

BAB 3 METODOLOGI PENELITIAN

3.1 Metodologi Penelitian

Metode penelitian yang digunakan pada penelitian ini terdiri dari beberapa tahap antara lain.

1. Studi literatur

Studi literatur menjadi tahap awal dari proses penelitian. Hal ini dilakukan dengan mencari, membaca, dan mempelajari beberapa sumber jurnal ilmiah atau karya tulis ilmiah lainnya. Tujuannya agar dapat meningkatkan pemahaman penulis terkait teori mengenai metode yang digunakan dalam penelitian ini.

2. Analisis Kebutuhan

Proses ini dilakukan bersamaan dengan studi literatur. Hal ini dilakukan agar penulis dapat memahami hal-hal apa saja yang diperlukan untuk membentuk sistem terkait penelitian berdasarkan informasi yang diperoleh selama studi literatur. Tahap ini termasuk juga dengan mengumpulkan dataset.

3. Perancangan dan Implementasi Sistem

Proses ini dilakukan dengan memahami dan menggambarkan alur logika program. Alur logika sistem digambarkan ke dalam bentuk flowchart. Setelah alur sistem tersusun secara sistematis, implementasi dapat dilakukan. Dimulai dengan menghitung nilai RFM, melakukan *preprocessing dataset* dan menerapkan metode normalisasi, standarisasi dan *k-means clustering* pada *preprocessed data*.

4. *Testing* dan Evaluasi

Dataset yang sudah diolah ditampilkan secara visual kemudian dilakukan pengujian untuk membandingkan hasil proses *data* dengan metode normalisasi dan metode standarisasi.

5. Dokumentasi

Hasil yang didapatkan kemudian dituliskan ke dalam laporan. Dokumentasi dilakukan secara bertahap mulai dari pendahuluan hingga kesimpulan.

Selama proses dokumentasi, konsultasi dengan dosen pembimbing untuk mendapatkan arahan agar laporan dapat tersusun dengan baik dan benar.

3.2 Gambaran Umum Sistem

Sistem yang dibuat merupakan secara umum merupakan implementasi teknik RFM dan K-Means untuk mengelompokkan dataset dalam membagi jenis customer. Penambahan teknik normalisasi data seperti *Standard Scaler* dan *MinMax Scaler* juga diperlukan agar data dengan skala yang bervariasi menjadi terdistribusi dengan jarak yang ditentukan. Langkah awal diawali dengan membaca dataset terlebih dahulu.

Setelah *dataset* dibaca oleh sistem, akan dilakukan memasuki flowchart RFM yang bertujuan untuk mendapatkan nilai *Monetary*, *Frequency*, *Recency* untuk digunakan pada normalisasi data dan mengelompokkan data menggunakan K-Means clustering.

Pada saat penampilan data, dilakukan *plotting* untuk melihat grafik perbandingan nilai RFM dan total cluster yang dibagi. Dari hasil *plotting*, dapat disimpulkan jenis customer dari potensi terbaik hingga terendah.

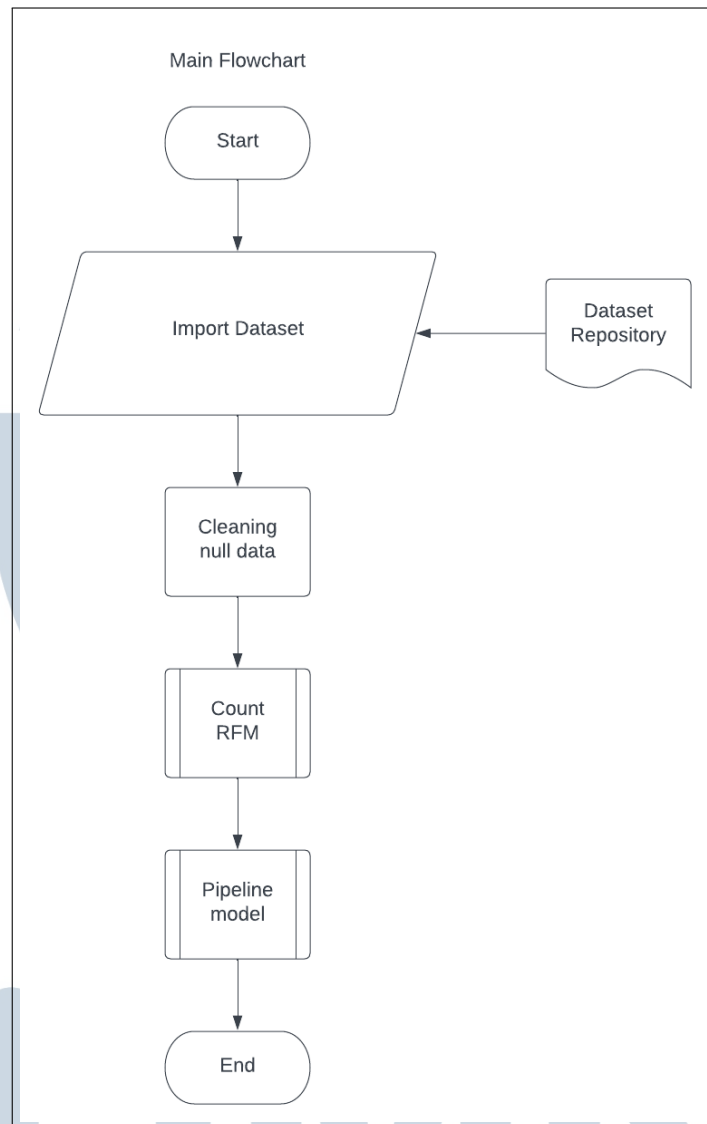
3.3 Flowchart

Proses berjalannya program dapat dijelaskan dalam bentuk alur diagram yang terdiri dalam beberapa bagian: *Main Flowchart*, *RFM Flowchart*, *Preprocessing Flowchart*, *K-Means Clustering Flowchart*.

3.3.1 Main Flowchart

Gambar 3.1 merupakan gambar diagram alur utama yang menjelaskan alur kerja program secara umum. Langkah pertama diawali dengan melakukan import library yang diperlukan untuk menjalankan program. Setelah itu, dilakukan import dataset dan penghapusan barisan data yang mengandung nilai *Null* pada dataset.

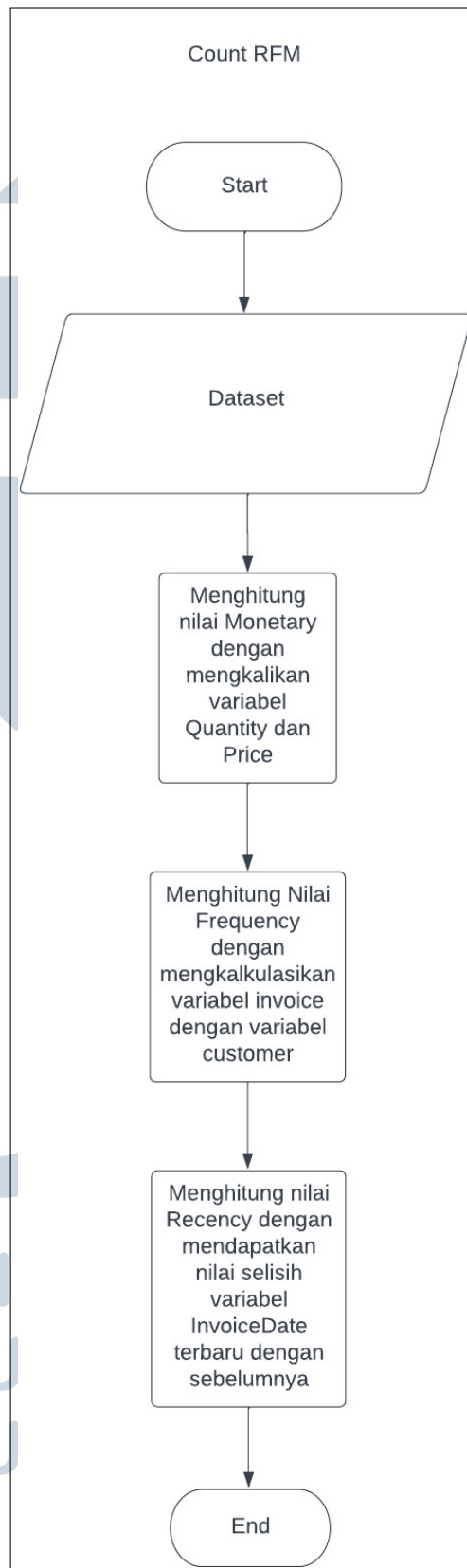
Dataset yang sudah dibersihkan dari nilai *Null* dilanjutkan dengan cari nilai RFM yang akan memasuki tahap preprocessing normalisasi. Setelah dataset dinormalisasi, akan memasuki tahap K-Means clustering. *Plotting* dilakukan untuk mendapatkan kesimpulan dari hasil *plotting cluster*.



Gambar 3.1. Main Flowchart

3.3.2 Count RFM Flowchart

Gambar 3.2 merupakan gambar *flowchart* perhitungan nilai RFM. Dimulai dari pembacaan *dataset* dan memulai perhitungan nilai *Monetary* terlebih dahulu berdasarkan jumlah barang dan harga per satuan barang yang dibeli oleh *customer*, kemudian dilanjutkan perhitungan nilai *Frequency* berdasarkan jumlah transaksi yang dilakukan *customer* dan terakhir mendapatkan nilai *Recency* berdasarkan perbedaan waktu terakhir kali *customer* melakukan transaksi.

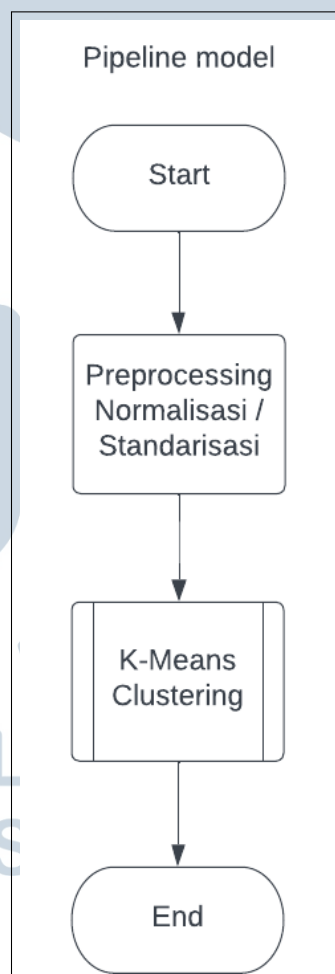


Gambar 3.2. Count RFM Flowchart

3.3.3 Pipeline Model Flowchart

Gambar 3.4 merupakan gambar diagram *flowchart pipeline model*. Dimulai dengan melewati proses *preprocessing* normalisasi untuk mengolah dataset. Data yang sudah dipreprocessing selanjutnya dapat digunakan untuk memulai serangkaian proses pengolahan data data yang sudah dilakukan *preprocessing* dalam algoritma *k-means clustering*.

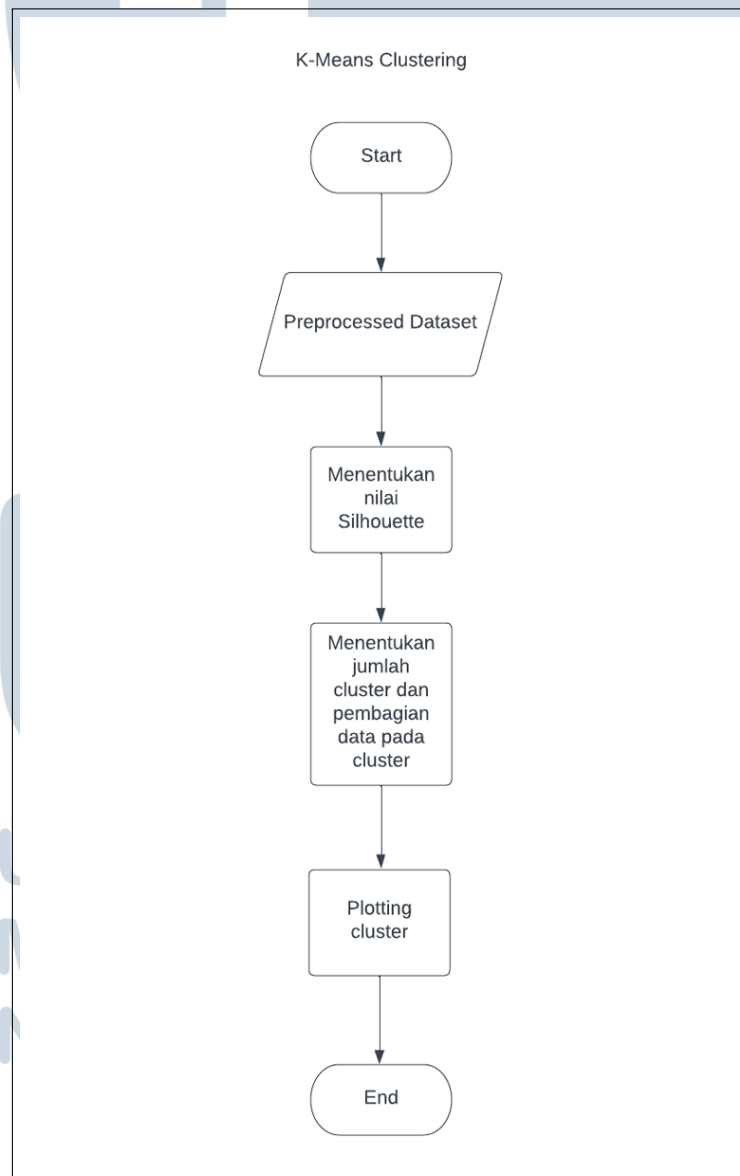
Setelah mendapatkan nilai hasil RFM (Recency Frequency Monetary) dilanjutkan dengan *preprocessing dataset* agar hasilnya dapat digunakan untuk proses K-Means *clustering*. Pada penelitian ini tahapan *preprocessing* digunakan dua metode *preprocessing* yaitu Normalisasi Min-Max dan Standarisasi. Setelah data diolah menggunakan dua metode *preprocessing* dilanjutkan dengan proses pengolahan data yang sudah dilakukan *preprocessing* dalam algoritma *k-means clustering*.



Gambar 3.3. Pipeline Model Flowchart

3.3.4 K-Means Clustering Flowchart

Gambar 3.5 merupakan gambar diagram alur untuk proses *k-means clustering*. Data yang sudah dinormalisasikan dari proses *preprocessing* dihitung menggunakan metode *silhouette analysis*. Metode *silhouette analysis* untuk mendapatkan nilai *silhouette score* yang akan menentukan tingkat kemiripan data tersebut dengan *cluster*. Untuk mendapatkan nilai *silhouette score* yang akurat akan dilakukan perhitungan menggunakan metode iterasi terhadap jangkauan *cluster*. Total *cluster* bergantung dari jumlah kelompok yang akan dibagikan data tersebut. Pada penelitian ini digunakan



Gambar 3.4. K-Means Clustering Flowchart