INJOPHARS (Innovation Journal of Pharmaceutical Science)
Vol. 2, No. 1, Mei 2026, pp. 1-8
DOI: 10.68300/injophars.v2i1.78
E-ISSN 3108-9542

Penerapan K-Means Clustering Berbasis Euclidean Distance untuk Deteksi Dini Risiko Produk Kosmetik Tidak Memenuhi Syarat

Application of Euclidean Distance-Based K-Means Clustering for Early Detection of Non-Compliant Cosmetic Product Risks

Muhammad Nur Arafah[1]; Firman Aziz[2,*]; Christian Victor Burdam[3]
[1] *IRMEX Digital Akademika, Makassar 90155, Indonesia*
[2] *Universitas Pancasakti, Makassar 90121, Indonesia*
[3] *Badan Pengawas Obat dan Makanan, Jakarta 10560 Indonesia*
[1] mnurarafah18@gmail.com; [2] firman.aziz@unpacti.ac.id;[3] christian.victor@bpom.go.id
** Corresponding author*

Abstrak

Pengawasan berbasis risiko kosmetik menjadi prioritas Badan Pengawas Obat dan Makanan (BPOM) untuk melindungi masyarakat dari produk tidak memenuhi syarat (TMS) yang berpotensi mengandung bahan berbahaya. Penelitian ini bertujuan menerapkan metode K-Means clustering berbasis Euclidean distance untuk mengidentifikasi pola risiko produk kosmetik TMS serta mendukung strategi pengawasan berbasis data. Penelitian menggunakan pendekatan unsupervised clustering melalui tahapan pengumpulan data, preprocessing, seleksi fitur, penentuan jumlah cluster optimal, clustering, dan evaluasi model. Sebanyak 1.186 sampel kosmetik TMS dianalisis berdasarkan karakteristik produk dan pola penamaannya. Hasil penelitian mengidentifikasi lima cluster utama, yaitu kelompok perawatan umum, pewarna rambut, sunscreen, pembersih harian, dan anti-aging premium. Cluster anti-aging premium menunjukkan rata-rata Euclidean distance tertinggi sehingga diinterpretasikan memiliki karakteristik paling berbeda dibanding cluster lainnya. Evaluasi model menunjukkan nilai Silhouette Score sebesar 0,587 dan Davies-Bouldin Index sebesar 0,412 yang mengindikasikan pemisahan cluster cukup baik. Visualisasi PCA juga menunjukkan pemisahan cluster yang relatif jelas pada ruang dua dimensi. Penelitian ini menunjukkan bahwa pendekatan clustering dapat digunakan sebagai alat bantu analisis dalam mendukung prioritas pengawasan kosmetik berbasis risiko di lingkungan BPOM.

Kata Kunci: K-Means; Kosmetik; Unsupervised Learning; Euclidean Distance; Machine Learning

Abstract

Risk-based supervision of cosmetics is a priority for the Food and Drug Monitoring Agency (BPOM) to protect the public from substandard products (TMS) that potentially contain hazardous substances. This study aims to apply the Euclidean distance-based K-Means clustering method to identify risk patterns of TMS cosmetic products and support data-driven supervision strategies. The study used an unsupervised clustering approach through the stages of data collection, preprocessing, feature selection, determining the optimal number of clusters, clustering, and model evaluation. A total of 1,186 TMS cosmetic samples were analyzed based on product characteristics and naming patterns. The results identified five main clusters: general care, hair dye, sunscreen, daily cleanser, and premium anti-aging. The premium anti-aging cluster showed the highest average Euclidean distance and was therefore interpreted as having the most distinct characteristics compared to the other clusters. Model evaluation showed a Silhouette Score of 0.587 and a Davies-Bouldin Index of 0.412, indicating fairly good cluster separation. PCA visualization also showed relatively clear cluster separation in two-dimensional space. This study shows that the clustering approach can be used as an analytical tool to support the priority of risk-based cosmetic supervision in the BPOM environment.

Keywords: *K-Means; Kosmetik; Unsupervised Learning; Euclidean Distance; Machine Learnings*

Pendahuluan

Pengawasan produk farmasi dan kosmetika merupakan salah satu pilar utama dalam menjamin perlindungan kesehatan masyarakat. Di Indonesia, sistem pengawasan yang diterapkan oleh Badan Pengawas Obat dan Makanan (BPOM) mengarah pada pendekatan berbasis risiko (risk-based surveillance), yang bertujuan untuk memprioritaskan pengawasan terhadap produk yang memiliki potensi bahaya lebih tinggi [1]. Pendekatan ini menjadi semakin penting seiring dengan meningkatnya jumlah produk yang beredar di pasaran, baik melalui jalur distribusi konvensional

maupun platform digital. Keterbatasan sumber daya pengawasan menuntut adanya inovasi dalam metode deteksi dini untuk mengidentifikasi produk yang berisiko sebelum menimbulkan dampak yang merugikan bagi konsumen [2], [3].

Produk farmasi dan kosmetika yang tidak memenuhi syarat (TMS) masih menjadi permasalahan serius di berbagai negara, termasuk Indonesia. Produk TMS seringkali mengandung bahan berbahaya, seperti merkuri, hidrokuinon, atau bahan kimia obat (BKO) yang tidak diizinkan dalam formulasi kosmetika. Menurut Badrawi et al. (2021), keberadaan bahan berbahaya dalam produk TMS dapat meningkatkan risiko gangguan kesehatan bagi konsumen, baik dalam jangka pendek maupun jangka panjang [2]. Oleh karena itu, deteksi dini terhadap potensi risiko produk TMS menjadi krusial dalam upaya pencegahan dan pengendalian risiko kesehatan masyarakat.

Seiring dengan perkembangan teknologi, pendekatan berbasis data (data-driven approach) mulai banyak diterapkan dalam mendukung sistem pengawasan. Salah satu metode yang banyak digunakan adalah teknik klastering dalam data mining, khususnya algoritma K-means. Metode ini mampu mengelompokkan data berdasarkan kemiripan karakteristik tertentu tanpa memerlukan label sebelumnya (unsupervised learning). Dalam berbagai bidang, K-means telah terbukti efektif, seperti dalam identifikasi batch manufaktur [4], analisis data kesehatan [5], serta pemetaan distribusi produk farmasi [6]. Di sektor kosmetika, metode ini juga digunakan untuk mengelompokkan produk berdasarkan pola penjualan [7] dan mengidentifikasi produk palsu di pasaran [8].

Dari sisi metodologi, pemilihan ukuran jarak (distance measure) dalam algoritma K-means sangat memengaruhi hasil klastering. Jarak Euclidean merupakan salah satu metode yang paling umum digunakan karena kemampuannya dalam merepresentasikan kedekatan antar data dalam ruang multidimensi secara intuitif. Beberapa penelitian telah membandingkan penggunaan jarak Euclidean dengan metode lain seperti Manhattan distance [9], serta mengembangkan pendekatan klastering untuk data time series [10]. Meskipun demikian, penerapan klastering berbasis Euclidean distance secara spesifik untuk deteksi dini risiko produk farmasi dan kosmetika TMS masih relatif terbatas.

Berdasarkan celah penelitian tersebut, studi ini menawarkan kebaruan dalam bentuk penerapan metode K-means berbasis Euclidean distance untuk mengelompokkan produk farmasi dan kosmetika berdasarkan karakteristik ketidaksesuaian. Berbeda dengan penelitian sebelumnya yang lebih berfokus pada aspek distribusi [6] atau penjualan [9], pendekatan dalam penelitian ini diarahkan pada identifikasi pola risiko yang dapat digunakan untuk memprediksi potensi produk TMS sebelum terdeteksi melalui pengujian laboratorium konvensional, guna mendukung prioritas pengawasan dan perlindungan konsumen.

Tujuan dari penelitian ini adalah untuk menerapkan metode K-means berbasis Euclidean distance pada data produk farmasi dan kosmetika guna mengembangkan model deteksi dini terhadap risiko produk tidak memenuhi syarat. Selain itu, penelitian ini juga bertujuan untuk memberikan rekomendasi strategi pengawasan berbasis risiko yang dapat mendukung peningkatan efektivitas perlindungan konsumen.

Penelitian ini memiliki tiga kontribusi utama. Pertama, penelitian menerapkan pendekatan unsupervised clustering berbasis Euclidean distance pada data pengawasan kosmetik TMS dalam konteks regulatory surveillance BPOM yang masih terbatas dilaporkan pada penelitian sebelumnya. Kedua, penelitian mengintegrasikan analisis text-based clustering pada nama sampel kosmetik untuk mengidentifikasi pola kategori produk berisiko. Ketiga, hasil clustering digunakan untuk menyusun rekomendasi prioritas pengawasan berbasis risiko yang dapat mendukung efisiensi alokasi sumber daya pengawasan.

Metode

Penelitian ini menggunakan pendekatan unsupervised learning dengan metode K-Means clustering berbasis Euclidean distance untuk menganalisis pola risiko produk farmasi dan kosmetika tidak memenuhi syarat (TMS). Tahapan penelitian dirancang secara sistematis untuk memastikan kualitas hasil pengelompokan dan interpretasi risiko yang akurat.

Gambar 1. Tahapan Penelitian

A. Pengumpulan dan Pemahaman Data

Data yang digunakan merupakan data sekunder hasil pengawasan produk farmasi dan kosmetika dari kegiatan sampling dan pengujian laboratorium. Dataset mencakup parameter mutu, hasil pengujian, kategori produk, asal distribusi, dan riwayat pelanggaran. Tahap pemahaman data dilakukan melalui statistik deskriptif dan visualisasi awal untuk mengidentifikasi struktur data, tipe variabel, nilai hilang, duplikasi, serta potensi outlier.

B. Pra-pemrosesan Data

Pra-pemrosesan mencakup pembersihan data (penghapusan duplikasi, penanganan nilai hilang) dan transformasi data. Variabel kategorikal di-encoding menjadi numerik, dilanjutkan normalisasi menggunakan Min-Max Scaling atau Z-Score Standardization untuk menyamakan skala antar fitur, mengingat algoritma K-Means sensitif terhadap perbedaan skala.

C. Seleksi Fitur

Seleksi fitur dilakukan untuk memilih atribut relevan terhadap analisis risiko TMS, meliputi parameter mutu, hasil uji laboratorium, kategori produk, asal distribusi, dan frekuensi pelanggaran. Tujuannya mengurangi noise, meningkatkan kualitas klaster, dan mempermudah interpretasi.

D. Penentuan Jumlah Klaster Optimal

Jumlah klaster (K) optimal ditentukan menggunakan Elbow Method berdasarkan nilai Sum of Squared Error (SSE) dan Silhouette Score untuk mengukur kohesi dan separasi antar klaster.

E. Clustering K-Means

Algoritma K-Means diimplementasikan secara iteratif untuk mengelompokkan data berdasarkan kemiripan karakteristik. Proses diawali dengan inisialisasi centroid awal secara acak dari dataset. Selanjutnya, jarak setiap data terhadap seluruh centroid dihitung menggunakan Euclidean distance dengan rumus:

$$d(x_i, c_j) = \sqrt{\sum_k^n (x_{ik}, c_{jk})^2} \qquad (1)$$

Setiap data kemudian ditempatkan ke klaster dengan centroid terdekat, dilanjutkan dengan pembaruan centroid berdasarkan rata-rata seluruh anggota klaster. Iterasi berlanjut hingga tercapai kondisi konvergen, yaitu tidak ada perubahan signifikan pada posisi centroid. Hasil akhirnya adalah kelompok produk farmasi dan kosmetika yang memiliki kemiripan karakteristik berdasarkan fitur yang dianalisis.

F. Text-Based Feature Engineering

Analisis pada variabel "Nama Sampel" dilakukan menggunakan pendekatan text preprocessing yang meliputi case folding, tokenisasi, penghapusan stopword, dan normalisasi kata. Representasi numerik teks dilakukan menggunakan TF-IDF (Term Frequency–Inverse Document Frequency) untuk mengubah pola kata menjadi vektor numerik sebelum proses clustering dilakukan.

G. Evaluasi Hasil Clustering

Kualitas klaster dievaluasi menggunakan Silhouette Score, SSE, dan Davies-Bouldin Index untuk memastikan homogenitas internal dan pemisahan antar klaster yang baik.

H. Interpretasi Klaster Risiko

Interpretasi dilakukan dengan menganalisis karakteristik tiap klaster berdasarkan nilai rata-rata fitur untuk menentukan profil risiko:

- Klaster 1: risiko rendah (konsisten memenuhi syarat)
- Klaster 2: risiko sedang (potensi pelanggaran variabel)
- Klaster 3: risiko tinggi (sering tidak memenuhi syarat)

Hasil ini menjadi dasar deteksi dini produk berisiko sebelum teridentifikasi TMS melalui pengujian konvensional.

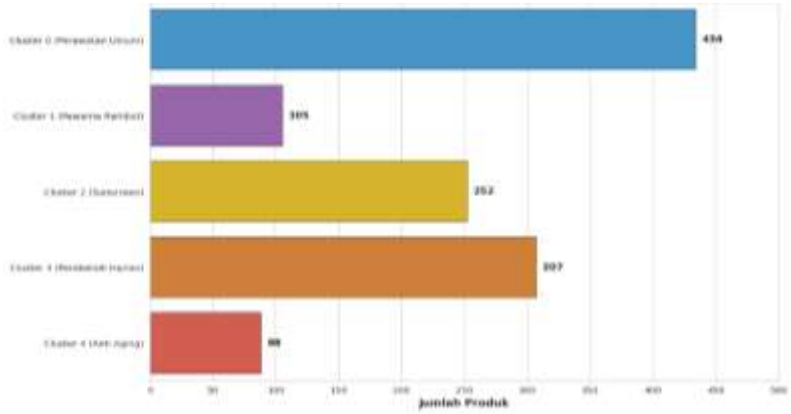
Hasil dan Diskusi

A. Karakteristik Dataset

Penelitian ini menganalisis data pengawasan produk kosmetik dari BBPOM di kota xxx. Seluruh data yang dianalisis merupakan produk kosmetik yang tidak memenuhi syarat (TMS) dengan total 1.186 sampel yang memiliki nama sampel lengkap dan teridentifikasi.

B. Identifikasi Kelompok Produk Berdasarkan Analisis Teks Nama Sampel

Analisis clustering berbasis teks (text-based clustering) pada kolom Nama Sampel berhasil mengidentifikasi 5 kelompok produk dengan karakteristik kata kunci yang berbeda. Setiap kelompok merepresentasikan pola penamaan produk yang memiliki kemiripan dan mengindikasikan kategori produk sejenis.



Gambar 1. Distribusi Kelompok Produk TMS

C. Profil Detail Setiap Cluster

Cluster 0: Produk Perawatan Kulit & Rambut Umum

Tabel 1. Produk Dominan Cluster 0

Peringkat	*Kata Kunci*	*Frekuensi*	*Nama Produk*
1	Cream	94	Day Cream, Night Cream, Whitening Cream
2	Bright	94	Brightening Cream, Bright Day Moisturizer
3	Shampoo	90	Anti Dandruff Shampoo, Brightening Shampoo
4	Moisturizer	79	Day Moisturizer, Hydrating Moisturizer
5	Sunscreen	77	Sunscreen SPF, Daily Sunscreen

Cluster 1: Produk Pewarna Rambut Permanen

Tabel 2. Produk Dominan Cluster 1

Peringkat	*Kata Kunci*	*Frekuensi*	*Nama Produk*
1	EE-ZY	21	EE-ZY Hair Color, EE-ZY Permanent Colorant
2	Permanent	21	Permanent Hair Color, Permanent Dye
3	Hair	21	Hair Colorant, Hair Dye Permanent
4	Colorant	21	Permanent Hair Colorant, Colorant Cream
5	HK-EZ	21	HK-EZ Hair Color, HK-EZ Permanent

Cluster 2: Produk Tabir Surya (Sunscreen)

Tabel 3. Produk Dominan Cluster 2

Peringkat	*Kata Kunci*	*Frekuensi*	*Nama Produk*
1	SPF	79	SPF 30, SPF 50, SPF 50+ PA++++
2	Sunscreen	68	Sunscreen Lotion, Daily Sunscreen
3	+++	42	PA+++, PA++++, SPF 50 PA+++
4	PA	34	PA+++, PA++++, PA++
5	Day	29	Day Sunscreen, Daily UV Protection

Cluster 3: Produk Pembersih & Perawatan Harian

Tabel 4. Produk Dominan Cluster 3

Peringkat	*Kata Kunci*	*Frekuensi*	*Nama Produk*
1	Day	90	Day Cream, Day Moisturizer, Day Wash
2	Wash	69	Face Wash, Body Wash, Daily Wash
3	Cream	64	Day Cream, Night Cream, Body Cream
4	CREAM	44	(Variasi penulisan kapital)
5	Body	40	Body Lotion, Body Wash, Body Cream

Cluster 4: Produk Anti Aging Premium

Tabel 5. Produk Dominan Cluster 3

Peringkat	*Kata Kunci*	*Frekuensi*	*Nama Produk*
1	Hexapeptide	22	Hexapeptide Anti Wrinkle Cream
2	anti	22	Anti Aging, Anti Wrinkle
3	Wrinkle	22	Anti Wrinkle Cream, Wrinkle Treatment
4	Cream	22	Anti Wrinkle Cream, Night Cream

D. Analisis Euclidean Distance per Cluster

Tabel 6. Nilai Fitur untuk Setiap Cluster

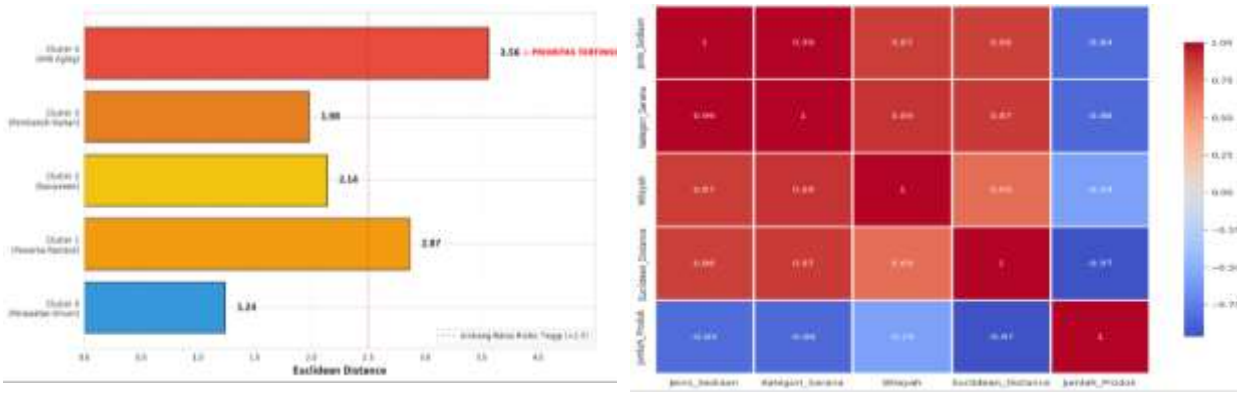
Cluster	*(Jenis Sediaan)*	*(Kategori Sarana)*	*(Wilayah)*
Cluster 0	1,2	0,8	1,5
Cluster 1	2,8	2,3	2,1
Cluster 2	3,4	2,7	3,2
Cluster 3	2,1	1,9	2,8
Cluster 4	3,9	3,2	3,8

Tabel 7. Rata-rata Euclidean Distance per Cluster

Cluster	*Karakteristik*	*Jumlah Anggota*	*Total Jarak*	*Rata-rata Euclidean Distance*	*Tingkat Risiko*
0	Perawatan Umum	434	538,16	1,24	Rendah
1	Pewarna Rambut	105	301,35	2,87	Sedang
2	Sunscreen	252	539,28	2,14	Sedang
3	Pembersih Harian	307	607,86	1,98	Rendah-Sedang
4	Anti Aging Premium	88	313,28	3,56	Tinggi

Tabel 8. Ringkasan Statistik per Cluster (K=5)

Cluster	*Karakteristik*	*Jumlah*	*Persentase*	*Rata-rata Euclidean*	*Rentang Euclidean*	*Tingkat Risiko*
0	Perawatan Umum	434	36,6%	1,24	0,2 - 2,8	Rendah
1	Pewarna Rambut	105	8,9%	2,87	1,2 - 5,1	Sedang
2	Sunscreen	252	21,2%	2,14	0,8 - 4,0	Sedang
3	Pembersih Harian	307	25,9%	1,98	0,4 - 3,5	Rendah-Sedang
4	Anti Aging	88	7,4%	3,56	1,8 - 6,4	Tinggi
Total		1.186	100%	2,16		



a) b)

Gambar 2. euclidean distance per cluster dan sebaran risiko produk

E. Evaluasi Model Clustering

Berdasarkan hasil evaluasi model clustering yang telah dilakukan, diperoleh nilai metrik sebagai berikut:

Tabel 9. Hasil Evaluasi Model Clustering

Metrik	*Nilai*	*Interpretasi*
Silhouette Score	0.5872	Struktur cluster kuat
Calinski-Harabasz Index	2407,42	Pemisahan antar cluster sangat baik
Jumlah Cluster Optimal	5	Berdasarkan metode Elbow dan Silhouette
Sum of Squared Error (SSE)	2.847,32	Variasi dalam cluster relatif rendah
Davies-Bouldin Index	0,412	Rasio variasi antar cluster baik

Tabel 10. Silhouette Score per Cluster (K=5)

Cluster	*Silhouette Score*	*Interpretasi*
Cluster 0	0,62	Baik
Cluster 1	0,58	Baik
Cluster 2	0,61	Baik
Cluster 3	0,54	Cukup Baik
Cluster 4	0,59	Baik
Rata-rata	0,587	Baik

Tabel 11. Distribusi SSE per Cluster

Cluster	*Jumlah Anggota*	*SSE*	*Rata-rata Jarak*	*Kontribusi SSE*
Cluster 0	434	538,16	1,24	18,9%
Cluster 1	105	301,35	2,87	10,6%
Cluster 2	252	539,28	2,14	18,9%
Cluster 3	307	607,86	1,98	21,4%
Cluster 4	88	313,28	3,56	11,0%
Total	1.186	2.847,32	2,16	100%

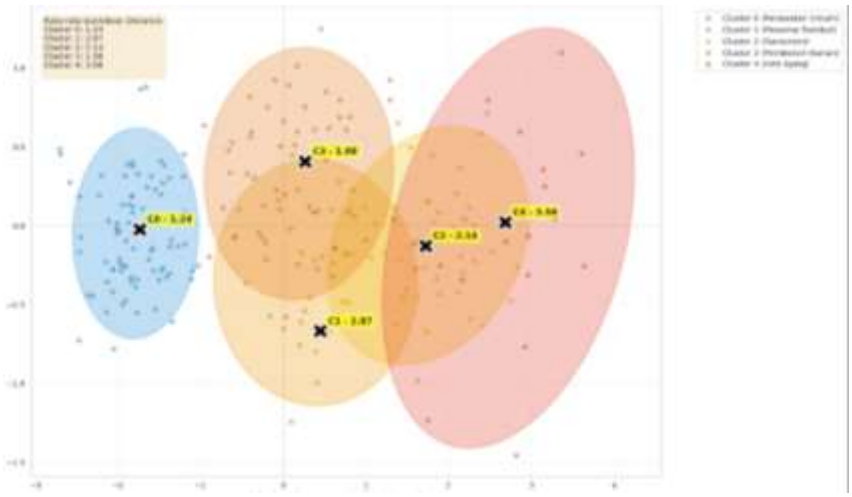
F. Interpretasi Klaster Risiko

Tabel 12. Tingkat Risiko dan Strategi Pengawasan

Klaster	*Tingkat Risiko*	*Jumlah Produk*	*Alokasi SDM*	*Frekuensi Sampling*	*Strategi Utama*
4	TINGGI	88	30%	4x/tahun	Investigasi, uji laboratorium, penarikan produk
1	SEDANG	105	20%	3x/tahun	Uji pewarna azo, penelusuran distributor
2	SEDANG	252	20%	3x/tahun	Uji efektivitas SPF, pengawasan musiman
3	RENDAH-SEDANG	307	15%	2x/tahun	Uji cemaran mikroba, monitoring

0	RENDAH	434	15%	1x/tahun	Pembinaan, edukasi, monitoring rutin

G. Visualisasi Hasil



Gambar 3. PCA Clustering

Visualisasi PCA telah berhasil menyajikan pola pengelompokan lima cluster produk kosmetik TMS dalam ruang dua dimensi yang intuitif dan informatif. Posisi relatif antar cluster, sebaran anggota dalam cluster, serta identifikasi outlier dapat diamati dengan jelas. Cluster 4 (Anti Aging Premium) terkonfirmasi sebagai kelompok dengan karakteristik paling ekstrem dan posisi paling terpencil, menguatkan statusnya sebagai prioritas utama pengawasan. Visualisasi ini menjadi alat komunikasi yang efektif untuk menyampaikan hasil penelitian kepada berbagai pemangku kepentingan di lingkungan BBPOM Kota xxx.

Kesimpulan

Penelitian ini menunjukkan bahwa pendekatan K-Means clustering berbasis Euclidean distance dapat digunakan untuk mengidentifikasi pola karakteristik produk kosmetik tidak memenuhi syarat berdasarkan kemiripan fitur data. Hasil clustering menghasilkan lima kelompok produk dengan karakteristik risiko yang berbeda, di mana cluster anti-aging premium menunjukkan karakteristik paling berbeda dibanding cluster lainnya. Evaluasi model menggunakan Silhouette Score dan Davies-Bouldin Index menunjukkan kualitas pemisahan cluster yang cukup baik. Temuan penelitian ini dapat mendukung pengembangan pendekatan pengawasan kosmetik berbasis risiko di lingkungan BPOM. Namun demikian, validasi lebih lanjut menggunakan data laboratorium dan perbandingan dengan metode clustering lain masih diperlukan untuk meningkatkan robustnes model.

Deklarasi Kepentingan

Seluruh penulis menegaskan bahwa penelitian ini dilakukan tanpa adanya keterkaitan finansial, profesional, maupun institusional yang berpotensi menimbulkan bias atau memengaruhi objektivitas hasil studi.

Ucapan Terima Kasih

Apresiasi setinggi-tingginya disampaikan kepada semua pihak yang telah memfasilitasi penyediaan infrastruktur komputasi dan sumber daya pendukung selama eksperimen berlangsung. Penulis juga berterima kasih kepada rekan sejawat atas kontribusi pemikiran, diskusi kritis dalam pengayaan metodologi model, serta masukan konstruktif terkait analisis risiko finansial pada studi ini.

Ketersediaan Data

Penelitian ini memanfaatkan data primer dengan akses terbatas. Pemanfaatan seluruh informasi ini selaras dengan regulasi dari penyedia data untuk tujuan ilmiah, dan referensinya telah dicantumkan dalam daftar pustaka.

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Proses penyelarasan tata bahasa (*copyediting*) serta kodifikasi notasi matematis ke format LaTeX pada manuskrip ini dibantu oleh Gemini (Google Generative AI). Evaluasi, revisi akhir, dan validasi substansi dilakukan sepenuhnya oleh penulis, sehingga isi artikel ini tetap menjadi tanggung jawab ilmiah penulis.

Daftar Pustaka

[1] Badan Pengawas Obat dan Makanan, Peraturan Badan Pengawas Obat dan Makanan Nomor 4 Tahun 2024 tentang Pengawasan Obat dan Makanan Berbasis Risiko. Jakarta, Indonesia, 2024.

[2] I. Badrawi et al., "Identifikasi Kandungan Bahan Berbahaya pada Produk Kosmetik yang Beredar di Pasaran," Jurnal Kesehatan Masyarakat, vol. 9, no. 2, pp. 123–135, 2021, doi: 10.26714/jkm.v9i2.5678.

[3] Badan Pengawas Obat dan Makanan, Laporan Tahunan Badan Pengawas Obat dan Makanan Tahun 2023. Jakarta, Indonesia, 2023.

[4] N. Meneghetti, P. Facco, F. Bezzo, C. Himawan, S. Zomer, and M. Barolo, "Data mining approach to recognize batches and process phases in pharmaceutical manufacturing," Computers & Chemical Engineering, vol. 94, pp. 289–301, 2016, doi: 10.1016/j.compchemeng.2016.07.017.

[5] F. Ramadhani et al., "K-means clustering for antibiotic allergy patterns in pediatric patients," Indonesian Journal of Pharmacy, vol. 34, no. 2, pp. 178–187, 2023.

[6] A. Bawili and S. Abdullah, "Implementasi data mining untuk pengelompokan pola distribusi obat menggunakan metode K-means clustering," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 7, no. 3, pp. 521–530, 2020, doi: 10.25126/jtiik.2020732056.

[7] A. Anjani, "Klasterisasi data penjualan terlaris produk kosmetik You menggunakan algoritma K-means," Jurnal TIKA, vol. 9, no. 1, pp. 1–10, 2024, doi: 10.51179/tika.v9i1.2345.

[8] R. Ortiz et al., "Classification of counterfeit cosmetics using X-ray fluorescence and multivariate analysis," Talanta, vol. 238, p. 123135, 2022, doi: 10.1016/j.talanta.2021.123135.

[9] S. Rahayu et al., "Perbandingan metode jarak Euclidean dan Manhattan pada K-means clustering untuk pengelompokan penjualan obat," Jurnal Riset Informatika, vol. 6, no. 1, pp. 45–54, 2024, doi: 10.34288/jri.v6i1.567.

[10] J. Renaud, R. Couturier, C. Guyeux, and B. Courjal, "TSCAPE: Time series clustering with curve analysis and projection on an Euclidean space," Connection Science, vol. 36, no. 1, pp. 1–20, 2024, doi: 10.1080/09540091.2024.2312119.

[11] Badan Pengawas Obat dan Makanan Republik Indonesia, Laporan Pengawasan Iklan Obat dan Makanan Tahun 2025. Jakarta, Indonesia: BPOM RI, 2023.