

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital yang pesat telah mengubah cara masyarakat dalam melakukan transaksi keuangan. Salah satu layanan keuangan yang semakin banyak digunakan adalah kartu kredit. Kartu kredit menawarkan kemudahan bagi pengguna untuk melakukan transaksi tanpa harus membawa uang tunai, sekaligus menawarkan keuntungan. Data terkini dari Asosiasi Kartu Kredit Indonesia (AKKI) menunjukkan bahwa jumlah kartu kredit beredar di Indonesia mencapai 18.732.074 keping pada Juni 2025, dengan peningkatan sebesar 182.388 keping dibandingkan tahun sebelumnya[1].

Namun, di balik kemudahan tersebut terdapat potensi risiko, salah satunya adalah gagal bayar atau disebut juga *credit default*. Risiko gagal bayar merupakan tantangan besar bagi lembaga keuangan karena dapat menimbulkan kerugian finansial dan mempengaruhi kestabilan operasional dan reputasi perusahaan. Menurut data statistik dari Otoritas Jasa Keuangan (OJK), rasio *Non Performing Loan gross* perbankan secara agregat tercatat sebesar 2,29% pada Mei 2025, yang meningkat dari 2,24% pada April 2025[2].

Salah satu penyebab tingginya tingkat gagal bayar adalah kurangnya pemahaman lembaga keuangan terhadap pola perilaku pengguna kartu kredit. Setiap nasabah memiliki karakteristik dan kebiasaan transaksi yang berbeda, seperti pola pembayaran, jumlah tagihan bulanan, dan pemanfaatan batas kredit. Ketika lembaga keuangan tidak mampu mengidentifikasi secara akurat pola perilaku tersebut, risiko pemberian kredit kepada nasabah dengan potensi gagal bayar akan semakin meningkat[3].

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan analisis data yang lebih canggih, salah satunya dengan memanfaatkan teknik *data mining* dan *clustering* untuk segmentasi perilaku nasabah. *Clustering* merupakan metode *unsupervised learning* yang efektif untuk mengelompokkan data berdasarkan

kemiripan karakteristik tanpa memerlukan label awal. Teknik *clustering* terbukti efektif dalam mengelompokkan nasabah berdasarkan data yang diperoleh menjadi sejumlah *cluster*, membantu pihak lembaga untuk memahami karakteristik nasabah[3].

Beberapa penelitian terdahulu telah menunjukkan bahwa teknik *clustering* terbukti efektif dalam segmentasi perilaku pengguna kartu kredit dengan tingkat akurasi baik [4]. Beberapa teknik *clustering* yang umum digunakan dalam segmentasi nasabah kartu kredit dan analisis risiko default antara lain *K-Means Clustering*, *Hierarchical Clustering*, serta *DBSCAN*. *Hierarchical Clustering* membentuk struktur hierarki berupa *dendrogram* yang memungkinkan visualisasi hubungan antar *cluster*, sedangkan *DBSCAN* unggul dalam mendeteksi cluster dengan bentuk tak beraturan serta mengidentifikasi *outlier* atau anomali secara otomatis, sehingga sering digunakan untuk deteksi *fraud* pada transaksi kartu kredit [5].

Meskipun demikian, dalam penelitian ini *K-Means Clustering* dipilih sebagai metode utama karena memiliki beberapa keunggulan dibandingkan metode lainnya. *K-Means* memiliki karakteristik yang sederhana dalam implementasi, memiliki kecepatan komputasi yang tinggi, serta skalabel terhadap dataset berukuran besar yang umum ditemui pada data historis kartu kredit [6]. Penelitian-penelitian terkini juga menunjukkan bahwa *K-Means* sering memberikan performa optimal pada segmentasi nasabah kartu kredit dibandingkan metode *hierarchical* yang lebih intensif secara komputasi atau *DBSCAN* yang lebih cocok untuk kasus deteksi anomali daripada segmentasi umum [7]. Keterbatasan *K-Means*, seperti sensitivitas terhadap inisialisasi *centroid* dan asumsi bentuk *cluster*, dapat diatasi dengan teknik validasi seperti *Elbow Method* dan *Silhouette Analysis*.

Penelitian ini dilakukan untuk mengembangkan model segmentasi perilaku pengguna kartu kredit serta menganalisis kemungkinan terjadinya *default* berdasarkan hasil klasterisasi menggunakan *K-Means Clustering*. Melalui analisis ini, diharapkan dapat diketahui kelompok pengguna dengan risiko gagal bayar

tinggi sehingga lembaga keuangan dapat mengambil langkah preventif dalam mengelola risiko kredit dan menyusun strategi kebijakan yang lebih tepat sasaran.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana menentukan jumlah klaster optimal pada segmentasi perilaku pengguna kartu kredit menggunakan *Elbow Method* dan *Silhouette Analysis* untuk mendapat hasil akurasi tertinggi menggunakan *K-Means Clustering*?
2. Bagaimana hasil segmentasi perilaku pengguna kartu kredit dapat digunakan untuk mengidentifikasi kelompok pengguna dengan risiko *default payment* yang tinggi?

1.3 Batasan Masalah

Agar penelitian ini terarah dan sesuai dengan tujuan yang ingin dicapai, maka ditetapkan beberapa batasan masalah sebagai berikut:

1. Data penelitian menggunakan sebuah dataset yang diperoleh dari Kaggle Credit Card Dataset.
2. Metode *clustering* yang digunakan dalam penelitian ini adalah K-Means Clustering.
3. Penentuan jumlah klaster optimal menggunakan *Elbow Method* dan *Silhouette score*.
4. Analisis prediksi *default payment next month* dilakukan berdasarkan hasil segmentasi tanpa menggunakan algoritma klasifikasi lanjutan.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai dalam penelitian ini adalah:

1. Mengimplementasikan algoritma K-Means Clustering untuk melakukan segmentasi perilaku pengguna kartu kredit.
2. Menentukan jumlah klaster optimal dengan menggunakan metode *Elbow* dan *Silhouette Analysis*.
3. Menganalisis hasil segmentasi pengguna untuk mengidentifikasi kelompok yang memiliki potensi *default*

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan referensi akademik bagi pengembangan penelitian di bidang *data mining*, khususnya penerapan algoritma *K-Means Clustering* untuk analisis perilaku pengguna.
2. Memberi kontribusi ilmiah terhadap pemanfaatan analisis *cluster* dalam pengelolaan risiko keuangan dan prediksi gagal bayar.
3. Membantu lembaga keuangan memahami pola perilaku pengguna kartu kredit untuk mendeteksi potensi gagal bayar lebih dini.
4. Menjadi dasar dalam pengambilan keputusan berbasis data (*data-driven decision making*) bagi strategi mitigasi risiko kredit dan kebijakan produk keuangan.

1.6 Sistematika Penulisan

Sistematika penulisan dalam laporan penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN : Berisi uraian mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA : Membahas teori-teori yang mendukung penelitian, termasuk konsep kartu kredit, risiko gagal bayar, dasar-dasar *data mining*, algoritma *K-Means Clustering*, serta penelitian terdahulu yang relevan.

BAB III METODE PENELITIAN : Menjelaskan tahapan penelitian, sumber data, variabel yang digunakan, proses *preprocessing* data, serta tahapan penerapan *K-Means Clustering* dan metode evaluasi kluster.

BAB IV HASIL DAN PEMBAHASAN : Menyajikan hasil implementasi algoritma *K-Means Clustering*, analisis hasil segmentasi, serta interpretasi hubungan antara kluster pengguna dan variabel *default payment next month*.

BAB V PENUTUP : Berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan penelitian selanjutnya.