Indonesian Research Journal on Education

e-ISSN: 2775 – 8672, p-ISSN: 2775 – 9482
Volume 6 Nomor 3 Tahun 2026
Halaman: 1 – 7

Web: https://irje.org/index.php/irje

RESEARCH ARTICLE

Penerapan Metode Data Mining Clustering terhadap Penjualan Produk Detergen pada Ritel Mart Menggunakan Algoritma K-Means

Nurul Aulya[1], Fitri Safnita[2]
Informatika, Fakultas Teknologi Industri, Universitas Bung Hatta, Padang, Indonesia
nurulaulya@bunghatta.ac.id[1], fitrisafnita@bunghatta.ac.id[2]

Kata Kunci:
Data Mining, Clustering, K- means, Produk Detergen

Abstrak

Data mining dapat membantu pengguna mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi penggunanya. Informasi yang dihasilkan kemudian diolah untuk menghasilkan yang biasanya berbentuk sebuah pola. kemudian pola tersebut diubah menjadi sebuah pengetahuan. salah satu metode data mining adalah Clustering yaitu pengelompokkan data berdasarkan karakteristik data tertentu kedalam karakteristik objek yang sama untuk menghasilkan pengelompokkan objek yang mirip satu sama lain dalam kelompok-kelompok. K-means merupakan salah satu algoritma Clustering yang bertujuan untuk membagi data menjadi beberapa kelompok. Pada penelitian ini, Algorritma K-means akan digunakan untuk pengelompokkan pembelian detergen pada sebuah retail mart yang terletak dikota Payakumbuh. Pengelompokkan ini akan menghasilkan detergen mana yang paling laku atau yang paling banyak dibeli dan juga detergen mana yang kurang diminati pembeli sehingga pemilik toko dapat mencari alternatif lain agar produk detergen yang kurang diminati menjadi laku dan banyak diminati oleh konsumen.

Key Word:
Data Mining, Clustering, K-means, Detergent Products

Abstract

Students who migrate to pursue higher education often face social and emotional adjustment challenges. Separation from family, environmental changes, and academic pressure may trigger loneliness. This study aims to examine the relationship between self-compassion and loneliness among migrant students at STIKes in Padang City. The study employed a quantitative correlational design. Purposive sampling was used with criteria including active students originating from outside Padang City and at least in their second semester. The instruments used were the Self-Compassion Scale based on Neff (2003) and the UCLA Loneliness Scale developed by Russell (1996). Data were analyzed using validity and reliability testing, normality and linearity tests, and Pearson Product Moment correlation analysis. The findings revealed a significant negative relationship between self-compassion and loneliness. Students with higher self-compassion tended to report lower levels of loneliness. These findings suggest that self-compassion functions as a protective factor against loneliness among migrant students. This study contributes to the development of self-compassion-based interventions in higher education settings.

Copyright © 2026 Nurul Aulya, Fitri Safnita
*This work is licensed under an **Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)***

PENDAHULUAN

Kondisi dunia yang memasuki era globalisasi membuat industri semakin kompetitif untuk menunjukan keunggulan akan produk dan jasa. Hal ini terjadi guna untuk memenuhi dan memuaskan kebutuhan dan keinginan konsumen yang beraneka ragam. Kondisi ini juga mengakibatkan terjadinya peningkatan jumlah kompetitor yang terlibat dalam berbagai industri strategis termasuk industri retail customer goods. Modernisasi yang terjadi mengakibatkan pasar modern menjadi pilihan utama bagi masyarakan untuk berbelanja kebutuhan rumah tangga dibanding pasar tradisional.

Pada Ritel Mart, produk detergen termasuk bahan pokok yang sering dibutuhkan konsumen dan mempunyai berbagai macam jenis dan merk. Pada produk detergen juga terdapat produk yang kurang diminati dan produk yang diminati oleh konsumen. Hal itu menyebabkan penempatan produk detergen menjadi hal yang utama bagi penjual agar dapat memudahkan akses konsumen untuk mengambil produk detergen yang akan dibeli. Penempatan produk dilakukan dengan cara adanya Clustering terhadap produk yang kurang diminati dan produk yang diminati di Ritel Mart tanpa memisahkan produk yang kurang diminati dengan produk yang diminati. Pemecahan masalah tersebut dapat dianalisis dengan metode Algoritma K-means pada data penjualan produk detergen di Ritel Mart.

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482

METODE

Penelitian ini dilakukan berdasarkan pada banyaknya mini market atau ritel mart yang tidak menyusun barang-barangnya dengan rapi dan sesuai dengan kelompoknya, sehingga barang yang dijualpun menjadi ada yang dan kurang laku. Untuk mengurangi kejadian seperti ini maka diperlukan clustering pada sebuah barang dengan menggunakan metode K-Means. Data dikumpulkan melalui wawancara pemilik ritel mart, pengkajian data penjualan produk detergen, dan dianalisis dengan baik.

HASIL DAN PEMBAHASAN

Data-data yang akan dianalisa system sesuai metode yang diterapkan kedalam data mining adalah berupa data penjualan produk detergen pada minimarket Ritel Mart. Data yang akan diproses nantinya diasumsikan sebanyak 20 produk detergen yang transaksi penjualannya dimulai dari 1 Agustus 2019 hingga akhir Oktober 2019.

Dalam melakukan proses *clustering*memerlukan data barang dan penjualan yang ada pada minimarket Ritel Mart. Data yang akan diolah adalah data penjualan yang dimulai dari 1 Agustus 2019 hingga akhir Oktober 2019. Langkah-langkah dalam melakukan clustering dalam metode *K-Means* dari data yang telah diperoleh adalah sebagi berikut: pertama, Menentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk. Data dimisalkan dengan M (X:Y), dimana M adalah nama barang/ produk, X adalah stok, dan Y adalah jumlah terjual. Kedua Bangkitkan *k centroid* (titik pusat *cluster*) awal secara acak atau random. Penentuan *cluster* awal diasumsikan sebagai berikut: Pusat C1= (100;29), Pusat C2= (50;15). Ketiga Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua objek yaitu Euclidean Distance yaitu:

$$M_n = \sqrt{(Mnx - Cnx)^2 + (Mny - Cny)^2}$$

Keempat, Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*nya. Kelima, Tentukan posisi *centroid* baru C*k* dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama. Keenam, Ulangi langkah 3. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua objek yaitu Euclidean Distance. Ketujuh, Ulangi langkah 4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*nya. Kedelapan, Proses perhitungan dihentikan, pusat *centroid* sama.

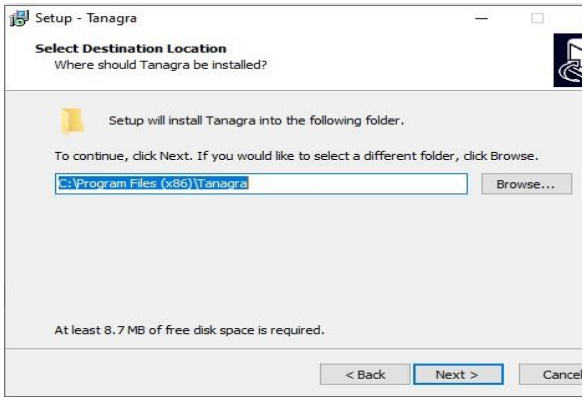
Kesimpulannya adalah dari proses *clustering* diatas dapat diketahui bahwa dalam pola komsumsi masyarakat dengan kelompok barang pada *cluster* 1 merupakan produk detergen yang kurang diminati oleh konsumen (produk detergen yang kurang laris/ kurang laku) di awal bulan Agustus hingga akhir bulan Oktober 2019. Sedangkan dalam pola komsumsi masyarakat dengan kelompok barang pada c*luster* 2 merupakan produk detergen yang paling diminati oleh konsumen (produk detergen yang laris/ yang laku) di awal bulan Agustus hingga akhir bulan Oktober 2019.

Implementasi adalah suatu tindakan atau pelaksana renacana yang telah disusun secara rinci dan jelas. Implementasi dilakukan untuk mengetahui bagaiamana suatu aplikasi yang telah dirancang dapat dijalankan dengan sebuah sistem atau program, apakah aplikasi tersebut dapat bermanfaat bagi penggunanya, implementasi juga digunakan untuk mengetahui batasan dari sistem yang dibutuhkan untuk menjalankan aplikasi telah dirancang tersebut. Tanagra 1.4.50 adalah software data mining yang diperlukan untuk keperluan akademik dan riset.

Langkah-langkah Instalisasi

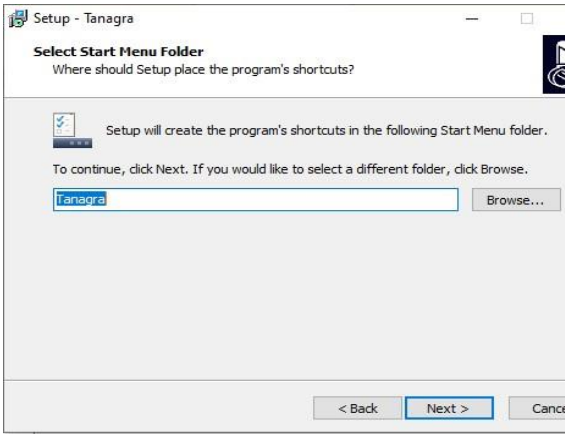


Gambar 1 Instalasi Software Tanagra 1.4.50

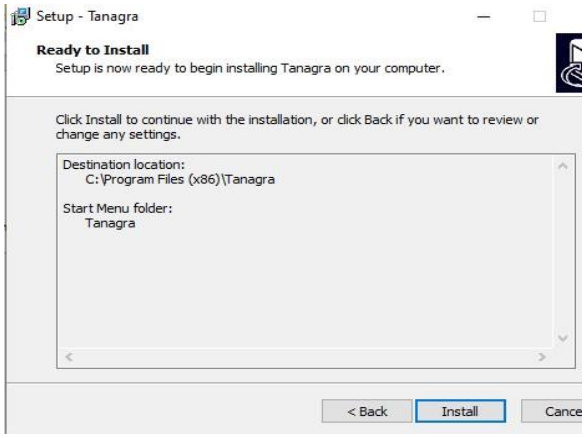


Gambar 2 Tampilan Tempat Lokasi Penyimpanan Installasi

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482



Gambar 3. Tampilan Nama Installasi



Gambar 4 Tampilan Pemilihan Untuk Installasi



Gambar 5. Tampilan Installasi Complete

Gambar 1 merupakan tampilan saat akan menginstal aplikasi tangra, pilih I accept the agreement lalu tekan next. Gambar 5 merupakan tampilan proses akhir saat menginstal tanagra. Tekan finish untuk mengakhiri dan proses installasi pun selesai.

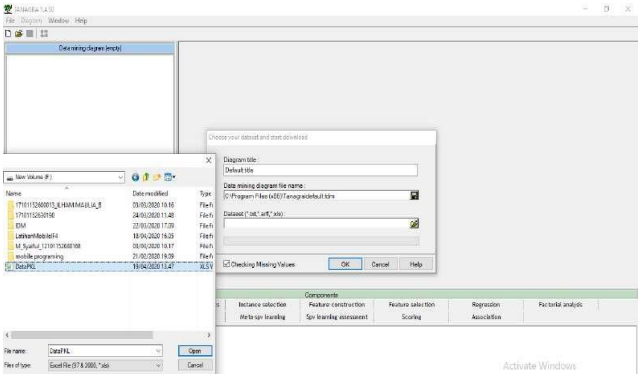
Disini penulis menggunakan software atau aplikasi Tanagra 1.4.50 dalam melakukan pengujian terhadap data yang diperoleh dan telah diolah. Perangkat yang dibutuhkan untuk mengisntal Tanagra 1.4.50 adalah Software data mining Tanagra 1.4.50, penyusunan hasil data produk detergen menggunakan Microsoft Excel. Penulis menggunakan spesifikasi sebagai berikut: 1) Laptop ASUS A456U Series; 2) Processor Intel i5-7200U; 4) Memory 4GB DDR5; 5) HDD 1TB

Pengujian dilakukan untuk menentukan apakah ada kesalahan yang terjadi atau yang tidak kita ketahui dan tidak sesuai dengan tujuan pengembangan yang diharapkan. Penulis menggunakan data penjualan produk detergen pada sebuah ritel mart. Variabel yang digunakan yaitu nama barang, jumlah dan stok. Data yang diinputkan yaitu data selama 3 bulan penjualan dan dipilih 20 detergent sebagai sampel. Setelah Tanagra 1.4.50 dijalankan, masukkan data dengan cara klik File lalu pilih New, kita dapat memilih tempat penyimpanan datanya untuk dimasukkan ke dalam Tanagra 1.4.50 lalu klik OK.

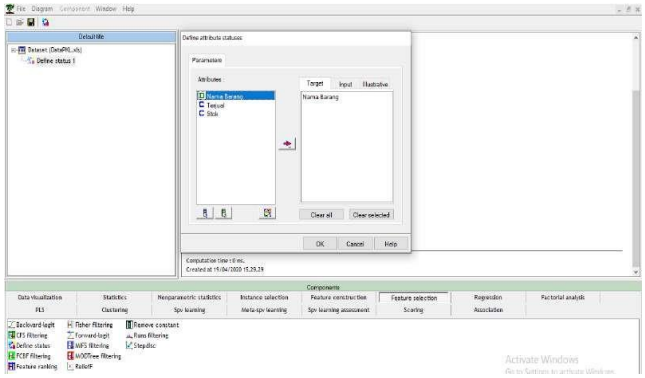
Tambahkan Define Status dengan cara tarik Dataset. Lalu klik kanan pada Define Status pilih parameter, masukan atribut nama barang pada target, masukan jumlah dan stok pada Input. Setelah klik OK, Tambahkan komponen More Univariate Cont Stat yang terletak di bagian Statistis. Tarik More Univariate Cont Stat lalu letakkan di Define status 1, setelah itu klik kanan pilih view. Tambahkan komponen Standarizedi yang terletak di Featur Contruction, tarik ke bagian More Univariate Cont Stat 1. Lalu klik kanan pada Standarize pilih view. Masukkan komponen Define Status baru ke dalam Standardize 1, klik kanan lalu pilih Parameters masukkan std_terjual_1 dan tsd_stok_1 pada input lalu tekan OK. Masukkan K-Means pada Define Status 2, lalu klik kanan dan pilih Parameters. Atur nilai cluster menjadi 2 grup. Pada DistanceNormalization ubah menjadi None lalu OK. Selanjutnya klik kanan pada KMeans dan pilih View. Ini merupakan tahap awal pada proses Clustering. Tambahkan View Dataset yang terletak pada Data Visualization, tarik View Dataset ke K-Means 1, klik kanan lalu pilih View. Masukkan Define Status baru(3) ke dalam K-Means1, setelah muncul klik kanan pada Define Status (3), lalu pilih Parameters. Masukkan nama barang,jumlah dan stok pada Input, juga masukkan Cluster_Kmeans_1. Tambahkan Group Characterization yang terletak pada Statistic. Tambahkan lagi Define Status ke dalam K-Means. Sehingga menampilkan Define Status 4. Klik kanan pada Define Status lalu pilih parameters, jadikan nama barang di

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482

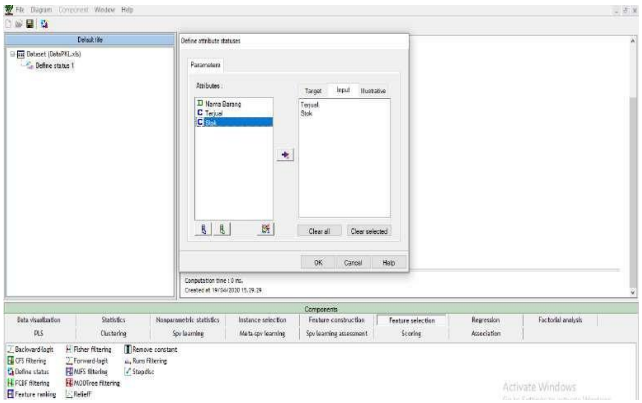
bagian target, dan Cluster_Kmeans_1 pada bagian Input. Tambahkan Contigency Chi-Square yang terletak pada NonParametic Statisticstab ke dalam Define Status 4. Sehingga nantinya menampilkan nama barang, cluster_1 dan cluster_2 serta Sum. Kita dapat menempatkan cluster menurut suatu variabel. Tambahkan Scatterplot yang terletak di Visualitation pada K-Means 1. Lalu klik kanan pilih View. Tambahkan Principal Componen Analysis yang terletak pada Factorila Analysis ke dalam K-Means 1. Tambahkan Scatterplot lagi, menjadi Scatterplot 2 ke dalam Pricipal Componen Analysis 1. Ini merupakan langkah terakhir dari analisis ini, kita akan mengekspor dataset dengan menambah kolom yang mengindikasikan setiap anggota cluster. Tambahkan Define Status 5 ke dalam K-Means 1, lalu klik kanan pilih parameter , masukkan nama barang, jumlah,stok dan cluster_Kmeans_1 pada bagian input dan tekan OK. Tambahkan Export Dataset yang terletak pada bagian Data Visualization ke dalam Define Status 5. Lalu klik kanan pilih parameter, pada Attributes Selection ubah kebagian Input Attributes dan tekan OK. Kita dapat menentukan tempat penyimpanan yang kita inginkan berupa hasil outputnya. Seperti pada gambar berikut:



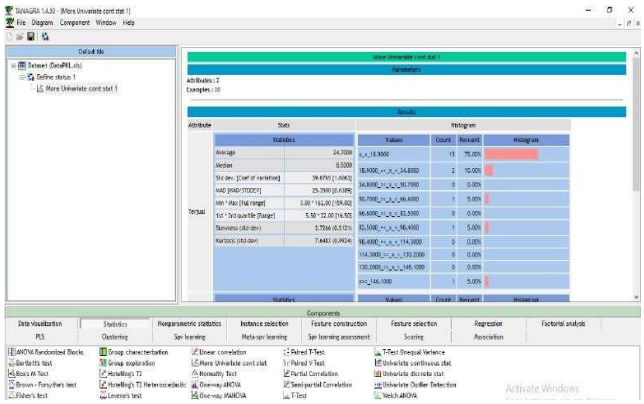
Gambar 6 Input Data



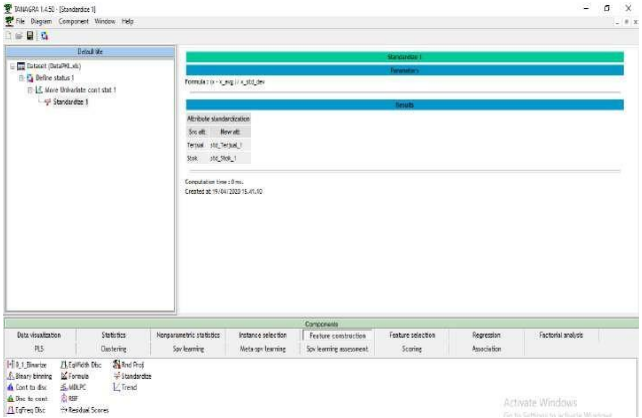
Gambar 7. Define status 1 Target



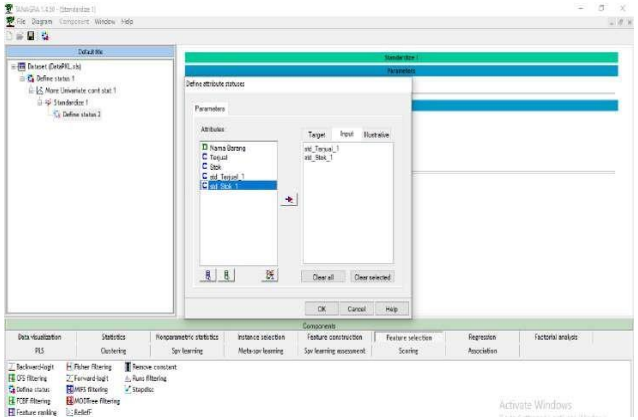
Gambar 8. Define Status 1 Input



Gambar 9. Tampilan Univariate

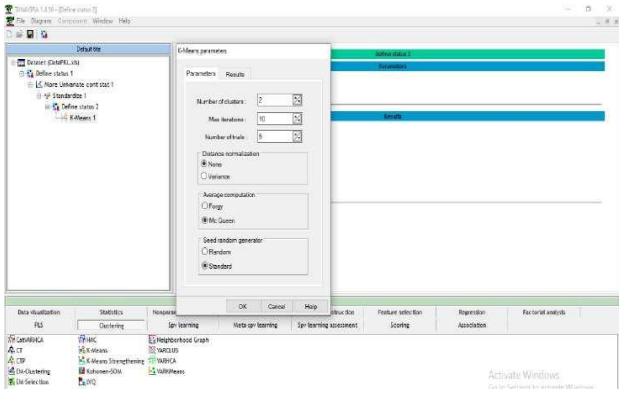


Gambar 10 Tampilan dari Standardize

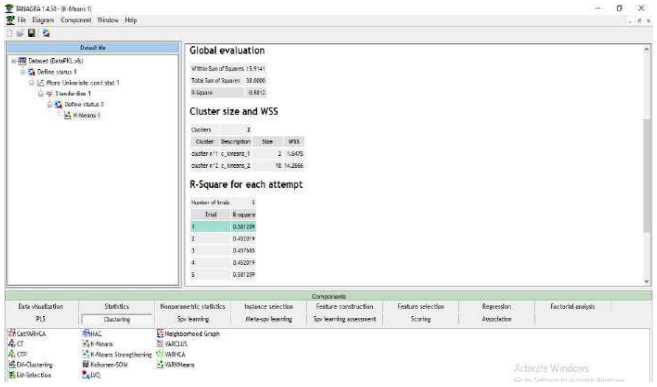


Gambar 11 Define Status 2 pada bagian Input

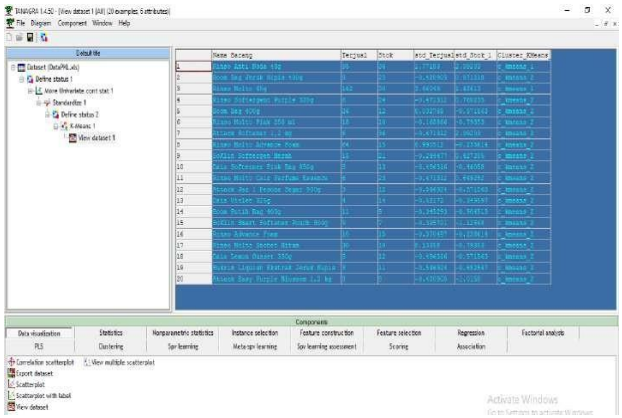
e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482



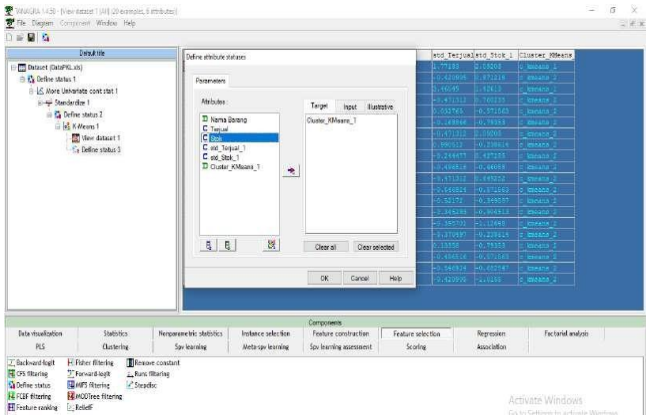
Gambar 12 Parameters K-Means



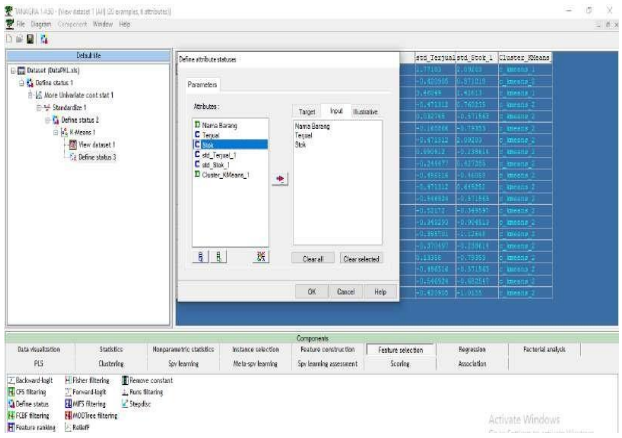
Gambar 13 Tampilan K-Means



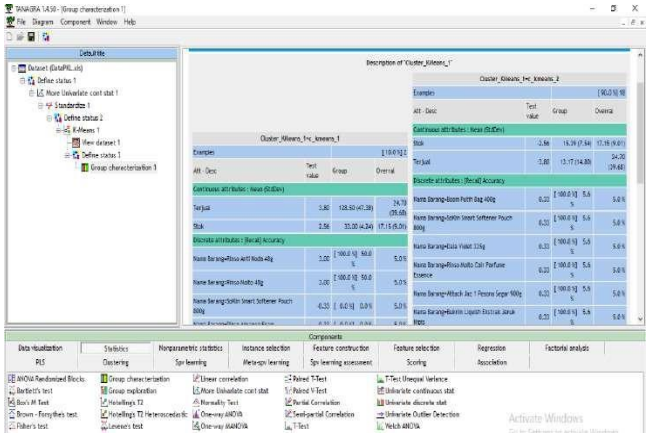
Gambar 14 Tampilan Dataset



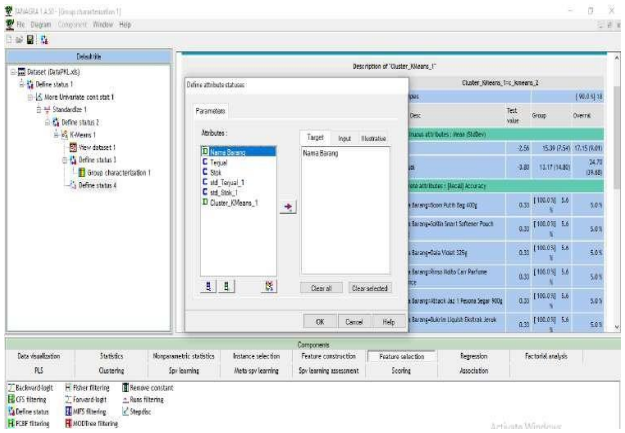
Gambar 15 Tampilan Define Status 3 bagian Target



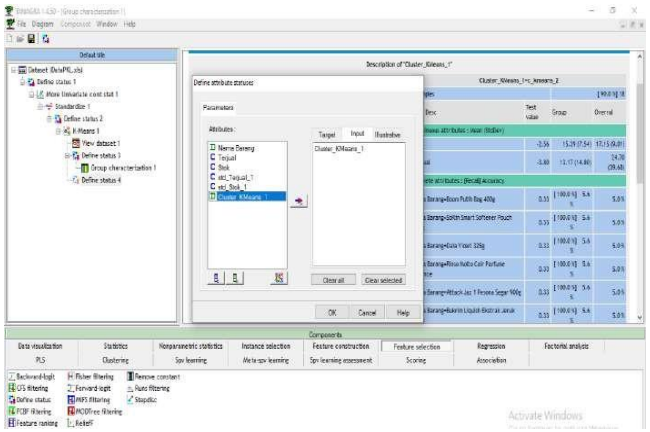
Gambar 16. Tampilan Define Status 3 bagian Input



Gambar 17. Tampilan Hasil Group Characterization

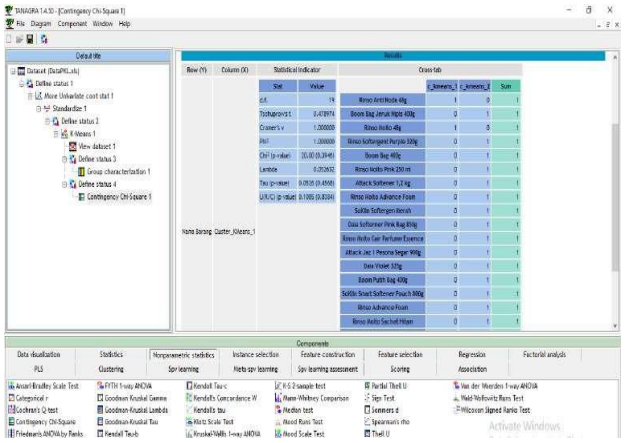


Gambar 18 Tampilan Define Status 4 pada bagian Target

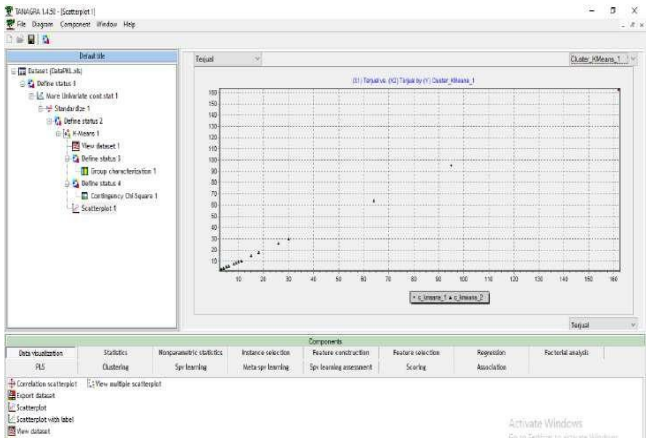


Gambar 19 Tampilan Define Status 4 pada bagian Input

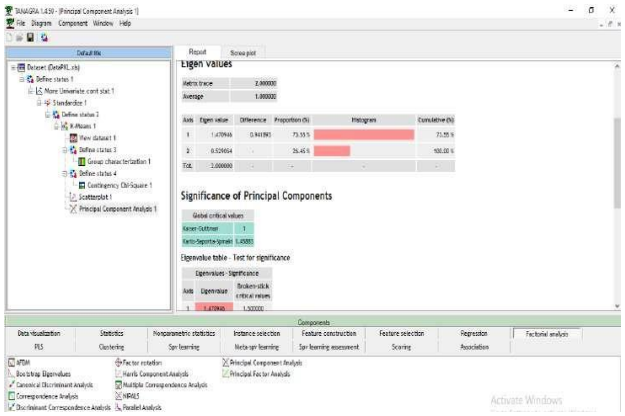
e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482



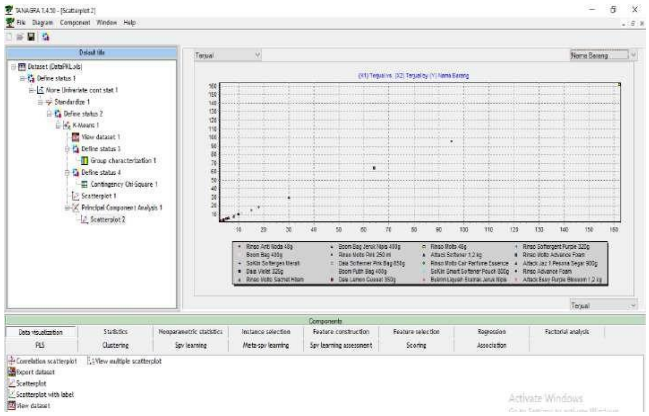
Gambar 20 Tampilan dari Contigency Chi-Square 1



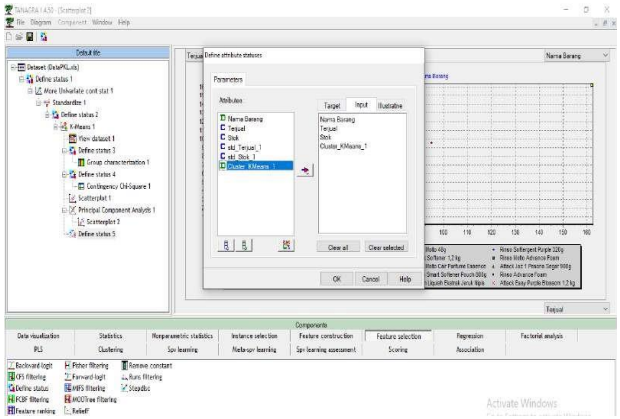
Gambar 21 Tampilan Grafik pada Scatterplot 1



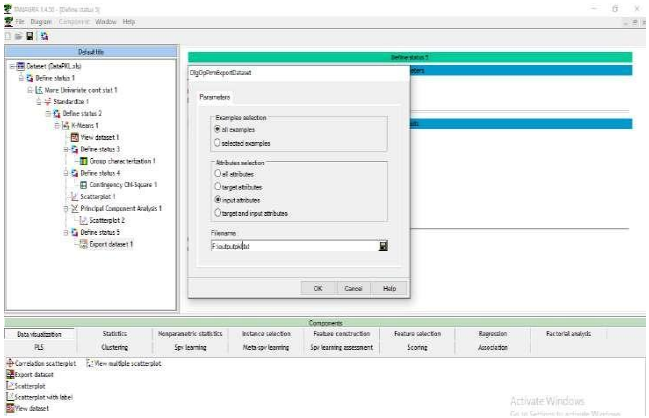
Gambar 22 Tampilan Pricipal Componen Analysis 1



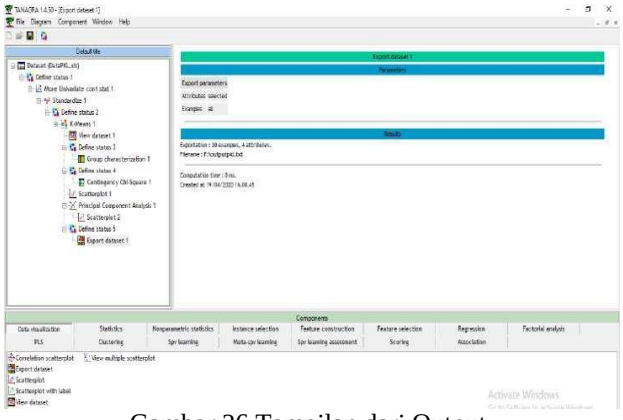
Gambar 23 Tampilan dari Scatterplot 2



Gambar 24 Tampilan Define Status 5 pada bagian Input



Gambar 25 Tampilan dari Export Dataset



Gambar 26 Tampilan dari Output

*outputpkl - Notepad

File Edit Format View Help

Nama Barang	Terjual	Stok	Cluster_KMeans_1
Rinso Anti Noda 48g	95	36	c_kmeans_1
Boom Bag Jeruk Nipis 400g	8	25	c_kmeans_2
Rinso Molto 48g	162	30	c_kmeans_1
Rinso Softergent Purple 320g	6	24	c_kmeans_2
Boom Bag 400g	26	12	c_kmeans_2
Rinso Molto Pink 250 ml	18	10	c_kmeans_2
Attack Softener 1,2 kg	6	36	c_kmeans_2
Rinso Molto Advance Foam	64	15	c_kmeans_2
SoKlin Softergen Merah	15	21	c_kmeans_2
Daia Softerner Pink Bag 850g	5	13	c_kmeans_2
Rinso Molto Cair Parfume Essence	6	23	c_kmeans_2
Attack Jaz 1 Pesona Segar 900g	3	12	c_kmeans_2
Daia Violet 325g	4	14	c_kmeans_2
Boom Putih Bag 400g	11	9	c_kmeans_2
SoKlin Smart Softener Pouch 800g	9	7	c_kmeans_2
Rinso Advance Foam	10	15	c_kmeans_2
Rinso Molto Sachet Hitam	30	10	c_kmeans_2
Daia Lemon Gusset 350g	5	12	c_kmeans_2
Bukrim Liquish Ekstrak Jeruk Nipis	3	11	c_kmeans_2
Attack Easy Purple Blossom 1,2 kg	8	8	c_kmeans_2

Gambar 27 Tampilan Output

KESIMPULAN

Dari uraian masalah diatas, serta berdasarkan analisa dan pembahasan dari bab-bab yang ada maka dapat diambil kesimpulan sebagai berikut: 1) Dengan Data Mining Clustering menggunakan Algoritma K-Means maka didapatkan pola konsumen yaitu adanya produk detergen yang paling diminati oleh konsumen dan produk detergen yang kurang diminati oleh konsumen; 2) Analisa Data Mining Clustering menggunkan Algoritma K-Means dapat membantu pihak ritel mart mengelompokkan detergen agar produk detergen yang kurang diminati dapat menjadi produk detergen yang diminati.

DAFTAR REFERENSI

Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan algoritma K-means untuk clustering data obat-obatan. *Jurnal Nasional Teknologi dan Sistem Informasi*, *5*(1), 17–24.

Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-means dalam pengelompokan penjualan produk pada Swalayan Fadhila. *Jurnal Media Infotama, 11*(2).

Nur, F., Zarlis, M., & Nasution, B. B. (2017). Penerapan algoritma K-means pada siswa baru sekolah menengah kejuruan untuk clustering jurusan. *InfoTekJar: Jurnal Nasional Informatika dan Teknologi Jaringan, 1*(2), 100–105.

Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan algoritma K-means dalam menentukan pegawai yang layak mengikuti assessment center untuk clustering program SDP. *Computer Engineering, Science and System Journal, 3*(1), 87–93.

Sibuea, M. L., & Safta, A. (2017). Pemetaan siswa berprestasi menggunakan metode K-means clustering. *JURTEKSI, 4*(1), 85–92.