SKRIPSI

IMPLEMENTASI DATA MINING TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA *CLUSTERING* DENGAN METODE *K-MEANS*

IMPLEMENTATION OF SALES TRANSACTION DATA MINING USING CLUSTERING ALGORITHM WITH K-MEANS METHOD

SARIPA HASIPAH

D0218516

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS SULAWESI BARAT

MAJENE

2025

LEMBAR PENGESAHAN

IMPLEMENTASI DATA MINING TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA CLUSTERING DENGAN METODE K-MEANS

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:

SARIPA HASIPAH
D0218516

Skripsi ini telah diujikan dan dinyatakan lulus
pada 25 Juli 2025
Telah diperiksa dan disetujui oleh:

Pembimbing I	Pembimbing II
Dr. Eng. Sulfayani, S.Si.,M.T	**Wawan Firgiawan, S.T., M.Kom**
NIP:198903172020122011	NIDK:8948080023

Dekan Fakultas Teknik,
Universitas Sulawesi Barat

Prof.Dr. Ir. Hafsah Nirwana, M.T
NIP: 196404051990032002

Ketua Program Studi
Informatika

Muh. Rafli Rasyid, S.Kom., M.T
NIP: 198808182022031006

LEMBAR PERSETUJUAN

SKRIPSI

IMPLEMENTASI DATA MINING TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA *CLUSTERING* DENGAN METODE *K-MEANS*

Telah dipersiapakan dan disusun oleh

SARIPA HASIPAH

D0218516

~~Telah dipertahankan di depan Tim Penguji~~

Pada tanggal 28 Mei 2025

Susunan Tim Penguji

Pembimbing I

Dr. Eng. Sulfayanti, S.Si.,M.T
NIP: 198903172020122011

Penguji I

Arnita Irianti, S.Si.,M.Si
NIP: 198708062018032001

Pembimbing II

Wawan Firgiawan, S.T., M.Kom
NIDK: 8948080023

Penguji II

Musyrifah, S.Pd., M.Pd
NIDN: 0014119302

Penguji III

Muh. Imam Quraisy, S.Kom.,M.Kom
NIDN: 0027019205

ABSTRAK

Saripa Hasipah. Impelementasi Data Mining Transaksi Penjualan Menggunakan Algoritma *Clustering* Dengan Metode *K-Means*. (Dibimbing oleh Dr. Eng. Sulfayani, S.Si.,M.T dan Wawan Firgiawan, S.T., M.Kom)

Perkembangan teknologi informasi yang pesat mendorong pemanfaatan data dalam pengambilan keputusan bisnis, salah satunya dengan penerapan data mining. Penelitian ini bertujuan untuk mengelompokkan data transaksi penjualan produk pada toko Mandar Sutera menggunakan algoritma *K-Means clustering*. Permasalahan utama yang dihadapi adalah kesulitan dalam menentukan jumlah stok barang berdasarkan tingkat penjualan. Metode *K-Means* digunakan karena kemampuannya dalam mengelompokkan data ke dalam beberapa *cluster* berdasarkan kemiripan atribut, yaitu stok awal, stok keluar, dan stok akhir. Penelitian ini menggunakan 29 data produk dan dilakukan proses *preprocessing*, penentuan jumlah *cluster* (k=3), perhitungan jarak menggunakan *Euclidean Distance*, serta evaluasi hasil dengan menghitung nilai centroid baru. Hasil pengelompokan menghasilkan tiga kategori penjualan: sangat laris, laris, dan kurang laris. Implementasi dilakukan menggunakan bahasa pemrograman *Python*. Pengujian dengan metode *Silhouette Score* 0,6151 menunjukkan bahwa *cluster* yang terbentuk memiliki kualitas cukup baik hasil tersebut menunjukkan bahwa *silhouette score* efektif dalam menilai kualitas pengelompokan dan menunjukkan *K-Means* memberikan hasil yang lebih baik. Penelitian ini diharapkan dapat menjadi acuan bagi pemilik usaha dalam mengelola stok produk secara lebih efektif dan efisien.

Kata Kunci: D*ata Min*ing, *K-Means Clustering*, *Python*, Mandar Sutera

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Kemajuan teknologi terus meningkat dari tahun ke tahun, salah satunya adalah internet. Internet menjadi alat yang efektif untuk memperoleh informasi dan berkomunikasi secara cepat dan tepat. Hal ini menyebabkan banyak pihak memanfaatkan internet untuk berbagai keperluan, termasuk dalam dunia bisnis (Aulia, 2021). Pemilik usaha kecil hingga besar memanfaatkan perkembangan teknologi internet sebagai sarana untuk mempromosikan produk atau melakukan iklan.

Clustering merupakan suatu proses pengelompokan sejumlah data atau objek ke dalam *cluster* (*group*) sehingga setiap dalam pada suatu *cluster* yang sama akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya (Handoko, dkk, 2020). Dalam konteks ini, algoritma *clustering K-Means* dalam data mining sangat cocok untuk mengatur persediaan, merancang strategi penjualan, serta mengelompokkan produk ke dalam beberapa kategori, yaitu penjualan paling laku, kurang laku, dan tidak laku (Nabila, dkk, 2021). Alasan penggunaan metode *K-Means* adalah untuk mempermudah toko dalam menganalisis serta mengelompokkan data, sehingga dapat menentukan persediaan produk dengan lebih baik dari data transaksi penjualan yang besar. Metode ini memungkinkan proses dilakukan secara cepat dan efisien. Oleh karena

itu, penelitian ini menerapkan analisis data *mining* menggunakan teknik *clustering* dengan algoritma *K-Means*.

Penelitian yang berkaitan dengan *clustering* penentuan stok barang pernah dilakukan oleh Fintri Indriyani dan Eni Irfiani dengan judul "*Clustering* Data Penjualan pada Toko Perlengkapan *Outdoor* Menggunakan Metode *K-Means*". Fokus masalah dalam penelitian ini adalah melakukan analisis penerapan Data *Mining* dalam mengelompokkan jumlah data penjualannya. Metode penyelesaian dengan menggunakan metode *K-Means* dengan penerapan data *mining*. Hasil dari penelitian ini dapat memaksimalkan manajemen dan mengatur stok barang penjualan pada Genta Corp untuk menjaga kepuasan pelanggan (Indriyani. F & Irfiani. E, 2019).

Persediaan produk memiliki peran penting bagi toko dalam mempermudah manajemen produk serta memberikan informasi tentang ketersediaan produk saat terjadi kekurangan (Srisulistiowati, D. B & Khaerudin. M, 2020).

Teknologi data *mining* untuk mengoptimalkan kinerja toko dalam memprediksi jumlah barang yang akan terjual berdasarkan minat konsumen, sehingga pemilik usaha dapat menyediakan produk sesuai dengan perkiraan yang akurat. Penerapan data *mining* dengan metode *clustering* digunakan untuk mengetahui minat konsumen dalam pembelian jenis produk yang mereka inginkan. Sebagai hasil dari pengelompokan data ini, toko dapat mengidentifikasi barang yang paling laris, laris, dan tidak laris, sehingga persediaan di gudang tidak menumpuk. Pengolahan data ini diharapkan memberikan solusi konkret bagi perusahaan dalam memahami barang mana yang paling laris, laris, serta yang

kurang laris. Salah satu metode yang dapat digunakan untuk mencapai tujuan ini adalah penerapan data *mining* (Parsaoran, dkk. 2019).

Permasalahan yang dihadapi pada Toko Mandar Sutera disebabkan oleh kesulitan dalam menentukan stok barang yang perlu dipenuhi sesuai dengan minat konsumen. Untuk mengatasi masalah ini, toko memerlukan metode dan sistem perencanaan stok yang lebih efisien, sehingga dapat menentukan produk mana yang harus disediakan dalam jumlah banyak, sedang, atau sedikit, guna menghindari kekurangan atau kelebihan stok pada jenis barang tertentu.

Solusi untuk mengatasi permasalahan ini adalah dengan memanfaatkan algoritma *K-Means clustering*. *K-Means* merupakan salah satu algoritma *clustering* yang paling sederhana dibandingkan dengan algoritma *clustering* lainnya. Kelebihan algoritma ini terletak pada kemudahannya untuk diterapkan dan dijalankan, kecepatannya yang relatif tinggi, mudah diadaptasi, serta menjadi salah satu algoritma yang paling sering digunakan dalam berbagai tugas data *mining* (Harahap. B, 2019).

Berdasarkan permasalahan yang telah diuraikan diatas penulis melakukan penelitian dengan judul **"**Implementasi Data *Mining* Transaksi Penjualan Menggunakan Algoritma C*lustering* dengan Metode *K-Means* dengan studi kasus pada Toko Mandar Sutera.

B. Rumusan Masalah

Berdasarkan dari uraian latar belakang masalah dapat dirumuskan permasalahan sebagai berikut :

1. Bagaimana cara mengelompokkan produk berdasarkan data jumlah transaksi penjualan menggunakan algoritma *K-Means* ?
2. Bagaimana efektivitas algoritma *K-Means* dalam mengelompokkan produk ?

C. Batasan Masalah

Agar penelitian ini tetap fokus, maka adapun batasan masalah dalam tugas akhir ini adalah sebagai berikut :

1. Penelitian ini menggunakan metode algoritma *K-Means clustering* dilakukan berdasarkan hasil data yang didapat di Toko Mandar Sutera.
2. Atribut atau kriteria yang digunakan adalah stok awal, stok keluar dan stok akhir yang terjual.
3. Melakukan pengelompokan data menjadi 3 *cluster*.

D. Tujuan Penelitian

Adapun tujuan ysng dicapai dalam penelitian ini sebagai berikut :

1. Mengimplementasikan algoritma *K-Means* untuk mengelompokkan produk berdasarkan jumlah data transaksi penjualan.
2. Mengetahui efektivitas algoritma *K-Means* dalam pengelompokan produk.

E. Manfaat Penelitian

Adapun manfaat dari penelitian ini sebagai berikut :

1. Implementasi data *mining* menggunakan algoritma *K-Means clustering* untuk menentukan tingkat penjualan di Toko Mandar Sutera.
2. Untuk mengelompokkan produk yang dijual pada Toko Mandar Sutera menjadi beberapa *cluster* untuk mengetahui produk mana yang paling diminati.

BAB II

TINJAUAN PUSTAKA

A. Data *Mining*

Data *mining* adalah proses ekstraksi informasi berguna dari data yang tersimpan dalam database. Data *mining* sendiri merupakan rangkaian proses yang bertujuan mencari informasi tambahan yang belum ada dalam *database* serta menemukan pola-pola data yang bisa diubah menjadi informasi penting melalui proses pemisahan dan pengenalan pola yang relevan atau menarik dari data yang ada. Data *mining* sebenarnya merupakan bagian dari proses pencarian pengetahuan dalam database, yang dikenal dengan istilah *Knowledge Discovery in Database* (KDD).

Data *mining* merupakan proses menemukan informasi dari suatu data yang tersimpan dalam suatu *database* atau *datasheet*. Pembuatan model dilakukan dengan proses menggunakan algoritma atau rumus tertentu. Proses data *mining* menggunakan berbagai teknik seperti teknik dalam proses statistik, matematika, dan *machine learning* yang digunakan dalam melakukan identifikasi dan mengolah berbagai data menjadi informasi yang bermanfaat (Arhami & Nasir, 2020).

Suatu konsep yang dipakai dalam menghasilkan suatu aturan dalam penemuan pengetahuan disebut data *mining*. Data *mining* atau tambang data merupakan metode, teknik, *artificial intelegent* dan mesin pembelajaran yang diekstrasi sehinga menghasilkan suatu pengetahuan dan informasi yang berguna

yang tersimpan dalam suatu *database* besar. Pada prinsipnya data *mining* mewarisi banyak aspek dan teknik bidang-bidang ilmu. Data *mining* bukanlah sesuatu hal yang baru, karena data *mining* merupakan akar dari berbagai bidang ilmu tersebut (Nasir. J, 2020).

Data *mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau Algoritma dalam data *mining* sangat bervariasi, Pemilihan metode atau Algoritma yang tepat sangat bergantung pada tujuan dan proses secara keseluruhan (Setiawan, 2020). Data *mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan *visualisasi* untuk pengenalan permasalahan pengambilan informasi dari *database* yang besar (Setiawan. R, 2016). Data *mining* menggunakan beberapa teknik, antara lain (Y. Asriningtias, 2014) :

1. *Association Discovery*

Association Discovery adalah teknik mempelajari sekumpulan data dana untuk menunjukkan hubungan antara kemunculan atribut-atribut dalam data.

2. *Clustering*

Clustering dapat juga digunakan untuk mendeteksi secara otomatis *cluster* dari *record-record* yang berdekatan dengan pengertian tertentu didalam keseluruhan variabel-variabel.

3. *Sequential Discovery*

Sequential discovery adalah teknik mencari pola-pola diantara peristiwa-peristiwa yang muncul dalam periode waktu.

4. *Classification*

Classification adalah proses pengumpulan data bersama sama yang didasarkan atas sekumpulan kesamaan yang awalnya telah ditentukan oleh seorang analis sebelum analisa dimulai.

5. *Neural Network*

Neural network merupakan sebuah metode khusus untuk pengendalian identifikasi pola yang digunakan pada *trend* perkiraan berdasarkan kebiasaan yang telah diketahui sebelumnya.

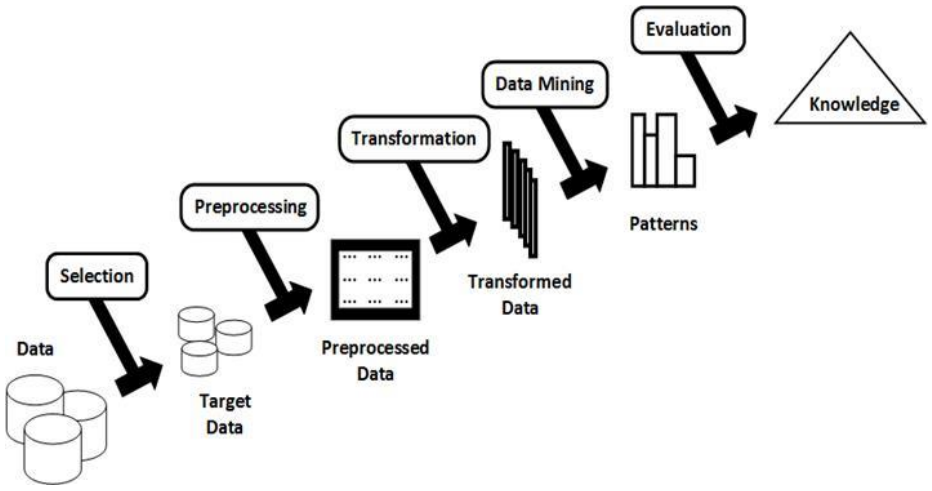
B. *Knowledge Discovery in Databases* (KDD)

Knowlegde Discovery in Databases (KDD) adalah sekumpulan proses untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan yang berguna.*Knowlegde Discovery in Databases (KDD)* terdiri dari serangkaian langkah perubahan, termasuk data *pre-processing* dan juga *post processing*. Data *preprocessing* merupakan langkah untuk mengubah data mentah menjadi format yang sesuai untuk tahap analisis berikutnya (I. Setiawan, 2018).

KDD berkaitan dengan teknik integrasi dan penemuan ilmiah. Interpretasi dan visualisasi pola yang dihasilkan dari kumpulan data KDD adalah proses *non-trivial* yang bertujuan untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, di mana pola tersebut harus sah, bermanfaat, dan dapat dipahami. Proses ini mencakup tahap pembersihan dan integrasi data. Proses tersebut digunakan untuk

menghapus data yang tidak konsisten dan mengandung *noise* dari berbagai basis data yang mungkin memiliki format atau platform yang berbeda, yang kemudian diintegrasikan ke dalam suatu data *warehouse* (Nurani, N., & Gani, H, 2017).

Knowledge Discovery in Database (KDD) suatu teknik pembentukan pola atau rule dalam informasi. Informasi yang dihasilkan didapatkan dari suatu data yang besar atau dikenal dengan tambang data yang disimpan dalam basis data yang awalnya belum diketahui dan menghasilkan suatu data yang potensial bermanfaat. Iterasi dalam data *mining* disebut proses KDD. Berikut langkah-langkah dari KDD (Januardi Nasir, 2020) :



Gambar 2.1. Tahapan dalam KDD (*Knowledge Discovery in Database*)

Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Zulfaa. I, dkk. 2020) :

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang

akan digunakan untuk proses data *mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/ Cleaning*

Sebelum proses data *mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses "memperkaya" data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih sehingga data tersebut sesuai untuk proses data *mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. Data *Mining*

Data *mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data *mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation/ Evaluation*

Pola informasi yang dihasilkan dari proses data *mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini

merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya. Melalui berbagai penjelasan diatas dapat ditarik kesimpulan bahwa data *mining* adalah sebuah proses yang memiliki peran penting dalam kehidupan sehari-hari dengan memahami definisi, fungsi, metode penerapannya sehingga mudah untuk diimplementasikan dalam kehidupan sehari-hari mulai dari telekomunikasi, asuransi, olahraga hingga keuangan. Tanpa data *mining* kemungkinan proses pengambilan keputusan dari suatu masalah akan lebih sulit dilakukan karena tidak ada data yang bisa menjadi dasar pertimbangan.

C. *Clustering*

Clustering disebut sebagai *segmention*. Metode ini mengidetifikasi kelompok dalam sebuah kasus yang didasarkan pada kelompok atribut yang memiliki kemiripan. Cara kerja *clustering* memisahkan sejumlah kelompok data berdasarkan ciri masing-masing, dimana objeknya dapat berupa orang, peristiwa dan lainya yang didistribusikan ke dalam kelompok sehingga terdapat beberapa tingkatan yang saling berhubungan antar *cluster*, kuat dan lemahnya antar anggota dari *cluster* yang berbeda terlihat pada anggota *cluster* yang sama.

Berbeda dengan klasifikasi, pengklasteran tidak melibatkan variabel target. *Clustering* tidak digunakan untuk mengklasifikasikan, mengestimasi, atau memprediksi nilai dari suatu target. Sebaliknya, pengklasteran digunakan untuk

membagi keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (Dharma. P.Y, dkk, 2021).

Clustering adalah salah satu teknik dalam data *mining* yang bertujuan untuk mengidentifikasi kelompok objek yang memiliki karakteristik serupa, yang dapat dipisahkan dari kelompok lainnya. Dengan demikian, objek dalam kelompok yang sama akan relatif lebih homogen dibandingkan dengan objek di kelompok yang berbeda. Jumlah kelompok yang dapat diidentifikasi bergantung pada jumlah dan variasi data objek. Tujuan dari pengelompokan sekumpulan data objek ke dalam beberapa kelompok yang memiliki karakteristik tertentu dan dapat dibedakan satu sama lain adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan. Model yang digunakan diasumsikan dapat memanfaatkan data dalam bentuk data interval, frekuensi, dan biner.

D. Algoritma *K-Means Clustering*

Algoritma *K-Means* terkenal karena kesederhanaannya dan kemampuannya untuk mengklaster data dalam jumlah besar serta menangani *outlier* dengan sangat cepat. Dalam algoritma ini, setiap data harus dimasukkan ke dalam *cluster* tertentu, dan dimungkinkan bahwa data yang berada dalam satu *cluster* pada tahap awal dapat berpindah ke *cluster* lain pada tahap berikutnya. Posisi pusat *cluster* akan terus dihitung ulang hingga semua data diklasifikasikan sesuai dengan pusat *cluster* masing-masing, dan akhirnya terbentuk posisi pusat *cluster* yang baru (Putri & Saepudin. S, 2021).

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* menetapkan nilai nilai *cluster* (k) secara acak, dimana nilai tersebut menjadi pusat dari *cluster* atau disebut sebagai *centroid*, *mean* atau *means*. Algoritma *K-Means* dalam implementasinya sangat mudah, cepat, mudah beradaptasi sederhana untuk diimplementasikan dan dijalankan, relatif cepat, dan mudah beradaptasi serta mempunyai kemampuan yang besar dalam mengolah data yang cukup besar dan waktu lebih efisien.

Tahapan dalam melakukan *clustering* dengan metode *K-Means* adalah sebagai berikut:

1. Pilih secara acak k buah data sebagai pusat *cluster*.
2. Pemberian nilai k pusat *cluster* ini dapat dilakukan dengan berbagai cara dan salah satunya adalah secara *random.* Pusat *cluster* diberi nilai awal dengan angka acak.
3. Alokasikan semua data/objek pada *cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antar data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster.* Jarak antara satu data dengan satu *cluster* tertentu akan menentukan suatu data akan masuk ke dalam *cluster* yang mana. Rumus menghitung jarak data ke setiap titik pusat *cluster* bisa menggunakan jarak *Euclidean Distance* yang dijelaskan dalam:

   $$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \ldots + (X_{ki} - X_{kj})^2}$$ **Persamaan (2.1)**

   Keterangan :

   D(i,j) =Jarak data ke i ke pusat *cluster* j

Xki =Data ke i pada atribut data ke k

Xkj = Titik pusat ke j pada atribut ke k

4. Kalkulasi lagi pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah *average* dari semua data/objek dalam *cluster* tertentu dan bisa juga menggunakan media dari *cluster* tersebut. Pusat *cluster* baru akan ditentukan bila semua data telah ditetapkan dalam *cluster* terdekat. Rumus menghitung titik pusat baru:

   $C_k = \frac{1}{nk} \sum d_l$ **Persamaan (2.2)**

   Keterangan :

   Nk = Jumlah dokumen dalam *cluster* k

   *d*1 = Dokumen dalam *cluster* k

5. Ulang lagi perintah bagi setiap objek untuk menggunakan pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses *cluster* selesai atau kembali ke langkah bagian 3 sampai pusat *cluster* tidak berubah lagi.

Berdasarkan penjelasan di atas, dapat disimpulkan bahwa *clustering* adalah metode analisis data yang sering digunakan dalam data mining, dengan tujuan mengelompokkan data berdasarkan karakteristik yang serupa ke dalam satu kelompok, dan data dengan karakteristik berbeda ke dalam kelompok lain. Konsep dasar metode *K-Means cluster*ing terdiri dari lima tahap: menentukan titik pusat *cluster*, menghitung jarak data ke setiap *cluster*, mengalokasikan data ke *cluster*, menentukan titik pusat *cluster* yang baru, dan terakhir memverifikasi titik pusat *cluster*.

Sebagai salah satu contoh penerapan algoritma *K-Means clustering* adalah sebuah toko ingin mengelompokkan produk yang dijual berdasarkan tingkat penjualan untuk membantu manajemen pendataan. Toko tersebut memiliki data penjualan barang 3 bulan terakhir, meliputi jumlah unit terjual, jumlah stok awal, dan jumlah stok yang tersisa. Adapun data yang meliputi barang A dengan 300 unit terjual, stok awal 500 unit, dan sisa stok 200 unit, barang B dengan 50 unit terjual, stok awal 100 unit, dan sisa stok 50 unit, barang C dengan 400 unit terjual, stok awal 400 unit, dan sisa stok 0 unit, dan barang D dengan 30 unit terjual, stok awal 150 unit, dan sisa stok 120 unit. Jumlah *cluster* ditentukan, misalnya K=3 (sangat laris,laris,kurang laris). Kemudian melakukan tahap *clustering* dengan metode *K-Means* dengan jumlah *cluster* yang ditentukan mendapatkan hasil akhir, barang dikelompokkan *cluster* 1 (sangat laris) pada barang C, *cluster* 2 (laris) pada barang A dan B, dan *cluster* 3 (kurang laris) pada barang D. Hasil *clustering* dapat membantu toko menentukan strategi, seperti meningkatkan stok barang yang sangat laris, menyesuaikan promosi untuk barang yang kurang laris, dan mengevaluasi strategi penjualan barang tertentu.

E. Penelitian Terkait

Penelitian terkait adalah upaya peneliti untuk mencari perbandingan terkait penelitian selanjutnya. Pada bagian ini peneliti memiliki beberapa penelitian terkait dengan perbandingan dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan kode *K-Means clustering* dalam klasifikasi penjualan di Mandar Sutera sebagai berikut :

1. Hani Prastiwi, Jeny Pricilia, Errissya Raswir pada tahun 2022 mengadakan penelitian berjudul **"Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode *K-Means Clustering*"**. Dalam penelitian ini dapat diambil kesimpulan bahwa Implementasi Data *RapidMiner* Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode *K-Means Clustering*. Implementasi metode *clustering* tersebut lebih efektif dan efesien sehingga dapat tersedianya barang dengan jenis dan jumlah yang sesuai kebutuhan konsumen. Melalui pencarian kesamaan dalam data, sesorang dapat mempresentasikan data yang sama dengan lebih sedikit simbol. Pada hasil pengujian dengan 20 data, *cluster* optimal menyumbangkan 17 data untuk *cluster* C1 dan 3 data untuk *cluster* C2.
2. Fintri Indriyani dan Eni Irfian pada tahun 2021 mengadakan penelitian berjudul **"*Clustering* Data Penjualan pada Toko Perlengkapan *Outdoor* Menggunakan Metode *K-Means*"**. Dalam penelitian ini dapat diambil kesimpulan bahwa data diolah dengan perhitungan manual menggunakan algorithma *K-Means* dan menggunakan *Software RapidMiner* sehingga didapatkan hasil akhir berupa tiga *cluster* dimana terdapat 2 jenis barang paling laris, 8 jenis barang yang cukup laris dan 18 jenis barang yang kurang laris. Penerapan metode *K-Means* dalam pengelompokan data penjualan pada Toko Genta Corp dapat menghasilkan rekomendasi barang yang laris, Kurang laris dan cukup laris. Sehingga data dijadikan rujukan bagi menajemen untuk mengatur stok barang agar toko tidak mengecewakan pelaggan karena barangyang ingin di beli tidak tersedia.

3. Agung Nugraha, Odi Nurdiawan, Gifthera Dwilestari pada tahun 2022 mengadakan penelitian berjudul **"Penerapan Data Mining Metode *K-Means Clustering* Untuk Analisa Penjualan Pada Toko Yana Sport"**. Dalam penelitian ini dapat diambil kesimpulan bahwa mendapatkan informasi atau pola dari penerapan algoritma *K-Means* dengan data penjualan terdapat sebanyak 99 item barang yang laris terjual dan terdapat 23 item barang yang tidak terjual sehingga pemilik dapat melakukan strategi penjualan dan pembelian ulang berdasarkan barang yang laris terjual.
4. Januardi Nasir pada tahun 2020 mengadakan penelitian berjudul **"Penerapan Data *RapidMiner Clustering* Dalam Mengelompokan Buku Dengan Metode *K-Means*"**. Dalam penelitian ini dapat disimpulkan bahwa Hasil yang diperoleh antara hitungan manual dan *RapidMiner* dari data peminjam buku adalah sama dengan hasil yang telah diproses maka didapatkan jumlah buku yang banyak dipinjam terdapat pada *cluster* 1 sebanyak 9 item, jumlah buku yang paling sedikit dipinjam terdapat pada *cluster* 2 sebanyak 15 item, jumlah buku yang cukup banyak dipinjam terdapat pada *cluster* 0 sebanyak 12 item. Penerapan data *mining* dengan metode *K-Means clustering* dapat diterapkan pada pengelompokan buku sehingga membantu pihak Perpustakaan sehingga dapat mengetahui buku mana yang sering dipinjam.
5. Suhandio Handoko, Fauziah , Endah Tri Esti Handayani pada tahun 2020 mengadakan penelitian berjudul **"Implementasi Data *Mining* Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode *K-Means Clustering*"**. Dalam penelitian ini dapat disimpulkan bahwa

hasil dari metode Algoritma *K-Means clustering* data *mining* didapatkan daerah penjualan produk yang tinggi,sedang, dan rendah. Daerah dengan penjualan produk yang rendah akan dilakukan promosi penjualan produk dan untuk daerah penjualan yang tinggi tidak diadakan promosi.

6. Sabrina Aulia Rahmah pada tahun 2020 mengadakan penlitian berjudul **"Klasterisasi Pola Penjualan Pestisida Menggunakan Metode *K-Means Clustering* (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)"**. Dalam penelitian ini dapat disimpulkan bahwa Metode *K-Means clustering* dapat diterapkan pada penjualan Pestisida di Toko Jaunda Tani, sehingga metode ini sangat membantu dalam mengelompokan pola penjualan selama satu Musim dan pengujian dengan *RapidMiner* 7.5 sangat efektif dan akurat karena nilai centroid yang didapat dengan perhitungan yang dilakukan secara manual dan aplikasi adalah sama.
7. Haditsah Annur pada tahun 2019 mengadakan penelitian berjudul **"Penerapan Data *Mining* Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode *K-Means Clustering* (Studi Kasus Toko Luxor Variasi Gorontalo)"**. Dalam penelitian ini dapar diambil kesimpulan bahwa hasil penelitian dan perhitungan maka menentukan strategi penjualan variasi mobil menggunakan metode *K-Means clustering* dapat diterapkan. dapat digunakan untuk memberi saran pertimbangan dalam menentukan strategi penjualan yaitu mengeliminasi produk dengan posisi *cluster* terbawah dan lebih memfokuskan pada produk dengan posisi *cluster* tertinggi.
8. Hendra Agustian Siregar , Azlan , Nur Yanti Lumban Gaol pada tahun 2023

mengadakan penelitian berjudul **"Penerapan Data *Mining* Pada Penjualan Rumah Makan Kasih Ibu Menggunakan Metode *K-Means Clustering*"**. Dalam penelitian ini dapat diambil kesimpulan bahwa sistem yang telah dibangun dapat digunakan untuk memberikan informasi terkait penjualan data menu rumah makan yang paling diminati, cukup diminati dan kurang diminati dalam penerapan data *mining Microsoft visual Studio* 2008 dan menggunakan *database Microsoft Access* 2007 sebagai penyimpanan basis data pada sistem.

9. Putri Ulil Fatma Aulia, Sudin Saepudin pada tahun 2021 mengadakan penelitian berjudul **"Penerapan Data *Mining K-Means Clustering* Untuk Mengelompokkan Berbagai Jenis Merk Laptop"**. Dalam penelitian ini dapat disimpulkan bahwa telah berhasil dilakukan pengelompokkan data laptop menggunakan algoritma *K-means clustering* menjadi 3 kelompok yaitu untuk kelompok 1 berjumlah 11 data, kelompok 2 berjumlah 6 data, kelompok 3 berjumlah 8 data. Berdasarkan pengujian yang dilakukan mampu memberikan hasil yang baik sesuai dengan perhitungan yang digunakan dan dapat membantu konsumen yang akan membeli. penelitian ini juga bisa memberi hubungan positif yang sangat signifikan antara persepsi konsumen terhadap merek, kualitas dengan minat membeli.
10. Ferlanda, Septi Andryana, Eri Mardiani pada tahun 2021 mengadakan penelitian yang berjudul **"Implementasi Data *Mining* Untuk Menentukan Persediaan Stok Obat Di Enok Menggunakan Metode *K-Means Clustering*"**. Dalam penelitian ini dapat disimpulkan bahwa penelitian ini adalah untuk menerapkan algoritma *K-Means*, dan data transaksi obat dari

Apotek Enok di berikan sebagai contoh tipikal. Hasil analisis untuk penelitian ini menggunakan 20 buah data. Pengumpulan data obat yang di lakukan dengan algortma *K-Means* diulang sebanyak tiga kali, kemudian didapatkan hasil kelompok yaitu kelompok 1 berisi 6 obat kerja lambat dan kelompok 2 terdapat 14 obat kerja cepat. Pencarian *cluster* ini menggunakan fitur *web* untuk menemukan produk lambat dan obat cepat.

DAFTAR PUSTAKA

Annur, H. (2019). Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode *K-Means Clustering* (Studi Kasus Toko Luxor Variasi Gorontalo). JURNAL INFORMATIKA UPGRIS Vol.5,No.1.https://doi.org/10.26877/jiu.v5i1.3091

Arhami, M., & Nasir, M. (2020). Data Mining Algoritma dan Implementasi Penerbit Andi. https://books.google.co.id/books/about/Data_Mining_Algoritma_dan_Implementasi.html?id=AtcCEAAAQBAJ&redir_esc=y

Aulia, N. , Indra, & Nur, N. (2021). Implementasi Data Mining menggunakan Algoritma Apriori untuk Menentukan Pola Pembelian Obat di Rumah Sakit. Journal of Computer and Information System (J-CIS), Vol 4(2), 1929. https://doi.org/10.31605/jcis.v4i2.1259

Aulia, P, U, F., & Saepudin, S. (2021). Penerapan Data Mining *K-Means Clustering* Untuk Mengelompokkan Berbagai Jenis Merk Laptop. Seminar Nasional Sistem Informasi dan Manajemen Informatika,

Vol.1 .https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/47

Aulia, S. (2021). Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). Djtechno: Jurnal Teknologi Informasi, 1(1) https://doi.org/10.46576/djtechn o.v1i1.964

Dharma Putra, Y., Sudarma, M., & Swamardika, I. B. A. (2021). Clustering History Data Penjualan Menggunakan Algoritma K-Means. Majalah Ilmiah Teknologi Elektro, 20(2), 195. https://doi.org/10.24843/mite.2021.v20i02.p03

Handoko, S., Fauziah, & Handayani, E, T, E. (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode *K-Means Clustering*. Jurnal Ilmiah Teknologi dan Rekayasa, Vol 25,No 1. http://dx.doi.org/10.35760/tr.2020.v25i1.2677

Harahap, B (2019). Penerapan Algoritma *K-Means* Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung).

I. Setiawan. (2018). Knowledge Discovery In Databases (KDD) Terhadap Customer Reviews Pada Situs E-Commerce.Program Studi Sistem Informasi, no. 09031281621045.

Indriany, F., & Irfiany, E. (2019). *Clustering* Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode *K-Means*. Jurnal Informatika, Vol.7, No. 2. 10.30595/juita.v7i2.5529

Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. Jurnal Teknologi Dan Sistem Informasi (JTSI), 2(2)

http://jim.teknokrat.ac.id/index.php /JTSI

Nasir, J. (2020). Penerapan Data Mining *Clustering* Dalam Mengelompokan Buku Dengan Metode *K-Means.* Jurnal SIMETRIS, Vol. 11 No. 2. https://doi.org/10.24176/simet.v11i2.5482

Nugraha, A. , Nurdiawan, O. , & Dwilestari, G. (2022). Penerapan Data Mining Metode *K-Means Clustering* Untuk Analisa Penjualan Pada Toko Yana Sport. Jurnal Mahasiswa Teknik Informatika,Vol.6 No.2 https://doi.org/10.36040/jati.v6i2.5755

Parsoaran, S.T., Toknady, F. K., & Feryanto. (2019). Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Menggunakan K-Means Clustering,(online) Vol. 2, No. 2. Jusikom Prima.

Prastiwi, H., Pricilia, J., & Raswir, E. (2022). Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode *K-Means Clustering*. Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM), Vol. 1 No. 2 10.33998/jakakom.2022.2.1.34

Rahmah, S, A. (2020). Klasterisasi Pola Penjualan Pestisida Menggunakan Metode *K-Means Clustering* (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). Journal of Information Technology Research, Vol. 1 No. 1.https://doi.org/10.46576/djtechno.v1i1.964

Setiawan, R., (2016). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi,Jurnal Lentera ICT ISSN: 2338-3143, Volume 3, Nomor 1

Siregar, H, A. , Azlan, & Gaol, N, Y, L. (2023). Penerapan Data Mining Pada

Penjualan Rumah Makan Kasih Ibu Menggunakan Metode *K-Means Clustering*. JURNAL SISTEM INFORMASI TGD, Vol. 2 No. 5. https://doi.org/10.53513/jursi.v2i5.8955

Srisulistiowati, D. B., & , Muhamad Khaerudin, S. R. (2020). Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode FpGrowth (Studi Kasus Toko Koperasi Sekolah Bina Mulia). Jurnal Sistem Informasi Universitas Suryadarma, 8(2). https://doi.org/10.35968/jsi.v8i2.739

Y. Asriningtias. (2014). APLIKASI DATA MINING UNTUK MENAMPILKAN INFORMASI TINGKAT KELULUSAN MAHASISWA, vol. 8, no. 1

Zulfaa,I., Rayuwatib, R., & Kokoa, K. (2020). Implementasi data mining untuk menentukan strategi penjualan buku bekas dengan pola pembelian konsumen menggunakan metode Apriori (studi kasus: Kota Medan), (Online) Vol 16 No.1 http://jurnal.untirta.ac.id/index.php/ju-tek