

Penerapan Algoritma K-Means Clustering Pada Penjualan Produk Skincare

Ruqoyyatul Widad¹, Zaehol Fatah²,

^{1,2}Sistem Informasi, Universitas Ibrahimy, Situbondo, Indonesia

Email : ruqoyyatulwidad08@gmail.com, zaeholfatah@gmail.com

Abstrak

Data mining dapat membantu perusahaan menentukan strategi pemasaran. Industri kecantikan, khususnya produk skincare, telah mengalami pertumbuhan yang signifikan di era digital saat ini. Dalam menghadapi persaingan yang semakin ketat, sangat penting bagi perusahaan untuk menganalisis data pelanggan guna mengoptimalkan strategi penjualannya. Penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* dalam pengelompokan produk *skincare* berdasarkan data penjualan dengan bantuan perangkat lunak *RapidMiner*.

Proses penelitian meliputi beberapa tahapan, mulai dari seleksi data, preprocessing, normalisasi, transformasi atribut, hingga penerapan algoritma *K-Means*. Hasil dari pengelompokan ini mengidentifikasi tiga kelompok produk *skincare* berdasarkan pola penjualan yang ada, memberikan wawasan berharga bagi pelaku bisnis dalam merumuskan strategi pemasaran yang lebih efektif. Penelitian ini membuktikan bahwa algoritma *K-Means* merupakan alat analisis yang andal untuk mendukung pengambilan keputusan dalam dunia bisnis.

Kata Kunci: *Data Mining, K-Means Clustering, Skincare, RapidMiner*

PENDAHULUAN

Di era digital saat ini, industri kecantikan, khususnya produk *skincare*, mengalami pertumbuhan yang sangat signifikan. Pertumbuhan ini dipicu oleh perubahan gaya hidup, meningkatnya kesadaran masyarakat mengenai pentingnya perawatan kulit, serta kemudahan dalam mengakses informasi melalui berbagai platform digital. Produk kosmetik telah menjadi bagian integral dalam kehidupan sehari-hari, terutama bagi wanita dari berbagai usia, mulai remaja hingga dewasa. Kosmetik dapat dibedakan menjadi beberapa kategori berdasarkan kegunaan dan fungsinya, seperti makeup atau riasan wajah, serta produk yang fokus pada perawatan seperti *skincare*, *bodycare*, *haircare*, *lipcare*, dan lainnya. Banyak orang kini menggunakan produk kecantikan untuk merawat kulit wajah mereka, dan perkembangan produk kecantikan lokal pun semakin pesat. Dalam persaingan yang ketat ini, perusahaan dituntut untuk dapat mengidentifikasi dan mengklasifikasikan pelanggan yang memiliki potensi besar

untuk membeli. Oleh karena itu, penting bagi perusahaan untuk mengumpulkan dan menganalisis data pelanggan guna meningkatkan tawaran mereka. (Clustering et al., 2024)

K-Means adalah salah satu metode pengelompokan data yang bertujuan untuk membagi data menjadi dua kelompok atau lebih. Metode ini mengelompokkan data sedemikian rupa sehingga data dengan karakteristik yang sama ditempatkan dalam satu kelompok, sementara data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. (Utami & Eka Putra, 2022)

Metode ini adalah teknik pengelompokan yang bertujuan untuk mengelompokkan objek sedemikian rupa sehingga jarak antara setiap objek dengan pusat kelompok dalam satu kelompok menjadi minimum. (Metisen & Sari, 2015)

Penelitian ini bertujuan untuk mengoptimalkan penjualan produk *skincare* dengan menerapkan algoritma *K-Means Clustering* pada dataset yang terdiri dari 100 sampel produk *skincare*, masing-

masing dengan karakteristik yang berbeda. Dengan menggunakan pendekatan ini, diharapkan dapat dihasilkan model pengelompokan yang lebih objektif. Hasil dari proses *clustering* ini akan membantu para penjual dalam membuat keputusan yang lebih tepat terkait strategi penjualan produk *skincare* mereka. (Informasi et al., 2024)

METODE

Bagian ini membahas dasar teori mengenai teknik pengumpulan data serta tahapan perancangan.

1. Data Mining

Proses pencarian dan penggalian informasi dari tumpukan data yang besar merupakan inti dari data mining. Tujuan utama pengolahan data ini adalah untuk menghasilkan informasi baru yang berharga. (Muliono & Sembiring, 2019) Proses analisis data dilakukan dari berbagai perspektif untuk menghasilkan informasi penting yang dapat dimanfaatkan guna meningkatkan keuntungan, mengurangi biaya pengeluaran, atau bahkan mencapai kedua tujuan tersebut. Secara teknis, data mining dapat diartikan sebagai proses dalam menemukan korelasi atau pola dari ratusan hingga ribuan bidang dalam sebuah basis data relasional yang besar. (Yunita, 2018)

2. Knowledge Discovery In Data Base

Knowledge Discovery in Databases (KDD) adalah sebuah proses analisis terstruktur yang bertujuan untuk memperoleh informasi yang akurat, baru, dan bermanfaat, serta untuk mengidentifikasi pola-pola dalam kumpulan data yang besar dan kompleks. Di dalam proses KDD, data mining (DM) berperan sebagai inti, di mana algoritma-algoritma tertentu digunakan untuk mengeksplorasi data, membangun model, dan menggali pola-pola yang belum terungkap. (Zanuardi & Suprayitno, 2018)

3. Clustering

Clustering adalah salah satu alat yang

digunakan dalam data mining. Tujuan dari *clustering* adalah untuk mengelompokkan objek-objek ke dalam kelompok yang disebut cluster. Cluster sendiri merupakan sekelompok objek data yang memiliki kesamaan. Dalam proses ini, data dikelompokkan ke dalam cluster yang sama, sedangkan setiap cluster dibedakan berdasarkan perbedaan dengan objek-objek dari cluster lainnya. (Virgo et al., 2020)

4. Algoritma K-Means

Algoritma *K-Means clustering* adalah metode klasifikasi yang baik dengan tingkat akurasi yang tinggi dibandingkan metode lain. Algoritma ini termasuk dalam kelompok Unsupervised learning dan digunakan untuk mengelompokkan data ke dalam beberapa bagian dengan sistem partisi. (Filki, 2022)

5. Rapid Miner

Rapid Miner adalah perangkat lunak sumber terbuka (*open source*) yang dirancang sebagai solusi untuk analisis prediktif. Dengan menggunakan berbagai teknik deskriptif dan prediktif, *Rapid Miner* memberikan wawasan berharga kepada penggunanya, membantu dalam mengambil keputusan yang tepat. Perangkat lunak ini menawarkan lebih dari 500 operator untuk data mining, termasuk untuk input, output, pra-pemrosesan data, dan visualisasi. Sebagai *software* yang berdiri sendiri untuk analisis data, *Rapid Miner* juga memiliki kemampuan integrasi dengan produk lain. Ditulis menggunakan bahasa Java, *Rapid Miner* dapat berfungsi di berbagai sistem operasi. (Bahtiar et al., 2023)

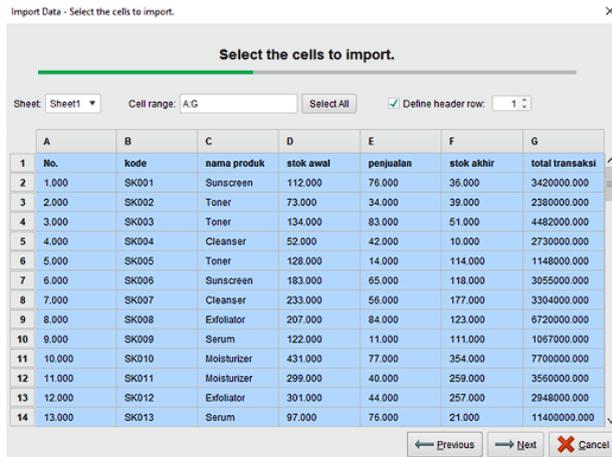
HASIL DAN PEMBAHASAN

Bagian ini berisikan mengenai hasil-hasil penerapan algoritma K-Means Clustering berdasarkan hasil penjualan produk *skincare*

1. Data Selection

Langkah pertama dalam menerapkan algoritma K-Means di

RapidMiner adalah dengan memilih "Blank Process", lalu "Import Data".

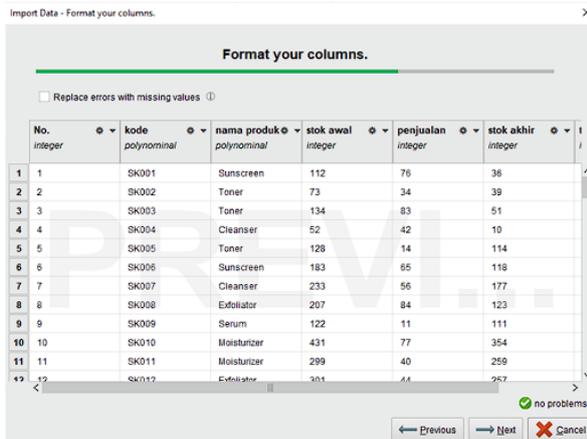


Gambar 1. Proses Import Data

2. Preprocessing Data

a. Cleaning Data

Langkah selanjutnya setelah pengumpulan data adalah tahap preprocessing yang melibatkan pembersihan data dengan fitur "Replace Error With Missing Value" untuk mengidentifikasi dan mengatasi masalah data. Gambar ilustrasi proses penggantian kesalahan.



Gambar 2. Proses Cleaning Replace Error With Missing Value

3. Tranformasi Data

Tahap selanjutnya adalah melakukan perubahan atribut pada data, yang mencakup penggunaan operator Nominal to Numerical dan operator Normalize.

a. Nominal to Numerical

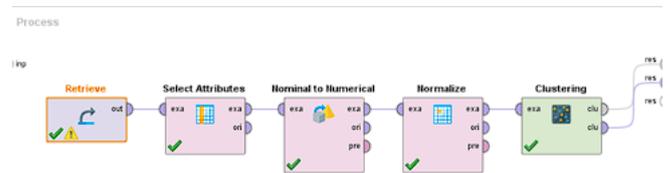
Pada parameter dari *Nominal ke Numerical*, atribut filter yang digunakan adalah *typeAll*, sedangkan tipe pengkodean yang diterapkan adalah *Unique Integer*.

b. Normalize

Pada parameter Normalisasi, atribut filter yang digunakan adalah tipe 'All', sementara metode yang diterapkan adalah *Z-Transformation*.

4. Implementasi Algoritma K-Means Clustering

Implementasi model *algoritma K-Means Clustering* pada *RapidMiner* adalah sebagai berikut:



Gambar 3. Implementasi Algoritma K-Means

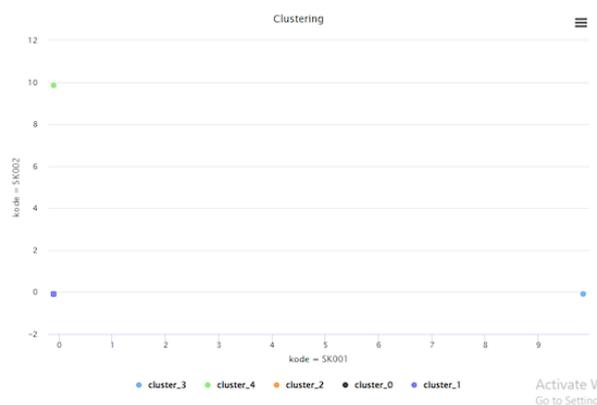
5. Hasil Komputasi Clustering

Pengelompokan produk *skincare* berdasarkan hasil penjualan produk *skincare* pada *RapidMiner* menghasilkan 3 kelompok atau 3 *cluter* model diantaranya sebagai berikut:

Cluster Model

Cluster 0: 12 items
 Cluster 1: 20 items
 Cluster 2: 19 items
 Cluster 3: 25 items
 Cluster 4: 23 items
 Total number of items: 99

Gambar 4. Hasil Komputasi Clustering (Cluter Model)



Gambar 5. Visualisasi Hasil Komputasi *Clustering*

SIMPULAN (PENUTUP)

Penelitian ini berhasil mengimplementasikan algoritma *K-Means Clustering* untuk menganalisis dan mengelompokkan penjualan produk *skincare* berdasarkan dataset yang tersedia. Proses *clustering* dilakukan melalui tahapan seleksi data, preprocessing (termasuk normalisasi dan transformasi atribut), hingga penerapan algoritma *K-Means* menggunakan perangkat lunak *RapidMiner*. Hasilnya, data produk *skincare* terbagi ke dalam tiga kelompok (*cluster*) berdasarkan karakteristik tertentu yang diidentifikasi.

Penelitian menunjukkan bahwa penggunaan algoritma *K-Means* memberikan manfaat signifikan bagi pelaku bisnis dalam memahami pola penjualan serta menentukan strategi pemasaran yang lebih tepat sasaran. Dengan pendekatan ini, perusahaan dapat lebih efisien dalam mengalokasikan sumber daya untuk meningkatkan penjualan produk *skincare*. Sebagai langkah pengembangan lebih lanjut, disarankan untuk memperluas dataset yang digunakan atau menggabungkan metode data mining lainnya guna mendapatkan hasil analisis yang lebih komprehensif.

UCAPAN TERIMA KASIH

Ucapan terima kasih diberikan kepada semua pihak yang telah membantu dalam penyusunan jurnal ini, terutama kepada pembimbing Bapak Zaehol Fatah,

M.Kom, keluarga dan sahabat-sahabat seperjuangan yang telah memberikan dukungan motivasi dan nasehat. Semoga jurnal ini dapat memberikan manfaat bagi pembaca.

DAFTAR PUSTAKA

- Bahtiar, D., Kunci, K., Abstrak, :, Penerima, P., Dana, B., Desa, S., & Jaya, W. (2023). Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma *K-Means Clustering*. *Scientia Sacra: Jurnal Sains*, 3(2), 29–39. <http://pijarpemikiran.com/index.php/Scientia>
- Clustering, A. K., Melakukan, U., Pelanggan, S., Toko, D., Glow, G., Bumirejo, B., Aisyah, S., Darmawan, W., Ismanto, B., Informasi, S., Pekalongan, K., Informatika, M., Pekalongan, K., Informatika, T., Pekalongan, K., & Clustering, K. (2024). *Algoritma K-Means Clustering Untuk Melakukan Segmentasi Pelanggan*. 20(2), 26–31.
- Filki, Y. (2022). Algoritma *K-Means Clustering* dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa. *Jurnal Informatika Ekonomi Bisnis*, 4, 166–171. <https://doi.org/10.37034/infv4i4.166>
- Informasi, S., Muasaroh, Y. I., & Fatah, Z. (2024). *Jurnal Advance Research Informatika Implementasi RapidMiner dalam Optimasi Pembentukan Kelas Unggulan Menggunakan K-Means Clustering*. 3(1), 66–72.
- 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), 110–118.
- Muliono, R., & Sembiring, Z. (2019). Data Mining Clustering Menggunakan Algoritma *K-Means* Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen. *Journal of*

- Computer Engineering System and Science*, 4(2), 2502–2714.
- Utami, N. W., & Eka Putra, I. G. J. (2022). Text Minig Clustering Untuk Pengelompokan Topik Dokumen Penelitian Menggunakan Algoritma K-Means Dengan Cosine Similarity. *Jurnal Informatika Teknologi Dan Sains*, 4(3), 255–259. <https://doi.org/10.51401/jinteks.v4i3.1907>
- Virgo, I., Defit, S., & Yuhandri, Y. (2020). Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering. *Jurnal Sistim Informasi Dan Teknologi*, 2, 23–28. <https://doi.org/10.37034/jsisfotek.v2i1.17>
- Yunita, F. (2018). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru. *Sistemasi*, 7(3), 238. <https://doi.org/10.32520/stmsi.v7i3.388>
- Zanuardi, A., & Suprayitno, H. (2018). Analisa Karakteristik Kecelakaan Lalu Lintas di Jalan Ahmad Yani Surabaya melalui Pendekatan Knowledge Discovery in Database. *Jurnal Manajemen Aset Infrastruktur & Fasilitas*, 2(1), 45–55. <https://doi.org/10.12962/j26151847.v2i1.3767>