



**Implementasi Metode K-Means Clustering Untuk Pengelompokan
Data Penjualan Pada Minimarket
Remaja Kampus Bengkulu**
*Implementation of the K-Means Clustering Method for Grouping Sales
Data at Minimarkets
Bengkulu Campus Youth*

Roki Aprinsa, Siswanto, & Ila Yati Beti

Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu

*Corresponding Email: rokiaprinsah@gmail.com

Abstrak

Minimarket Remaja Kampus merupakan salah satu jenis usaha di bidang penjualan keperluan sehari-hari. Untuk pengambilan keputusan dalam menentukan jumlah persediaan produk yang dapat disesuaikan dengan permintaan pasar Minimarket Remaja Kampus belum menggunakan sistem dan masih dihitung secara manual. Oleh karena itu dilakukan penelitian ini dengan tujuan untuk mengimplementasikan metode k-means clustering dalam pengelompokan data penjualan pada minimarket Remaja Kampus Bengkulu. Sehingga dengan mudah dapat menentukan dan mengklasifikasikan penjualan produk yang tinggi, sedang dan rendah. Implementasi sistem menggunakan bahasa pemrograman PHP dan Database MySQL dan metode yang digunakan dalam penelitian ini adalah metode waterfall. Setelah dilakukan proses K-means pada Minimarket Remaja Kampus dengan uji data 15 data maka didapatkan 3 cluster barang yaitu cluster 1 sebagai cluster penjualan tinggi dengan 7 barang, cluster 2 dengan penjualan sedang sebanyak 4 barang dan 4 barang di cluster penjualan rendah. Berdasarkan hasil pengolahan 278 data penjualan barang bulan Desember 2021 di Minimarket Remaja Kampus dengan menggunakan Metode k-Means clustering, maka didapatkan hasil pengelompokan tingkat penjualan produk di Minimarket Remaja Kampus Bengkulu sebanyak 3 cluster. Yaitu cluster 1 kelompok dengan tingkat penjualan produk tinggi dengan Jumlah 54 data produk, cluster 2 dengan tingkat penjualan produk sedang dengan 165 jenis produk dan cluster 3 dengan tingkat penjualan produk rendah dengan 51 jumlah produk. Berdasarkan cluster data, maka dapat dijadikan acuan oleh Minimarket Remaja Kampus untuk persediaan produk bulan berikutnya. Yang mana cluster produk yang memiliki tingkat penjualan tinggi memiliki jumlah pemesanan yang tinggi atau stabil seperti sebelumnya. Kemudian cluster produk dengan tingkat penjualan rendah, maka jumlah persediaan produk untuk berikutnya dikurangi agar tidak terjadi penumpukan produk di gudang dan mengalami kadaluarsa.

Kata Kunci: Data Mining, Clustering K-Means, Minimarket Remaja Kampus

Abstract

Campus Youth Minimarket is one type of business in the field of selling daily necessities. For decision making in determining the amount of product inventory that can be adjusted to market demand, the Campus Youth Minimarket has not used the system and is still calculated manually. Therefore, this research was conducted with the aim of implementing the K-means Clustering method in grouping sales data at the Bengkulu Campus Youth minimarket. So that it can easily determine and classify high, medium and low product sales. The implementation of the system uses the PHP programming language and MySQL database and the method used in this research is the waterfall method. After the K-means process was carried out at the Campus Youth Minimarket with 15 data data tests, 3 clusters of goods were obtained, namely cluster 1 as a high sales cluster with 7 items, cluster 2 with moderate sales of 4 items and 4 items in a low sales cluster. Based on the results of processing 278 data on sales of goods in December 2021 at the Campus Youth Minimarket using the K-Means Clustering Method, the results of the grouping of product sales levels at the Bengkulu Campus Youth Minimarket were 3 clusters. Namely cluster 1 group with a high level of product sales with a total of 54 product data, cluster 2 with a moderate level of product sales with 165 types of products and cluster 3 with a low level of product sales with 51 total products. Based on the data cluster, it can be used as a reference by the Campus Youth Minimarket for the following month's product inventory. Which product clusters that have a high level of sales have a high or stable number of orders as before. Then product clusters with low sales levels, then the amount of product inventory for the next is reduced so that there is no accumulation of products in the warehouse and experiencing expiration.

Keywords: Data Mining, K-Means Clustering, Remaja Kampus Minimarket

PENDAHULUAN

Dalam dunia bisnis yang selalu dinamis dan penuh persaingan para pelakunya harus selalu memikirkan cara-cara untuk terus survive mengembangkan skala bisnis mereka. Untuk mencapai hal itu, ada tiga kebutuhan bisnis yang dapat dilakukan, yaitu penambahan jenis maupun peningkatan kapasitas produk, pengurangan biaya operasional perusahaan, dan peningkatan efektifitas pemasaran serta keuntungan. Agar bisa memenuhi kebutuhan-kebutuhan bisnis di atas banyak cara yang dapat ditempuh salah satunya adalah dengan melakukan analisis data perusahaan.

Dalam rangka menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, pimpinan perusahaan maupun manajemen dalam suatu perusahaan tersebut dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut. Pihak eksekutif perusahaan mengharapkan adanya teknologi yang mampu menghasilkan suatu informasi yang siap digunakan untuk membantu mereka dalam mengambil keputusan strategis perusahaan. Mereka ingin mengetahui produk apa yang harus ditingkatkan persediaan berikutnya, seberapa besar pencapaian hasil yang diperoleh oleh perusahaan. Untuk memenuhi kebutuhan-kebutuhan di atas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan data perusahaan (Data Mining).

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar. Dan Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

Minimarket Remaja Kampus merupakan salah satu jenis usaha di bidang penjualan keperluan sehari-hari. Di mana swalayan ini setiap harinya harus memenuhi kebutuhan konsumen dan dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan jumlah persediaan produk yang dapat disesuaikan dengan permintaan pasar. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut.



Berdasarkan kebutuhan di atas, penulis memberi alternatif bantuan yang diwujudkan dalam penelitian yang berjudul “Implementasi metode K-mean Clustering untuk pengelompokan data penjualan pada minimarket Remaja Kampus Bengkulu”.

METODE PENELITIAN

Minimarket Remaja Kampus didirikan oleh bapak Saherman Jamal pada tahun 1989. Minimarket Remaja Kampus beralamat di Jalan Wr. Supratman No. 01 Kota Bengkulu. Awal mulanya Minimarket Remaja Kampus merupakan warung kecil-kecilan di Sukamerindu, dan berkat tekad dan semangat Minimarket Remaja Kampus mengalami kemajuan yang sangat pesat, sehingga sampai saat ini Minimarket Remaja Kampus sudah memiliki 6 orang karyawan dengan bangunan ruko 3 tingkat.

Penelitian dilakukan secara langsung dengan turun ke lokasi target penelitian dengan mengamati secara menyeluruh sistem yang telah berjalan dan yang akan dibangun nantinya. Melakukan Tanya jawab secara langsung dengan pemilik minimarket Remaja Kampus yaitu Bapak Saherman Jamal. Pengumpulan data yang dilakukan oleh peneliti dengan mengumpulkan data yang berasal dari bahan pustaka seperti buku tentang data mining, Clustering dan Algoritma K-Means. Studi pustaka juga diperoleh dengan jurnal dari internet.

Dari hasil pra penelitian yang dilakukan di minimarket Remaja Kampus, didapatkan sistem yang sudah berjalan dan digunakan saat ini masih manual yaitu menghitung sisa produk satu per satu, tentunya hal tersebut membutuhkan waktu yang lama dan sering terjadinya kesalahan. Karena jika dilihat dari perhitungan produk yang tersisa, hal tersebut tentunya akan sulit karena banyaknya produk yang laku setiap harinya dan banyak pula produk yang masuk setiap harinya.

Dalam penelitian ini, pemrosesan data dalam penelitian ini diambil dari minimarket Remaja Kampus Bengkulu yaitu data penjualan produk pada tahun 2020. Data yang digunakan dalam permasalahan ini ada 2 variabel, dimana variabel-variabel tersebut adalah *stock* barang dan barang terjual, sedangkan obyek-obyek yang diproses dalam penelitian ini kurang lebih 200 jenis produk. Sementara yang dijadikan sampel adalah 50 data.



Kemudian dari 50 data sample di atas, diambil 15 data untuk dijadikan contoh perhitungan. Jarak tiap objek (jenis barang) ke masing-masing *centroid* menggunakan rumus kolerasi antar dua objek yaitu *Euclidean Distance*.

$$M_n = (x, y)$$

Keterangan :

Mn: Data ke-

X : Jumlah Stok

Y : Jumlah Barang Terjual

Langkah pertama dalam menentukan cluster adalah mengambil titik centroid secara random. Menurut Johan (2013:23), rumus menghitung jarak data ke titik centroid dihitung dengan menggunakan rumus *Euclidean Distance*, yaitu:

$$D_{(i,j)} = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2}$$

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

Yaitu dengan mencari jarak masing-masing objek (data) ke titik pusat cluster (*centroid*). Dari hasil perhitungan jarak dengan rumus ini akan berpengaruh pada penempatan setiap data ke cluster, adalah sebagai berikut:

1. Proses *cluster* secara hirarki dengan menggunakan metode *K-means* mampu menunjukkan keterkaitan setiap produk.
2. Hasil analisis *cluster_K-means* dari 15 jenis produk yang dapat dikelompokkan menjadi tiga *cluster_k-means*, yaitu:
 - a. *Cluster_k-means* 1, adalah kelompok produk penjualan tinggi, yaitu lebih dari separuh jumlah barang.
 - b. *Cluster_k-means* 2, adalah kelompok produk penjualan sedang
 - c. *Cluster_k-means* 3, adalah kelompok produk penjualan rendah

Dari kekurangan sistem yang sedang berjalan tersebut maka dalam penelitian ini diterapkanlah analisis data mining menggunakan *Clustering* menggunakan algoritma *K-Means*. Sehingga dengan mudah menentukan dan mengklasifikasikan penjualan produk yang laku dan kurang laku. Sehingga pemesanan barang yang kurang laku dapat dikurangi.



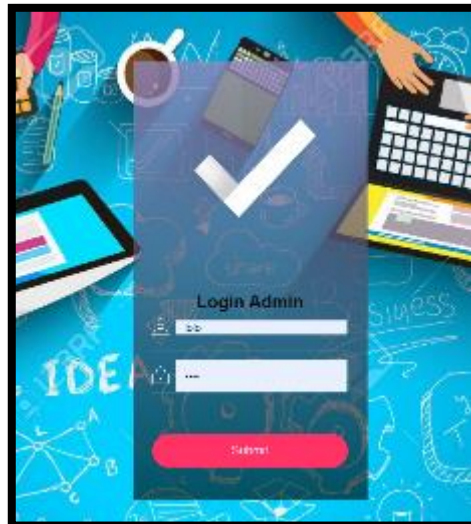
Struktur *database* merupakan kumpulan dari data-data beserta tipenya yang merupakan komponen penting dalam membuat suatu program. Struktur database merupakan struktur data yang saling berhubungan satu sama lain sehingga sangat diperlukan dalam menjalankan program dan juga menyimpan data dalam suatu sistem database. Desain *user interface* merupakan desain tampilan sistem yang akan dibuat. Rancangan home merupakan halaman yang akan tampil pertama kali setelah admin login ke sistem. Pada halaman ini terdapat form yang digunakan untuk menginput data penjualan ke sistem. Pada halaman proses mining *K-Means Clustering*, terdapat form yang digunakan untuk menginput jumlah cluster dan nilai centroid awal yang digunakan. Kemudian klik *Button* "Proses". Pada halaman Menu Anggota Cluster terdapat informasi anggota masing-masing Cluster. Pengujian yang dilakukan dalam penelitian ini adalah Pengujian *black box* adalah pengujian aspek fundamental sistem tanpa memperhatikan struktur logika internal perangkat lunak. Metode ini digunakan untuk mengetahui apakah perangkat lunak berfungsi dengan benar. Pengujian *black box* merupakan metode perancangan data uji yang didasarkan pada spesifikasi perangkat lunak yang dibuat.

HASIL DAN PEMBAHASAN

Hasil Program dan Pembahasan

Analisis data penjualan pada minimarket Remaja Kampus menggunakan data mining dengan metode clustering dan algoritma k-means ini diimplementasikan dengan menggunakan bahasa pemrograman PHP dan Database MySQL. Adapun rancangan tampilan masing-masing menu adalah sebagai berikut:

Tampilan Halaman *Login Admin*. Tampilan *login* merupakan tampilan awal pada halaman menu. Pada tampilan menu utama atau *login* ini admin harus melakukan *login* untuk dapat mengakses halaman berikutnya. Berikut ini tampilan halaman *login* yang terlihat pada Gambar 1.



Gambar 1. Tampilan Menu Login Admin

Tampilan Menu Home. Rancangan home merupakan halaman yang akan tampil pertama kali setelah admin login ke sistem. Rancangan halaman home dapat dilihat pada gambar berikut:



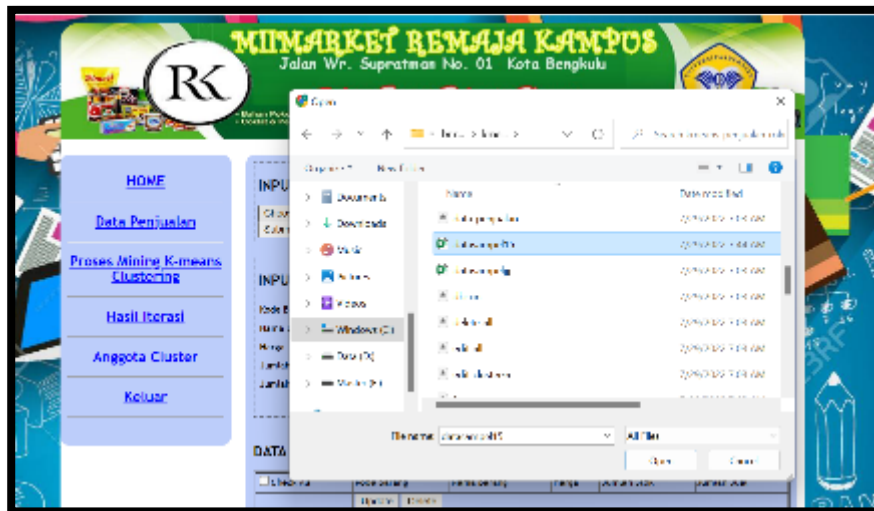
Gambar 2. Tampilan Menu Home

Tampilan Halaman Data Penjualan. Pada halaman ini terdapat *form* yang digunakan untuk menginput data penjualan ke sistem. Rancangan halaman data penjualan dapat dilihat pada gambar berikut:



Gambar 3. Tampilan Data Penjualan

Data penjualan yang diinput ke sistem berformat *.csv. Untuk menginput data klik *button* "Browse". Maka akan muncul tampilan seperti gambar berikut:



Gambar 4. Tampilan Menu Input Data

Kemudian, setelah data yang akan diolah di pilih, maka klik *button* "Submit". Maka akan tampil data penjualan yang akan diolah. Seperti gambar berikut :



Gambar 5. Tampilan Menu Data Penjualan

Tampilan Proses Mining K-Means Clustering. Pada halaman ini terdapat form yang digunakan untuk menginput jumlah *cluster* dan nilai centroid awal yang digunakan. Kemudian klik *Button* “Simpan”. Rancangan halaman proses mining *K-Means Clustering* dapat dilihat pada gambar berikut:



Gambar 6 Tampilan Menu Proses K-Means

Tampilan Hasil Iterasi. Pada halaman ini terdapat informasi hasil iterasi proses *K-Means Clustering*. Rancangan halaman hasil iterasi dapat dilihat pada gambar berikut:



Gambar 7. Tampilan Hasil Iterasi

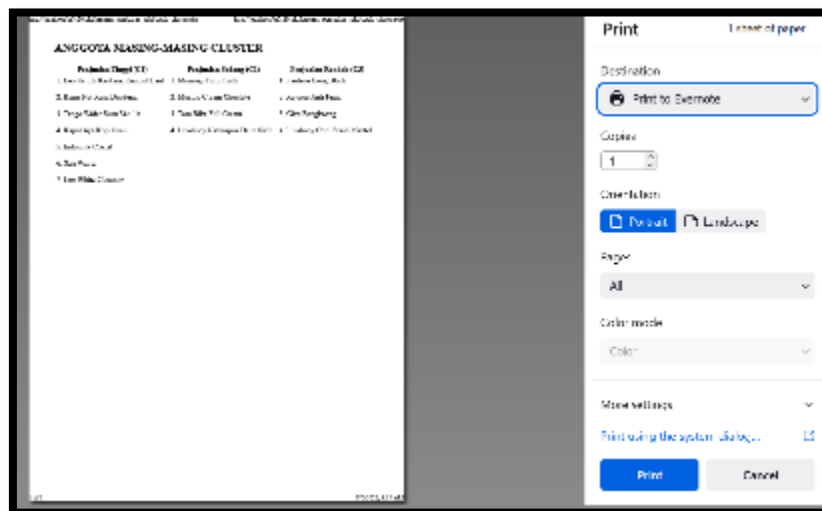


Tampilan Menu Anggota Cluster. Pada halaman ini terdapat informasi anggota masing-masing *Cluster*. Rancangan halaman anggota *cluster* dapat dilihat pada gambar berikut:



Gambar 8. Tampilan Anggota Cluster

Kemudian hasil tersebut dapat dicetak seperti pada gambar berikut:



Gambar 9. Tampilan Cetak Anggota Cluster

Pengujian Data

Pengujian Black Box. Metode pengujian yang digunakan adalah pengujian *black box*. Pengujian *black box* adalah pengujian aspek fundamental sistem tanpa memperhatikan struktur logika internal perangkat lunak. Metode ini digunakan untuk mengetahui apakah perangkat lunak berfungsi dengan benar. Pengujian *black box*

merupakan metode perancangan data uji yang didasarkan pada spesifikasi perangkat lunak yang dibuat. Adapun pengujian *Black Box* yang dilakukan adalah sebagai berikut:

Tabel 1. Pengujian Login Admin

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
User name dan password: terisi dengan benar	Akan menampilkan form admin	Menampilkan form utama admin	[✓] diterima [] ditolak
User name dan password kosong atau user name atau password salah.	Akan menampilkan pesan "password salah!!"	Akan menampilkan pesan "password salah!!"	[✓] diterima [] ditolak

Tabel 2 Pengujian Menu Home

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
Menu Home	Tampilan menu home	Dapat menampilkan menu home	[✓] diterima [] ditolak

Tabel 3 Pengujian Penginputan Data Penjualan

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
Import Data penjualan dari Excell	Data ditampilkan	Dapat menampilkan data penjualan	[✓] diterima [] ditolak

Tabel 4. Pengujian Proses

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
Proses Data Mining	Menampilkan hasil proses data mining dengan algoritma K-Means Clustering	Dapat Menampilkan hasil proses data mining dengan algoritma K-Means Clustering	[✓] diterima [] ditolak

Tabel 5 .Pengujian Hasil Iterasi

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
Hasil iterasi	Data ditampilkan	Dapat menampilkan hasil iterasi dari proses yang dilakukan	[✓] diterima [] ditolak

Tabel 6. Pengujian Hasil Iterasi

Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
Anggota Cluster	Data ditampilkan	Dapat menampilkan anggota cluster produk penjualan tinggi, sedang dan rendah	[✓] diterima [] ditolak

Berdasarkan hasil pengujian dengan kasus uji *sample* di atas dapat ditarik kesimpulan bahwa secara fungsional perangkat lunak sudah berjalan dan sesuai dengan yang diharapkan bahwa tidak ada redundansi pada sistem.

Hasil Pengujian Kuisisioner. Pengujian ini merupakan pengujian yang dilakukan secara objektif dengan kata lain diuji secara langsung ke lapangan yaitu dengan pembagian kuisisioner kepada pengguna terkait penggunaan sistem. Dari hasil kuisisioner akan dilakukan perhitungan statistik sebagai berikut:



Tabel 7. Pengolahan Kuisioner

No	Pernyataan	SS	S	KS	TS
1	Sistem ini dapat memasukkan data penjualan produk	3	2	0	0
2	Sistem ini mudah digunakan	2	3	0	0
3	Dengan adanya sistem ini maka dapat memberikan kemudahan bagi pihak minimarket dalam menentukan kelompok produk yang laku tinggi, sedang dan rendah	3	2	0	0
4	Dengan adanya sistem ini maka data penjualan produk dapat dikelompokkan	2	2	1	0
5	Sistem ini layak digunakan	1	4	0	0

Berdasarkan hasil kuisioner 5 orang responden di atas, maka persentasen jawaban masing-masing pertanyaan adalah sebagai berikut:

Tabel 8. Pengolahan data dari hasil kuisioner

No	Pernyataan	SS	S	KS	TS
1	Pertanyaan 1	60%	40%	0%	0%
2	Pertanyaan 2	40%	60%	0%	0%
3	Pertanyaan 3	60%	40%	0%	0%
4	Pertanyaan 4	40%	40%	20%	0%
5	Pertanyaan 5	60%	40%	0%	0%
Rata-rata		52%	44%	4%	0%

Berdasarkan pembahasan kuisioner, hasil rata-rata jawaban responden yang mencapai 52% menjawab sangat setuju, 44% menjawab setuju, 4 % menjawab kurang setuju dan hanya 0% responden yang menjawab tidak setuju. Maka dari hasil pengolahan data tersebut dapat disimpulkan bahwasanya sistem ini sudah layak untuk digunakan. Karena lebih dari setengah responden menyetujui sistem ini untuk layak digunakan.

SIMPULAN

Berdasarkan hasil observasi 278 sampel penjualan produk pada bulan Desember 2021 di Minimarket Remaja Kampus dengan menggunakan Metode *K-Means Clustering*, maka didapatkan hasil pengelompokan tingkat penjualan produk di Minimarket Remaja Kampus Bengkulu sebanyak 3 *cluster*. Yaitu *cluster* 1 kelompok dengan tingkat penjualan produk tinggi dengan Jumlah 54 data produk, jika dipersentasekan sebanyak 18%, *cluster* 2 dengan tingkat penjualan produk sedang dengan 165 jenis produk dengan persentase sebesar 55% dan 17% untuk *cluster* 3 dengan tingkat penjualan produk rendah dengan 51 jumlah produk. Sehingga dapat disimpulkan bahwa dari *cluster* produk yang dihasilkan dapat dinyatakan lebih dari setengah produk di Minimarket Remaja Kampus masuk ke

dalam kategori laris atau penjualan yang tinggi. Berdasarkan pengelompokan data tersebut, maka dapat dijadikan acuan oleh Minimarket Remaja Kampus untuk persediaan produk bulan berikutnya. Yang mana *cluster* produk yang memiliki tingkat penjualan tinggi memiliki jumlah pemesanan yang tinggi atau stabil seperti sebelumnya. Kemudian *cluster* produk dengan tingkat penjualan rendah, maka jumlah persediaan produk untuk berikutnya dikurangi agar tidak terjadi penumpukan produk di gudang dan mengalami kadaluarsa

DAFTAR PUSTAKA

- [1] Afrisawati. (2013). "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means." Vol V No.3.
- [2] Budanis Dwi Meilani dan Nofi Susanti, (2014), "Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naïve Bayes." Ed. Jurnal LINK Vol 21/No.2.
- [3] Deka Dwinavinta, et.al. (2014). "Klasterisasi Judul Buku dengan Menggunakan Metode K-Means." Ed. Seminar Nasional Aplikasi Teknologi Informasi (SNATI).
- [4] Ernie Kurniawan, dkk. (2013). "Penerapan Algoritma K-means untuk Clustering Dokumen E-jurnal STMIK GI MDP." STMIK GI MDP.
- [5] Johan Oscar Ong. (2013). "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Presiden University." Ed. Jurnal Ilmiah Teknik Industri, Vol.12, No.1.
- [6] Kurniawan, Rulianto. 2015. Joomla untuk Orang Awam. Palembang. Maxikom
- [7] Rosa, A.S dan M. Shalahuddin. 2014. Rekayasa Perangkat Lunak. Informatika Bandung: Bandung.
- [8] Tutik Khotimah. (2014). "Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means." Ed. Jurnal Simetris, Vol 5 No 1.
- [9] Utami, F. H., & Hutahaean, J. Cloud 2017. Computing Data Mining Didalam Dunia Pendidikan.
- [10] Wahyu, Supriyanto, dan Ahmad Muhsin. 2018. Teknologi Informasi Perpustakaan. Kanikus: Yogyakarta.
- [11] Romadansyah Siagian, dkk, (2021). E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model. Journal Of Informatics and Telecommunication Engineering, Vol 5, No 1
- [12] D. Alfiandi, E. Ernawati, dan E. P. Purwandari, "Implementasi K-Means Clustering dan Pemetaan Pemukiman Kumuh di Kota Bengkulu Berbasis Web", *Rekursif*, vol. 6, no. 2, Nov 2018.
- [13] U. Indriani, "PENDEKATAN K-MEANS CLUSTERING TERHADAP REKOMENDASI FILM UNTUK DITONTON", *Jurnal Sistem Informasi Kaputama (JSIK)*, vol. 2, no. 1, 2018.
- [14] S. Suryadi, "PENERAPAN METODE CLUSTERING K-MEANS UNTUK PENGELOMPOKAN KELULUSAN MAHASISWA BERBASIS KOMPETENSI", *INFORMATIKA*, vol. 6, no. 1, 2018.
- [15] Asroni, dan R. Adrian, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang", *76 JURNAL ILMIAH S ERESTA TEKNIKA*, vol. 18, no. 1, 2015.