

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemudahan akses penggunaan dompet digital pada perangkat seluler pintar telah memberikan perubahan signifikan terhadap pola transaksi masyarakat. Kehadiran layanan pembayaran digital memungkinkan pengguna melakukan transaksi secara lebih cepat, praktis, dan efisien. Namun, kemudahan tersebut juga menimbulkan tantangan baru dalam pengelolaan keuangan pribadi karena meningkatnya jumlah transaksi digital yang dilakukan secara rutin. Tanpa adanya pengelolaan keuangan yang baik, pengguna dapat mengalami kesulitan dalam mengontrol pola pengeluaran dan mengevaluasi kondisi finansialnya. Padahal, stabilitas ekonomi individu sangat dipengaruhi oleh kemampuan dalam menerapkan sistem pengelolaan keuangan pribadi yang terencana dan terstruktur [1]. Oleh karena itu, diperlukan suatu sistem pengelolaan keuangan yang mampu membantu pengguna melakukan pencatatan, pemantauan, serta analisis terhadap aktivitas keuangan sehingga pengguna dapat memperoleh informasi yang lebih baik mengenai kondisi finansialnya [2].

Ekosistem pembayaran di Indonesia mengalami transformasi yang signifikan seiring dengan meningkatnya penggunaan layanan pembayaran digital, khususnya dompet digital. Indonesia menjadi salah satu pasar pembayaran digital yang berkembang pesat di Asia Tenggara dengan berbagai *platform* seperti GoPay, DANA, ShopeePay, OVO, dan LinkAja yang telah digunakan dalam aktivitas transaksi masyarakat sehari-hari [3]. Selain itu, implementasi *Quick Response Code Indonesian Standard* (QRIS) turut mempercepat adopsi transaksi non-tunai dengan

meningkatnya jumlah transaksi digital di berbagai sektor [4]. Perkembangan tersebut menunjukkan bahwa masyarakat semakin terbiasa menggunakan layanan digital dalam melakukan aktivitas keuangan, sehingga menghasilkan jumlah data transaksi elektronik yang semakin besar dari waktu ke waktu.

Di balik peningkatan penggunaan transaksi digital tersebut, muncul permasalahan baru dalam pengelolaan keuangan pribadi. Pengguna sering kali memiliki lebih dari satu akun dompet digital untuk memenuhi kebutuhan transaksi yang berbeda, sehingga informasi saldo dan riwayat transaksi tersebar pada berbagai platform. Setiap aplikasi dompet digital umumnya hanya menyediakan informasi transaksi pada ekosistemnya masing-masing, sehingga pengguna perlu membuka beberapa aplikasi secara terpisah untuk memperoleh gambaran mengenai keseluruhan aktivitas keuangannya. Kondisi tersebut menyebabkan pengguna mengalami kesulitan dalam melakukan pemantauan keuangan secara menyeluruh karena tidak tersedianya informasi transaksi yang terintegrasi dalam satu sistem.

Sebagai alternatif, sebagian pengguna masih memanfaatkan metode pencatatan manual maupun aplikasi pencatat keuangan sederhana untuk mengelola transaksi pribadi. Namun, pendekatan tersebut memiliki beberapa keterbatasan karena proses pencatatan sangat bergantung pada kedisiplinan pengguna, membutuhkan waktu tambahan, serta memiliki risiko terjadinya transaksi yang tidak tercatat. Akibatnya, informasi keuangan yang diperoleh dapat menjadi kurang lengkap dan kurang akurat dalam menggambarkan kondisi finansial pengguna [5].

Selain permasalahan pencatatan, banyaknya transaksi digital yang tersimpan pada berbagai *platform* belum sepenuhnya dimanfaatkan sebagai sumber informasi untuk memahami perilaku finansial pengguna. Sebagian pengguna hanya

menjadikan riwayat transaksi sebagai arsip pembayaran tanpa melakukan evaluasi terhadap pola pengeluaran, kategori transaksi, maupun kebiasaan finansial yang terbentuk. Akibatnya, pengguna dapat mengalami kesulitan dalam mengevaluasi kondisi finansialnya sehingga proses pengambilan keputusan keuangan belum sepenuhnya didukung oleh hasil analisis terhadap data transaksi yang dimiliki. Kondisi tersebut berkaitan dengan hasil Survei Nasional Literasi dan Inklusi Keuangan (SNLIK) tahun 2024 yang menunjukkan bahwa tingkat inklusi keuangan masyarakat Indonesia mencapai 80,51%, sedangkan tingkat literasi keuangan berada pada angka 66,46% [6]. Perbedaan tingkat tersebut menunjukkan bahwa meningkatnya akses masyarakat terhadap layanan keuangan digital belum sepenuhnya diikuti dengan kemampuan dalam memahami, mengelola, serta memanfaatkan informasi keuangan secara optimal.

Berdasarkan permasalahan tersebut, diperlukan suatu sistem yang mampu melakukan *Monitoring* transaksi secara *real-time* dengan memanfaatkan informasi transaksi digital yang diterima melalui perangkat pengguna. *Monitoring* transaksi diperlukan agar setiap aktivitas keuangan dapat tercatat secara lebih efektif tanpa sepenuhnya bergantung pada proses input manual. Data transaksi yang terkumpul selanjutnya dapat dimanfaatkan tidak hanya sebagai riwayat keuangan, tetapi juga sebagai sumber informasi untuk melakukan analisis terhadap pola pengeluaran pengguna. Salah satu pendekatan yang dapat digunakan adalah *profiling* finansial berdasarkan karakteristik transaksi dengan metode machine learning. Pada penelitian sebelumnya, metode *K-Means Clustering* digunakan untuk melakukan pengelompokan pengguna berdasarkan pola transaksi keuangan sehingga dapat menghasilkan kategori perilaku finansial tertentu [7]. Melalui pendekatan tersebut,

sistem dapat memberikan gambaran mengenai karakteristik finansial pengguna berdasarkan pola pengeluaran yang terbentuk. Hasil pengelompokan tersebut selanjutnya dapat diinterpretasikan menjadi Profil Finansial pengguna berdasarkan karakteristik masing-masing kluster.

Selain menghasilkan profil finansial, sistem pengelolaan keuangan modern juga perlu memiliki kemampuan dalam memberikan dukungan keputusan secara lebih cerdas. Konsep *Financial Intelligent Assistant* memungkinkan sistem tidak hanya berfungsi sebagai media pencatatan transaksi, tetapi juga mampu memberikan *insight* keuangan berdasarkan hasil analisis data pengguna. Berdasarkan penelitian terkait *AI Based Personal Finance Management System*, penerapan kecerdasan buatan pada pengelolaan keuangan dapat digunakan untuk menghasilkan *insight* keuangan personal, memberikan peringatan terhadap pola pengeluaran tertentu, serta membantu pengguna dalam melakukan pengendalian keuangan secara lebih efektif [8]. Dengan demikian, sistem dapat berperan sebagai pendukung keputusan yang membantu pengguna memahami kondisi finansialnya melalui informasi dan *insight* yang dihasilkan dari data transaksi. Selain analisis perilaku finansial, informasi transaksi juga dapat dimanfaatkan untuk membantu pengguna memantau pengeluaran terhadap batas anggaran yang telah ditentukan. Pemantauan tersebut dapat memberikan peringatan ketika pengeluaran telah mendekati atau mencapai batas anggaran sehingga pengguna dapat mengetahui kondisi penggunaan anggarannya secara lebih cepat.

Meskipun berbagai penelitian telah mengembangkan aplikasi personal *finance management*, *expense tracker*, serta sistem analisis perilaku keuangan berbasis machine learning, sebagian besar penelitian tersebut masih memiliki fokus

yang terbatas pada fungsi tertentu. Beberapa penelitian berfokus pada pencatatan transaksi dan pengelolaan anggaran, sedangkan penelitian lainnya lebih menitikberatkan pada analisis pola perilaku finansial atau pemberian rekomendasi berbasis kecerdasan buatan. Berdasarkan kajian penelitian terdahulu, masih terdapat keterbatasan dalam mengintegrasikan *Monitoring* transaksi digital dari berbagai sumber *E-Wallet*, *profiling* finansial menggunakan metode *K-Means Clustering*, serta *Financial Intelligent Assistant* yang mampu menghasilkan *insight* dan peringatan pengeluaran dalam satu sistem yang terintegrasi [7], [8]. Keterbatasan tersebut menjadi *research gap* yang mendasari pengembangan aplikasi FINSIA.

Berdasarkan permasalahan dan *research gap* tersebut, penelitian ini mengembangkan aplikasi FINSIA (*Financial Intelligent Assistant*) sebagai sistem pendukung keputusan keuangan pribadi yang mengintegrasikan *Monitoring* transaksi digital melalui pembacaan notifikasi transaksi, *profiling* finansial menggunakan metode *K-Means Clustering* yang diinterpretasikan ke dalam Profil Finansial pengguna, serta penyediaan *insight* keuangan dan *spending alert* berdasarkan hasil analisis perilaku transaksi pengguna. Melalui integrasi tersebut, FINSIA diharapkan mampu mengurangi ketergantungan terhadap pencatatan transaksi secara manual, menyediakan gambaran keuangan dari berbagai sumber transaksi digital, membantu pengguna memahami karakteristik perilaku finansialnya, mengendalikan pengeluaran secara lebih bijaksana, serta mendukung proses pengambilan keputusan keuangan yang lebih efektif dan berkelanjutan.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dipaparkan, rumusan masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana merancang dan membangun aplikasi FINSIA yang mampu melakukan *Monitoring* transaksi keuangan secara *real-time* melalui pembacaan notifikasi transaksi digital pada perangkat *Android*?
2. Bagaimana menerapkan metode *K-Means Clustering* pada aplikasi FINSIA untuk melakukan *profiling* finansial berdasarkan pola transaksi pengguna?
3. Bagaimana mengembangkan fitur *Financial Intelligent Assistant* pada aplikasi FINSIA yang mampu memberikan *insight* keuangan dan *spending alert* berdasarkan hasil analisis perilaku transaksi pengguna?

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini berfokus pada pengambilan informasi transaksi dari notifikasi aplikasi dompet digital GoPay, DANA, dan ShopeePay pada perangkat Android menggunakan *Accessibility Service* dan *Notification Listener Service*.
2. Sistem yang dikembangkan hanya melakukan *Monitoring* dan pencatatan informasi transaksi digital yang diterima melalui perangkat pengguna, serta tidak melakukan integrasi langsung dengan server atau API resmi penyedia layanan dompet digital.
3. Penerapan algoritma *K-Means Clustering* memanfaatkan data transaksi yang telah dikumpulkan dan diolah dengan variabel masukan berupa total pengeluaran bulanan, jumlah transaksi, serta persentase alokasi pengeluaran pada tiga kategori utama, yaitu *Persen Kebutuhan*, *Persen Hiburan*, dan *Persen Belanja*.
4. Hasil *profiling* perilaku keuangan pengguna pada penelitian ini dibatasi menjadi tiga kategori profil finansial, yaitu *Hemat*, *Stabil*, dan *Konsumtif*.

5. Aplikasi ini hanya berfungsi sebagai sistem *Monitoring*, analisis, dan pendukung keputusan keuangan pribadi, sehingga tidak melakukan transaksi pembayaran, pemindahan saldo, maupun perubahan data pada aplikasi dompet digital pengguna.

1.4 Tujuan dan Manfaat

Berikut adalah tujuan dan manfaat penelitian ini antara lain:

1.4.1 Tujuan Penelitian

Adapun tujuan dari pelaksanaan penelitian ini adalah:

1. Merancang dan membangun aplikasi FINSIA yang mampu melakukan *Monitoring* transaksi keuangan secara *real-time* melalui pembacaan notifikasi transaksi digital pada perangkat *Android*.
2. Mengimplementasikan metode *K-Means Clustering* untuk melakukan *profiling* finansial pengguna berdasarkan karakteristik pola transaksi sehingga dapat menghasilkan kategori perilaku keuangan pengguna.
3. Mengembangkan fitur *Financial Intelligent Assistant* yang mampu memberikan *insight* keuangan serta *spending alert* berdasarkan hasil analisis pola transaksi pengguna.

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah memberikan kemudahan bagi pengguna dalam melakukan pengelolaan keuangan pribadi melalui aplikasi yang mampu melakukan *Monitoring* transaksi digital secara *real-time* dan mengurangi ketergantungan terhadap pencatatan transaksi secara manual. Melalui penerapan metode *K-Means Clustering*, aplikasi FINSIA dapat membantu pengguna memahami karakteristik perilaku finansial berdasarkan pola transaksi yang

terbentuk melalui proses *profiling* finansial. Selain itu, fitur *Financial Intelligent Assistant* mampu menyajikan *insight* keuangan serta peringatan pengeluaran (*spending alert*) sehingga pengguna dapat mengetahui kondisi pola pengeluaran dan penggunaan anggarannya sebagai informasi pendukung dalam melakukan evaluasi keuangan.

1.5 Tinjauan Pustaka

Tinjauan pustaka ini menguraikan penelitian yang berkaitan dengan pengembangan sistem pengelolaan keuangan pribadi (*personal finance*), otomatisasi pencatatan transaksi, pemantauan notifikasi *real-time*, pengelompokan perilaku belanja (*clustering*), serta penggunaan asisten keuangan berbasis kecerdasan buatan (*AI assistant*). Penelitian terdahulu digunakan untuk mengetahui perkembangan penerapan teknologi informasi dalam domain finansial pribadi dan menentukan posisi penelitian yang dilakukan.

1.5.1 Pelacakan Keuangan dan Sistem Penganggaran Berbasis Web

Pengembangan *Personal Finance Tracker* berbasis web memanfaatkan Angular untuk antarmuka pengguna, *ASP.NET Core* untuk *backend* API, serta *SQL Server* sebagai basis data. Fitur utamanya meliputi sistem penganggaran (*budgeting*), peringatan batas pengeluaran (*budget alerts*), dan visualisasi data keuangan menggunakan *Chart.js*. Inovasi ini terbukti membantu pengguna mengurangi ketergantungan pada pencatatan tradisional. Namun, karena berbasis aplikasi web, sistem ini memiliki keterbatasan teknis yaitu tidak dapat membaca notifikasi transaksi dari dompet digital pada ponsel pintar secara otomatis, sehingga seluruh data transaksi masih harus dimasukkan secara manual oleh pengguna [2].

1.5.2 Pelacakan Pengeluaran Cerdas Menggunakan OCR dan API Perbankan

Pengembangan aplikasi mobile *Smart Expense Tracker* menggunakan *frontend React/Node.js*, basis data *MySQL/Firebase*, serta pustaka *machine learning TensorFlow* dan *Scikit-Learn* untuk analisis prediksi. Aplikasi ini menawarkan pelacakan pengeluaran, kategorisasi transaksi otomatis, dan visualisasi data. Untuk meminimalkan input manual, sistem ini memanfaatkan teknologi *Tesseract OCR* untuk membaca struk belanja serta integrasi API (*Plaid* dan *Razorpay*) untuk impor data perbankan. Namun, aplikasi ini belum mendukung fitur pembacaan notifikasi transaksi otomatis di latar belakang (*background*) ponsel. Selain itu, meskipun memiliki fitur analisis prediksi pengeluaran, aplikasi ini belum dilengkapi dengan algoritma klasterisasi perilaku belanja pengguna (*user profiling*) maupun *Financial Intelligent Assistant* yang menghasilkan *insight* berdasarkan karakteristik perilaku pengguna [9].

1.5.3 Pelacakan Pengeluaran Bersama Multi-Akun

Pengembangan dan evaluasi aplikasi *Moneyge It* dirancang khusus untuk pelacakan pengeluaran bersama (*shared expense tracking*). Aplikasi mobile yang dibangun menggunakan *React Native*, *SQLite*, dan *Firestore* ini memungkinkan pengguna mencatat transaksi dari berbagai sumber akun keuangan (*cash*, *E-Wallet*, dan rekening bank) serta membagikan catatan tersebut dengan kolaborator. Hasil pengujian menunjukkan bahwa fitur pencatatan multi-akun sangat membantu pengguna memantau aliran dana dan meningkatkan kesadaran keuangan mereka. Namun, pencatatan transaksi pada aplikasi ini masih harus dimasukkan secara manual oleh pengguna, dan

sistem ini belum dilengkapi dengan pemetaan profil perilaku belanja pengguna secara otomatis [10].

1.5.4 Pemantauan Transaksi Otomatis *Real-Time* pada Aplikasi Mobile

Pengembangan aplikasi *Mobile Point of Sale (mPOS)* berbasis *React Native* yang terintegrasi dengan *Firebase Realtime Database* dirancang untuk mendeteksi perubahan data transaksi dan mutasi stok secara *real-time* melalui notifikasi otomatis dengan sinkronisasi latensi rendah. Meskipun terbukti meningkatkan efisiensi operasional dan meminimalkan input manual, sistem ini dirancang khusus untuk transaksi kasir dan inventaris bisnis UMKM, bukan untuk kebutuhan pengelolaan keuangan pribadi (*personal finance*). Sistem ini juga belum memiliki modul *profiling* perilaku pengeluaran ataupun *Financial Intelligent Assistant* untuk memberikan insight kepada pengguna individu [11].

1.5.5 Pengelompokan Perilaku Keuangan Menggunakan *K-Means Clustering*

Penerapan algoritma *K-Means Clustering* digunakan untuk mengelompokkan kebiasaan belanja Generasi Z. Penelitian ini membagi perilaku pengeluaran menjadi beberapa kategori kluster (seperti hemat, moderat, dan konsumtif) serta mengevaluasi kualitas kluster menggunakan metrik *Davies-Bouldin Index (DBI)*, *Silhouette Score*, dan *Calinski-Harabasz Index*. Namun, riset ini masih menggunakan dataset statis dari Kaggle (*Gen Z Money Spending*) yang berisikan data survei dari 1.700 responden. Proses klasterisasi tersebut belum dihubungkan dengan sistem pencatatan transaksi otomatis secara *real-time* maupun *Financial Intelligent Assistant* dalam satu aplikasi utuh [12].

1.5.6 Asisten Keuangan Digital Personal Berbasis AI dan OCR

Pengembangan asisten keuangan digital personal berbasis AI dibangun menggunakan *Next.js*, basis data *Supabase*, *Prisma ORM*, dan *Clerk* untuk sistem autentikasi. Sistem ini menggunakan teknologi *OCR Tesseract* untuk membaca struk belanja, asisten *virtual (chatbot)* untuk konsultasi keuangan, serta pengiriman email notifikasi. Meskipun mampu memberikan saran keuangan personal, namun pengguna harus memotret struk terlebih dahulu agar data dapat diekstrak oleh *OCR*. Sistem ini juga belum menggunakan hasil klasterisasi perilaku belanja pengguna (*K-Means*) sebagai dasar pembentukan profil finansial, serta belum mampu mendeteksi transaksi *E-Wallet* secara *real-time* melalui notifikasi sistem [13].

Dari berbagai penelitian di atas, belum terdapat sistem terpadu yang menggabungkan pembacaan transaksi *E-Wallet* secara otomatis, pengelompokan profil finansial menggunakan *K-Means Clustering* berdasarkan riwayat transaksi aktual, serta penyediaan insight keuangan dan peringatan pengeluaran dalam satu aplikasi. Celah penelitian (*research gap*) tersebut menjadi dasar pengembangan aplikasi FINSIA yang mengintegrasikan pemantauan transaksi secara *real-time*, *profiling finansial*, *Financial Intelligent Assistant*, dan *spending alert*.

Tabel 1. 1 Gap Penelitian

No	Tahun	Pembeda	Peneliti Saat Ini
1	2026	Sistem berbasis web dengan proses input data transaksi manual	Sistem dikembangkan dalam <i>platform mobile (Android)</i> dengan fitur otomatisasi

No	Tahun	Pembeda	Peneliti Saat Ini
		secara penuh tanpa pelacakan notifikasi <i>E-Wallet</i> secara otomatis (<i>real-time</i>).	<i>Monitoring</i> transaksi dompet digital melalui pembacaan notifikasi secara <i>real-time</i> .
2	2025	Belum mendukung pembacaan notifikasi transaksi otomatis di latar belakang (<i>background</i>), algoritma klasterisasi (<i>clustering</i>), dan <i>Financial Intelligent Assistant</i> .	Mengintegrasikan pembacaan notifikasi <i>E-Wallet</i> secara otomatis di latar belakang, metode <i>K-Means Clustering</i> untuk <i>profiling</i> finansial, serta <i>Financial Intelligent Assistant</i> untuk menghasilkan <i>insight</i> keuangan
3	2025	Pencatatan transaksi masih dilakukan secara manual dan belum menyediakan fitur <i>profiling</i> perilaku keuangan pengguna secara otomatis.	Pencatatan transaksi diproses secara otomatis melalui deteksi notifikasi sistem dan dilengkapi dengan <i>profiling</i> finansial menggunakan <i>K-Means Clustering</i> serta <i>Financial Intelligent Assistant</i> untuk memberikan <i>insight</i> keuangan.
4	2024	Sistem dirancang untuk operasional bisnis UMKM (<i>Point of Sale</i>), bukan	Aplikasi dirancang khusus untuk manajemen keuangan pribadi individu dengan integrasi modul

No	Tahun	Pembeda	Peneliti Saat Ini
		pengelolaan keuangan pribadi (<i>personal finance</i>), serta belum mendukung <i>profiling</i> pengguna.	<i>profiling</i> finansial dan <i>Financial Intelligent Assistant</i> untuk memberikan <i>insight</i> keuangan.
5	2025	Klasterisasi didasarkan pada data kuesioner/survei statis (Kaggle), bukan riwayat data transaksi aktual dari aplikasi pencatatan keuangan secara <i>real-time</i> .	Klasterisasi <i>K-Means</i> diterapkan secara dinamis langsung pada data riwayat transaksi aktual pengguna yang ditangkap melalui notifikasi perangkat harian.
6	2026	Masih bersifat semi-otomatis (mengandalkan pemindaian struk/OCR) serta tidak mengintegrasikan klasterisasi <i>K-Means</i> sebagai dasar pembentukan Profil Finansial.	Menggunakan pendeteksian notifikasi dompet digital untuk mengotomatisasi pencatatan transaksi serta menggunakan hasil klasterisasi <i>K-Means</i> sebagai dasar pembentukan Profil Finansial pengguna.

1.6 Data Penelitian

Penelitian ini memanfaatkan data transaksi keuangan sebagai dasar dalam pengembangan fitur utama aplikasi FINSIA, yaitu *Monitoring* transaksi secara *real-time* melalui pembacaan notifikasi transaksi dompet digital, *profiling* finansial menggunakan metode *K-Means Clustering*, serta penyediaan *insight* keuangan dan *spending alert*. *Profiling* finansial dilakukan berdasarkan hasil analisis pola

transaksi pengguna, sedangkan *spending alert* digunakan untuk memberikan peringatan berdasarkan penggunaan anggaran yang telah ditentukan.

1.6.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini dikelompokkan menjadi dua jenis utama yaitu sebagai berikut:

- a. Otomatis: Penangkapan notifikasi transaksi masuk menggunakan *Notification Listener* serta pemindaian saldo secara berkala menggunakan *Accessibility Service* ketika pengguna membuka aplikasi *E-Wallet* DANA, GoPay, dan ShopeePay.
- b. Manual: Pencatatan transaksi secara manual oleh pengguna melalui aplikasi, khususnya untuk transaksi tunai (*cash*).

Struktur dan atribut data transaksi yang tersimpan pada tabel *public.transaksi* di dalam *database Supabase* ditunjukkan pada Tabel 1.2. Contoh data di bawah ini mencakup transaksi dari *E-Wallet* DANA, GoPay, ShopeePay, serta transaksi manual

Tabel 1. 2 Struktur Data Transaksi Primer pada *Supabase*

user_id	ewallet_id	type	amount	category	date	note
e01e1367-0d5a-4a07-b6e1-5fcd2f978401	dana	pemasukan	5	Lainnya	2026-06-28 14:14:58.724059+00	Rp5 telah diterima dari RENITA NUR HIKMAH 

user_id	ewallet_id	type	amount	category	date	note
69085392-b291-40df-8a26-a0dd2897176c	manual	pengeluaran	5000	Makanan & Minuman	2026-07-11 18:29:16.236927+00	bakso 5000
c63a7cbb-9d20-4e5f-8024-24848e87d08b	shopee pay	pengeluaran	35380	Lainnya	2026-07-15 22:06:42.741122+00	Penarikan dana sejumlah Rp35.380 ke berhasil ditransfer ke rekening bank-mu.
69085392-b291-40df-8a26-a0dd2897176c	dana	pemasukan	5	Lainnya	2026-07-11 18:28:25.706653+00	Rp5 telah diterima dari RENITA NUR HIKMAH 
e01e1367-0d5a-4a07-b6e1-5fcd2f978401	manual	pengeluaran	50000	Belanja	2026-06-29 09:18:23.397734+00	beli sepatu 50000
0e0c3879-2ec7-4be6-83c7-9f2c647a183a	manual	pemasukan	600000	Lainnya	2026-07-09 04:54:28.596105+00	gaji sampingan
e01e1367-0d5a-4a07-b6e1-5fcd2f978401	manual	pengeluaran	500000	Lainnya	2026-06-28 14:44:34.731001+00	pajak 500000

user_id	ewallet_id	type	amount	category	date	note
c63a7cbb-9d20-4e5f-8024-24848e87d08b	shopee pay	pemasukan	35380	Belanja	2026-07-15 22:06:42.707397+00	Shopee sedang memproses penarikan dana sejumlah Rp35.380 ke SeaBank. SeaBank akan langsung memproses penarikan danamu. Kamu akan menerima dana secara <i>real-time</i> .
0e0c3879-2ec7-4be6-83c7-9f2c647a183a	manual	pengeluaran	10000	Transportasi	2026-07-22 06:15:16.750354+00	bensin 10000

1.6.2 Pengolahan dan Integrasi Data

Data transaksi yang telah dikumpulkan selanjutnya diintegrasikan dan diproses menggunakan Google Colab untuk menghasilkan dataset yang digunakan dalam proses *profiling* finansial pada aplikasi FINSIA. Pengolahan data dilakukan melalui beberapa tahapan, yaitu integrasi data, validasi dataset, klasifikasi kategori transaksi, rekayasa fitur (*feature engineering*), dan proses *K-Means Clustering*.

1. Integrasi Data

Data transaksi diperoleh dari beberapa sumber, yaitu data transaksi manual serta data transaksi dari dompet digital DANA, GoPay, dan

ShopeePay. Masing-masing sumber data terlebih dahulu diproses agar memiliki struktur atribut yang seragam.

Data dari sumber yang berbeda kemudian digabungkan menjadi satu *dataframe* utama. Dataset tersebut memiliki atribut:

- a) tanggal, yaitu tanggal transaksi;
- b) deskripsi, yaitu keterangan atau deskripsi transaksi;
- c) kategori, yaitu kategori transaksi;
- d) tipe, yaitu tipe transaksi;
- e) nominal, yaitu nilai transaksi; dan
- f) sumber, yaitu sumber data transaksi.

Setelah proses integrasi, dilakukan pemeriksaan jumlah data, struktur kolom, distribusi tipe transaksi, distribusi kategori, duplikasi data, rentang tanggal, dan distribusi sumber data untuk memastikan dataset dapat digunakan pada tahap pengolahan selanjutnya.

2. Klasifikasi Kategori Transaksi

Data transaksi yang telah terintegrasi kemudian diklasifikasikan berdasarkan deskripsi transaksi menggunakan metode *rule-based classification*. Tahap ini dilakukan karena data dari berbagai sumber memiliki bentuk deskripsi transaksi yang berbeda.

Sebelum dilakukan klasifikasi, deskripsi transaksi dinormalisasi melalui fungsi `normalisasi_deskripsi()`. Proses normalisasi meliputi:

- a) mengubah teks menjadi huruf kecil (*lowercase*);
- b) menghilangkan spasi yang tidak diperlukan;
- c) melakukan normalisasi Unicode;

- d) menghilangkan tanda diakritik;
- e) mengganti karakter tertentu seperti &, /, -, dan _;
- f) menghilangkan karakter selain huruf, angka, dan spasi; serta
- g) menghilangkan spasi berlebih.

Hasil normalisasi disimpan pada atribut `deskripsi_bersih`.

Selanjutnya, sistem melakukan pencocokan kata kunci pada deskripsi transaksi. Kata kunci dikelompokkan ke dalam tujuh kategori, yaitu:

- a) Makanan & Minuman, yang mencakup transaksi makanan dan minuman;
- b) Transportasi, yang mencakup transportasi, bahan bakar, parkir, dan perawatan kendaraan;
- c) Pendidikan, yang mencakup kebutuhan kuliah, tugas, ujian, dan kebutuhan pendidikan lainnya;
- d) Tagihan, yang mencakup kas, kos, kebutuhan komunikasi, dan transaksi terkait tagihan;
- e) Hiburan, yang mencakup bioskop, permainan, karaoke, konser, wisata, dan kegiatan rekreasi;
- f) Belanja, yang mencakup pembelian barang dan kebutuhan ritel maupun perdagangan elektronik; dan
- g) Lainnya, untuk transaksi yang tidak memenuhi aturan pada kategori sebelumnya.

Pencocokan dilakukan menggunakan fungsi `mengandung_keyword()`. Untuk kata kunci yang terdiri dari satu kata digunakan pencocokan berbasis batas kata (*word boundary*), sedangkan kata kunci yang terdiri dari beberapa kata

dicocokkan berdasarkan kemunculan frasa pada deskripsi yang telah dinormalisasi.

Beberapa aturan khusus diterapkan sebelum pencocokan kategori umum. Transaksi yang mengandung "uang saku" dan "transfer bulanan" dikategorikan sebagai Pendidikan, sedangkan transaksi yang mengandung "kas" dikategorikan sebagai Tagihan.

Hasil klasifikasi kemudian disimpan kembali pada atribut kategori dan digunakan sebagai dasar pengolahan data berikutnya.

3. Validasi Hasil Klasifikasi

Hasil *rule-based classification* kemudian diperiksa melalui audit distribusi kategori. Audit dilakukan dengan menghitung jumlah transaksi dan persentase masing-masing kategori serta total nominal transaksi pada setiap kategori.

Selain pemeriksaan distribusi, dilakukan pengujian terhadap sejumlah deskripsi transaksi yang telah ditentukan untuk memastikan bahwa aturan klasifikasi menghasilkan kategori yang sesuai. Contoh pengujian meliputi transaksi seperti *Nonton film bioskop*, *Bayar kos*, *Bensin*, *Kas*, *Indomaret*, *Main game*, *Nasi kepala ayam*, *Telco*, *by.U*, *Karaoke*, dan *Undar Kapal*.

Transaksi yang masih masuk ke kategori Lainnya juga diaudit dengan melihat deskripsi, nominal, sumber, serta distribusi kemunculannya. Hasil audit digunakan untuk mengetahui apakah terdapat transaksi yang seharusnya dapat dikategorikan ke dalam kategori tertentu berdasarkan aturan yang telah dibuat.

4. Rekayasa Fitur (*Feature Engineering*)

Setelah proses klasifikasi selesai, data transaksi pengeluaran digunakan untuk membentuk data agregasi bulanan. Kolom tanggal terlebih dahulu dikonversi menjadi format tanggal dan kolom nominal dikonversi menjadi numerik.

Selanjutnya, periode transaksi dibentuk berdasarkan bulan dan tahun menggunakan atribut `bulan_tahun`. Data kemudian dikelompokkan berdasarkan periode tersebut untuk menghasilkan karakteristik pengeluaran pengguna setiap bulan.

Fitur yang digunakan sebagai atribut *K-Means* terdiri atas lima fitur, yaitu:

- a) Total Pengeluaran, yaitu jumlah seluruh nominal transaksi pengeluaran dalam satu bulan.
- b) Jumlah Transaksi, yaitu jumlah transaksi pengeluaran yang terjadi dalam satu bulan.
- c) Persen Kebutuhan, yaitu persentase pengeluaran kategori Makanan & Minuman, Transportasi, Pendidikan, dan Tagihan terhadap total pengeluaran bulanan.
- d) Persen Hiburan, yaitu persentase pengeluaran kategori Hiburan terhadap total pengeluaran bulanan.
- e) Persen Belanja, yaitu persentase pengeluaran kategori Belanja terhadap total pengeluaran bulanan.

Perhitungan persentase dilakukan dengan membagi total nominal masing-masing kelompok kategori dengan total pengeluaran pada bulan yang sama, kemudian dikalikan 100.

Hasil agregasi disimpan dalam `df_bulanan`, kemudian disalin ke dalam `df_kmeans` sebagai dataset yang digunakan pada proses *clustering*. Kelima fitur tersebut ditentukan dalam variabel `fitur_kmeans_cols` dan digunakan sebagai fitur masukan pada tahap *preprocessing*.

5. *Preprocessing* dan Klasterisasi *K-Means*

Data fitur bulanan yang telah melalui tahap analisis kelayakan selanjutnya dipersiapkan untuk proses *K-Means Clustering*. Tahap *preprocessing* dilakukan untuk memastikan seluruh fitur memiliki format numerik dan berada pada skala yang sesuai untuk proses perhitungan jarak.

Pertama, dilakukan penanganan *outlier* menggunakan metode *Interquartile Range* (IQR). Nilai IQR dihitung berdasarkan kuartil pertama ((`Q_1`)) dan kuartil ketiga ((`Q_3`)) menggunakan persamaan:

$$\text{IQR} = Q_3 - Q_1$$

Batas bawah dan batas atas ditentukan menggunakan persamaan:

$$\text{Batas\ Bawah} = Q_1 - 1,5(\text{IQR})$$

$$\text{Batas\ Atas} = Q_3 + 1,5(\text{IQR})$$

Nilai fitur yang berada di luar batas tersebut tidak dihapus dari dataset, tetapi dibatasi (*clipping*) pada nilai batas bawah atau batas atas. Dengan demikian, jumlah observasi tetap dipertahankan dalam proses *clustering*.

Setelah penanganan *outlier*, dilakukan normalisasi menggunakan *Min-Max Scaling* dengan `MinMaxScaler()`. Normalisasi dilakukan karena kelima

fitur memiliki rentang nilai yang berbeda. Proses ini mengubah nilai setiap fitur ke dalam rentang 0 sampai 1 sehingga perbedaan skala antarfitur dapat dikurangi. Persamaan *Min-Max Scaling* adalah:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

dengan (x) merupakan nilai asli, $(x_{\min})$ merupakan nilai minimum fitur, $(x_{\max})$ merupakan nilai maksimum fitur, dan (x') merupakan nilai hasil normalisasi.

Hasil normalisasi kemudian disimpan dalam variabel X_{scaled} dan digunakan sebagai masukan pada proses *K-Means Clustering*.

Selanjutnya, dilakukan pengujian jumlah kluster menggunakan *K-Means Clustering* dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) pada beberapa nilai K . WCSS menunjukkan jumlah kuadrat jarak setiap data terhadap *centroid* pada *cluster* tempat data tersebut berada. Hasil pengujian tersebut digunakan sebagai dasar dalam menentukan jumlah kluster. Pada penelitian ini digunakan $K = 3$.

Proses *K-Means* dilakukan dengan mengelompokkan data berdasarkan kemiripan karakteristik lima fitur yang digunakan. Implementasi *K-Means* menggunakan parameter $K_OPTIMAL=3$, $init='random'$, $n_init=10$, dan $random_state=42$. Selanjutnya Adapun Tahapan algoritma meliputi:

- a) Inisialisasi *centroid*, yaitu menentukan titik pusat awal untuk tiga kluster.
- b) Perhitungan jarak, yaitu menghitung jarak setiap data terhadap *centroid* menggunakan jarak *Euclidean*.

$$d(x,c) = \sqrt{\sum_{j=1}^p (x_j - c_j)^2}$$

dengan $d(x,c)$ merupakan jarak antara data x dan *centroid* c , x_{j_j} merupakan nilai fitur ke- j pada data, c_{j_j} merupakan nilai fitur ke- j pada *centroid*, dan p merupakan jumlah fitur yang digunakan.

- c) Penentuan klaster, yaitu memasukkan data ke klaster dengan jarak terdekat terhadap *centroid*.
- d) Pembaruan *centroid*, yaitu menghitung kembali posisi *centroid* berdasarkan rata-rata data yang menjadi anggota setiap klaster.
- e) Iterasi, yaitu mengulangi proses perhitungan jarak, penentuan klaster, dan pembaruan *centroid* hingga model mencapai kondisi konvergen atau memenuhi batas iterasi yang ditentukan.

Parameter $n_init=10$ digunakan agar proses *K-Means* dijalankan dengan beberapa inisialisasi *centroid*, sedangkan $random_state=42$ digunakan untuk menjaga konsistensi hasil proses *clustering*.

Hasil proses *K-Means* menghasilkan label *cluster* untuk setiap observasi pada dataset. Label tersebut disimpan pada atribut ***Cluster*** dalam data hasil *clustering*.

Selanjutnya, karakteristik setiap *cluster* dianalisis berdasarkan nilai rata-rata dari lima fitur yang digunakan, yaitu Total Pengeluaran, Jumlah Transaksi, Persen Kebutuhan, Persen Hiburan, dan Persen Belanja. Hasil analisis karakteristik *cluster* digunakan sebagai dasar untuk melakukan interpretasi Profil Finansial.

Profil Finansial kemudian dikelompokkan menjadi tiga kategori, yaitu Hemat, Stabil, dan Konsumtif. Penamaan profil tersebut merupakan hasil interpretasi terhadap karakteristik masing-masing *cluster* dan bukan

merupakan label yang dihasilkan secara langsung oleh algoritma *K-Means*. Oleh karena itu, nomor *cluster* yang dihasilkan, seperti *Cluster 0*, *Cluster 1*, dan *Cluster 2*, tidak secara langsung menunjukkan kategori Profil Finansial tertentu.

Hasil akhir dari proses pengolahan dan klasterisasi digunakan sebagai dasar dalam menghasilkan Profil Finansial pada aplikasi FINSIA. Profil tersebut menggambarkan pola pengeluaran berdasarkan karakteristik transaksi pada periode yang dianalisis.

6. Analisis dan Interpretasi Profil Finansial

Hasil *clustering* selanjutnya dianalisis berdasarkan karakteristik masing-masing klaster. Interpretasi dilakukan dengan melihat nilai fitur dan *centroid* yang dihasilkan oleh model.

Label atau nama profil finansial tidak ditentukan sebelum proses *clustering*, tetapi ditentukan berdasarkan karakteristik hasil klaster. Dengan demikian, setiap klaster menggambarkan kelompok perilaku pengeluaran yang memiliki karakteristik relatif serupa berdasarkan fitur yang digunakan.

Hasil *clustering* kemudian digunakan sebagai dasar pembentukan profil finansial pada aplikasi FINSIA.

7. Integrasi Hasil Profil Finansial ke FINSIA

Selain hasil *clustering*, data transaksi juga digunakan sebagai dasar penyajian informasi pendukung keputusan finansial, seperti pengeluaran terbesar, perubahan pola pengeluaran, dan pola perilaku pengguna. Sistem juga menyediakan fitur pengelolaan anggaran yang memberikan peringatan berdasarkan batas pengeluaran yang telah ditentukan. Kemudian jika

pengeluaran mencapai batas maka fitur tersebut berubah menjadi warna merah yang berarti *Over Budget*.

Fitur pengelolaan anggaran berjalan secara terpisah dari proses *K-Means Clustering*. Fitur batas pengeluaran (*Over Budget*) tidak digunakan sebagai fitur masukan dalam proses *clustering* sehingga tidak mengubah hasil Profil Finansial pengguna. Dengan demikian, *K-Means* digunakan untuk menentukan karakteristik Profil Finansial berdasarkan lima fitur transaksi, sedangkan fitur anggaran digunakan untuk memantau kondisi pengeluaran terhadap batas anggaran yang ditentukan pengguna.

1.7 Alat Penelitian

Dalam merancang, mengimplementasikan, serta menguji aplikasi FINSIA, dibutuhkan dukungan perangkat keras dan perangkat lunak dengan spesifikasi kebutuhan sebagai berikut:

Tabel 1. 3 Kebutuhan Alat Penelitian

No	Perangkat Keras Alat Penelitian	
	Nama Alat	Fungsi
1.	Laptop	Sebagai Perangkat utama yang digunakan untuk membangun aplikasi dan model <i>Machine Learning</i>
2.	Ruang Penyimpanan (SSD & Cloud/G.Drive)	Sebagai Penyimpanan file seperti kebutuhan dataset model
3.	Smartphone	Sebagai perangkat pendukung untuk melakukan tes dan <i>debugging</i> dalam proses pengembangan aplikasi
4.	Kabel USB	Sebagai komponen untuk menghubungkan <i>smartphone</i> dengan laptop saat dilakukan tes aplikasi

No	Perangkat Lunak Alat Penelitian	
	Nama Alat	Fungsi
1.	VS Code & Android Studio	Sebagai aplikasi kode editor untuk pengembangan aplikasi
2.	Google Collab	Sebagai <i>Platform</i> untuk membuat pengembangan model <i>machine learning</i>
3.	Figma	Sebagai <i>Platform</i> untuk merancang desain tampilan antarmuka aplikasi