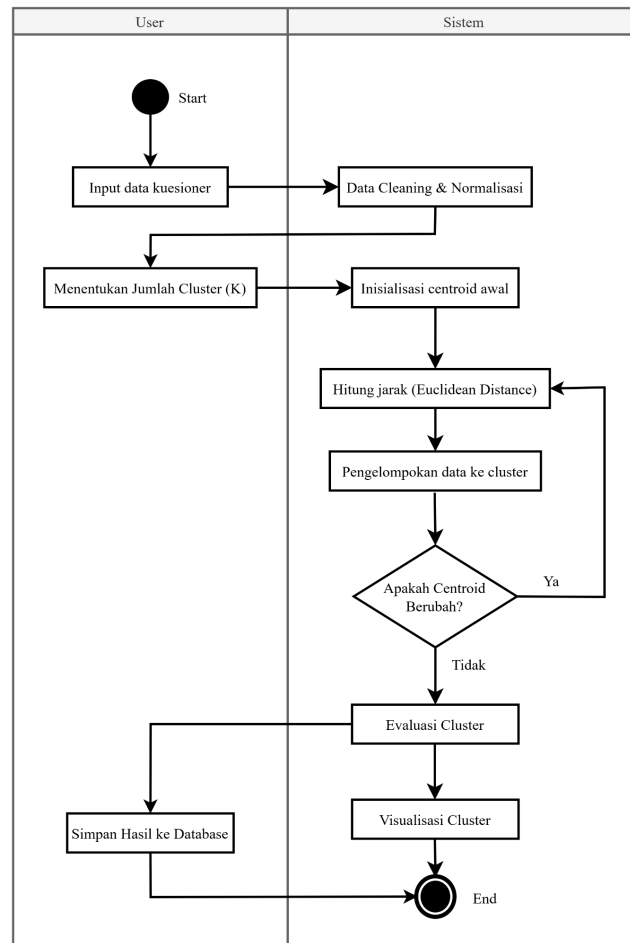
h. Reset / Hapus Data

Pengguna dapat membersihkan seluruh data baik dalam session state aplikasi maupun dalam database. *Use case* ini digunakan apabila pengguna ingin memulai proses analisis dari awal.

Use Case Diagram memberikan gambaran menyeluruh mengenai fungsi-fungsi yang tersedia dalam sistem serta hubungan antar proses. Melalui diagram ini, batasan sistem, peran pengguna, serta alur interaksi dapat dipahami dengan jelas sehingga mempermudah tahap implementasi dan pengembangan sistem secara keseluruhan.

3.5.3 Activity Diagram

Activity Diagram digunakan untuk menggambarkan alur aktivitas atau proses yang terjadi di dalam sistem secara berurutan. Diagram ini menjelaskan tahapan proses analisis pengelompokan penggunaan *e-wallet* mulai dari input data hingga interpretasi hasil *clustering*.



Gambar 3.3 *Activity Diagram*

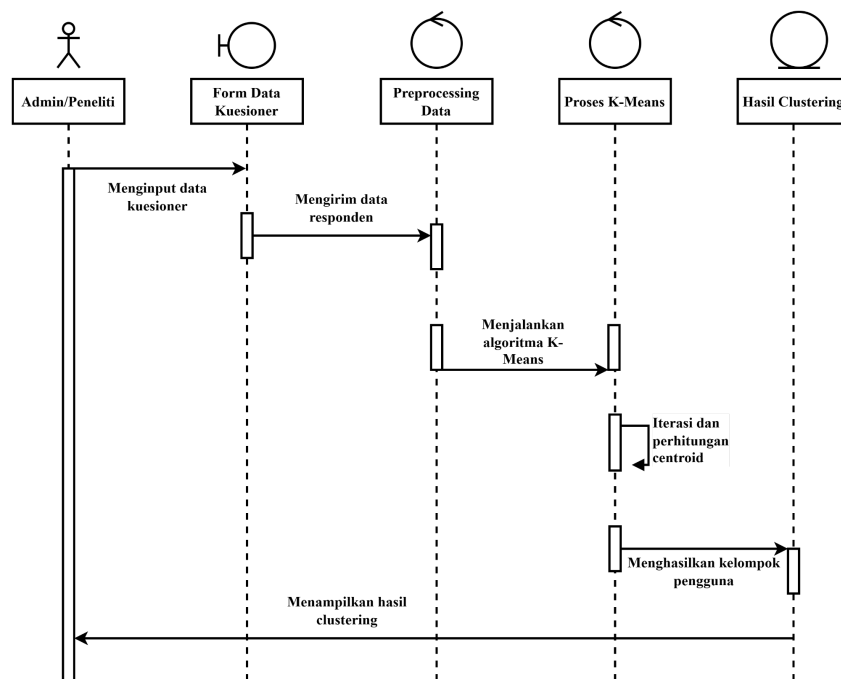
Alur aktivitas dalam sistem dimulai dari proses input data kuesioner mahasiswa, kemudian dilanjutkan dengan tahapan pra-pemrosesan data yang mencakup pembersihan data dan normalisasi. Setelah data siap, sistem menentukan jumlah *cluster* (K) dan melakukan inisialisasi *centroid* awal. Selanjutnya, sistem menghitung jarak data ke masing-masing *centroid*

menggunakan metode *Euclidean Distance* dan mengelompokkan data ke *centroid* terdekat.

Proses ini dilakukan secara iteratif hingga *centroid* tidak mengalami perubahan. Setelah proses *clustering* selesai, sistem melakukan hasil *cluster* dan menampilkan visualisasi serta interpretasi hasil pengelompokan. *Activity Diagram* ini membantu menjelaskan alur kerja sistem secara logis dan mudah dipahami.

3.5.4 Sequence Diagram

Sequence Diagram digunakan untuk menggambarkan urutan interaksi antar objek dalam sistem berdasarkan waktu (*sequence*). Diagram ini memperlihatkan bagaimana pesan dikirim dan diterima antar objek ketika suatu proses dijalankan.



Gambar 3.4 Sequence Diagram

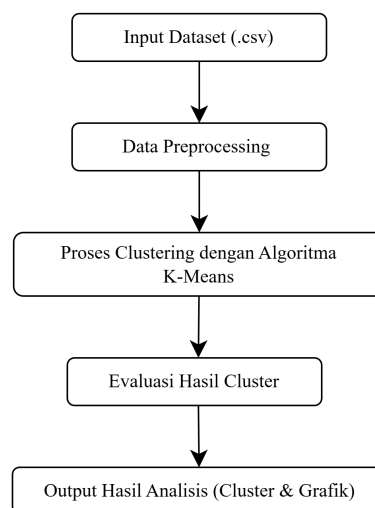
Pada penelitian ini, *sequence diagram* menggambarkan urutan proses ketika pengguna menjalankan analisis *clustering*, mulai dari input data,

pemanggilan proses *clustering*, perhitungan algoritma *K-Means*, hingga sistem menghasilkan dan menampilkan hasil cluster. *Sequence diagram* memberikan gambaran detail mengenai aliran komunikasi antar komponen sistem secara kronologis.

Penggunaan *sequence diagram* dalam penelitian ini bertujuan untuk memperjelas dinamika proses sistem dan memastikan bahwa alur komunikasi antar objek telah sesuai dengan rancangan sistem.

3.5.5 Gambaran Umum dan Alur Kerja Sistem

Sistem analisis pengelompokan perilaku penggunaan transaksi *cashless* dengan algoritma *K-Means Clustering* dirancang sebagai sistem pendukung analisis yang membantu peneliti dalam mengelompokkan mahasiswa berdasarkan pola penggunaan *e-wallet*. Sistem ini menerima data hasil kuesioner mahasiswa sebagai input utama kemudian diolah untuk menghasilkan informasi berupa kelompok perilaku penggunaan *e-wallet* sebagai output.



Gambar 3.5 Alur Kerja Sistem

Secara umum, alur kerja sistem dimulai dari proses input data kuesioner yang dilakukan oleh admin atau peneliti. Data yang dimasukkan merupakan data primer yang diperoleh dari hasil pengisian kuesioner oleh mahasiswa. Selanjutnya, sistem melakukan tahap pra-pemrosesan data yang meliputi pembersihan data dan normalisasi untuk memastikan data berada dalam kondisi yang layak dianalisis.

Setelah tahap pemrosesan selesai, sistem melanjutkan ke proses pengelompokan data menggunakan algoritma *K-Means clustering*. Pada tahap ini, admin menentukan jumlah *cluster* (K) yang digunakan, kemudian sistem melakukan inisialisasi *centroid* awal dan menjalankan proses perhitungan jarak serta pembaruan *centroid* secara iteratif hingga diperoleh *cluster* yang stabil.

Hasil dari proses *clustering* kemudian dievaluasi menggunakan metode evaluasi *cluster* untuk memastikan kualitas pengelompokan data. Selanjutnya, sistem menampilkan hasil pengelompokan dalam bentuk tabel dan visualisasi yang memudahkan peneliti dalam melakukan interpretasi terhadap karakteristik masing-masing kelompok.

Dengan alur kerja tersebut. Sistem ini mampu memberikan gambaran yang jelas mengenai pola perilaku penggunaan *e-wallet* di kalangan mahasiswa Kota Medan dan mendukung proses analisis data secara sistematis dan terstruktur.

3.6 Penerapan Algoritma *K-Means Clustering*

3.6.1 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh melalui penyebaran kuesioner kepada mahasiswa di Kota Medan. Kuesioner disebarakan secara daring menggunakan Google Form kepada

responden yang aktif menggunakan layanan transaksi *cashless* berbasis *e-wallet*. Jumlah responden yang berhasil dikumpulkan dalam penelitian ini sebanyak 213 responden.

Secara umum, data awal yang terkumpul masih berbentuk data kategorikal, pilihan ganda (*checkbox*), pilihan rentang nominal, serta skala Likert, sehingga belum dapat langsung digunakan dalam proses perhitungan algoritma *K-Means*.

Adapun indikator perilaku penggunaan *e-wallet* yang digunakan dalam penelitian ini meliputi :

1. Frekuensi penggunaan *e-wallet*
2. Jenis aplikasi *e-wallet*
3. Tujuan transaksi *e-wallet*
4. Nominal transaksi *e-wallet*
5. Kemudahan penggunaan *e-wallet*
6. Kepercayaan terhadap keamanan *e-wallet*

Dari keenam indikator tersebut, hanya indikator yang bersifat numerik dan mempresentasikan intensitas perilaku yang digunakan sebagai variabel proses *clustering*, yaitu :

1. Frekuensi penggunaan
2. Nominal transaksi
3. Kemudahan penggunaan
4. Kepercayaan terhadap keamanan

Indikator jenis aplikasi dan tujuan transaksi tidak digunakan dalam perhitungan *K-Means* karena bersifat nominal, namun tetap dimanfaatkan sebagai

informasi pendukung dalam analisis karakteristik *cluster* pada tahap pembahasan hasil penelitian.

Sebelum dilakukan proses *clustering*, seluruh data kuesioner dilakukan tahap pra pengolahan (*preprocessing*) menggunakan *Microsoft Excel*. Tahap ini meliputi konversi data kategorikal dan skala Likert dalam bentuk numerik, serta perhitungan nilai rata-rata untuk indikator yang terdiri dari beberapa pernyataan. Hasil dari tahap ini berupa dataset numerik yang siap digunakan sebagai input algoritma *K-Means Clustering*.

Penelitian ini menggunakan data asli yang diperoleh dari 213 responden kuesioner. Namun peneliti hanya menampilkan 10 data responden sebagai sampel untuk keperluan simulasi perhitungan manual algoritma *K-Means Clustering*. Hal ini dilakukan agar proses perhitungan dapat dijelaskan secara sistematis dan mudah dipahami.

Tabel 3.7 Tabel Data Responden dalam Bentuk Numerik

Mahasiswa	Frekuensi Penggunaan (X1)	Nominal Transaksi (X2)	Kemudahan Penggunaan (X3)	Kepercayaan Keamanan (X4)
Data 1	4	4	4	3
Data 2	1	1	4	3.5
Data 3	3	3	4	3
Data 4	3	3	4	4
Data 5	3	3	4	4
Data 6	4	2	5	3.8333
Data 7	3	4	4.1667	4
Data 8	3	4	4.1667	3.1667
Data 9	4	2	4.8333	4.3333

Data 10	2	3	4	4
---------	---	---	---	---

3.6.2 Tahapan Perhitungan *K-Means Clustering*

Tahapan perhitungan algoritma *K-Means Clustering* pada penelitian ini dilakukan secara manual untuk memberikan pemahaman mengenai proses pengelompokan data. Perhitungan manual ini menggunakan sebagian data penelitian sebagai contoh, sedangkan pengolahan data secara keseluruhan dilakukan menggunakan sistem analisis clustering berbasis K-Means yang dikembangkan dalam penelitian ini. Adapun tahapan perhitungannya adalah sebagai berikut :

1. Menentukan jumlah *Cluster*

Jumlah *cluster* yang digunakan dalam penelitian ini ditetapkan sebanyak tiga *cluster* ($K=3$). Penentuan jumlah *cluster* ini bertujuan untuk mengelompokkan perilaku penggunaan transaksi *cashless* mahasiswa ke dalam tiga kelompok utama, yaitu tingkat penggunaan rendah, sedang, dan tinggi berdasarkan indikator yang telah ditentukan.

2. Menentukan pusat *cluster* awal (*centroid* awal)

Centroid awal ditentukan dengan memilih tiga responden secara acak dari data yang telah dinormalisasi dan dikonversi ke dalam bentuk numerik. *Centroid* awal dipilih dari Data 1, Data 4, Data 7 sehingga diperoleh *centroid* awal C1, C2, dan C3. Setiap *centroid* mempresentasikan nilai awal dari masing-masing *cluster* yang akan digunakan dalam proses perhitungan jarak.

$$C1 = (4, 4, 4, 3)$$

$$C2 = (3, 3, 4, 4)$$

$$C3 = (3, 4, 4.1667, 4)$$

Nilai *centroid* awal ini digunakan sebagai titik pusat awal untuk menghitung jarak setiap data terhadap masing-masing *cluster*.

3. Menghitung jarak *Euclidean* dan pembentukan iterasi pertama

Setiap data dihitung jaraknya terhadap masing-masing *centroid* menggunakan metode *Euclidean Distance*. Perhitungan jarak ini bertujuan untuk mengetahui kedekatan setiap data terhadap pusat *cluster*. Hasil perhitungan jarak tersebut kemudian dibandingkan, dan setiap data ditempatkan ke dalam *cluster* dengan jarak terdekat. Berdasarkan proses ini, terbentuk pembagian *cluster* awal pada iterasi pertama.

Rumus *Euclidean Distance* :

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

a. Contoh perhitungan jarak (Data 1 = 4, 4, 4, 3)

Jarak ke C1

$$d = \sqrt{(4 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (3 - 3)^2} = \sqrt{0} = 0$$

Jarak ke C2

$$d = \sqrt{(4 - 3)^2 + (4 - 3)^2 + (4 - 4)^2 + (3 - 4)^2} = \sqrt{3} = 1.723$$

Jarak ke C3

$$\begin{aligned} d &= \sqrt{(4 - 3)^2 + (4 - 4)^2 + (4 - 4.1667)^2 + (3 - 4)^2} \\ &= \sqrt{2.028} = 1.424 \end{aligned}$$

b. Contoh perhitungan jarak (Data 2 = 1, 1, 4, 3.5)

Jarak ke C1

$$d = \sqrt{(1 - 4)^2 + (1 - 4)^2 + (4 - 4)^2 + (3.5 - 3)^2} = \sqrt{18.25} = 4.272$$

Jarak ke C2

$$d = \sqrt{(1 - 3)^2 + (1 - 3)^2 + (4 - 4)^2 + (3.5 - 4)^2} = \sqrt{8.25} = 2.872$$

Jarak ke C3

$$d = \sqrt{(1 - 3)^2 + (1 - 4)^2 + (4 - 4.1667)^2 + (3.5 - 4)^2}$$

$$= \sqrt{13.2778} = 3.644$$

Hasil perhitungan awal (iterasi 1) :

Tabel 3.8 Hasil Perhitungan Awal (iterasi 1)

Data	d ke C1 (4, 4, 4, 3)	d ke C2 (3, 3, 4, 4)	d ke C3 (3, 4, 4.1667, 4)	Cluster Terdekat
Data 1	0.000	1.723	1.424	C1
Data 2	4.272	2.871	3.644	C2
Data 3	1.414	1.000	1.424	C2
Data 4	1.732	0.000	1.014	C2
Data 5	1.732	0.000	1.014	C2
Data 6	2.386	1.740	2.3892	C2
Data 7	1.424	1.014	0.000	C3
Data 8	1.028	1.312	0.833	C3
Data 9	2.545	1.676	2.357	C2
Data 10	2.449	1.000	1.424	C2

4. Kelompokkan berdasarkan centroid terdekat

Berdasarkan hasil perhitungan jarak Euclidean pada iterasi pertama, setiap data dikelompokkan ke dalam *cluster* dengan jarak terdekat terhadap centroid awal.

Cluster 1 (C1) : Data 1

Cluster 2 (C2) : Data 2, 3, 4, 5, 6, 9, 10

Cluster 3 (C3) : Data 7,8

Selanjutnya dilakukan pembaruan centroid berdasarkan rata-rata nilai masing-masing variabel pada setiap *cluster* untuk dilanjutkan ke iterasi berikutnya.

5. Menghitung centroid baru

Setelah dilakukan pengelompokan pada iterasi pertama, selanjutnya dihitung centroid baru dengan mengambil nilai rata-rata dari masing-masing variabel pada setiap *cluster*. *Centroid* baru ini digunakan sebagai titik pusat pada iterasi berikutnya untuk menghitung ulang jarak setiap data terhadap *cluster*.

$$C_{baru} = \frac{\sum X}{n} = \frac{\sum \text{nilai variabel}}{\text{jumlah data dalam cluster}}$$

a. *Centroid* baru *Cluster 1 (C1)*

Tabel 3.9 Tabel Data Responden Data 1

Data	X1	X2	X3	X4
1	4	4	4	3

$$C1 (X1) = \frac{4}{1} = 4$$

$$C1 (X1) = \frac{4}{1} = 4$$

$$C1 (X1) = \frac{4}{1} = 4$$

$$C1 (X1) = \frac{3}{1} = 3, \text{ Centroid baru C1 tetap sama}$$

b. *Centroid* baru *Cluster* 2 (C2)

Tabel 3.10 Tabel Data Responden Data 2, 3, 4, 5, 6, 9, 10

Data	X1	X2	X3	X4
2	1	1	4	3.5
3	3	3	4	3
4	3	3	4	4
5	3	3	4	4
6	4	2	5	3.8333
9	4	2	4.8333	4.3333
10	2	3	4	4

$$C2(X1) = \frac{\sum X1}{7} = \frac{20}{7} = 2.8571$$

$$C2(X2) = \frac{\sum X2}{7} = \frac{17}{7} = 2.4286$$

$$C2(X3) = \frac{\sum X3}{7} = \frac{20}{5} = 4.2619$$

$$C2(X4) = \frac{\sum X4}{5} = \frac{26.6666}{7} = 3.8095$$

c. *Centroid* baru *Cluster* 3 (C3)

Tabel 3.11 Tabel Data Responden Data 7 dan 8

Data	X1	X2	X3	X4
7	3	4	4.1667	4
8	3	4	4.1667	3.1667

$$C3(X1) = \frac{\sum X1}{2} = \frac{6}{2} = 3$$

$$C3(X2) = \frac{\sum X2}{2} = \frac{8}{2} = 4$$

$$C3(X3) = \frac{\Sigma X3}{2} = \frac{8.3334}{2} = 4.1667$$

$$C3(X4) = \frac{\Sigma X4}{4} = \frac{7.1667}{2} = 3.5833$$

Tabel 3.12 Centroid Baru

Centroid	X1	X2	X3	X4
C1	4.0	4.0	4.0	3.0
C2	2.4	2.6	4.0	3.7
C3	3.5	3.0	4.583	3.833

6. Pembentukan iterasi kedua

Menghitung kembali jarak *Euclidean* semua data terhadap centroid baru.

a. Contoh perhitungan jarak (Data 1 = 4, 4, 4, 3)

Jarak ke C1

$$d = \sqrt{(4 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (3 - 3)^2} = \sqrt{0} = 0$$

Jarak ke C2

$$\begin{aligned} d &= \sqrt{(4 - 2.8571)^2 + (4 - 2.4286)^2 + (4 - 4.2619)^2 + (3 - 3.8095)^2} \\ &= \sqrt{4.4974} = 2.1207 \end{aligned}$$

Jarak ke C3

$$\begin{aligned} d &= \sqrt{(4 - 3)^2 + (4 - 4)^2 + (4 - 4.1667)^2 + (3 - 3.5834)^2} \\ &= \sqrt{1.3682} = 1.1697 \end{aligned}$$

Sehingga hasil perhitungan iterasi 2 :

Tabel 3.13 Hasil Perhitungan iterasi 2

Data	d ke C1 (4, 4, 4, 3)	d ke C2 (3, 3, 4, 4)	d ke C3 (3, 4, 4.1667, 4)	Cluster Terdekat
Data 1	0.000	2.1207	1.1697	C1
Data 2	4.272	2.3780	3.6104	C2
Data 3	1.4142	1.0348	1.0934	C2
Data 4	1.7321	0.6721	1.0141	C2
Data 5	1.7321	0.000	1.0141	C2
Data 6	2.3863	1.4269	2.4063	C2
Data 7	1.4240	1.7033	0.4167	C3
Data 8	1.0285	1.5660	0.4167	C3
Data 9	2.5451	1.4398	2.2602	C2
Data 10	2.4495	0.9427	1.4143	C2

Berdasarkan hasil perhitungan pada iterasi kedua, setiap data kembali dihitung jaraknya terhadap centroid terbaru pada masing-masing cluster. Data kemudian dikelompokkan ke dalam cluster yang memiliki jarak paling kecil terhadap centroid. Iterasi K-Means dihentikan ketika centroid sudah tidak mengalami perubahan atau keanggotaan cluster telah stabil.

Hasil pengelompokan menunjukkan bahwa tidak terdapat perubahan anggota cluster dibandingkan dengan iterasi sebelumnya. Oleh karena itu, proses iterasi pada algoritma K-Means dihentikan pada iterasi ke-2 karena telah mencapai kondisi konvergen.

3.6.3 Iterasi dan Pembentukan *Cluster*

Setelah dilakukan perhitungan jarak dan pembentukan *cluster* pada iterasi pertama, proses *K-Means Clustering* dilanjutkan dengan iterasi kedua

menggunakan centroid baru yang diperoleh dari hasil iterasi sebelumnya. Pada tahap ini, setiap data kembali dihitung jaraknya terhadap centroid baru dan dikelompokkan ke dalam *cluster* dengan jarak terdekat.

Hasil pengelompokan pada iterasi kedua menunjukkan bahwa tidak terdapat perubahan keanggotaan data pada beberapa *cluster* dibandingkan dengan iterasi pertama. Kondisi ini menandakan bahwa proses pengelompokan telah mencapai kondisi stabil (konvergen), di mana tidak ada lagi perpindahan data antar *cluster*.

Secara teori, algoritma *K-Means* akan terus melakukan iterasi hingga kondisi konvergen tercapai, yaitu ketika centroid tidak lagi mengalami perubahan atau keanggotaan *cluster* sudah tetap. Berdasarkan hasil perhitungan manual yang dilakukan hingga iterasi kedua, proses clustering pada penelitian ini telah mencapai kondisi konvergen.

Dengan demikian, hasil iterasi manual tersebut menunjukkan bahwa proses pengelompokan data menggunakan algoritma K-Means Clustering telah menghasilkan pembentukan cluster yang stabil.

3.6.4 Evaluasi Cluster Menggunakan WCV dan BCV

Evaluasi kualitas hasil pengelompokan dilakukan untuk memastikan bahwa cluster yang terbentuk oleh algoritma *K-Means* benar-benar mencerminkan struktur data yang ada. Pada penelitian ini, evaluasi dilakukan menggunakan dua ukuran, yaitu *Within-Cluster Variation* (WCV) dan *Between-Cluster Variation* (BCV). Perhitungan nilai WCV dan BCV ini dilakukan dengan menggunakan 30 data awal responden sebagai input. Penggunaan data asli tanpa normalisasi dimaksudkan untuk menghitung jarak *Euclidean* langsung terhadap centroid

sehingga nilai variasi yang diperoleh dapat mencerminkan struktur numerik dataset secara apa adanya.

1. Centroid Setiap Cluster

Perhitungan centroid dilakukan dengan menghitung rata-rata nilai setiap variabel pada masing-masing cluster.

Tabel 3.14 Centroid Cluster Final

Cluster	X1	X2	X3	X4
Cluster 1	4.4000	2.9333	4.5445	3.8222
Cluster 2	2.3333	2.3333	3,9861	3.7361
Cluster 3	2.0000	3.6667	3.7778	3.3333

Centroid ini digunakan sebagai dasar perhitungan jarak internal (WCV) maupun jarak antar cluster (BCV)

2. Perhitungan WCV (*Within-Cluster Variation*)

WCV dihitung sebagai total jarak kuadrat antara setiap anggota cluster dengan centroidnya.

a. Centroid Cluster 1

Cluster 1 centroid (hasil perhitungan otomatis sistem)

$$\mu_1 = (4.400, 2.9333, 4.5445, 3.8222)$$

b. Pilih 1 Sampel Data dalam Cluster 1 (misalnya data Responden 1 (R1))

$$x_i = (4, 2, 4.1667, 4)$$

c. Hitung Selisih Data dengan Centroid

$$x_i - \mu_i = (4 - 4.4000 - 2.9333, 4.1667 - 4.5445, 4 - 3.8222)$$

$$= (-0.4000, -0.9333, -0.3778, 0.1778)$$

d. Hitung Kuadrat Selisih

$$(-0.4000)^2 = 0.1600$$

$$(-0.9333)^2 = 0.8711$$

$$(-0.3778)^2 = 0.1427$$

$$(-0.1778)^2 = 0.0316$$

e. Jumlahkan untuk Mendapatkan Jarak Kuadrat Data ke Centroid

$$\|x_i - \mu_k\|^2 = 0.1600 + 0.8711 + 0.1427 + 0.0316 = 1.2054$$

f. Ulangi untuk Semua Anggota Cluster

Proses di atas dilakukan untuk 18 anggota Cluster 1, lalu dijumlahkan :

$$WCV_1 = \sum_{i=1}^{18} \|x_i - \mu_k\|^2 = 28.9140$$

Tabel 3.15 WCV Per Cluster

Cluster	Jumlah Anggota	WCV
Cluster 1	18	28.9140
Cluster 2	9	13.7125
Cluster 3	3	13.5833
WCV Total		56.2098

Nilai ini menunjukkan bahwa seluruh cluster memiliki tingkat kekompakan yang cukup tinggi karena nilai WCV berada pada angka yang relatif rendah dibandingkan rentang skala variabel (1-5).

3. Perhitungan BCV (*Between-Cluster Variation*)

BCV dihitung dengan mengukur jarak antara centroid tiap cluster dengan centroid global.

a. Centroid Cluster 1

Cluster 1 centroid (hasil perhitungan otomatis sistem)

$$\mu_1 = (4.400, 2.9333, 4.5445, 3.8222)$$

b. Centroid Global

$$\mu = (3.4333, 2.7000, 4.1550, 3.7889)$$

c. Hitung Selisih Centroid terhadap Centroid Global

$$\begin{aligned} \mu_1 - \mu_i &= (4.4000 - 3.4333, 2.9333 - 2.7000, \\ &\quad 4.5445 - 4.1550, 3.8222 - 3.7889) \\ &= (-0.4000, -0.9333, -0.3778, 0.1778) \end{aligned}$$

d. Hitung Kuadrat Selisih

$$(0.9667)^2 = 0.9345$$

$$(0.2333)^2 = 0.0544$$

$$(0.3895)^2 = 0.1517$$

$$(0.0333)^2 = 0.0011$$

e. Jumlahkan

$$\begin{aligned} \|\mu_1 - \mu\|^2 &= 0.9345 + 0.0544 + 0.1517 + 0.0011 \\ &= 1.1417 \end{aligned}$$

f. Kalikan dengan Jumlah Anggota Cluster

Cluster 1 berisi 18 anggota, sehingga :

$$BCV_1 = 18 \times 1.1417 = 17.8221$$

Tabel 3.16 BCV Per Cluster

Cluster	BCV
Cluster 1	17.8221
Cluster 2	11.0157
Cluster 3	14.0657
BCV Total = 42.9035	

Nilai BCV yang tinggi menunjukkan bahwa centroid cluster memiliki perbedaan karakteristik yang jelas.

4. Jarak Antar Centroid (*Pairwise Distance*)

Tabel 3.17 Jarak Antar Centroid

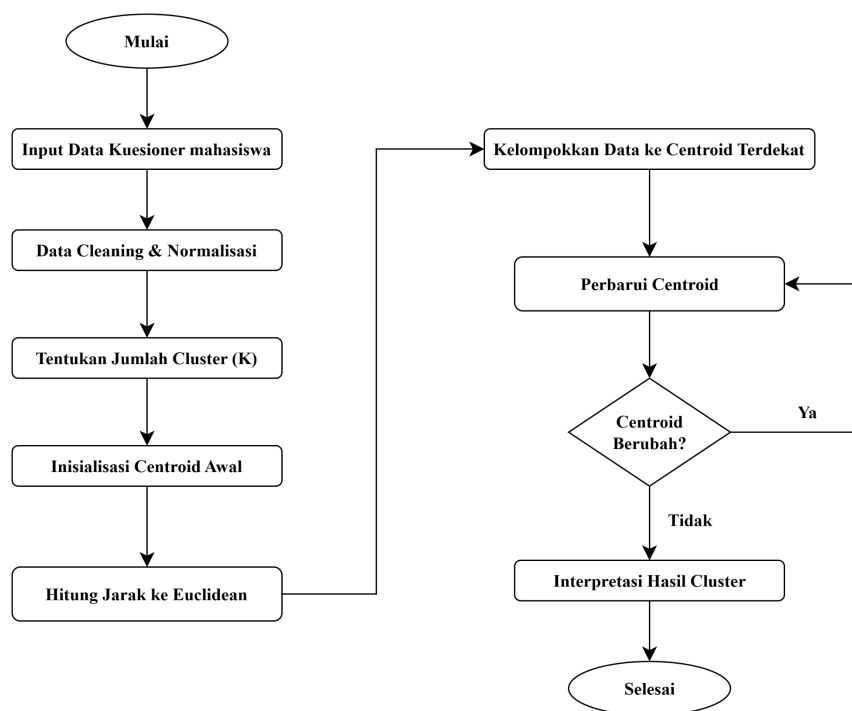
Perbandingan	Jarak
Cluster 1 ↔ Cluster 2	2.8143
Cluster 1 ↔ Cluster 3	2.9587
Cluster 2 ↔ Cluster 3	1.9911

Jarak antar centroid yang semuanya berada di atas 1.9 menunjukkan bahwa ketiga cluster berada pada posisi yang terpisah secara signifikan.

Evaluasi kualitas cluster dilakukan dengan menghitung nilai *Within-Cluster Variation* (WCV) dan *Between-Cluster Variation* (BCV) berdasarkan centroid akhir setiap cluster. Hasil perhitungan menunjukkan bahwa WCV total bernilai 56.2098 dan BCV total bernilai 42.9035. Nilai ini mengindikasikan bahwa setiap cluster memiliki kekompakan internal yang baik dan pemisahan antar cluster yang jelas.

3.7 Flowchart Penelitian

Flowchart penelitian merupakan diagram alir yang menggambarkan urutan tahapan kegiatan penelitian dari awal hingga akhir. Dalam penelitian ini, *flowchart* menjelaskan tahapan-tahapan penelitian mulai dari pengumpulan data, pengolahan, penerapan algoritma *K-Means Clustering*, serta interpretasi hasil pengelompokan perilaku mahasiswa pengguna *e-wallet*. Adapun tahapan dalam *Flowchart* terdiri atas :



Gambar 3.6 *Flowchart* Tahapan Penelitian

Berikut ini adalah penjelasan tiap-tiap tahapan dalam *flowchart* tersebut :

1. Mulai
2. Input Data Kuesioner Mahasiswa

Data dikumpulkan dari hasil pengisian kuesioner online berbasis indikator perilaku penggunaan *e-wallet*.

3. *Data Cleaning* dan Normalisasi

- a. Menghapus data duplikat kosong
- b. Mengubah nilai skala Likert menjadi numerik (1-5)
- c. Melakukan normalisasi agar semua variabel memiliki skala yang sebanding.

4. Menentukan Jumlah *Cluster* (K)

Biasanya K ditentukan antara 2-5 untuk segmentasi perilaku.

5. Inisialisasi *Centroid* Awal

Centroid awal dipilih secara acak dari sekumpulan data awal setiap *cluster*.

6. Hitung Jarak *Euclidean* Tiap Data ke *Centroid*

Setiap data mahasiswa dihitung jaraknya terhadap tiap *centroid* menggunakan rumus *Euclidean*.

7. Kelompokkan Data ke *Centroid* Terdekat

Data Mahasiswa dimasukkan ke *cluster* dengan jarak terdekat dari *centroid*.

8. Perbarui *Centroid* (*Recalculate Centroid*)

Hitung ulang posisi *centroid* baru dengan mengambil rata-rata dari seluruh dalam *cluster* tersebut.

9. Cek Konvergensi (*Decision Symbol*)

Jika posisi *centroid* tidak berubah lagi (stabil) maka lanjut ke tahap evaluasi. Jika masih berubah, ulangi proses dari langkah ke-6 (perhitungan jarak)

10. Interpretasi Hasil *Cluster*

Setiap cluster diinterpretasikan untuk mengidentifikasi karakteristik perilaku mahasiswa, misalnya :

Cluster 1 : Pengguna *e-wallet* intensif

Cluster 2 : Pengguna moderat

Cluster 3 : Pengguna rendah

11. Selesai

Flowchart ini diikuti oleh diagram algoritma penelitian yang menjelaskan langkah-langkah komputasi dalam menerapkan algoritma *K-Means Clustering*. Kedua bagian ini saling terkait dan memberikan gambaran menyeluruh mengenai proses analisis perilaku penggunaan *e-wallet* mahasiswa Kota Medan.

3.8 Rancangan Antarmuka Sistem

Rancangan antarmuka sistem ini bertujuan untuk menyediakan sarana interaksi antara pengguna dan sistem analisis pengelompokan perilaku penggunaan transaksi *cashless*. Antarmuka dirancang agar pengguna dapat melakukan input data, menjalankan proses *clustering* menggunakan algoritma *K-Means*, serta melihat dan menganalisis hasil pengelompokan secara mudah dan sistematis.

Sistem yang dirancang pada penelitian ini merupakan sistem analisis data berbasis lokal yang digunakan untuk keperluan penelitian, sehingga tidak mencakup fitur autentikasi pengguna, manajemen hak akses, maupun penyimpanan data berbasis server online.

3.8.1 Prinsip Desain Antarmuka

Rancangan antarmuka sistem ini memiliki empat prinsip sederhana, yakni:

1. Sederhana (*Simplicity*)

Antarmuka hanya memuat fitur yang relevan dengan proses analisis *clustering*.

2. Fungsional (*Functionality*)

Setiap elemen antarmuka memiliki fungsi yang jelas dalam mendukung proses penelitian.

3. Keterbacaan (*Readability*)

Informasi dan hasil analisis ditampilkan dalam bentuk tabel grafik yang mudah dipahami.

4. Kesesuaian dengan Metode

5. Antarmuka dirancang sesuai tahapan algoritma *K-Means Clustering*.

3.8.2 Struktur Umum Antarmuka Sistem

Secara umum, sistem terdiri dari beberapa halaman utama yang saling terhubung, yaitu halaman input data, halaman pra-pemrosesan data, halaman proses *clustering*, halaman hasil *clustering*, dan halaman visualisasi hasil.

Berikut rancangan layar yang akan dibuat :

1. Rancangan Layar 1 - Halaman Input Data

Pada halaman input data difungsikan untuk memasukkan data kuesioner mahasiswa dan menjadi titik awal proses analisis. Elemen antarmuka meliputi tombol upload data (*Excel/CSV*), tabel pratinjau data, dan informasi jumlah data.

Sistem Analisis Pengelompokan Penggunaan E-Wallet dengan K-Means Clustering																																			
	Input Data <input type="button" value="Upload Data"/>																																		
Input Data	Unggah file data dalam format Excel atau CSV untuk melanjutkan analisis pengelompokan.																																		
Pra-Pemrosesan	Pratinjau Data <table border="1"> <thead> <tr> <th>Mahasiswa</th> <th>Frekuensi</th> <th>Nominal Transksi</th> <th>Jenis E-Wallet</th> <th>Kepercayaan</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>B</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>C</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>D</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>E</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> </tbody> </table>					Mahasiswa	Frekuensi	Nominal Transksi	Jenis E-Wallet	Kepercayaan	A	xxx	xxx	xxx	xxx	B	xxx	xxx	xxx	xxx	C	xxx	xxx	xxx	xxx	D	xxx	xxx	xxx	xxx	E	xxx	xxx	xxx	xxx
Mahasiswa	Frekuensi	Nominal Transksi	Jenis E-Wallet	Kepercayaan																															
A	xxx	xxx	xxx	xxx																															
B	xxx	xxx	xxx	xxx																															
C	xxx	xxx	xxx	xxx																															
D	xxx	xxx	xxx	xxx																															
E	xxx	xxx	xxx	xxx																															
Proses Clustering	Total Data: 100		Total Data: 100																																
Hasil Clustering																																			
Visualisasi Cluster																																			

Gambar 3.7 Rancangan Layar 1

2. Rancangan Layar 2 - Halaman Pra-pemrosesan Data

Sistem Analisis Pengelompokan Penggunaan E-Wallet dengan K-Means Clustering																																			
	Pra-Pemrosesan <table border="1"> <thead> <tr> <th>Cleaning Data</th> <th>Normalisasi Data</th> </tr> </thead> <tbody> <tr> <td>Melakukan pembersihan data, seperti menghapus data duplikat dan data yang tidak valid <input type="button" value="Proses Cleaning"/></td> <td>Mengubah data ke dalam skala yang seragam guna meningkatkan akurasi analisis. <input type="button" value="Proses Normalisasi"/></td> </tr> </tbody> </table>					Cleaning Data	Normalisasi Data	Melakukan pembersihan data, seperti menghapus data duplikat dan data yang tidak valid <input type="button" value="Proses Cleaning"/>	Mengubah data ke dalam skala yang seragam guna meningkatkan akurasi analisis. <input type="button" value="Proses Normalisasi"/>																										
Cleaning Data	Normalisasi Data																																		
Melakukan pembersihan data, seperti menghapus data duplikat dan data yang tidak valid <input type="button" value="Proses Cleaning"/>	Mengubah data ke dalam skala yang seragam guna meningkatkan akurasi analisis. <input type="button" value="Proses Normalisasi"/>																																		
Input Data																																			
Pra-Pemrosesan																																			
Proses Clustering	Data yang Telah Diproses <table border="1"> <thead> <tr> <th>Mahasiswa</th> <th>Frekuensi</th> <th>Nominal Transksi</th> <th>Jenis E-Wallet</th> <th>Kepercayaan</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>B</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>C</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>D</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> <tr> <td>E</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> <td>xxx</td> </tr> </tbody> </table>					Mahasiswa	Frekuensi	Nominal Transksi	Jenis E-Wallet	Kepercayaan	A	xxx	xxx	xxx	xxx	B	xxx	xxx	xxx	xxx	C	xxx	xxx	xxx	xxx	D	xxx	xxx	xxx	xxx	E	xxx	xxx	xxx	xxx
Mahasiswa	Frekuensi	Nominal Transksi	Jenis E-Wallet	Kepercayaan																															
A	xxx	xxx	xxx	xxx																															
B	xxx	xxx	xxx	xxx																															
C	xxx	xxx	xxx	xxx																															
D	xxx	xxx	xxx	xxx																															
E	xxx	xxx	xxx	xxx																															
Hasil Clustering	Total Data: 100		Total Data: 100																																
Visualisasi Cluster																																			

Gambar 3.8 Rancangan Layar 2

Berfungsi untuk melakukan *data cleaning* dan melakukan normalisasi data. Elemen antarmuka mencakup tombol proses data, informasi metode normalisasi, dan tabel hasil data setelah diproses. Tahapan ini penting karena algoritma *K-Means* sensitif terhadap perbedaan skala data.

3. Rancangan Layar 3 - Halaman Proses *Clustering*

Halaman ini berfungsi untuk menjalankan algoritma *K-Means*. Elemen antarmuka meliputi input nilai jumlah *cluster* (K), tombol proses *Clustering*, informasi jumlah iterasi, dan status proses *clustering*. Di halaman inilah algoritma *K-Means Clustering* diterapkan untuk melakukan pengelompokan perilaku penggunaan *e-wallet* mahasiswa.

Sistem Analisis Pengelompokan Penggunaan E-Wallet dengan K-Means Clustering																																		
	Proses Clustering																																	
	Data Preprocessing																																	
	Jumlah Cluster (K):																																	
Input Data	<input type="text"/>																																	
Pra-Pemrosesan	<input type="button" value="Proses Clustering"/>																																	
Proses Clustering	Hasil Proses																																	
Hasil Clustering	<table border="1"> <thead> <tr> <th>Mahasiswa</th> <th>Frekuensi</th> <th>Normalisasi Nominal Transaksi</th> <th colspan="2">Jenis E-Wallet</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>0.xx</td> <td>0.xx</td> <td colspan="2">xxx</td> </tr> <tr> <td>B</td> <td>0.xx</td> <td>0.xx</td> <td colspan="2">xxx</td> </tr> <tr> <td>C</td> <td>0.xx</td> <td>0.xx</td> <td colspan="2">xxx</td> </tr> <tr> <td>D</td> <td>0.xx</td> <td>0.xx</td> <td colspan="2">xxx</td> </tr> <tr> <td>E</td> <td>0.xx</td> <td>0.xx</td> <td colspan="2">xxx</td> </tr> </tbody> </table>				Mahasiswa	Frekuensi	Normalisasi Nominal Transaksi	Jenis E-Wallet		A	0.xx	0.xx	xxx		B	0.xx	0.xx	xxx		C	0.xx	0.xx	xxx		D	0.xx	0.xx	xxx		E	0.xx	0.xx	xxx	
Mahasiswa	Frekuensi	Normalisasi Nominal Transaksi	Jenis E-Wallet																															
A	0.xx	0.xx	xxx																															
B	0.xx	0.xx	xxx																															
C	0.xx	0.xx	xxx																															
D	0.xx	0.xx	xxx																															
E	0.xx	0.xx	xxx																															
Visualisasi Cluster	Total Data: 100		Total Data: 100																															

Gambar 3.9 Rancangan Layar 3

4. Rancangan Layar 4 - Halaman Hasil *Clustering*

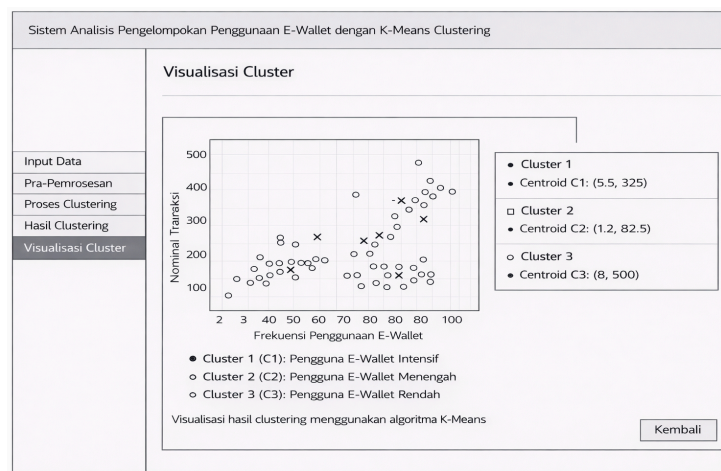
Sistem Analisis Pengelompokan Penggunaan E-Wallet dengan K-Means Clustering																													
	Hasil Clustering																												
	Hasil Pengelompokan																												
	Lakukan pembersihan data (data cleaning) dan normalisasi skala agar data siap digunakan dalam proses cluster sering.																												
Input Data	<input type="button" value="Proses Clustering"/>																												
Pra-Pemrosesan	Hasil Proses																												
Proses Clustering	<table border="1"> <thead> <tr> <th>Mahasiswa</th> <th>Frekuensi</th> <th>Nominal Transaksi</th> <th>Cluster</th> <th>Cluster</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>xxx</td> <td>0.xx</td> <td>C1</td> <td>C1</td> </tr> <tr> <td>B</td> <td>xxx</td> <td>0.xx</td> <td>C2</td> <td>C2</td> </tr> <tr> <td>C</td> <td>xxx</td> <td>0.xx</td> <td>C2</td> <td>C2</td> </tr> <tr> <td>D</td> <td>xxx</td> <td>0.xx</td> <td>C2</td> <td>C3</td> </tr> </tbody> </table>				Mahasiswa	Frekuensi	Nominal Transaksi	Cluster	Cluster	A	xxx	0.xx	C1	C1	B	xxx	0.xx	C2	C2	C	xxx	0.xx	C2	C2	D	xxx	0.xx	C2	C3
Mahasiswa	Frekuensi	Nominal Transaksi	Cluster	Cluster																									
A	xxx	0.xx	C1	C1																									
B	xxx	0.xx	C2	C2																									
C	xxx	0.xx	C2	C2																									
D	xxx	0.xx	C2	C3																									
Hasil Clustering	<ul style="list-style-type: none"> - Jumlah anggota Cluster 1 (C1): 40 - Jumlah anggota Cluster 2 (C2): 35 - Jumlah anggota Cluster 3 (C3): 25 			<input type="button" value="Lihat Visualisasi"/>																									
Visualisasi Cluster																													

Gambar 3.10 Rancangan Layar 4

Halaman ini berfungsi untuk menampilkan hasil pengelompokan data. Elemen antarmuka meliputi tabel hasil *cluster*, label *cluster* tiap mahasiswa, dan jumlah anggota tiap cluster.

5. Rancangan Layar 5 - Halaman Visualisasi *Cluster*

Halaman ini berfungsi untuk membantu interpretasi hasil. Elemen antarmuka meliputi grafik *scatterplot*, warna beda untuk setiap *cluster*, dan penanda *centroid* sehingga pembaca dapat memahami pola data secara visual.



Gambar 3.11 Rancangan Layar 5

Rancangan antarmuka sistem ini disusun untuk mendukung seluruh tahapan analisis pengelompokan perilaku pengguna transaksi *cashless* menggunakan algoritma *K-Means Clustering*, sehingga hasil penelitian dapat diperoleh secara sistematis, terstruktur, dan mudah dianalisis.

3.9 Pengujian Sistem

Pengujian sistem dilakukan untuk memastikan bahwa sistem analisis pengelompokan perilaku penggunaan transaksi *cashless* menggunakan algoritma *K-Means Clustering* dapat berjalan sesuai dengan tujuan penelitian. Pengujian ini

bertujuan untuk mengetahui apakah setiap fungsi dalam sistem telah bekerja dengan baik, mulai dari proses input data hingga penyajian hasil clustering dan visualisasi.

Metode pengujian yang digunakan dalam penelitian ini adalah *Black Box Testing*, yaitu metode pengujian yang berfokus pada pengujian sistem berdasarkan masukan (*input*) dan keluaran (*output*) tanpa memperhatikan struktur internal program. Metode ini dipilih karena sistem yang dibangun bersifat sederhana dan difokuskan pada fungsionalitas sebagai alat bantu analisis *data mining*, sehingga pengujian difokuskan pada keberhasilan fungsi-fungsi utama sistem.

Terdapat lima objek pengujian dalam sistem yang meliputi seluruh fungsi utama sistem, yaitu :

1. Fungsi input data kuesioner
2. Fungsi pra-pemrosesan data
3. Fungsi *Clustering* menggunakan algoritma *K-Means*
4. Fungsi penampilan hasil *clustering*
5. Fungsi visualisasi *cluster*

3.9.1 Skenario dan Tahapan Pengujian Sistem

Pengujian sistem dilakukan dengan menjalankan setiap fungsi utama sistem dan membandingkan hasil yang diperoleh dengan hasil yang diharapkan.

Tahapan pengujian meliputi :

1. Pengujian Input Data

Pengujian dilakukan mengunggah data kuesioner mahasiswa ke dalam sistem. Sistem diuji untuk memastikan data dapat dimuat dan ditampilkan dengan benar tanpa kesalahan format.

2. Pengujian Pra-Pemrosesan Data

Pengujian dilakukan dengan menjalankan proses pembersihan dan normalisasi data. Sistem diuji untuk memastikan data berhasil diubah ke dalam bentuk numerik dan berada pada skala yang seragam sehingga siap digunakan dalam proses *clustering*.

3. Pengujian Proses *Clustering*

Pengujian dilakukan dengan menentukan jumlah *cluster* (K) dan menjalankan algoritma *K-Means*. Sistem diuji untuk memastikan data dapat dikelompokkan ke dalam sejumlah *cluster* sesuai dengan nilai K yang ditentukan.

4. Pengujian Hasil *Clustering*

Pengujian dilakukan dengan menampilkan tabel hasil pengelompokan. Sistem diuji untuk memastikan setiap data memiliki label *cluster* yang sesuai dan dapat ditampilkan dengan jelas.

5. Pengujian Visualisasi *Cluster*

Pengujian dilakukan dengan menampilkan grafik visualisasi *cluster*. Sistem diuji untuk memastikan grafik dan *centroid cluster* dapat ditampilkan sehingga memudahkan interpretasi hasil *clustering*.

Tabel 3.18 Tabel Pengujian Sistem (*Black Box System*)

No	Fitur yang Diuji	Skenario Pengujian	Input	Output yang Diharapkan
1	Input Data	Mengunggah file data kuesioner mahasiswa	<i>File Excel/CSV</i>	Data berhasil ditampilkan dalam tabel
2	Pra-pemrosesan Data	Menjalankan proses <i>data cleaning</i> dan normalisasi	Data mentah	Data ternormalisasi dan siap

				dianalisis
3	Proses <i>Clustering</i>	Menjalankan algoritma <i>K-Means</i> dengan nilai K tertentu	Nilai K	Data terbagi menjadi ke dalam beberapa <i>cluster</i>
4	Hasil <i>Clustering</i>	Menampilkan tabel hasil pengelompokan	Data <i>cluster</i>	Label <i>cluster</i> tampil untuk setiap data
5	Visualisasi <i>Cluster</i>	Menampilkan grafik hasil <i>cluster</i>	Data <i>cluster</i>	Grafik <i>cluster</i> dan <i>centroid</i> tampil

Sistem dinyatakan berhasil apabila seluruh fungsi utama dapat dijalankan tanpa mengalami kesalahan, serta hasil *clustering* dan visualisasi dapat ditampilkan dan diinterpretasikan sesuai dengan tujuan penelitian. Pengujian sistem ini dilakukan untuk memastikan sistem dapat menjalankan analisis pengelompokan perilaku penggunaan transaksi *cashless* mahasiswa Kota Medan menggunakan algoritma *K-Means Clustering* dengan baik.

BAB IV

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi Sistem

Implementasi sistem merupakan tahap penerapan rancangan sistem yang telah dibuat pada tahap perancangan ke dalam bentuk aplikasi yang dijalankan. Pada tahap ini, sistem yang telah dirancang pada Bab III diimplementasikan menjadi sebuah aplikasi yang mampu melakukan proses analisis pengelompokan perilaku penggunaan transaksi *e-wallet* di kalangan mahasiswa Kota Medan menggunakan algoritma *K-Means Clustering*.

Sistem pada penelitian ini dikembangkan dengan bahasa pemrograman *Python* dengan memanfaatkan *framework Streamlit* sebagai antarmuka aplikasi berbasis web. Melalui sistem ini, pengguna dapat melakukan proses pengolahan data, menjalankan algoritma *K-Means Clustering* untuk melakukan pengelompokan data, serta menampilkan hasil clustering yang diperoleh. Selain itu, hasil pengelompokan data yang dihasilkan oleh sistem juga disimpan ke dalam *database MySQL* yang dikelola melalui *phpMyAdmin* dengan menggunakan *XAMPP* sebagai server lokal.

Dengan adanya implementasi sistem ini, proses analisis pengelompokan data dapat dilakukan secara lebih cepat dan terstruktur dibandingkan dengan proses perhitungan secara manual.

4.2 Tampilan Antarmuka Sistem

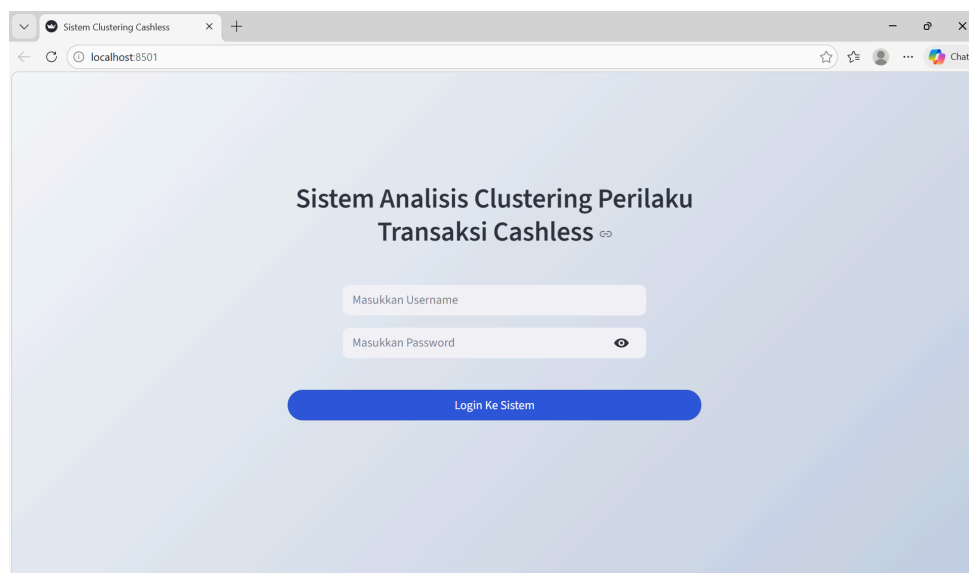
Tampilan antarmuka sistem merupakan bagian yang berfungsi sebagai media interaksi antara pengguna dengan sistem yang telah dibangun. Melalui antarmuka ini, pengguna dapat menjalankan berbagai fungsi yang tersedia dalam

aplikasi seperti melakukan pengolahan data, menjalankan proses clustering menggunakan algoritma *K-Means*, serta melihat hasil pengelompokan data yang dihasilkan oleh sistem.

Sistem pada penelitian ini dikembangkan menggunakan *framework Streamlit* sehingga dapat dijalankan melalui web browser. Setelah dijalankan menggunakan perintah `python -m streamlit run app.py`, aplikasi secara otomatis terbuka pada browser dan menampilkan halaman utama sistem.

4.2.1 Halaman Utama Sistem

Halaman utama sistem merupakan tampilan awal yang muncul ketika aplikasi dijalankan. Pada halaman ini, halaman login merupakan halaman awal yang pertama kali dijalankan. Halaman ini berfungsi sebagai mekanisme keamanan untuk membatasi akses pengguna terhadap sistem. Pada halaman ini, pengguna diwajibkan memasukkan *username* dan *password* yang valid agar dapat mengakses fitur-fitur yang tersedia pada aplikasi.

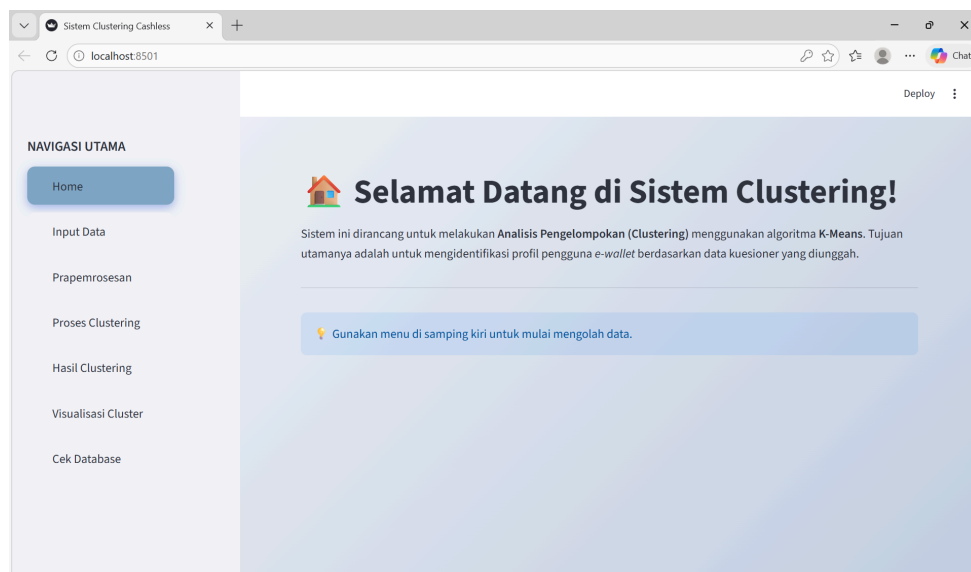


Gambar 4.1 Halaman Utama Sistem

Antarmuka halaman login dirancang dengan tampilan sederhana dan terpusat pada layar agar memudahkan pengguna dalam melakukan proses autentikasi. Dengan adanya mekanisme sistem login ini, sistem dapat menjaga keamanan akses terhadap proses pengolahan data yang dilakukan pada aplikasi.

4.2.2 Halaman *Home*

Halaman *Home* merupakan halaman yang ditampilkan setelah pengguna berhasil melakukan proses login. Halaman ini berfungsi sebagai halaman yang memberikan gambaran umum mengenai sistem yang digunakan untuk melakukan analisis pengelompokan perilaku penggunaan *e-wallet* menggunakan algoritma *K-Means Clustering*.

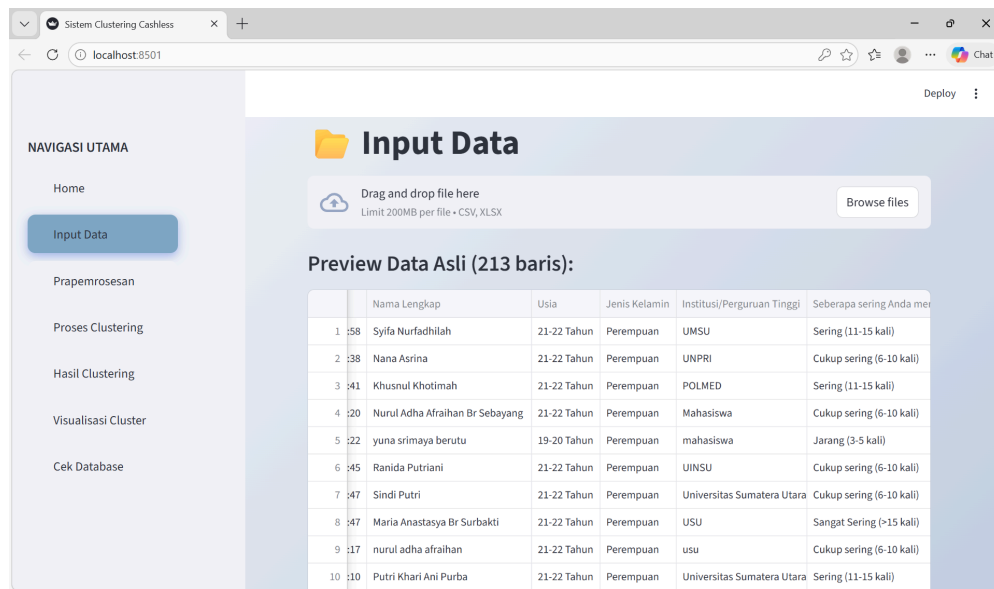


Gambar 4.2 Tampilan Halaman *Home*

Terdapat juga *sidebar* yang berisi menu navigasi utama sistem. *Sidebar* tersebut digunakan untuk memudahkan pengguna dalam berpindah dari satu fitur ke fitur lainnya.

4.2.3 Halaman Input Data

Halaman input data digunakan untuk memasukkan data penelitian ke dalam sistem. Pada halaman ini pengguna dapat mengunggah file data hasil kuesioner dalam format *CSV* atau *XLSX*. Setelah file berhasil diunggah, sistem akan membaca data tersebut dan menampilkannya dalam bentuk tabel sehingga pengguna dapat melihat kembali data yang telah dimasukkan ke dalam sistem.



Gambar 4.3 Tampilan Input Data

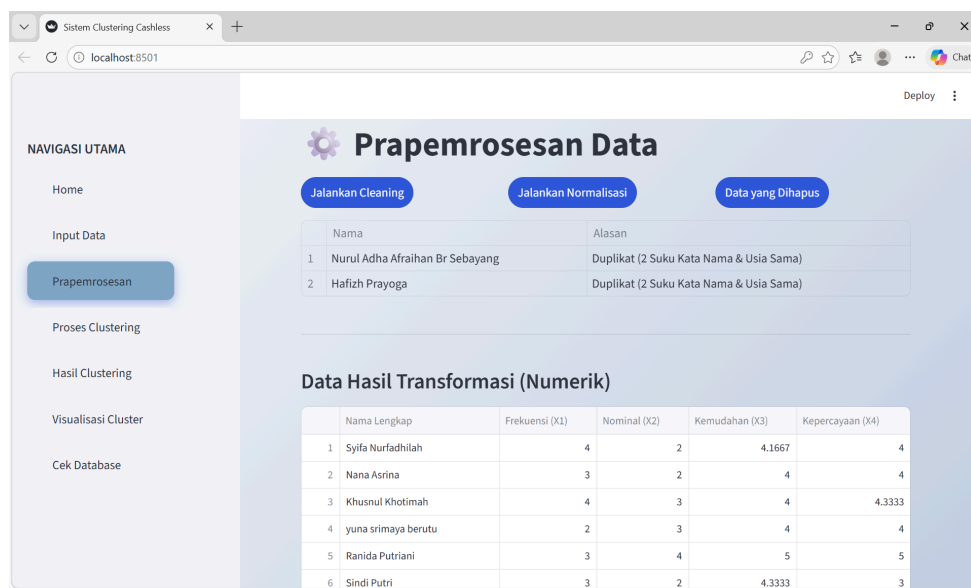
4.2.4 Halaman Pemrosesan Data

Halaman pemrosesan data digunakan untuk mempersiapkan data sebelum dilakukan proses clustering. Pada halaman ini terdapat beberapa tahapan pengolahan data yang dilakukan secara otomatis oleh sistem, yaitu proses pembersihan data (*data cleaning*) dan normalisasi data.

Proses data cleaning dilakukan untuk menghapus data yang tidak dibutuhkan dalam proses clustering serta mengubah jawaban kuesioner ke dalam bentuk skala numerik menggunakan skala Likert. Setelah proses pembersihan data selesai, sistem kemudian akan melakukan proses normalisasi data menggunakan

metode MinMaxScaler. Proses normalisasi ini bertujuan untuk mengubah rentang nilai data menjadi skala 0 sampai 1 sehingga setiap variabel memiliki kontribusi yang seimbang dalam proses perhitungan jarak pada algoritma K-Means.

Selain itu, halaman ini juga menampilkan tabel yang berisi data yang dihapus oleh sistem selama proses pembersihan data berlangsung. Hal ini bertujuan untuk memberikan transparansi terhadap proses prapemrosesan data yang dilakukan oleh sistem.



Nama	Alasan
1 Nurul Adha Afraihan Br Sebayang	Duplikat (2 Suku Kata Nama & Usia Sama)
2 Hafizh Prayoga	Duplikat (2 Suku Kata Nama & Usia Sama)

Data Hasil Transformasi (Numerik)					
	Nama Lengkap	Frekuensi (X1)	Nominal (X2)	Kemudahan (X3)	Kepercayaan (X4)
1	Syifa Nurfadhilah	4	2	4.1667	4
2	Nana Asrina	3	2	4	4
3	Khusnul Khotimah	4	3	4	4.3333
4	yuna srimaya berutu	2	3	4	4
5	Ranida Putriani	3	4	5	5
6	Sindi Putri	3	2	4.3333	3

Gambar 4.4 Tampilan Prapemrosesan Data

4.2.5 Halaman Proses *Clustering*

Halaman proses *clustering* digunakan untuk menjalankan algoritma *K-Means Clustering* terhadap data yang telah melalui tahap pemrosesan. Pada halaman ini pengguna dapat menentukan jumlah cluster yang akan digunakan dalam proses pengelompokan data.

Setelah pengguna menentukan jumlah cluster yang diinginkan dan menekan tombol proses, sistem akan menjalankan algoritma K-Means untuk mengelompokkan data berdasarkan kesamaan karakteristik yang dimiliki. Hasil

pengelompokan data kemudian ditampilkan oleh sistem beserta nilai Silhouette Score yang digunakan sebagai indikator untuk mengukur kualitas hasil clustering yang dihasilkan.

Nilai Silhouette Score tersebut digunakan untuk mengevaluasi seberapa baik data terkelompok dalam cluster yang terbentuk, sehingga dapat membantu dalam menentukan jumlah cluster yang optimal dalam proses analisis data.

	Nama Lengkap	Frekuensi (X1)	Nominal (X2)	Kemudahan (X3)	Kepercayaan (X4)	Cluster
1	Syifa Nurfadilah	4	2	4.1667	4	Cluster 1
2	Nana Asrina	3	2	4	4	Cluster 2
3	Khusnul Khotimah	4	3	4	4.3333	Cluster 1
4	yuna srimaya berutu	2	3	4	4	Cluster 2
5	Ranida Putriani	3	4	5	5	Cluster 1
6	Sindi Putri	3	2	4.3333	3	Cluster 2

Gambar 4.5 Tampilan Proses *Clustering*

4.2.6 Halaman Hasil *Clustering*

Halaman hasil *clustering* menampilkan data responden yang telah dikelompokkan ke dalam beberapa cluster berdasarkan hasil perhitungan algoritma *K-Means*. Pada halaman ini pengguna dapat melihat tabel data yang telah dilengkapi dengan label cluster untuk setiap responden.

Hasil Akhir

Default (Data Awal) ▼

Default (Data Awal)

Berdasarkan Cluster

Berdasarkan Nama

2	Nana Asrina	3	2	4	4	Cluster 2
3	Khusnul Khotimah	4	3	4	4.3333	Cluster 1
4	yuna srimaya berutu	2	3	4	4	Cluster 2
5	Ranida Putriani	3	4	5	5	Cluster 1
6	Sindi Putri	3	2	4.3333	3	Cluster 2
7	Maria Anastasya Br Surbakti	5	5	5	4	Cluster 1
8	nurul adha afraihan	3	2	4.5	4	Cluster 2
9	Putri Khari Ani Purba	4	2	4.6667	3.6667	Cluster 1
10	sinthia atikah	2	2	3.5	3.1667	Cluster 2

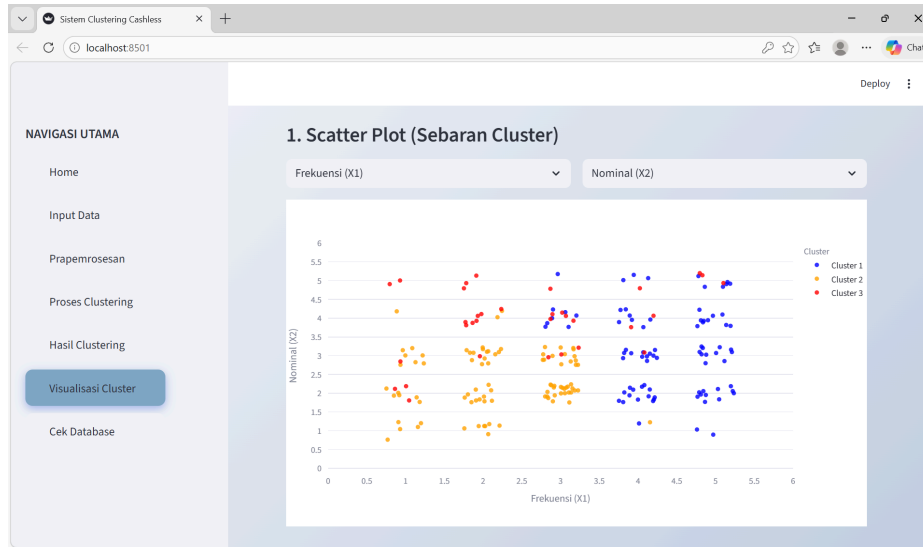
Gambar 4.6 Tampilan Hasil *Clustering*

Selain itu, sistem juga menyediakan fitur pengurutan data yang memungkinkan pengguna untuk menampilkan data berdasarkan beberapa kriteria, seperti urutan data awal kuesioner, pengelompokan berdasarkan cluster, maupun urutan abjad nama responden. Fitur ini bertujuan untuk memudahkan pengguna dalam melakukan analisis terhadap hasil pengelompokan data yang dihasilkan oleh sistem.

4.2.7 Halaman Visualisasi Cluster

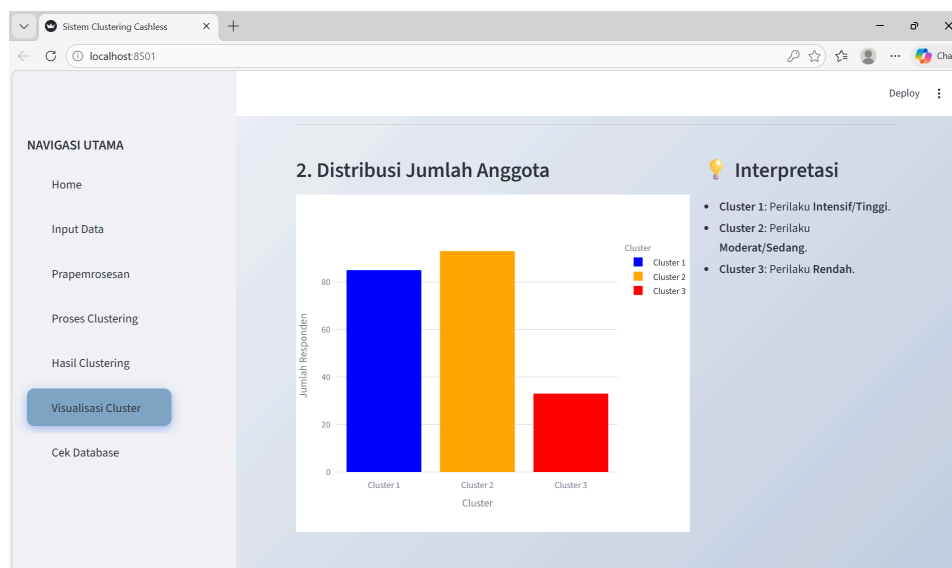
Halaman visualisasi cluster digunakan untuk menampilkan hasil analisis clustering dalam bentuk visualisasi grafik yang interaktif. Pada halaman ini terdapat beberapa jenis visualisasi yang dapat digunakan untuk membantu memahami karakteristik dari setiap cluster yang terbentuk.

Visualisasi pertama adalah *scatter plot* yang menampilkan sebaran data berdasarkan variabel yang dipilih oleh pengguna pada sumbu X dan sumbu Y. Visualisasi ini digunakan untuk melihat pola penyebaran data serta hubungan antar variabel dalam setiap cluster.



Gambar 4.7 Visualisasi Grafik *Scatter Plot*

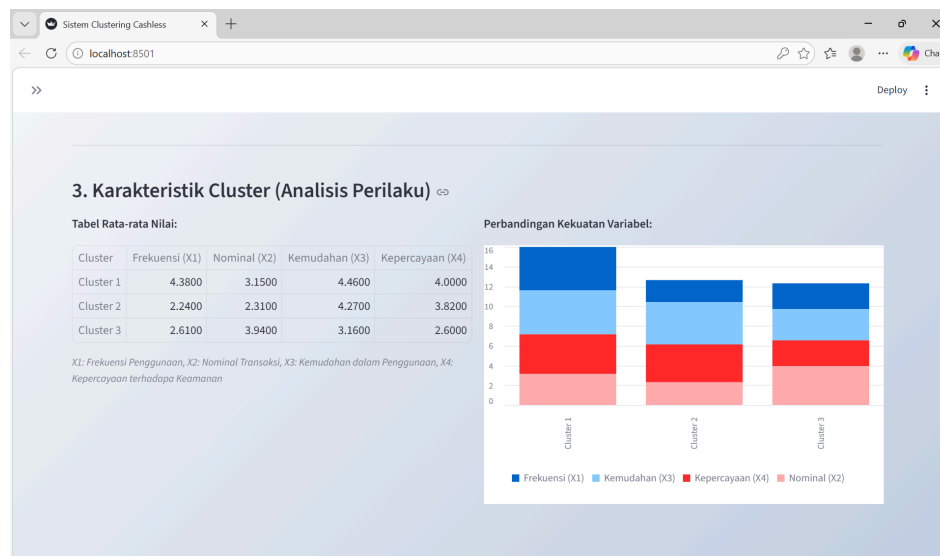
Visualisasi kedua adalah *bar chart* yang menunjukkan jumlah anggota pada setiap cluster. Grafik ini digunakan untuk mengetahui distribusi jumlah responden dalam masing-masing kelompok yang terbentuk.



Gambar 4.8 Visualisasi *Bar Chart* dan Interpretasi

Selain itu, halaman ini juga menampilkan tabel yang berisi nilai rata-rata dari setiap variabel pada masing-masing cluster. Informasi ini digunakan untuk

mengidentifikasi karakteristik dari setiap cluster sehingga dapat membantu dalam proses interpretasi hasil *clustering*.



Gambar 4.9 Visualisasi *Profiling Cluster*

Berdasarkan hasil visualisasi sistem pada Gambar 4.9, diperoleh nilai rata-rata variabel pada masing-masing cluster yang digunakan untuk mengidentifikasi karakteristik perilaku pengguna *e-wallet*. Variabel yang dianalisis meliputi frekuensi penggunaan, nominal transaksi, kemudahan penggunaan, serta kepercayaan terhadap keamanan *e-wallet* yang digunakan dalam proses pengelompokan menggunakan algoritma *K-Means Clustering*.

4.2.8 Halaman Cek Database

Halaman cek database digunakan untuk menampilkan data hasil clustering yang telah disimpan pada database sistem. Halaman ini terhubung langsung dengan database *MySQL* yang dikelola melalui *phpMyAdmin* menggunakan *XAMPP* sebagai server lokal.

Riwayat Clustering

Data di phpMyAdmin (211 baris):

	Nama Lengkap	Frekuensi (X1)	Nominal (X2)	Kemudahan (X3)	Kepercayaan (X4)	Cluster
1	Syifa Nurfadhilah	4	2	4.1667	4	Cluster 1
2	Nana Asrina	3	2	4	4	Cluster 2
3	Khusnul Khotimah	4	3	4	4.3333	Cluster 1
4	yuna srimaya berutu	2	3	4	4	Cluster 2
5	Ranida Putriani	3	4	5	5	Cluster 1
6	Sindi Putri	3	2	4.3333	3	Cluster 2
7	Maria Anastasya Br Surbakti	5	5	5	4	Cluster 1
8	nurul adha afraihan	3	2	4.5	4	Cluster 2
9	Putri Khari Ani Purba	4	2	4.6667	3.6667	Cluster 1
10	sinthia atikah	2	2	3.5	3.1667	Cluster 2

Buttons: Hapus Semua Data, Ke Home

Gambar 4.10 Tampilan Data yang Telah di *Clustering* pada Sistem

Menampilkan baris 0 - 210 (total 211, Pencarian dilakukan dalam 0,0009 detik.)

```
SELECT * FROM `clustering_cashless`
```

Tampilkan semua | Jumlah baris: 500 | Saring baris: Cari di tabel ini

Nama Lengkap	Frekuensi (X1)	Nominal (X2)	Kemudahan (X3)	Kepercayaan (X4)	Cluster
Syifa Nurfadhilah	4	2	4.16666666666667	4	Cluster 1
Nana Asrina	3	2	4	4	Cluster 2
Khusnul Khotimah	4	3	4	4.33333333333333	Cluster 1
yuna srimaya berutu	2	3	4	4	Cluster 2
Ranida Putriani	3	4	5	5	Cluster 1
Sindi Putri	3	2	4.33333333333333	3	Cluster 2
Maria Anastasya Br Surbakti	5	5	5	4	Cluster 1
nurul adha afraihan	3	2	4.5	4	Cluster 2
Putri Khari Ani Purba	4	2	4.66666666666667	3.66666666666667	Cluster 1
sinthia atikah	2	2	3.5	3.16666666666667	Cluster 2
tasya putri cahaya	5	1	3.66666666666667	4	Cluster 1
Mawar Dini	1	3	4.83333333333333	5	Cluster 2
Rizki Rayhan Damank	3	4	4	3.83333333333333	Cluster 3
Halifah Prayoga	2	3	4	4	Cluster 2
wiji lestari sitorus	3	2	3.66666666666667	3.66666666666667	Cluster 2
alqis hansapiah	4	4	4	3	Cluster 1

Gambar 4.11 Data *Clustering* yang Tersimpan pada *phpMyAdmin*

Setiap hasil proses *clustering* yang dilakukan pada sistem akan secara otomatis disimpan ke dalam *database* yang bertujuan untuk menjaga riwayat hasil analisis dapat diakses kembali apabila diperlukan pada waktu selanjutnya. Pada halaman ini juga tersedia fitur untuk menghapus seluruh data yang tersimpan pada sistem apabila diperlukan.

Dengan adanya fitur penyimpanan data ini, sistem tidak hanya mampu melakukan proses analisis clustering tetapi juga dapat menyimpan dan mengelola hasil analisis tersebut secara terstruktur.

4.3 Hasil Proses *Clustering*

Proses pengelompokan data pada penelitian ini dilakukan menggunakan algoritma *K-Means Clustering* yang diimplementasikan pada sistem yang telah dikembangkan. Data yang digunakan dalam proses clustering merupakan data hasil kuesioner responden yang telah melalui tahap prapemrosesan, yaitu proses pembersihan data (*data cleaning*) dan normalisasi.

Variabel yang digunakan dalam proses clustering meliputi frekuensi penggunaan *e-wallet*, nominal transaksi, kemudahan penggunaan, serta kepercayaan terhadap keamanan *e-wallet*. Keempat variabel tersebut dipilih karena memiliki nilai numerik yang dapat diolah dalam proses penghitungan jarak pada algoritma *K-Means*.

Pada penelitian ini jumlah cluster yang digunakan adalah $K = 3$, sehingga sistem akan mengelompokkan data responden ke dalam tiga kelompok berdasarkan tingkat kemiripan karakteristik perilaku penggunaan *e-wallet*. Setelah proses *clustering* dijalankan melalui sistem, seluruh data responden berhasil dikelompokkan ke dalam tiga cluster.

Hasil pengelompokan data responden ke dalam cluster dapat dilihat pada tabel berikut.

Tabel 4.1 Hasil Pengelompokan Data Responden

Kode Responden	Frekuensi	Nominal	Kemudahan	Kepercayaan	Cluster
R1	4	4	4.8333	3	Cluster 1

R2	1	1	4	3.5	Cluster 2
R3	3	3	4	3	Cluster 2
R4	3	3	4	4	Cluster 2
R5	3	3	4	4	Cluster 2
R6	4	2	5	3.8335	Cluster 1
R7	3	4	4.1667	4	Cluster 1
R8	3	4	4.1667	3.1665	Cluster 3
R9	4	2	4.8333	4.3333	Cluster 1
R10	2	3	4	4	Cluster 2

Selanjutnya untuk melihat distribusi jumlah anggota pada setiap cluster, dapat dilihat pada tabel berikut.

Tabel 4.2 Jumlah Anggota pada Setiap Cluster

Cluster	Jumlah Responden
Cluster 1	85 Responden
Cluster 2	93 Responden
Cluster 3	33 Responden

Berdasarkan tabel tersebut dapat diketahui bahwa sistem berhasil mengelompokkan seluruh data responden ke dalam tiga cluster yang berbeda. Setiap cluster memiliki jumlah anggota yang bervariasi sesuai dengan karakteristik perilaku penggunaan *e-wallet* dari masing-masing responden.

Hasil pengelompokan ini selanjutnya divisualisasikan dalam bentuk grafik sebaran data serta tabel rata-rata nilai variabel pada setiap cluster yang ditampilkan pada halaman visualisasi sistem. Visualisasi tersebut bertujuan untuk

mempermudah proses analisis dalam memahami perbedaan karakteristik perilaku pengguna *e-wallet* pada masing-masing cluster.

4.4 Analisis Karakteristik Cluster

Setelah proses clustering dilakukan menggunakan algoritma *K-Means*, diperoleh tiga kelompok responden yang memiliki karakteristik perilaku penggunaan *e-wallet* yang berbeda. Karakteristik setiap cluster dianalisis berdasarkan nilai rata-rata variabel yang digunakan dalam proses clustering, yaitu frekuensi penggunaan, nominal transaksi, kemudahan penggunaan, serta kepercayaan terhadap keamanan *e-wallet*.

Selain itu, indikator yang bersifat nominal seperti jenis aplikasi *e-wallet* yang digunakan serta tujuan transaksi juga digunakan sebagai informasi pendukung dalam memahami karakteristik masing-masing cluster.

4.4.1 Karakteristik Cluster 1

Cluster 1 merupakan kelompok responden yang memiliki tingkat penggunaan *e-wallet* yang relatif tinggi. Hal ini terlihat dari nilai rata-rata frekuensi penggunaan serta nominal transaksi yang cenderung lebih tinggi dibandingkan cluster lainnya. Selain itu, responden pada cluster ini juga menunjukkan tingkat kemudahan penggunaan dan kepercayaan terhadap keamanan *e-wallet* yang cukup baik.

Berdasarkan hasil tersebut, responden dalam cluster ini dapat dikategorikan sebagai pengguna aktif *e-wallet* yang telah terbiasa memanfaatkan layanan pembayaran digital dalam berbagai aktivitas transaksi sehari-hari.

Jika ditinjau dari indikator pendukung, sebagian besar responden dalam cluster ini menggunakan aplikasi *e-wallet* untuk berbagai kebutuhan seperti

pembayaran makanan, belanja online, maupun pembayaran layanan transportasi. Hal ini menunjukkan bahwa *e-wallet* telah menjadi bagian penting dalam aktivitas transaksi responden pada kelompok ini.

4.4.2 Karakteristik Cluster 2

Cluster 2 merupakan kelompok responden yang memiliki tingkat penggunaan *e-wallet* pada kategori sedang. Nilai rata-rata frekuensi penggunaan dan nominal transaksi pada cluster ini berada pada tingkat menengah dibandingkan cluster lainnya.

Responden pada cluster ini masih menggunakan *e-wallet* dalam aktivitas transaksi, namun intensitas penggunaannya tidak setinggi kelompok pengguna aktif. Selain itu, tingkat kemudahan penggunaan dan kepercayaan terhadap keamanan *e-wallet* juga berada pada kategori cukup.

Berdasarkan indikator pendukung, responden pada cluster ini umumnya menggunakan *e-wallet* untuk transaksi tertentu saja, seperti pembayaran makanan atau pembelian pulsa dan paket data. Hal ini menunjukkan bahwa penggunaan *e-wallet* pada kelompok ini bersifat situasional, yaitu hanya digunakan pada kondisi atau kebutuhan tertentu.

4.4.3 Karakteristik Cluster 3

Cluster 3 merupakan kelompok responden yang memiliki tingkat penggunaan *e-wallet* yang relatif rendah. Hal ini terlihat dari nilai rata-rata frekuensi penggunaan serta nominal transaksi yang lebih rendah dibandingkan cluster lainnya. Selain itu, responden pada cluster ini juga menunjukkan tingkat kemudahan penggunaan dan kepercayaan terhadap keamanan *e-wallet* yang

cenderung lebih rendah. Kondisi ini dapat mempengaruhi intensitas penggunaan *e-wallet* dalam aktivitas transaksi sehari-hari.

Berdasarkan indikator pendukung, sebagian responden pada cluster ini menggunakan *e-wallet* hanya untuk transaksi tertentu atau dalam frekuensi yang jarang. Hal ini menunjukkan bahwa responden pada cluster ini masih belum sepenuhnya memanfaatkan layanan *e-wallet* sebagai metode pembayaran utama. Dengan demikian, kelompok ini dapat dikategorikan sebagai pengguna *e-wallet* dengan intensitas rendah.

4.4.4 Informasi Pendukung Berdasarkan Indikator Nominal

Selain variabel numerik yang digunakan dalam proses clustering, penelitian ini turut memanfaatkan dua indikator nominal sebagai informasi pendukung, yaitu jenis aplikasi *e-wallet* dan tujuan transaksi *e-wallet*. Walaupun kedua indikator tersebut tidak dapat diolah langsung dalam perhitungan K-Means, keduanya tetap berperan penting dalam memperkuat interpretasi karakteristik setiap kelompok pengguna yang terbentuk.

Berdasarkan distribusi responden dalam masing-masing cluster, diperoleh pola pendukung sebagai berikut:

1. Cluster 1 – Pengguna Multiplatform dengan Tujuan Transaksi yang Luas

Cluster 1 memperlihatkan kecenderungan dominasi penggunaan DANA, ShopeePay, dan GoPay, baik digunakan secara tunggal maupun dalam bentuk kombinasi dua hingga tiga aplikasi. Dari sisi tujuan transaksi, kelompok ini paling banyak melakukan pembelian makanan/minuman, belanja online, transportasi online, serta transfer antar pengguna. Pola ini

menunjukkan bahwa pengguna dalam cluster ini memanfaatkan e-wallet untuk berbagai kebutuhan harian dan bersifat lebih aktif secara fungsional.

2. Cluster 2 – Pengguna Fokus dengan Aplikasi Terbatas dan Kebutuhan Transaksi Dasar

Cluster 2 didominasi oleh pengguna DANA sebagai aplikasi utama, disertai sebagian kecil penggunaan ShopeePay. Tujuan transaksi dalam cluster ini cenderung bersifat sederhana, terutama untuk pembelian makanan/minuman, transfer, dan sebagian kecil belanja online. Kelompok ini tampak memprioritaskan penggunaan yang praktis, dengan pilihan aplikasi yang terbatas dan tujuan transaksi yang tidak terlalu beragam.

3. Cluster 3 – Pengguna Variatif tetapi Tidak Dominan pada Aplikasi Tertentu

Cluster 3 menunjukkan pola yang lebih beragam dari sisi aplikasi, meliputi DANA, ShopeePay, dan GoPay, baik digunakan tunggal maupun kombinasi. Dari sisi tujuan transaksi, kelompok ini memanfaatkan e-wallet untuk belanja online, makanan/minuman, transportasi online, dan beberapa transaksi kebutuhan ringan lainnya. Tidak ada dominasi aplikasi atau tujuan tertentu, sehingga kelompok ini menggambarkan pengguna dengan pola konsumsi yang fleksibel namun tidak intensif.

Informasi berbasis indikator nominal ini membantu memberikan gambaran perilaku yang lebih menyeluruh terkait pola konsumsi digital mahasiswa. Jenis aplikasi yang digunakan memberikan konteks terkait preferensi platform, sedangkan tujuan transaksi memperlihatkan kecenderungan penggunaan e-wallet dalam aktivitas sehari-hari. Kedua informasi tersebut memperkuat interpretasi

karakteristik cluster yang terbentuk dan memberikan sudut pandang tambahan yang tidak terlihat melalui variabel numerik.

4.5 Analisis Hasil Sistem

Setelah proses clustering dilakukan menggunakan algoritma *K-Means Clustering*, sistem berhasil mengelompokkan 30 data responden ke dalam tiga cluster yang memiliki karakteristik perilaku penggunaan *e-wallet* yang berbeda. Pengelompokan dilakukan berdasarkan empat variabel numerik, yaitu frekuensi penggunaan *e-wallet*, nominal transaksi, kemudahan penggunaan, dan kepercayaan terhadap keamanan.

Secara keseluruhan, sistem dapat mengelompokkan data ke dalam tiga kelompok dengan pola perilaku yang cukup terpisah satu sama lain. Hal ini terlihat dari terbentuknya centroid yang berbeda secara signifikan pada setiap cluster, sehingga pola karakteristik yang muncul dapat dianalisis dengan jelas.

Selain menampilkan hasil pengelompokan, sistem juga dilengkapi dengan nilai *Silhouette Score* sebagai indikator kualitas clustering. Nilai ini mengukur tingkat kedekatan data dalam cluster yang sama serta tingkat pemisahan antar cluster. Semakin tinggi *Silhouette Score*, semakin baik struktur pengelompokan yang terbentuk. Hasil yang ditampilkan sistem menunjukkan bahwa pemilihan tiga cluster sudah cukup relevan dan stabil.

Analisis hasil clustering juga diperkuat dengan perhitungan *Within-Cluster Variation* (WCV) dan *Between-Cluster Variation* (BCV) yang ditampilkan pada lampiran. Nilai WCV yang rendah pada ketiga cluster mengindikasikan bahwa jarak antar data dalam cluster relatif kecil sehingga homogenitas data dalam cluster cukup baik. Sebaliknya, nilai BCV yang tinggi menunjukkan bahwa

centroid tiap cluster memiliki jarak yang signifikan terhadap pusat data keseluruhan, sehingga setiap cluster memiliki pola perilaku yang berbeda secara struktural.

Jarak antar centroid yang besar dengan nilai perbedaan rata-rata lebih dari 1.9 pada beberapa variabel penting menegaskan bahwa pemisahan antar cluster cukup jelas dan tidak terjadi tumpang tindih antar karakteristik kelompok. Dengan demikian, hasil pengelompokan yang dihasilkan sistem dapat digunakan sebagai dasar interpretasi perilaku penggunaan *e-wallet* mahasiswa secara akurat.

Secara keseluruhan, evaluasi ini menunjukkan bahwa implementasi algoritma *K-Means* pada sistem yang dikembangkan mampu bekerja secara optimal, menghasilkan pengelompokan yang konsisten, serta memberikan informasi yang bermanfaat untuk proses analisis. Sistem ini menjadikan proses pengelompokan data lebih cepat, akurat, dan efisien dibandingkan perhitungan manual.

4.6 Kelebihan dan Keterbatasan Sistem

4.6.1 Kelebihan Sistem

Sistem yang dikembangkan dalam penelitian ini memiliki beberapa kelebihan, antara lain :

1. Mampu melakukan proses clustering secara otomatis menggunakan algoritma *K-Means* sehingga proses pengelompokan data responden dapat dilakukan dengan lebih cepat dan efisien.
2. Memiliki fitur prapemrosesan data, seperti proses *cleaning* dan normalisasi data, sehingga data yang digunakan dalam proses *clustering* menjadi lebih terstruktur dan siap diolah.

3. Menyediakan visualisasi hasil clustering, sehingga pengguna dapat melihat distribusi data pada setiap cluster dengan lebih mudah.
4. Hasil *clustering* tersimpan secara otomatis dalam *database*, sehingga data dapat disimpan dan digunakan kembali untuk analisis selanjutnya.
5. Sistem memiliki antarmuka yang sederhana dan mudah digunakan, sehingga pengguna dapat menjalankan proses analisis tanpa memerlukan pemahaman teknis yang mendalam.

4.6.2 Keterbatasan Sistem

Selain memiliki beberapa kelebihan, sistem yang dikembangkan dalam penelitian ini juga memiliki beberapa keterbatasan, antara lain:

1. Sistem hanya menggunakan satu metode clustering, yaitu algoritma *K-Means*, sehingga belum dilakukan perbandingan dengan metode clustering lainnya.
2. Hasil *clustering* sangat dipengaruhi oleh jumlah cluster (K) yang ditentukan di awal, sehingga pemilihan nilai K dapat mempengaruhi hasil pengelompokan data.
3. Sistem hanya mengolah variabel yang bersifat numerik dalam proses *clustering*, sedangkan variabel yang bersifat nominal hanya digunakan sebagai informasi pendukung dalam analisis hasil.
4. Data yang digunakan dalam penelitian ini masih terbatas pada responden mahasiswa di Kota Medan, sehingga hasil *clustering* belum tentu dapat merepresentasikan perilaku seluruh pengguna *e-wallet* secara umum.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai analisis pengelompokan perilaku penggunaan transaksi *cashless* berbasis *e-wallet* di kalangan mahasiswa Kota Medan menggunakan algoritma *K-Means Clustering*, maka dapat diperoleh beberapa kesimpulan sebagai berikut:

1. Algoritma K-Means berhasil digunakan untuk membentuk perilaku penggunaan e-wallet berdasarkan variabel frekuensi penggunaan (X1), nominal transaksi (X2), kemudahan penggunaan (X3), dan kepercayaan terhadap keamanan (X4). Sistem yang dikembangkan mampu menjalankan seluruh tahapan algoritma, mulai dari prapemrosesan data, inisialisasi centroid, perhitungan jarak, pembaruan centroid, hingga menghasilkan cluster yang stabil (konvergen). Proses ini terotomasi melalui aplikasi yang dibangun, sehingga mempermudah analisis terhadap data kuesioner mahasiswa.
2. Hasil pengelompokan membentuk tiga kelompok (cluster) dengan karakteristik perilaku yang berbeda berdasarkan nilai rata-rata tiap variabel. Karakteristik masing-masing cluster adalah sebagai berikut:

- a. Cluster 1

Memiliki nilai rata-rata frekuensi penggunaan sebesar 4.38, nominal transaksi 3.15, kemudahan penggunaan 4.46, dan kepercayaan 4.00. Cluster ini menggambarkan kelompok mahasiswa dengan intensitas penggunaan relatif tinggi, merasa

sangat mudah dalam menggunakan e-wallet, dan memiliki tingkat kepercayaan yang baik terhadap keamanan layanan.

b. Cluster 2

Memiliki rata-rata frekuensi penggunaan 2.24, nominal transaksi 32.31, kemudahan penggunaan 4.27, dan kepercayaan 3.82. Cluster ini mewakili kelompok mahasiswa yang jarang menggunakan e-wallet, namun memiliki nilai kepercayaan dan kemudahan yang sangat tinggi, serta nominal transaksi yang cenderung lebih besar. Kelompok ini dapat dikategorikan sebagai pengguna pasif namun percaya dan terbiasa dengan e-wallet.

c. Memiliki rata-rata frekuensi penggunaan 2.61, nominal transaksi 3.94, kemudahan penggunaan 3.16, dan kepercayaan 2.60. Cluster ini merupakan kelompok dengan tingkat penggunaan rendah pada semua variabel, baik dari sisi frekuensi, nominal transaksi, maupun persepsi kemudahan dan kepercayaan. Kelompok ini menggambarkan pengguna dengan intensitas paling rendah.

3. Sistem yang dikembangkan mampu menampilkan visualisasi dan penyimpanan hasil cluster secara terstruktur, sehingga mendukung proses interpretasi perilaku pengguna e-wallet.
4. Fitur visualisasi seperti scatter plot, distribusi anggota cluster, dan tabel karakteristik cluster memudahkan pengguna dalam memahami perbedaan antar kelompok. Selain itu, penyimpanan otomatis ke dalam basis data MySQL memungkinkan hasil analisis terdokumentasi dengan baik dan dapat digunakan untuk keperluan lanjutan.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya, yaitu sebagai berikut:

1. Bagi peneliti selanjutnya, disarankan untuk menambah variabel penelitian agar hasil pengelompokan lebih komprehensif. Variabel seperti tingkat literasi keuangan digital, pengalaman penggunaan e-wallet, atau tingkat kebutuhan transaksi dapat memberikan gambaran perilaku pengguna yang lebih mendalam. Selain itu, variabel non-numerik dapat dipertahankan sebagai informasi tambahan, misalnya jenis e-wallet atau tujuan transaksi, untuk memperkaya interpretasi karakteristik cluster.
2. Sistem dapat dikembangkan dengan menambahkan fitur pengujian evaluasi cluster yang lebih beragam. Saat ini sistem telah menampilkan nilai Silhouette Score sebagai ukuran kualitas cluster, namun penelitian selanjutnya dapat menambahkan metode evaluasi lain seperti Davies–Bouldin Index atau Calinski–Harabasz Index agar penentuan jumlah cluster optimal dapat dianalisis lebih akurat.
3. Proses prapemrosesan data dapat dikembangkan menjadi lebih fleksibel dan adaptif. Misalnya, menyediakan opsi normalisasi lain (*Z-Score*, *Robust Scaler*), atau menyediakan fitur deteksi outlier sehingga kualitas data yang masuk ke proses clustering semakin baik dan minim bias.
4. Sistem dapat diperluas menjadi platform analisis data terintegrasi. Pengembangan lanjutan dapat mencakup pembuatan dashboard interaktif berbasis web, integrasi API dengan sistem pengumpulan data otomatis,

serta fitur pelaporan hasil clustering dalam bentuk file PDF atau Excel yang dapat diunduh oleh pengguna.

5. Disarankan untuk melakukan pengujian dengan jumlah responden yang lebih besar dan beragam. Jumlah sampel yang lebih besar akan membuat hasil cluster lebih stabil dan representatif terhadap populasi mahasiswa di Kota Medan. Selain itu, penelitian lintas perguruan tinggi dapat memberikan insight yang lebih luas terkait pola penggunaan e-wallet di kalangan mahasiswa.
6. Sistem penyimpanan data dapat ditingkatkan pada sisi keamanan dan skalabilitas. Penggunaan database server yang lebih kuat seperti PostgreSQL atau sistem penyimpanan berbasis cloud dapat dipertimbangkan apabila sistem ini akan dikembangkan untuk skala instansi atau digunakan oleh lebih banyak pengguna.

DAFTAR PUSTAKA

- Amin, F., Yarisma, F. W., Setiawati, I., & Ilham, R. M. (2025). CLUSTER ANALYSIS OF DANA AND GOPAY E-WALLET SERVICE REVIEWS USING TF-IDF AND K-MEANS/K-MEDOIDS: AN ECONOMIC AND MANAGEMENT STRATEGY PERSPECTIVE. *STABILITY: Journal of Management and Business*, 8(2). <https://doi.org/10.26877/r7hyad25>
- Bank Indonesia. (2019). *Blueprint Sistem Pembayaran Indonesia*. Bank Indonesia.
<https://www.bi.go.id/id/publikasi/kajian/Documents/Blueprint-Sistem-Pembayaran-Indonesia-2030.pdf>
- Bank Indonesia. (2023). *Atas Inovasi dalam Efisiensi Operasional, QRIS Mengukir Prestasi di Asia Pasifik*.
https://www.bi.go.id/id/publikasi/ruang-media/news-release/Pages/sp_2531823.aspx
- Bank Indonesia. (2024a). *Kanal Layanan Sistem Pembayaran Ritel*. Bank Indonesia.
<https://www.bi.go.id/id/fungsi-utama/sistem-pembayaran/ritel/kanal-layanan/default.aspx#Dompel-Elektronik>
- Bank Indonesia. (2024b). *Laporan Kebijakan Moneter Triwulan IV 2024*. Bank Indonesia.
<https://www.bi.go.id/id/publikasi/laporan/Pages/Laporan-Kebijakan-Moneter-Triwulan-IV-2024.aspx>
- Goh, M. L., Chow, M. M., Gan, S.-W., & Tang, C. X. (2025). Predicting E-wallet Continuation Behaviour Among University Students: Testing an Integrated

- Theoretical Framework. *Journal of Logistics, Informatics and Service Science*. <https://doi.org/10.33168/JLISS.2025.0217>
- Han, J., Kamber, M., & Pei, J. (2022). *Data Mining: Concepts and Techniques (4th ed.)*. <https://www.sciencedirect.com/book/9780128117606/data-mining-concepts-and-techniques>
- Huh, I., & Gim, J. (2025). Exploration of Likert scale in terms of continuous variable with parametric statistical methods. *BMC Medical Research Methodology*, 25(1), 218. <https://doi.org/10.1186/s12874-025-02668-1>
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210. <https://doi.org/10.1016/j.ins.2022.11.139>
- Lee, W. Q., & Thoo, A. C. (2022). Extended Technology Acceptance Model for Digital Wallet Usage. *International Journal of Academic Research in Business and Social Sciences*, 12(8), Pages 1490-1503. <https://doi.org/10.6007/IJARBSS/v12-i8/14416>
- Mauritius, T., & Binsar, F. (2020). *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. <https://mmsi.binus.ac.id/2020/09/18/cross-industry-standard-process-for-data-mining-crisp-dm/>
- Maylawati, D. S., Darmalaksana, W., & Ramdhani, M. A. (2018). Systematic Design of Expert System Using Unified Modelling Language. *IOP*

Conference Series: Materials Science and Engineering, 288, 012047.
<https://doi.org/10.1088/1757-899X/288/1/012047>

Muis, A., Syafwan, H., Arfianto, A. Z., Simanjuntak, M. S., Riyandi, A., Trisnawan, A. B., Ramdhan, W., Triana, H., Saputra, M. H., Handoko, W., Naswin, A., & Suhadi, S. (2025). *DATA MINING: Konsep, Metode, dan Aplikasi*. Faaslib Serambi Media.

Nababan, A. A., Hutabarat, W. B., Simanjuntak, R. P., & Fajri, M. Y. (2025). Analisis Pengaruh Penggunaan e-wallet Terhadap Pola Konsumsi Mahasiswa Universitas Negeri Medan di Era Digital: Penelitian. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 4(1), 7082–7089.
<https://doi.org/10.31004/jerkin.v4i1.2943>

Narulita, S., Nugroho, A., & Abdillah, M. Z. (2024). Diagram Unified Modelling Language (UML) untuk Perancangan Sistem Informasi Manajemen Penelitian dan Pengabdian Masyarakat (SIMLITABMAS). *Bridge : Jurnal Publikasi Sistem Informasi Dan Telekomunikasi*, 2(3), 244–256.
<https://doi.org/10.62951/bridge.v2i3.174>

Paramita, E. D., & Cahyadi, E. R. (2024). The Determinants of Behavioral Intention and Use Behavior of QRIS as Digital Payment Method Using Extended UTAUT Model. *Indonesian Journal of Business and Entrepreneurship*. <https://doi.org/10.17358/ijbe.10.1.132>

Ramadanti, H. R. A. S., Nawir, J., & Marlina. (2021). ANALISIS PERILAKU KEUANGAN GENERASI Z PADA CASHLESS SOCIETY: Literasi Keuangan, Gaya Hidup, Perilaku Keuangan, Cashless Society. *Jurnal Visionida*, 7(2), 96–109. <https://doi.org/10.30997/jvs.v7i2.4488>

- Ramli, F. A. A., & Hamzah, M. I. (2021). Mobile payment and e-wallet adoption in emerging economies: A systematic literature review. *Journal of Emerging Economies and Islamic Research*. <https://doi.org/10.24191/jeeir.v9i2.13617>
- Rokeman, N. R. M. (2024). Likert Measurement Scale in Education and Social Sciences: Explored and Explained. *EDUCATUM Journal of Social Sciences*, (10), 77–88. <https://doi.org/10.37134/ejoss.vol10.1.7.2024>
- Suryani, E., Ginting, Julia Kristina Br, & Tarigan, D. A. (2024). The role of digital finance in increasing financial access for Medan Area University (students:case study of Payment Gateways Implemetation). *Internasional Journal of Economic, Agribisnis and Development Studies*, 1(1), 1–4. <https://doi.org/10.65142/ijeads.v1i1.2>

LAMPIRAN

Lampiran 1 Surat Penetapan Dosen Pembimbing



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/BAN-PT/Ak.Ppy/PT/2024
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax (061) 6625474 - 6631003

 <http://fki.umsu.ac.id>
 fki@umsu.ac.id
 [umsu.medan](https://www.facebook.com/umsu.medan)
 [umsu.medan](https://www.instagram.com/umsu.medan)
 [umsu.medan](https://twitter.com/umsu.medan)
 [umsu.medan](https://www.youtube.com/umsu.medan)

PENETAPAN DOSEN PEMBIMBING
PROPOSAL/SKRIPSI MAHASISWA
NOMOR : 1030/IL3-AU/UMSU-09/F/2025

Assalamu'alaikum Warahmatullahi Wabarakatuh

Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, berdasarkan Persetujuan permohonan judul penelitian Proposal / Skripsi dari Ketua / Sekretaris.

Program Studi : Sistem Informasi
Pada tanggal : 06 November 2025

Dengan ini menetapkan Dosen Pembimbing Proposal / Skripsi Mahasiswa.

Nama : Despita Saragih
NPM : 2209010191
Semester : VII (Tujuh)
Program studi : Sistem Informasi
Judul Proposal / Skripsi : Analisis Pengelompokan Perilaku Penggunaan Transaksi Cashless di Kalangan Mahasiswa Kota Medan Menggunakan Algoritma K-Means Clustering

Dosen Pembimbing : Dr. Al-Khowarizmi, M.Kom

Dengan demikian di izinkan menulis Proposal / Skripsi dengan ketentuan

1. Penulisan berpedoman pada buku panduan penulisan Proposal / Skripsi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU
2. Pelaksanaan Sidang Skripsi harus berjarak 3 bulan setelah dikeluarkannya Surat Penetapan Dosen Pembimbing Skripsi.
3. **Proyek Proposal / Skripsi** dinyatakan " **BATAL** " bila tidak selesai sebelum Masa Kadalua tanggal : **06 November 2026**
4. Revisi judul.....

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Ditetapkan di : Medan
 Pada Tanggal : 15 Jumadil Awwal 1447 H
 06 November 2025 M




Dr. Al-Khowarizmi, M.Kom.
NIDN : 0127099201

Cc. File






Lampiran 2 Lembar Validasi Ahli Instrumen Kuesioner Penelitian

LEMBAR VALIDASI AHLI (EXPERT JUDGMENT) INSTRUMEN KUESIONER PENELITIAN

A. Identitas Penelitian

Judul Penelitian : Analisis Pengelompokan Perilaku Penggunaan Transaksi Cashless di Kalangan Mahasiswa Kota Medan Menggunakan Algoritma K-Means Clustering.
 Nama Peneliti : Despita Saragih
 NPM : 2209010191
 Program Studi : Sistem Informasi

B. Identitas Pakar/Validator

Nama Pakar : Agus Sani
 Bidang Keahlian : Manajemen strategis pada prospektif perilaku.
 Institusi : Universitas Muhammadiyah Sumatera Utara (UMSU)
 Jabatan : Ka. Prodi Manajemen
 Tanggal Validasi : 29/01/2026

C. Petunjuk Pengisian

Bapak/Ibu dimohon untuk memberikan penilaian terhadap setiap butir pernyataan kuesioner dengan memberi tanda centang (✓) pada kolom yang sesuai, serta memberikan saran atau catatan perbaikan apabila diperlukan.

Keterangan penilaian :

1. Relevan : pernyataan sesuai dengan indikator yang diukur
2. Jelas : pernyataan mudah dipahami dan tidak ambigu
3. Perlu Revisi : Pernyataan perlu diperbaiki redaksinya

D. Tabel Validasi Pernyataan

Indikator : Kemudahan Penggunaan E-Wallet

No	Pernyataan	Relevan	Jelas	Perlu Revisi	Keterangan
1	Aplikasi e-wallet mudah dipelajari oleh pengguna baru	✓			

2	Saya dapat menggunakan e-wallet tanpa bantuan orang lain	✓			
3	Proses pembayaran menggunakan e-wallet mudah dipahami	✓			
4	Menu dan fitur dalam aplikasi e-wallet tersusun dengan jelas	✓			
5	Saya jarang mengalami kesulitan saat menggunakan e-wallet	✓			
6	Secara keseluruhan, e-wallet mudah digunakan	✓			

Indikator : Kepercayaan Terhadap Keamanan E-Wallet

No	Pernyataan	Relevan	Jelas	Perlu Revisi	Keterangan
1	Saya percaya e-wallet aman digunakan untuk transaksi keuangan	✓			
2	Data pribadi saya terlindungi saat menggunakan e-wallet	✓			
3	Saya merasa aman menyimpan saldo dalam e-wallet	✓			

4	Sistem keamanan e-wallet dapat mencegah penyalahgunaan akun	✓			
5	Saya percaya transaksi e-wallet terlindungi dari penipuan	✓			
6	Secara keseluruhan, saya mempercayai keamanan e-wallet	✓			

E. Kesimpulan Validator

Berdasarkan hasil penilaian, instrumen kuesioner penelitian dinyatakan :

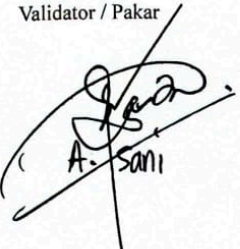
- Layak digunakan tanpa revisi
 Layak digunakan dengan revisi
 Tidak Layak digunakan

Catatan/Saran Umum :

.....

Demikian lembar validasi ini dibuat untuk digunakan sebagaimana mestinya.

Medan,
 Validator / Pakar

()
 A. Sani

Lampiran 3 Berita Acara Bimbingan Skripsi



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/III/2019
 Pusat Administrasi: Jalan Mukhtar Baerli No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
<http://www.umsu.ac.id> ✉ info@umsu.ac.id [umsumedan](#) [umsumedan](#) [umsumedan](#) [umsumedan](#)

Berita Acara Pembimbingan Skripsi

Nama Mahasiswa : Despita Saragih Program Studi : Sistem Informasi
 NPM : 2209010191 Judul Penelitian : Analisis Pola Pengelompokan Perilaku Penggunaan Transaksi Cashless di Kalangan Mahasiswa Medan Menggunakan Algoritma K-means Clustering

Nama Dosen Pembimbing : Dr. Al-Khowarizmi, M.Kom

Tanggal Bimbingan	Hasil Evaluasi	Paraf Dosen
Jum'at, 27 Feb 2026	Konultasi terkait saran perbaikan Skripsi berdasarkan masukan dosen Penguji pada saat Seminar proposal.	
Senin, 9 Mar 2026	Melakukan exhibition sistem yang telah dikembangkan serta meminta masukan dan saran untuk penyempurnaan sistem	
Kamis, 12 Mar 2026	Bimbingan penyusunan Skripsi, serta perbaikan dengan penambahan pada bagian kesimpulan.	
12 Maret 2026	Acc Sidang	

Medan, 12 Maret 2026

Diketahui oleh :
 Ketua Program Studi

(Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom)

Disetujui oleh :
 Dosen Pembimbing

(Dr. Al-Khowarizmi, M.Kom)



Lampiran 4 Hasil Cek Turnitin (Plagiarisme)

abstrak---bab-v_1778660444441.pdf

ORIGINALITY REPORT

22%
SIMILARITY INDEX

PRIMARY SOURCES

1	repository.umsu.ac.id Internet	290 words — 2%
2	eprints.upj.ac.id Internet	104 words — 1%
3	repository.ub.ac.id Internet	82 words — 1%
4	docplayer.info Internet	71 words — < 1%
5	dspace.uui.ac.id Internet	52 words — < 1%
6	Rahmat Hidayat Sanusi, Audy A. Kenap, Efraim R. S. Moningkey. "Implementasi Algoritma Dijkstra Untuk Menentukan Rute Terpendek Perjalanan Wisata Di Kabupaten Muna", RIGGS: Journal of Artificial Intelligence and Digital Business, 2026 Crossref	48 words — < 1%
7	repository.itpln.ac.id Internet	47 words — < 1%
8	repository.its.ac.id Internet	47 words — < 1%