CLUSTERING PADA DATA TRANSAKSI PENJUALAN MINIMARKET MENGGUNAKAN ALGORITMA K-MEANS

TUGAS AKHIR

Radhie Wiradika
41518010018

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

http://digilib.mercubuana.ac.id/

CLUSTERING PADA DATA TRANSAKSI PENJUALAN MINIMARKET MENGGUNAKAN ALGORITMA K-MEANS

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Radhie Wiradika

41518010018

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

http://digilib.mercubuana.ac.id/

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM	: 41518010018
Nama	: Radhie Wiradika
Judul Tugas Akhir	: Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 04 Juli 2022

Radhie Wiradika

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Radhie Wiradika
NIM	:	41518010018
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 04 Juli 2022

UNIVERSITAS
MERCU BUANA

METERAI TEMPEL
086AJX931500102

Radhie Wiradika

http://digilib.mercubuana.ac.id/

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Radhie Wiradika
NIM	:	41518010018
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓
		Jurnal Nasional Terakreditasi	✓		
		Jurnal International Tidak Bereputasi		Diterima	
		Jurnal International Bereputasi			
	Disubmit/dipublikasikan di :	Nama Jurnal	: Jurnal Informatika: Jurnal Pengembangan IT		
		ISSN	: 2477-5126		
		Link Jurnal	: https://ejournal.poltektegal.ac.id/index.php/informatika/index		
		Link File Jurnal Jika Sudah di Publish	:		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 04 Juli 2022

METERAI TEMPEL
2EDAJX931500127

Radhie Wiradika

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010018
Nama	:	Radhie Wiradika
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022

(Rahmat Budiarto, Dr. Prof)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010018
Nama	:	Radhie Wiradika
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022

(Rusherlata, S.Kom, MT)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518010018
Nama	:	Radhie Wiradika
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022

(Dr. Harwikarya, MT)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM	:	41518010018
Nama	:	Radhie Wiradika
Judul Tugas Akhir	:	Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 27 Juli 2022

Menyetujui,

(Yaya Sudarya Triana, Ph.D.)
Dosen Pembimbing

Mengetahui,

(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika

(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

KATA PENGANTAR

Segala puji dan syukur kehadirat Allah SWT atas berkah, rahmat dan hidayah-nya yang senantiasa dilimpahakan kepada penulis, sehingga bisa menyelasaikan skripsi dengan judul "Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means" sebagai syarat untuk menyelesaiakan Program Sarjana (S1) pada Program Sarjana Fakultas Ilmu Komputer Jurusan Informatika Universitas Mercu Buana.

Dalam penyusunan skripsi ini banyak hambatan serta rintangan yang penulis hadapi namun pada akhirnya dapat melaluinya berkat adanya bimbingan dan bantuan dari berbagai pihak baik secara moral maupu spiritual. Untuk itu pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua yang telah memberikan support serta kasih sayang hingga saya bisa menyelesaikan tugas akhir ini.
2. Bapak Yaya Sudarya Triana, Ph.D selaku dosen pembimbing saya yang telah membantu dan memberikan arahan kepada saya dalam menyelesaikan tugas akhir ini.
3. Bapak Emil R. Kaburuan Ph.D selaku Kepala Program Studi Informatika yang telah memberikan arahan dan bimbingannya selama perkuliahan.
4. Bapak Wawan Gunawan S.Kom.,MT selaku Koordinator Tugas AkhirProdi Informatika yang telah memberikan arahan selama periode tugas akhir
5. Bapak Sabar Rudiarto, M.Kom. selaku dosen pembimbing akademik saya yang telah membimbing selama perkuliahan hingga sampai pada tugas akhir.
6. Ibu Endang Kusminarti, SE, yang sudah memperbolehkan saya menggunakan data transaksi penjualan di Koperasi PT. Kawasan Berikat Nusantara untuk penelitian ini.
7. Teman-teman Informatika 2018, sahabat dan kerabat yang selalu memberikan support dan waktu untuk selalu mengingatkan dalam menyelesaikan tugas akhir ini.

Penulis menyadari bahwa penyusunan Jurnal Tugas Akhir ini masih jauh dari kesempurnaan dan masih terdapat kekurangan dan kelemahan, walaupun demikian, penulis telah berusaha semaksimal mungkin untuk mendapatkan hasil

http://digilib.mercubuana.ac.id/

yang terbaik. Akhir kata penulis Akhir kata, penulis berharap jika ada kritik atau saran apapun yang sifatnya membangun bagi saya dengan senang hati akan saya terima.

Jakarta, 04 Juli 2022

Radhie Wiradika

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

DAFTAR ISI

HALAMAN SAMPUL i
HALAMAN JUDUL i
LEMBAR PERNYATAAN ORISINALITAS ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR ... iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR iv
LEMBAR PERSETUJUAN PENGUJI v
LEMBAR PENGESAHAN viii
ABSTRAK ix
ABSTRACT x
KATA PENGANTAR xi
DAFTAR ISI xiii
NASKAH JURNAL 1
KERTAS KERJA 7
BAB 1. LITERATUR REVIEW 8
BAB 2. ANALISIS DAN PERANCANGAN 15
BAB 3. SOURCE CODE 18
BAB 4. DATASET 23
BAB 5. TAHAPAN EKSPERIMEN 26
BAB 6. HASIL SEMUA EKSPERIMEN 31
DAFTAR PUSTAKA 37
LAMPIRAN DOKUMEN HAKI 38
LAMPIRAN KORESPONDENSI 41
LAMPIRAN PERSETUJUAN DOSEN PEMBIMBING TA 42
LAMPIRAN PERSETUJUAN DOSEN PA 43

http://digilib.mercubuana.ac.id/

NASKAH JURNAL

Clustering Pada Data Transaksi Penjualan *Minimarket* Menggunakan Algorima *K-Means*

Radhie Wiradika[1], Yaya Sudarya Triana[2*)]
[1]Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercubuana, Jakarta
[2]Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercubuana, Jakarta
email: [1]radhiewiradika@gmail.com, [2*)]yaya.sudarya@mercubuana.ac.id

Abstract – Marketing strategy is carried out to achieve the targets in which are listed decisions regarding the target market, product placement, marketing mix, and thelevel of marketing costs required are listed. Therefore, it is necessary to carry out a data mining process on sales transaction data to obtain useful information for businessmen. This study aims to determine the selling, medium and unsold clusters and prove elbow performance to produce optimal clusters on sales transaction data using the K-Means algorithm as a data clustering method. Cluster optimization is obtained from elbow methods executed with Kaggle using the Python programming language. It can be concluded that the result of cluster k=3 with the caption that cluster 1 is the number of clusters that sell best total 2 items. Then cluster 0 with the caption cluster 0 is an unsold cluster with a total of 1196 items. Finally, cluster 2 with a description of the number of medium-sized clusters is 40 items, and shows that the Elbow method works very well in producing optimal clusters, which is found at k = 3 with the SSE difference value is 7.37E + 13 with k test = 9. And the results of the goods that sell well are purchased by consumers, namely: GG FILTER 12 and S MILD 16 BARU.

Keywords – Clustering, K-Means, Sales Transactions, Elbow.

Abstrak – Strategi pemasaran dilakukan untuk mencapai sasaran yang di dalamnya tercantum keputusan-keputusan mengenai target pasar, penempatan produk, bauran pemasaran, serta tingkat biaya pemasaran yang diperlukan. Maka dari itu, perlu dilakukan proses data mining pada data transaksi penjualan untuk mendapatkan informasi yang bermanfaat bagi para pembisnis. Penelitian ini bertujuan untuk menentukan cluster laku, laku sedang dan tidak lakunya serta membuktikan kinerja Elbow untuk menghasilkan cluster yang optimal pada data transaksi penjualan menggunakan algoritma K-Means sebagai metode pengelompokan data. Optimasi cluster diperoleh dari metode Elbow yang dieksekusi dengan Kaggle yang menggunakan bahasa pemrograman Python. Dapat disimpulkan bahwa hasil cluster k=3 dengan keterangan bahwa cluster 1 adalah jumlah cluster yang paling laku berjumlah 2 item. Kemudian cluster 0 dengan keterangan cluster 0 adalah cluster yang tidak laku dengan jumlah 1196 item. Terakhir cluster 2 dengan keterangan jumlah cluster laku sedang yaitu 40 item, serta memperlihatkan bahwa metode Elbow bekerja dengan sangat baik dalam menghasilkan cluster yang optimal yaitu terdapat pada k=3 dengan nilai selisih SSE adalah 7.37E+13 dengan k uji=9. Dan hasil barang yang laku dibeli konsumen yaitu : GG FILTER 12 dan S MILD 16 BARU.

***Kata Kunci – Clustering, K-Means* , Transaksi Penjualan , Elbow.**

*) **penulis korespondensi**: Yaya Sudarya Triana
Email: yaya.sudarya@mercubuana.ac.id

I. PENDAHULUAN

Siklus bisnis berkembang sangat cepat seiring dengan perubahan teknologi [1] & [2]. Strategi pemasaran dilakukan untuk mencapai sasaran yang di dalamnya tercantum keputusan-keputusan mengenai target pasar, kebijakan dan pedoman yang mengarahkan operasi pemasaran perusahaan, terutama sebagai respons dari korporasi terhadap lingkungan dan sifat pasar yang dinamis [3]. Pendekatan pokok yang digunakan oleh unit bisnis di dalam mencapai sasaran yang di dalamnya tercantum keputusan-keputusan mengenai target pasar, penempatan produk, bauran pemasaran, serta tingkat biaya pemasaran yang diperlukan. Untuk itu diperlukan suatu analisa dalam bisnis penjualan menggunakan Data Mining [4]. Data Mining merupakan metode yang menarik untuk mengekstraksi pola dari volume data yang sangat besar. Pendekatan yang menarik adalah pendekatan yang tidak sepele, implisit, diketahui sebelumnya, dan membantu. Rencananya harus sederhana untuk dipahami[5].

Data transaksi penjualan merupakan salah satu hal yang dapat digunakan untuk pengambilan keputusan bisnis. Sebagian besar data transaksi penjualan tidak digunakan kembali, transaksi penjualan yang tidak digunakan dengan baik dapat digali dan diolah kembali menjadi informasi yang berguna, data yang digunakan berasal dari Koperasi PT. Kawasan Berikat Nusantara, dengan menggunakan teknik data mining yaitu clustering [6]. Clustering merupakan metode yang terlibat dengan mengumpulkan hal-hal informasi menjadi beberapa kelompok sehingga masing-masing kelompokmempunyai persamaan antar lainnya [7]. Salah satu pendekatan yang digunakan dalam mengembangkan metode clustering yaitu metode K-Means, dimana metode

http://digilib.mercubuana.ac.id/

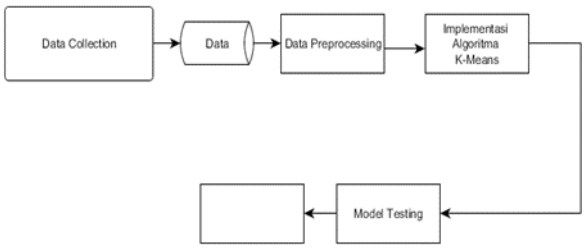
ini merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data ke dalam bentuk dua atau lebih kelompok(cluster) yang berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama [8]. Dengan ini data transaksi penjualan dapat diolah kembali sehingga bisa menghasilkan suatu pola pembelian konsumen. Pola pembelian konsumen inilah yang nantinya akan membantu owner untuk membuat suatu keputusan bisnis.

Melihat permasalahan di atas, penulis tertarik melakukan penelitian dengan memanfaatkan metode Data Mining yaitu clustering, untuk menentukan cluster terdapat 3 cluster yaitu : Cluster Laku, Laku sedang, dan Tidak Laku. Oleh karna itu analisa tentang data transaksi penjualan minimarket, bisa memperkirakan banyak atau tidaknya barang keluar yang bisa dikembalikan.

II. PENELITIAN TERKAIT

Penelitian yang melakukan clustering menggunakan algoritma K-Means telah banyak dilakukan. Sedangkan penelitian ini melakukan clustering dengan data transaksi penjualan Koperasi PT. Kawasan Berikat Nusantara. Pada penelitian terkait mengelopokan penjualan data bahan bangunan dengan metode K-Means Clustering. Hasil dari algoritma K-Means didapatkan bahan bangunan yang laris dibeli konsumen yaitu : Batu Ultra Tipis,Besi 4 Pas, Cat Bravo 5kg, Gembok 40 Risk, Keni 3 inci, Keramik Alberta Brown 40/40 [1]. Mengelompokan wilayah kelayakan tanam jagung di Kabupaten Lampung Selatan dengan algoritma K-Means Clustering. Hasil dari algoritma K-Means dinyatakan wilayah dengan jumlah panen jagung terbanyak adalah Penengahan dengan nilai centroid 2 79448.30257. lalu yang kedua wilayah dengan jumlah panen jagung terendah adalah Candipuro dengan nilai centroid 2 1.424.036868 [9]. Mengelompokan data penjualan batubara periode Januari 2015 hingga Agustus 2016 di PT Global Bangkit Utama menggunakan algoritma K-Means dengan metode Elbow. Hasil Clustering yang didapatkan ialah terdapat persamaan hasil cluster pada k-7 yang menunjukkan bahwa cluster tersebut merupakan cluster yang terbaik [10]. Mengelompokan data penjualan produk pada swalayan Fadhila menggunakan metode K-Means. Hasil dapat cluster 1, terdiri dari 4 produk tidak laris dan 6 produk yang laris di Cluster 2. Perbandingan pengujian menggunakan Tanagra dan XLMiner,pengujian menggunakan SPSS menghasilkan jumlah produk dalam cluster 1 dan cluster 2 yang berbeda. Namun, data produk yang di-clusterkan tetap sama. Yaitu 4produk yang tidak laris dan 6 produk yang laris [11].

III. METODE PENELITIAN



A. *Data Collection*

Pada data collection diambil dari Koperasi PT. Kawasan Berikat Nusantara yang berlokasi di Cilincing Jakarta

TANGGAL_TRANSAKSI	JENIS_TRANSAKSI	JENIS_TRANSAKSI	NO_JENIS_TRANSAKSI	DATA_TRANSAKSI	NAMA_TRANSAKSI
2/2/2022 12:00:	TUNAI	TUNAI	02	T03260	AMPLOP POLOS 90 BESAR
2/2/2022 12:00:	TUNAI	TUNAI	02	T00638	GG FILTER 12
2/2/2022 12:00:	TUNAI	TUNAI	02	T02755	INDOCAFE CF MIX 30X20
2/2/2022 12:00:	TUNAI	TUNAI	02	T02062	KERTAS FOLIO BERGARIS
2/2/2022 12:00:	TUNAI	TUNAI	02	T01227	S MILD 16 BARU
2/2/2022 12:00:	TUNAI	TUNAI	02	T02856	AQUA GELAS220ML/DUS
2/2/2022 12:00:	TUNAI	TUNAI	02	T00406	ADES AIR 1500 ML
2/2/2022 12:00:	TUNAI	TUNAI	02	T00768	GG SIGNATURE 12
2/2/2022 12:00:	TUNAI	TUNAI	02	T02856	AQUA GELAS220ML/DUS
2/2/2022 12:00:	TUNAI	TUNAI	02	T03621	AROMA BOLD 16 FILTER
2/2/2022 12:00:	TUNAI	TUNAI	02	T00648	MARLBORO RED
2/2/2022 12:00:	TUNAI	TUNAI	02	T00766	GULAKU KUNING 1 KG
2/2/2022 12:00:	TUNAI	TUNAI	02	T00816	AQUA 1500ML

Utara. Data yang digunakan diambil dari bulan Februari

JUMLAH_TRANSAKSI	SATUAN_BARANG	TOTAL_TERJUAL	HARGA_JUAL	TOTAL_HARGA	HARGA_BELI	TOTAL_HARGA_BELI
2	Pcs	190	500	1000	380	620
1	Pcs	18600	21000	21000	18600	2400
1	Pcs	1021	2000	2000	1021	979
1	Pack	2250	3000	3000	2250	750
1	Pcs	24100	27000	27000	24100	2900
1	Dus	28600	33000	33000	28600	4400
1	Botol	3799	5000	5000	3799	1201
1	Pcs	16700	18500	18500	16700	1800
1	Dus	28600	33000	33000	28600	4400
1	Pcs	13800	15000	15000	13800	1200
2	Pcs	30550	34000	68000	61100	6900
2	Pcs	13000	14500	29000	26000	3000
2	Botol	3917	6000	12000	7833	4167

- Mei 2022 dengan total data sebanyak 4.492 data.Untuk data yang digunakan adalah data transaksi penjulan.

B. *Preprocessing*

Setelah didapatkan data transaksi penjualan dengan kondisi yang perlu dilakukan penyeleksian ulang, maka sesuai dengan hasil kajian pustaka diperlukan adanya tahap preprocessing data sebelum dilakukan tahap pengimplementasikan pada sistem analisis. Adapun tahapan preprocessing data yang dilakukan sebagai berikut ini :

http://digilib.mercubuana.ac.id/

a. *Data Reduction*

Data yang ada terdapat 14 atribut, namun peneliti tidak menggunakan keseluruhan atribut, atribut yang digunakan hanya atribut yang sesusai berdasarkan melihat data dari laku, laku sedang, dan tidak laku. sehingga atribut yang digunakan hanya 3 atribut saja untuk digunakan pada pengimplementasian algoritma. Adapun atribut yang digunakan yaitu NAMA_TRANSAKSI, JUMLAH_TRANSAKSI, TOTAL_TERJUAL.

b. *Data Cleansing*

Penelitian ini menggunakan data sebanyak 4.492 data. Pada proses cleansing data peneliti menggabungkan serta menjumlahkan nilai seluruh data yang duplikat. Pada tahap ini data yang dihasilkan sebanyak 1.239 data.

c. Data Scaling

Dilakukan salah satu dari banyak metode pada data scaling dengan metode min-max scaler. metode ini mengubah nilai pada dataset kedalam skala atau range yang lebih kecil yang bernilai 0 sampai dengan 1. Atribut yang di scaling adalah JUMLAH_TRANSAKSI dan TOTAL_TERJUAL.

C. Algoritma K-Means

Setelah jumlah K-Means ditentukan, data yang tersisa ditentukan menggunakan K-Means berdasarkan K.

Menurut Andre dan Wahyu [12] , Algoritma K-Means merupakan salah satu jenis teknik analisis data yang melakukan pengolahan data dengan sistem partisi. Secara umum algoritma K-Means memiliki langkah-langkah seperti berikut:

1. Tentukan k sebagai jumlah cluster yang akan dibentuk
2. Tentukan k Centroid (titik pusat cluster) awal secara acak. Adapun rumus yang digunakan untuk menghitung centroid cluster ke-i berikutnya sebagai berikut:

$$\frac{\sum_{i=1}^{n} x^i}{n} \; ; i = 1, 2, 3, \dots n \qquad (1)$$

Dimana :

v : centroid pada cluster

xi : objek ke-i

n : banyaknya objek / jumlah objek yang menjadi anggota cluster

3. Hitung jarak setiap objek ke masing-masing centroid dari masing-masing cluster. Untuk menghitung jarak antara objek dengan centroid dapat menggunakan Euclidean Distance.

$d\ (x, y) = \|x - y\| = \sqrt{(\Sigma_(i=1\ )^n\ (x_i-y_i\ )^2\ )}; i = 1, 2, 3, \dots, n$ (2)

Dimana :

xi : objek x ke-i

yi : data y ke-i

n : banyaknya objek

4. Alokasikan masing-masing objek ke dalam centroid yang paling dekat.
5. Lakukan iterasi, kemudian tentukan posisi centroid baru dengan menggunakan persamaan (1)
6. Ulangi langkah 3 jika posisi centroid baru tidak sama [12].

D. Metode Elbow

Model akan diuji pada data yang telah diproses menggunakan metode elbow.

1. Tujuan dari metode Elbow adalah untuk mengumpulkan informasi dengan membandingkan hasil dari banyak cluster yang akan digunakan untuk menghitung metrik tertentu yang akan datang; nilai cluster yang dihasilkan kemudian akan digunakan sebagai data model untuk cluster terbaik. Selanjutnya, hasil proses perhitungan akan digunakan untuk membandingkan jumlah cluster. Berikut ini merupakan rumus SSE :

$$SSE = \sum_{k-1}^{k} \sum_{xigSk} \|Xi - Ck\|^2$$

Keterangan :

Xi : nilai atribut dari data ke-i

Ck : nilai atribut titik pusat Cluster ke-i

Sum of Square Error (SSE) adalah rumus yang digunakan untuk membandingkan perbedaan antara data yang telah diperoleh dari sumber selain model yang sedang dipertimbangkan sebelumnya. SSE sering digunakan sebagai bantuan dalam menentukan ukuran cluster terbaik [10].

IV. HASIL DAN PEMBAHASAN

Pada bagian ini akan menampilkan hasil semua eksperimen yang telah dilakukan.

A. Data Uji *(Training)*

http://digilib.mercubuana.ac.id/

Data Transaksi Penjualan merupakan dataset yang terdiri dari 1239 data yang sudah di preprocessing manual menggunakan Microsodt Excel. Data ditampilkan pada Table 6.1.

TABEL 1
DATA TRANSAKSI PENJUALAN

NAMA_TRANSAKSI		JUMLAH_TRANSAKSI	TOTAL_TERJUAL
1	AMPLOP POLOS 90 BESAR	194	760
2	GG FILTER 12	116	6204600
3	INDOCAFE CF MIX 30X20	235	32883
4	KERTAS FOLIO BERGARIS	2	4500
5	S MILD 16 BARU	416	6019809
6	AQUA GELAS220ML/DUS	9	170600
7	ADES AIR 1500 ML	242	471730
8	GG SIGNATURE 12	29	467600
9	AROMA BOLD 16 FILTER	33	454100
…	…	…	…
1239	SILVER QUEN BITES 35G	54	74

- *Data Scaling MinMax Scaler*

  Tahap selanjutnya peneliti melakukan scaling menggunakan MinMaxScaler pada atribut JUMLAH_TRANSAKSI dan TOTAL_TERJUAL dan membuat atribut baru dengan nama total_penjualan_cluster dan jumlah_transaksi_cluster. Berikut hasil data yang telah di scaling pada gambar 2.

Gbr 2 . Hasil Data Scaling MinMax Scaler

	NAMA_TRANSAKSI	JUMLAH_TRANSAKSI	TOTAL_TERJUAL
0	AMPLOP POLOS 90 BESAR	194	760
1	GG FILTER 12	116	6204600
2	INDOCAFE CF MIX 30X20	235	32883
3	KERTAS FOLIO BERGARIS	2	4500
4	S MILD 16 BARU	416	6019809
5	AQUA GELAS220ML/DUS	9	170600
6	ADES AIR 1500 ML	242	471730

total_penjualan_cluster	jumlah_transaksi_cluster
0.000122	0.130405
1.000000	0.077703
0.005300	0.158108
0.000725	0.000676
0.970217	0.280405
0.027496	0.005405
0.076029	0.162838

B. Algoritma K-Means Clustering

Pada data transaksi penjualan diujikan melalui algoritma K-Means untuk menentukan cluster laku, laku sendang dan tidak laku. Maka penulis menentukan jumlah cluster uji yaitu k=3. Jumlah anggota pengelompokan k=3 ditampilkan pada Tabel 2 dan gambar 2 berikut.

TABEL 2
PENGELOMPOKAN K=3

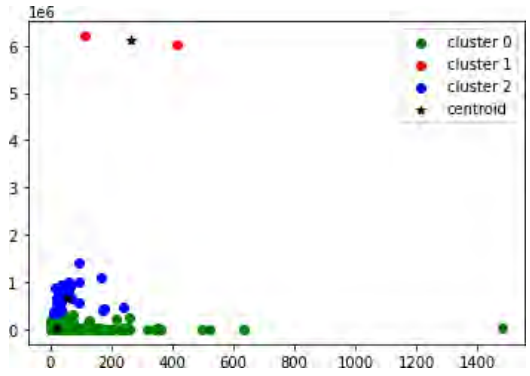
Jumlah k=3	**Jumlah anggota Cluster**
Cluster 0	1196 data
Cluster 1	2 data
Cluster 2	40 data

- Visualisasi hasil k=3

Gbr 2 Tampilan Hasil Klasterisasi dan centroid

http://digilib.mercubuana.ac.id/

Pengolahan Data

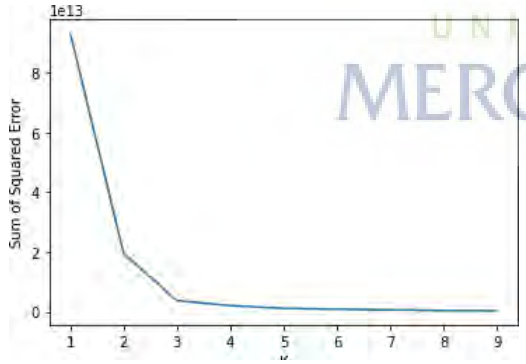


Bedasarkan dari hasil cluster k=3 dilakukan tahapan model testing untuk menguji jumlah cluster k=3 itu terbaik atau tidak, dapat dilihat pada Tabel 3 dan Gambar 3.

TABEL 3 HASIL SSE PADA TIAP CLUSTER

cluster (k)	hasil sse	selisih
2	9.32281E+13	…
3	1.94899E+13	7.37E+13
4	3.69303E+12	1.58E+13
5	2.01763E+12	1.68E+12
6	1.11542E+12	9.02E+11
7	8.0762E+11	3.08E+11
8	5.96382E+11	2.11E+11
9	3.83269E+11	2.13E+11
10	2.90176E+11	9.31E+10

Gbr 3 Tampilan Hasil Perhitungan SSE



Dapat dilihat hasil proses perhitungan SSE terhadap 1239 data Transaksi Penjualan maka hasil yang mengalami penurunan yang paling besar adalah pada k=3 dengan nilai selisih sebesar 7.37E+13.

- Informasi Pengelompokan dan Hasil Data Barang menurut Cluster

  Berdasarkan Tabel 3, nilai selisih tertinggi terdapat pada k=3 sebesar 7.37E+13. Nilai ini menunjukkan bahwa cluster optimal terdapat pada k=3. Dilihat dari hasil pengelompokan k=3 yang ditampilkan mulai dari cluster 0, cluster 1 dan cluster 2, masing-masing memiliki jumlah anggota cluster adalah 1236 data, 2 data, 40 data. Informasi yang diperoleh pada Tabel 4.

TABEL 4
INFORMASI CLUSTER

Cluster	Informasi
0 (1192 data)	Cluster ini merupakan data penjualan barang dengan rating yang kurang, dalam arti tidak laku.
1 (2 data)	Cluster ini merupakan data penjualan barang dengan rating yang yang sangat tinggi, dalam arti laku.
2 (40 data)	Cluster ini merupakan data penjualan barang dengan rating yang sedang, dalam arti laku sedang.

Gbr 4 Tampilan Hasil Barang Yang Telah Di Cluster

	NAMA_TRANSAKSI	JUMLAH_TRANSAKSI	TOTAL_TERJUAL	cluster
0	AMPLOP POLOS 90 BESAR	194	760	0
1	GG FILTER 12	116	6204600	1
2	INDOCAFE CF MIX 30X20	235	32883	0
3	KERTAS FOLIO BERGARIS	2	4500	0
4	S MILD 16 BARU	416	6019809	1
5	AQUA GELAS220ML/DUS	9	170600	0
6	ADES AIR 1500 ML	242	471730	2
7	GG SIGNATURE 12	29	467600	2
8	AROMA BOLD 16 FILTER	33	454100	2
9	MARLBORO RED	32	808950	2

Berdasarkan hasil yang ada pada Gambar 4, diperoleh barang yang laku dibeli konsumen yaitu : GG FILTER 12 dan S MILD 16 BARU.

V. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan, K-Means berhasil membentuk pengelompokan dengan anggota cluster yang tepat dan dioperasikan menggunakan pemrograman Python pada data Transaksi Penjualan. Melalui eksekusi Python, informasi yang dihasilkan sesuai dengan keadaan yang ada karena muncul dari cluster yang optimal hasil perhitungan SSE dengan nilai selisih tertinggi berada pada k=3 dengan selisih 7.37E+13. Dapat disimpulkan bahwa hasil cluster k=3 dengan keterangan bahwa cluster 1 adalah jumlah cluster yang paling laku berjumlah 2 item. Kemudian cluster 0 dengan keterangan cluster 0 adalah cluster yang tidak laku dengan jumlah 1196 item. Terakhir cluster 2 dengan keterangan jumlah cluster laku sedang yaitu 40 item. Dan hasil barang yang laku dibeli konsumen yaitu : GG FILTER 12 dan S MILD 16 BARU.

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

DAFTAR PUSTAKA

[1] F. Hadi and Y. Diana, "Pengklusteran Penjualan Bahan Bangunan Menggunakan Algoritma K-Means," JOISIE Journal Of Information System And Informatics Engineering, vol. 4, no. 1, pp. 22–28, 2020.

[2] G. Gunadi and D. Indra Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Baru Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-Growth) : Studi Kasus Percetakan PT. Gramedia," 2012.

[3] M. Ridwan Basalamah, I. Athia, and M. Jannah, "Strategi Pemasaran Yang Tepat Guna Dalam Menghadapi Persaingan Pasar UMKM," 2018.

[4] W. A. Triyanto, "Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk," Jurnal SIMETRIS, vol. 6, 2015.

[5] M. Stmik and T. Dharma, "Analisis Data Mining Pada Strategi Penjualan Produk PT Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering," J-SISKO TECH Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD, vol. 2, no. 1, pp. 32–41, 2019.

[6] I. Djamaludin and A. Nursikuwagus, "Analisis Pola Pembelian Konsumen Pada Transaksi Penjualan Menggunakan Algoritma Apriori," Jurnal SIMETRIS, vol. 8, 2017.

[7] K. Handoko and L. Sabda Lesmana Universitas Putera Batam, Data Mining Pada Jumlah Penumpang Menggunakan Metode Clustering. 2018.

[8] Y. Darmi, A. Setiawan, J. Bali, K. Kampung Bali, K. Teluk Segara, and K. Bengkulu, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," 2016.

[9] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency," in Journal of Physics: Conference Series, Jan. 2021, vol. 1751, no. 1. doi: 10.1088/1742-6596/1751/1/012038.

[10] A. T. Rahman, "Coal Trade Data Clusterung Using K-Means (Case Study PT. Global Bangkit Utama)."

[11] M. Benri, H. Metisen, and S. Latipa, "Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila," 2015.

[12] A. Winarta and W. J. Kurniawan, "Optimasi Cluster K-Means Menggunakan Metode Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python," Jurnal Teknik Informatika Kaputama (JTIK), vol. 5, no. 1, 2021.

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

KERTAS KERJA

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul "Clustering Pada Data Transaksi Penjualan Minimarket Menggunakan Algoritma K-Means". Kertas kerja berisi semua material hasil penelitan Tugas Akhir yang tidak dimuat/atau disertakan di artikel jurnal. Di dalam kertas kerja ini disajikan: literature review, dataset yang digunakan, source code, dan hasil eksperimen secara keseluruhan.

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/