

PENERAPAN ALGORITMA K-MEANS UNTUK KLASTERISASI PARFUM BRAND LOKAL INDONESIA BERBASIS WEBSITE

M. Krisna Audy Arwishak, Eko Heri Susanto, Nurlaily Vendyansyah

Teknik Informatika, Institut Teknologi Nasional Malang

Jalan Raya Karanglo km 2 Malang, Indonesia

audyarwishak10@gmail.com

ABSTRAK

Industri parfum lokal di Indonesia mengalami pertumbuhan pesat dengan beragam pilihan aroma, harga, konsentrasi, dan ukuran. Namun, konsumen sering menghadapi kesulitan dalam menentukan produk yang sesuai dengan preferensi mereka, sementara produsen juga kurang memahami segmentasi pasar secara objektif. Penelitian ini bertujuan menerapkan algoritma K-Means clustering untuk mengelompokkan parfum merek lokal berdasarkan variabel harga, konsentrasi, ukuran, dan tipe. Sistem klasterisasi parfum dibangun berbasis website dengan menggunakan teknologi framework Laravel agar hasil klasterisasi dapat diakses dengan mudah dan interaktif. Dataset yang digunakan dalam penelitian ini diambil dari platform Kaggle yang berisi 570 data parfum brand lokal Indonesia. Proses klasterisasi menghasilkan tiga kelompok utama, yaitu C1 untuk kategori rendah sejumlah 81 data, C2 untuk kategori sedang sejumlah 466 data, dan C3 untuk kategori tinggi sejumlah 22 data. Hasil penelitian menunjukkan bahwa sistem klasterisasi mampu membantu memberikan gambaran yang lebih jelas mengenai pola segmentasi produk parfum lokal sehingga dapat menjadi referensi baik bagi konsumen dalam memilih produk maupun bagi produsen dalam merumuskan strategi pemasaran yang lebih tepat sasaran. Pengujian sistem dilakukan menggunakan metode blackbox, dan hasilnya menunjukkan tingkat keberhasilan 100% di mana seluruh fungsi berjalan sesuai dengan yang diharapkan. Selain itu, pengujian validasi menggunakan silhouette score menunjukkan bahwa hasil klasterisasi berada pada tingkat yang cukup baik, sehingga dapat disimpulkan bahwa metode K-Means efektif diterapkan dalam analisis data parfum lokal Indonesia.

Kata kunci : Data mining, k-means, klasterisasi, parfum lokal

1. PENDAHULUAN

Industri parfum lokal di Indonesia telah mengalami perkembangan pesat dalam beberapa tahun belakangan, hal ini terlihat dengan kemunculan berbagai merek baru yang menawarkan beragam pilihan aroma, harga, ukuran, dan konsentrasi. Variasi produk ini, di satu sisi, memberikan lebih banyak pilihan bagi konsumen, tetapi di sisi lain dapat menimbulkan kebingungan saat mereka harus mengambil keputusan dalam pembelian. Konsumen sering kali mengalami hambatan dalam mencocokkan pilihan aroma dengan harga dan tingkat konsentrasi yang diinginkan, sementara produsen menghadapi kesulitan dalam memahami segmen pasar guna mengembangkan strategi pemasaran yang efektif [2]. Oleh sebab itu, dibutuhkan pendekatan analitis yang dapat mengategorikan produk parfum berdasarkan karakteristik utama, sehingga memudahkan konsumen dan pelaku industri dalam memahami profil pasar secara objektif.

Algoritma K Means dipilih dalam studi ini karena cepat dan efisien dalam mengolah dataset yang memiliki ribuan entri, serta dapat menghasilkan kluster yang mudah untuk dipahami. Meskipun demikian, K Means memiliki sejumlah kekurangan, seperti ketergantungan terhadap jumlah kluster awal (k), pemilihan centroid, dan bentuk distribusi data yang tidak berbentuk bulat [5]. Untuk menangani permasalahan tersebut, penelitian ini menggabungkan metode Elbow dan Silhouette Score dalam menentukan jumlah kluster yang optimal, disertai

dengan pemrosesan data awal yang mencakup normalisasi dan deteksi outlier untuk meningkatkan akurasi klusterisasi.

Berdasarkan isu yang telah dibahas, studi ini bertujuan untuk menerapkan algoritma K-Means dalam mengelompokkan parfum dari merek lokal Indonesia berdasarkan variabel harga, ukuran, konsentrasi, dan tipe, serta mengembangkan aplikasi web dengan menggunakan Laravel sebagai antarmuka pengguna. Oleh karena itu, diharapkan sistem klasterisasi ini dapat membantu konsumen dalam memilih produk yang sesuai dengan preferensi mereka dan memberikan pemahaman yang lebih jelas mengenai segmentasi pasar bagi produsen, sehingga mendukung pengambilan keputusan yang berbasis data di industri parfum lokal.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Segmentasi pasar menggunakan algoritma K-Means sangat efektif dalam mengelompokkan konsumen parfum berdasar pada preferensi aroma, seperti floral, oriental, dan citrus.[6] Penelitian ini menunjukkan bahwa metode ini dapat mengidentifikasi tiga segmen utama berdasarkan data survei pelanggan, yang meliputi preferensi aroma, kebiasaan pembelian, harga, dan kemasan. Hasil dari klasterisasi ini memungkinkan perusahaan untuk merancang strategi pemasaran yang lebih akurat dan meningkatkan personalisasi produk sesuai dengan karakteristik setiap segmen.

Penggunaan algoritma K-Means untuk mengkategorikan produk parfum di platform e-commerce berdasarkan faktor harga, tingkat popularitas, dan ulasan dari pelanggan. Dengan kumpulan data yang mencakup lebih dari 300 variasi parfum, pendekatan ini menghasilkan kelompok harga yang rendah, menengah, dan tinggi. Sebagai hasilnya, sistem klasifikasi ini mempermudah konsumen dalam menemukan produk dan memungkinkan e-commerce untuk memberikan rekomendasi serta strategi promosi yang lebih efektif sesuai dengan segmen pelanggan.[2]

Melaksanakan algoritma K-Means untuk mengelompokkan produk parfum di situs web berdasarkan harga, bahan, dan pendapat pengguna. Informasi diperoleh dari situs penjualan kosmetik yang menawarkan ribuan produk. Hasil dari analisis kluster mengelompokkan parfum ke dalam empat kategori utama: parfum premium, kelas menengah, untuk remaja, dan parfum ekonomis. Pendekatan ini mendukung strategi pemasaran yang lebih terarah dan memberikan pengalaman berbelanja yang lebih pribadi bagi para konsumen.[13]

2.2. Data Mining



Gambar 1. Data Mining

Data Mining merupakan langkah untuk menemukan pola yang tidak terlihat dalam kumpulan data yang besar melalui beberapa fase, yaitu pembersihan, penggabungan, pemilihan, analisis, dan penilaian. Proses ini bertujuan untuk menghasilkan informasi yang bermanfaat bagi pengambilan keputusan [12].

2.3. Clustering



Gambar 2. Clustering

Clustering adalah metode pengelompokan data yang tidak memiliki label, yang dilakukan berdasarkan kesamaan atribut tertentu, dengan tujuan untuk menemukan kelompok yang memiliki makna. Metode ini bermanfaat untuk melihat data lebih dalam dan sebagai langkah awal sebelum melakukan klasifikasi. Klasifikasi membantu mengidentifikasi pola tersembunyi dalam data dan menjadi dasar untuk pengembangan model analisis yang lebih lanjut.[7]

2.4. Algoritma K-Means



Gambar 3. K-Means

Algoritma K-Means adalah metode clustering non-hierarkis yang membagi data ke dalam beberapa cluster berdasarkan kemiripan karakteristik, dengan tujuan meminimalkan variasi dalam cluster dan memaksimalkan variasi antar cluster. Prosesnya dimulai dengan penentuan jumlah cluster (k), pemilihan centroid awal secara acak, perhitungan jarak antar objek menggunakan Euclidean Distance, pengalokasian objek ke centroid terdekat, dan pembaruan posisi centroid hingga konvergen[11].

Algoritma K-Means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat *cluster* dan proses pencarian anggota dari tiap-tiap *cluster*. Proses algoritma K-Means :

- Tentukan k sebagai jumlah *cluster* yang dibentuk
- Tentukan k *centroid* (titik pusat *cluster*) awal secara random

Penentuan *centroid* awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung *centroid cluster* ke- i berikutnya, digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^n x_i}{n}; i = 1, 2, 3, \dots, n \quad (1)$$

Dimana :

v : *centroid* pada *cluster*

x_i : objek ke- i

n : banyaknya objek/jumlah objek yang menjadi anggota *cluster*.

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

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

Dimana :

x_i : objek x ke-i

y_i : objek y ke-I

n : banyaknya objek

- d. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat. Untuk melakukan pengalokasian objek kedalam masing-masing cluster pada saat iterasi secara umum dapat dilakukan dengan cara hard k-means dimana secara tegas setiap objek dinyatakan sebagai anggota *cluster* dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *cluster* tersebut.
- e. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan 1
- f. Ulangi langkah c jika posisi *centroid* baru tidak sama

2.5. Website



Gambar 4. Website

Website merupakan sekumpulan halaman yang menyimpan informasi digital, seperti tulisan, gambar, suara, atau video, yang dapat diakses melalui internet. Halaman ini disusun menggunakan HTML dan ditampilkan oleh browser web sehingga dapat diakses dan dibaca oleh semua orang[1].

2.6. Database

Database merupakan sarana untuk menyimpan data secara teratur, sehingga dapat diakses, dikelola, dan dimanfaatkan dengan mudah. Basis data terdiri dari tabel yang berisi baris (rekaman) dan kolom (bidang), di mana bidang menyimpan jenis data dan rekaman menyimpan data lengkap mengenai suatu entitas. Struktur ini memungkinkan pengaturan data dengan cara yang logis dan efisien[8].

2.7. PHP

PHP merupakan sebuah bahasa pemrograman untuk web yang diciptakan oleh Rasmus Lerdorf pada tahun 1994 dengan tujuan untuk mencatat pengunjung di situs pribadinya. PHP merupakan bahasa pemrograman yang dijalankan di sisi server, diolah di server, dan hasilnya ditampilkan di peramban. PHP merupakan bahasa pemrograman yang mudah dipahami, tidak memerlukan biaya, dan digunakan untuk mengembangkan aplikasi web yang dinamis dan interaktif[10].

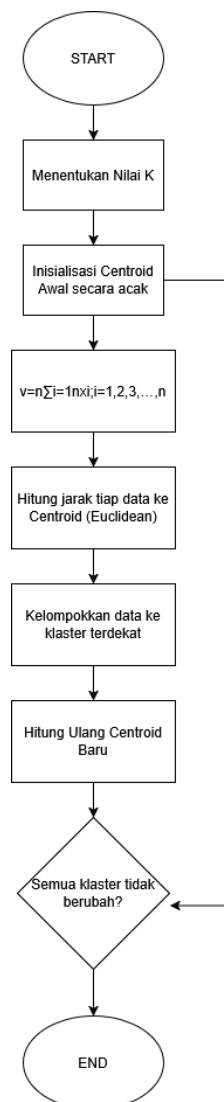
3. METODE PENELITIAN

Untuk memenuhi tujuan dalam penelitian ini, diperlukan beberapa langkah – langkah sebagai berikut:

3.1. Metode Pengumpulan Data

Pengumpulan data dalam penelitian ini menggunakan data sekunder. Data diperoleh dari website www.kaggle.com tahun 2022 hingga 2024 dengan jumlah data 570. Data yang dikumpulkan telah mencakup informasi Brand, Tahun, Variant, Ukuran, Harga, dan tipe.

3.2. Flowchart Metode



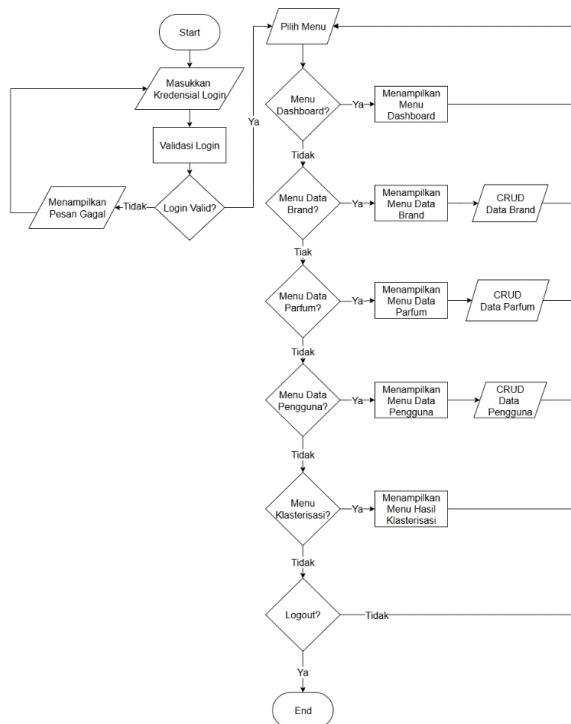
Gambar 5. Flowchart sistem klusterisasi parfum brand lokal

Flowchart algoritma K-Means dimulai dengan menentukan jumlah kluster (k) dan menginisialisasi centroid awal secara acak, lalu menghitung jarak Euclidean setiap data ke centroid untuk mengelompokkan data ke kluster terdekat. Selanjutnya, centroid baru dihitung ulang sebagai rata-

rata data dalam setiap kluster. Proses literatif ini berlanjut dengan membandingkan posisi centroid baru dengan centroid lama, jika terdapat perubahan, langkah perhitungan jarak dan pengelompokan diulang hingga centroid tidak lagi berubah atau mencapai batas iterasi, menandakan algoritma telah mencapai konvergensi dan proses selesai.

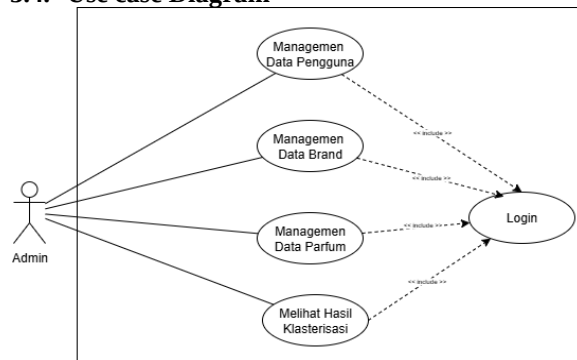
3.3. Flowchart Sistem

Diagram alur sistem dimulai dari proses login dengan memasukkan username dan password yang kemudian divalidasi oleh sistem. Jika valid, admin diarahkan ke halaman utama untuk memilih menu seperti Dashboard, Data Brand, Data Parfum, Data Pengguna, dan Klasterisasi. Masing-masing menu menampilkan submenu sesuai fungsinya, termasuk fitur CRUD untuk data brand, parfum, dan pengguna, serta hasil klasterisasi data parfum menggunakan algoritma K-Means.



Gambar 6. Flowchart sistem klasterisasi parfum brand lokal

3.4. Use case Diagram

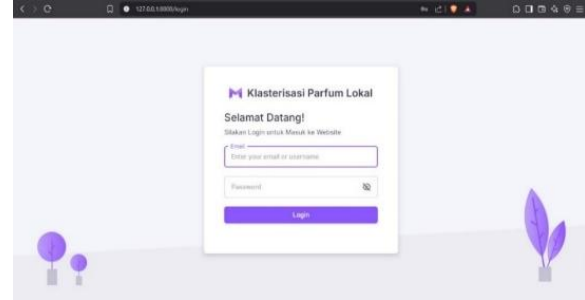


Gambar 7. Use case diagram

Gambar 7 Diagram use case menggambarkan peran Admin dalam sistem, mulai dari login, mengelola data pengguna, brand, dan parfum, hingga input dataset dan data uji. Data uji diproses dengan metode K-Means untuk menghasilkan hasil klasterisasi yang ditampilkan pada menu hasil. Seluruh fitur hanya dapat diakses setelah proses login.

4. HASIL DAN PEMBAHASAN

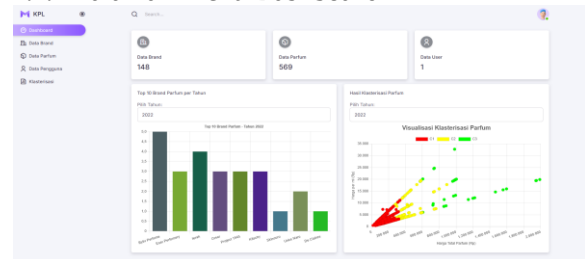
4.1. Halaman Menu Login



Gambar 8. Menu Login

Pada gambar 8 merupakan desain halaman login pada website, dimana pengguna atau admin memasukkan username dan password terlebih dahulu sebelum masuk ke dalam website.

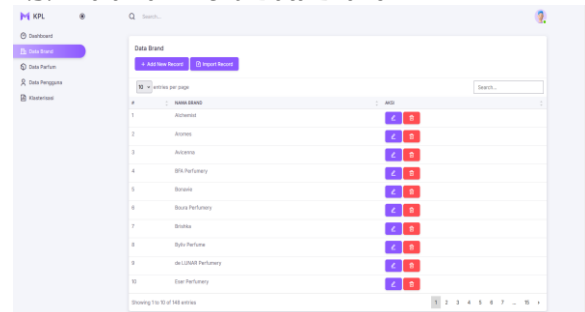
4.2. Halaman Menu Dashboard



Gambar 9. Menu Dashboard

Gambar 9 menampilkan menu dashboard yang merangkum jumlah data Brand, Parfum, Pengguna, dan Hasil Klasterisasi. Terdapat grafik batang Top 10 Brand Parfum per tahun dan grafik scatter hasil klasterisasi berdasarkan Harga Jual dan Jumlah Penjualan, yang memudahkan analisis pola dan segmentasi parfum.

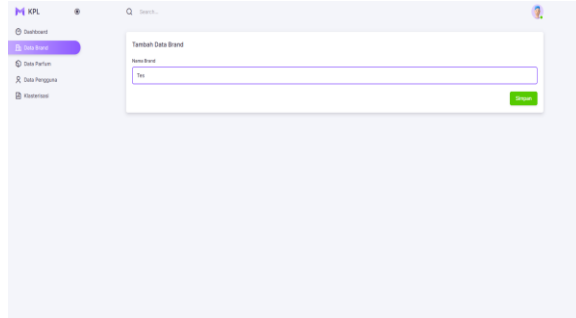
4.3. Halaman Menu Data Brand



Gambar 10. Menu Data Brand

Gambar 10 menunjukkan menu data Brand yang menampilkan daftar Brand Parfum serta fitur untuk menambah, mengedit, dan menghapus data brand.

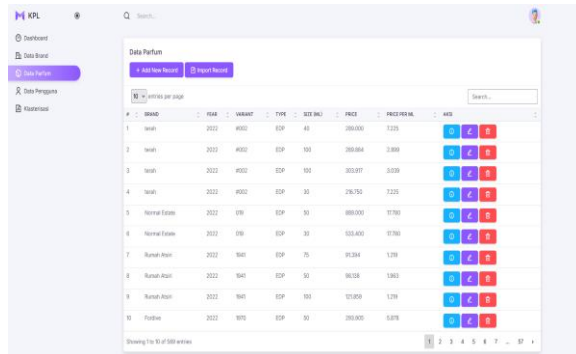
4.4. Halaman Tambah Data Brand



Gambar 11. Halaman Tambah Data Brand

Gambar 11 menunjukkan halaman Tambah Data Brand, di mana admin dapat memasukkan nama brand parfum melalui form sederhana dan menyimpannya ke database dengan tombol Simpan.

4.5. Halaman Menu Data Parfum



Gambar 12. Menu Halaman Data Parfum

Gambar 12 menampilkan menu data Parfum dengan fitur untuk menambah, mengedit, dan menghapus data parfum.

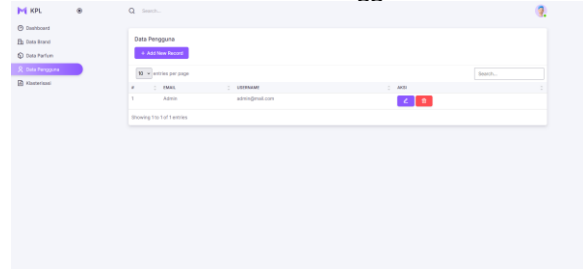
4.6. Halaman Tambah Data Parfum



Gambar 13. Halaman Tambah Data Parfum

Gambar 13 menunjukkan halaman Tambah Data Parfum yang memungkinkan admin mengisi detail produk parfum seperti Brand, Tahun, Variant, Type, Ukuran, Harga, dan Top Notes untuk disimpan ke dalam sistem.

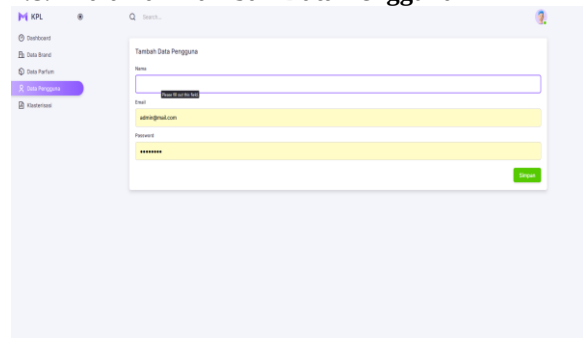
4.7. Halaman Menu Data Pengguna



Gambar 14. Halaman Menu Data Pengguna

Gambar 14 menampilkan halaman Data Pengguna yang berisi tabel akun pengguna, dengan opsi edit dan hapus untuk setiap data melalui ikon pensil dan tempat sampah.

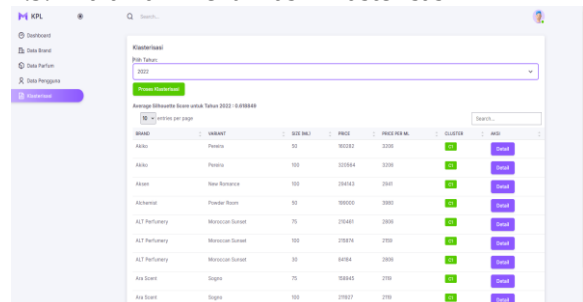
4.8. Halaman Tambah Data Pengguna



Gambar 15. Halaman Tambah Data Pengguna

Gambar 15 menunjukkan halaman Tambah Data Pengguna, di mana admin mengisi Nama, Email, dan Password untuk menambahkan akun baru. Data wajib diisi sebelum disimpan dengan tombol Simpan.

4.9. Halaman Menu Hasil Klasterisasi



Gambar 16. Halaman Menu Hasil Klasterisasi

Gambar 16 Halaman Hasil Klasterisasi menampilkan hasil pengelompokan data parfum berdasarkan data tahun yang dipilih, dalam hal ini tahun 2022. Berlaku untuk tahun 2023 dan 2024 Pengguna dapat memilih tahun dan menekan tombol Proses Klasterisasi untuk memulai proses pengelompokan data parfum.

4.10. Pengujian Blackbox

Pengujian *blackbox* sistem *clustering* Parfum Lokal Indonesia bertujuan untuk mengamati apakah hasil *input* yang dilakukan sesuai dengan *output* yang

diharapkan pada *web browser Mozilla Firefox* versi 139.0.1 (64-bit), dan *Microsoft Edge* versi Version 137.0.3296.52 (64-bit).

Tabel 1. Pengujian Blackbox

Fitur	Kasus Uji	Hasil yang diharapkan	Hasil yang Didapat	Kesimpulan	
				Mozilla Firefox	Microsoft Edge
Login	Memasukan <i>username</i> dan <i>password</i> salah	Muncul pesan <i>username</i> atau <i>password</i> salah	Sistem menolak dan mengirim notifikasi <i>username</i> dan <i>password</i> salah	✓	✓
	Memasukan <i>username</i> dan <i>password</i> benar	Menuju pada halaman <i>dashboard</i>	Sistem menerima akses dan langsung menuju ke menu <i>dashboard</i>	✓	✓
Halaman Dashboard	Memasukan informasi utama seperti data brand, data parfum, data pengguna dan Hasil Klasterisasi	Menampilkan informasi pada <i>dashboard</i> sesuai dengan data yang tersimpan di <i>database</i>	Sistem menampilkan data informasi sesuai dengan data yang tersimpan di <i>database</i>	✓	✓
Halaman Data Brand	Menambahkan data Brand	Muncul pesan “data berhasil di simpan”	Sistem menampilkan notifikasi “data berhasil di simpan”	✓	✓
	Mengedit data brand	Ketika di edit pada tabel data brand berubah	Data masuk kedalam tabel nama brand	✓	✓
	Menghapus data brand	Muncul pesan “apakah anda yakin akan hapus?”	Sistem berhasil menampilkan notifikasi “apakah anda yakin akan hapus?”	✓	✓
Halaman Data Parfum	Menambahkan data Parfum	Muncul pesan “data berhasil di simpan”	Sistem menampilkan notifikasi “data berhasil di simpan”	✓	✓
	Mengedit data Parfum	Ketika di edit pada tabel data parfum berubah	Data masuk kedalam tabel data parfum	✓	✓
	Menghapus data parfum	Muncul pesan “apakah anda yakin akan hapus?”	Sistem berhasil menampilkan notifikasi “apakah anda yakin akan hapus?”	✓	✓
Halaman Data Pengguna	Menambahkan Data Pengguna	Muncul pesan “data berhasil di simpan”	Sistem menampilkan notifikasi “data berhasil di simpan”	✓	✓
	Mengedit Data Pengguna	Ketika di edit pada tabel jenis data pengguna	Data masuk kedalam tabel data pengguna	✓	✓
	Menghapus Data Pengguna	Muncul pesan “apakah anda yakin akan hapus?”	Sistem berhasil menampilkan notifikasi “apakah anda yakin akan hapus?”	✓	✓
Halaman Klasterisasi	Menekan <i>button</i> proses klasterisasi data	Menampilkan hasil <i>cluster</i> yang sudah dihitung	Sistem menerima <i>direct</i> apabila tabel perhitungan sudah terisi	✓	✓
	Memilih menu pilih tahun	Menampilkan pilih tahun	Sistem berhasil memunculkan <i>cluster</i> setiap tahunnya	✓	✓
Persentase				100%	100%

Pada tabel pengujian blackbox diatas dilakukan pengujian pada 6 halaman, yaitu halaman Login, Dashboard, Data Brand, Data Parfum, Data Pengguna

dan Klasterisasi yang berhasil dilakukan pada 2 browser yaitu browser Mozzila dan Microsoft Edge dengan hasil persentase pengujian 100%.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan implementasi sistem yang telah dilakukan, dapat disimpulkan bahwa penerapan algoritma K-Means untuk klasterisasi parfum brand lokal Indonesia berjalan dengan baik dan efektif. Sistem mampu mengelompokkan 570 data parfum ke dalam tiga klaster yaitu rendah, sedang, dan tinggi pada tahun 2022, 2023, dan 2024 berdasarkan variabel harga, ukuran, dan variant dengan Average Score 0,5413. Hasil perhitungan sistem telah diverifikasi melalui pengujian manual menggunakan Microsoft Excel dan menunjukkan kecocokan hasil. Selain itu, pengujian blackbox pada web browser Mozilla Firefox dan Microsoft Edge menunjukkan bahwa seluruh fitur sistem berfungsi 100% sesuai dengan harapan. Adapun saran untuk penelitian selanjutnya, pengembangan sistem dapat dilakukan dengan menambahkan variabel baru seperti rating pengguna, aroma top notes, atau tingkat ketahanan parfum agar klasterisasi lebih komprehensif atau mencakup berbagai sudut pandang variabel yang relevan.

DAFTAR PUSTAKA

- [1] Abdulloh, R. (2021). *Easy & Simple Web Programming*. Jakarta: PT Elex Media Komputindo.
- [2] Fadhillah, R., & Putra, Y. (2022). Penerapan Algoritma K-Means untuk Klasterisasi Produk dalam E-Commerce. *Jurnal Sistem Informasi*, 12(1), 33–42.
- [3] Gustientiedina, M. A. (2023). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru. *Jurnal Nasional Teknologi dan Sistem Informasi*, 17–24.
- [4] Junirianto, E. (2020). *Pemrograman Web dengan Framework Laravel*. Ponorogo: Wade Group.
- [5] Maulana, A., Dana, R. D., & Nuris, N. D. (2024). Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Data Kerusakan Rumah Akibat Bencana Alam Di Kabupaten Cirebon. *JATI*, 8(2), 1417–1424.
- [6] Rahmawati, A., & Suryanto, H. (2021). Analisis Segmentasi Konsumen Parfum Menggunakan Algoritma K-Means. *Jurnal Teknologi Informasi dan Komputer*, 10(2), 45–55.
- [7] Nugraha, R., & Sari, M. (2023). Clustering Data Preferensi Konsumen Parfum Menggunakan Algoritma K-Means. *Jurnal Ilmu Komputer dan Informatika*, 11(2), 21–30.
- [8] Rossa, A. S., & Shalahuddin, M. (2020). *Rekayasa Perangkat Lunak: Terstruktur dan Berorientasi Objek*. Bandung: Informatika Bandung.
- [9] Sarwono, J. (2020). *Bikin Website Itu Mudah*. Jakarta: MediaKita.
- [10] Solichin, A. (2021). *Pemrograman Web dengan PHP dan MySQL*. Yogyakarta: Universitas Budi Luhur.
- [11] Sri Julyantari, N., Budiarta, I. K. B., & Putri, N. M. D. K. (2021). Implementasi K-Means untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih). *Jurnal Janitra Informatika dan Sistem Informasi*, 1(2), 92–101.
- [12] Enterprise, J. (2020). *MySQL Untuk Pemula*. Jakarta: PT Elex Media Komputindo.
- [13] W. Purba, W., Siawin, H. (2022). Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi PHK Dengan Algoritma K-Means Clustering. *JUSIKOM PRIMA*, 2(1), 85–90.
- [14] Wibowo, A., & Kartika, D. (2020). Pengelompokan Produk Kosmetik Berbasis Web Menggunakan K-Means. *Jurnal Data Mining dan Kecerdasan Buatan*, 8(3), 78–89.
- [15] Yuliana, D., & Pratama, R. (2023). Penerapan Algoritma K-Means Untuk Klasterisasi Data Penjualan Barang Pada Minimarket. *Jurnal Teknologi Informasi dan Sains*, 14(1), 55–63.