

Implementasi metode K-Means dalam klasterisasi laptop terbaik

Mila Amarila Prameswari

Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim Malang
e-mail: 200605110080@student.uin-malang.ac.id

Kata Kunci:

K-means; laptop; klastering

Keywords:

K-means; laptop; Clustering

ABSTRAK

Klastering merupakan metode yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki karakteristik serupa. Dalam konteks ini, klastering digunakan untuk mengelompokkan laptop berdasarkan spesifikasi dan harga guna membantu konsumen dalam memilih laptop yang sesuai dengan kebutuhan mereka. Dalam penelitian ini, metode klastering K-means digunakan untuk mengelompokkan

laptop menjadi beberapa klaster yang memiliki karakteristik serupa. Klastering dilakukan dengan mempertimbangkan beberapa faktor penting seperti penyimpanan, performa, portabilitas, layar touchscreen, ukuran, dan harga. Tahap awal penelitian melibatkan preprocessing data, pembagian data menjadi data training dan data testing, dan proses algoritma K-means untuk melakukan clustering pada fitur-fitur terkait dataset laptop. Evaluasi model dilakukan menggunakan metode Elbow untuk menentukan jumlah kelompok (kluster) yang optimal. Dalam kasus ini dilakukan analisis untuk membandingkan jenis laptop yang bisa direkomendasikan kepada konsumen berdasarkan keperluan dan biaya yang tersedia. Hasil klastering memberikan perbandingan laptop berdasarkan spesifikasi dan harga. Klaster 2 cocok untuk yang butuh penyimpanan besar, performa baik, dan harga terjangkau. Klaster 3 cocok untuk portabilitas, layar touchscreen, dan performa baik dengan ukuran kecil. Klaster 4 cocok untuk laptop kompak, performa baik, dan harga terjangkau.

ABSTRACT

Clustering is a method used to group data into groups that have similar characteristics. In this context, clustering is used to group laptops based on specifications and price to assist consumers in choosing a laptop that suits their needs. In this study, the K-means clustering method is used to group laptops into several clusters that have similar characteristics. Clustering is done by considering several important factors such as storage, performance, portability, touchscreen display, size, and price. The initial stages of the research involved preprocessing the data, dividing the data into training data and testing data, and processing the K-means algorithm to cluster the features related to the laptop dataset. Model evaluation was carried out using the Elbow method to determine the optimal number of clusters. In this case an analysis was carried out to compare the types of laptops that can be recommended to consumers based on the needs and costs available. Clustering results provide a comparison of laptops based on specifications and prices. Cluster 2 is suitable for those who need large storage, good performance, and affordable prices. Cluster 3 lends itself to portability, a touchscreen display, and good performance at a small size. Cluster 4 is suitable for compact laptops, good performance and affordable prices.

Pendahuluan

Dalam era kemajuan digital yang terus berkembang dengan pesat, penggunaan laptop telah menjadi salah satu kebutuhan utama bagi banyak individu di berbagai



This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.

Copyright © 2023 by Author. Published by Universitas Islam Negeri Maulana Malik Ibrahim Malang.

bidang. Namun, dengan tersedianya berbagai merek, spesifikasi, dan harga yang berbeda-beda, proses memilih laptop terbaik yang sesuai dengan kebutuhan dan preferensi individu dapat menjadi tugas yang menantang. Untuk membantu pengguna dalam mengatasi tantangan ini, teknik klastering menggunakan algoritma K-means telah diusulkan sebagai metode efektif untuk mengelompokkan laptop berdasarkan karakteristik dan kualitasnya (Guodong Liu, 2023).

Klastering merupakan teknik analisis data yang digunakan untuk mengelompokkan objek-objek serupa ke dalam kelompok atau klaster berdasarkan kesamaan karakteristik (Sinaga & Yang, 2020). Dalam konteks laptop, klastering dapat membantu mengidentifikasi kelompok laptop yang memiliki spesifikasi, performa, atau fitur serupa. Salah satu algoritma klastering yang umum digunakan dan terbukti efektif adalah algoritma K-means.

Metode K-means bekerja dengan mengelompokkan objek-objek dalam himpunan data ke dalam k klaster, di mana k merupakan jumlah klaster yang telah ditentukan sebelumnya (Ahmed, 2020). Algoritma ini mengukur jarak antara objek-objek dalam himpunan data dengan pusat-pusat klaster yang ada, lalu mengelompokkan objek-objek tersebut ke dalam klaster dengan pusat terdekat. Proses ini berlanjut hingga mencapai konvergensi, di mana jarak antara objek-objek dalam klaster tidak mengalami perubahan yang signifikan (Liu, 2022).

Dalam konteks klastering laptop, variabel-variabel yang digunakan untuk mengelompokkan laptop dapat mencakup spesifikasi seperti kecepatan prosesor, kapasitas penyimpanan, ukuran layar, RAM, kartu grafis, dan faktor-faktor lainnya. Dengan menerapkan algoritma K-means pada data laptop yang ada, kita dapat membentuk kelompok-kelompok laptop yang memiliki karakteristik serupa. Hal ini memudahkan pengguna untuk memilih laptop yang sesuai dengan kebutuhan mereka berdasarkan kelompok yang terbentuk (Kamalia, 2022).

Pada penelitian sebelumnya dengan judul "Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering". Pemilihan laptop pada penelitian ini menggunakan collaborative filtering, sehingga dibutuhkan data informasi rating dari setiap user yang telah memberi rating terhadap suatu item atau laptop. Dari data rating tersebut nantinya akan dihitung nilai kemiripannya menggunakan algoritma adjusted cosine similarity dan melakukan bobot perhitungan prediksi menggunakan algoritma weighted sum (Wijaya & Alfian, 2018).

Lalu pada penelitian yang lain dengan judul "Penerapan Data Mining Rekomendasi Laptop Menggunakan Algoritma Apriori". Didalamnya menggunakan metode Algoritma apriori yang termasuk jenis aturan asosiasi (Association Rules) dalam menentukan itemset dari transaksi penjualan laptop yang kemudian dikelompokkan dalam tabel itemset₁, tabel itemset₂, tabel nilai support, tabel nilai confidence dan mendapatkan pola penjualan laptop. Dibantu dengan data mining, hasil akhir dari sistem ini menginformasikan data transaksi sebuah penjualan laptop (Ulil, 2021).

Dalam penelitian yang akan dilakukan ini, peneliti akan menerapkan metode K-means yang berbeda dengan penelitian-penelitian sebelumnya serta dataset yang dipakai juga berbeda. Selanjutnya akan digunakan sejumlah variabel spesifikasi yang

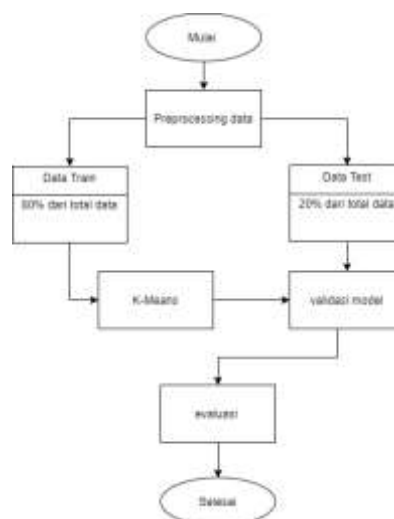
relevan untuk mengelompokkan laptop-laptop tersebut. Hasil klastering akan memungkinkan pengguna untuk mendapatkan pemahaman yang lebih baik tentang berbagai kelompok laptop yang ada, sehingga mereka dapat membuat keputusan yang lebih informasi dalam memilih laptop terbaik yang sesuai dengan kebutuhan dan preferensi mereka.

Dengan memanfaatkan teknik klastering menggunakan algoritma K-means, diharapkan bahwa penelitian ini akan memberikan kontribusi dalam memudahkan pengguna dalam memilih laptop terbaik yang sesuai dengan kebutuhan mereka. Selain itu, hasil klastering juga dapat memberikan wawasan yang berharga bagi produsen laptop dan penjual untuk mengidentifikasi tren dan preferensi pasar, serta memperbaiki strategi pemasaran mereka.

Pembahasan

Pada penelitian ini akan dilaksanakan dengan menggunakan serangkaian alur penelitian agar dapat memperoleh hasil yang tepat dan akurat. Alur penelitian tersebut diantaranya adalah sebagai berikut:

Gambar 1.1 Alur Penelitian



Gambar 2. Tahapan Kegiatan dalam Penelitian

Preprocessing Data

Dataset yang digunakan pada penelitian ini adalah Dataset Laptop Price. Pada dataset ini berisi informasi tentang atribut apa saja yang digunakan sebagai fitur untuk melakukan klastering pada dataset tersebut. Dataset tersebut memiliki 13 fitur, yaitu id laptop, Company, A Product, A TypeName, Inches, ScreenResolution, CPU, RAM, Memory, GPU, Operation System, Weight, dan Price_Euro.

Pembagian Data

Pada tahap ini dilakukan proses pembagian data, data akan dibagi menjadi dua bagian, yaitu data training dan data testing.

Pada data training dilakukan proses untuk melatih dataset menjadi sebuah model machine learning. Model yang dibentuk akan mempelajari pola dan hubungan antar data dan melakukan penyesuaian parameter untuk melakukan prediksi yang tepat. Setelah dilakukan data training maka akan dilanjutkan pada proses data testing yang menguji kinerja dari model yang telah dibuat di data training yang tidak pernah dilihat sebelumnya, yang membantu dalam memahami seberapa baik model yang telah dibuat dapat menggeneralisasi dan melakukan prediksi yang tepat dan akurat pada data baru. Umumnya data training dan data testing dibagi menjadi 2 bagian dengan proporsi yang berbeda, 80% untuk data training dan 20% lainnya untuk data testing (Das, 2023).

K Means dan Validasi Model

Pada tahap ini, algoritma K-Means akan melakukan clustering untuk mengelompokkan beberapa fitur yang memiliki kesamaan. Serta nantinya akan dilakukan identifikasi pada dataset yang telah diproses sebelumnya (Sulistiyawati & Supriyanto, 2021). Dalam proses validasi model ini digunakan metode Elbow. Metode tersebut melibatkan plot jumlah kelompok (K) pada sumbu horizontal dan jumlah kesalahan (inertia) pada sumbu vertikal. Inertia ini merupakan jumlah total dari jarak antara titik data dengan pusat kelompok terdekat. Pada plot tersebut dapat dicari titik dimana penurunan inertia menurun secara signifikan. Titik tersebut menunjukkan jumlah kelompok yang optimal (Niu, 2021).

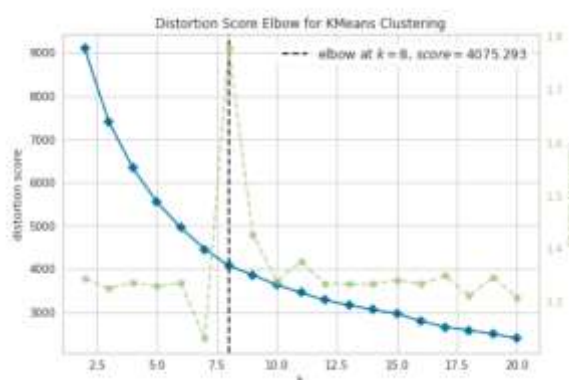
Evaluasi

Pada tahap evaluasi ini, terjadi proses penilaian dan pengukuran kualitas dari hasil clustering yang dihasilkan oleh algoritma K-Means. Tujuan utama dari evaluasi ini adalah untuk mengukur kemampuan algoritma K-Means dalam mengelompokkan data ke dalam kelompok yang berbeda. Serta menentukan hasil clustering tersebut sesuai dengan ekspektasi atau tujuan analisis (Setiawan, 2023).

Hasil

Berdasarkan proses dan alur yang telah dilakukan pada pembahasan diatas dapat dihasilkan beberapa output sebagai berikut:

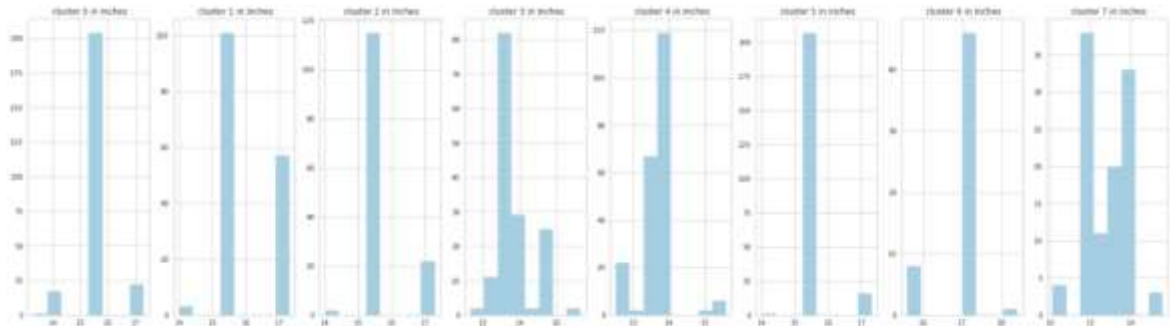
Gambar 3.2 Grafik Distortion Score Elbow



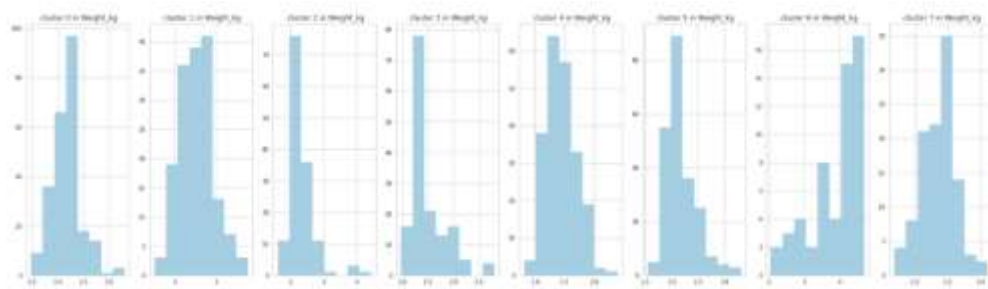
Gambar 4. Pembagian Kluster

Berdasarkan grafik yang tertera diatas, dapat diamati bahwa pembagian kluster yang terbaik yaitu sebanyak 8 kluster dan merupakan jumlah kluster yang paling optimum untuk digunakan. Berikut merupakan visualisasi dari masing-masing kluster dengan tiap kolomnya.

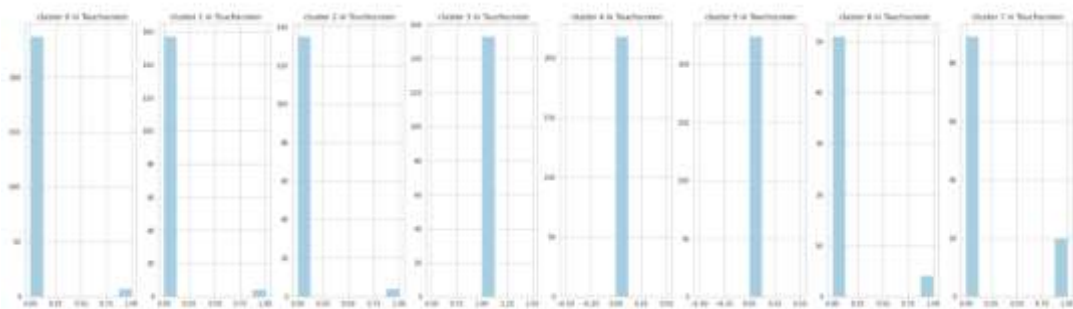
Gambar 5.3 – 1.11 Visualisasi Kluster



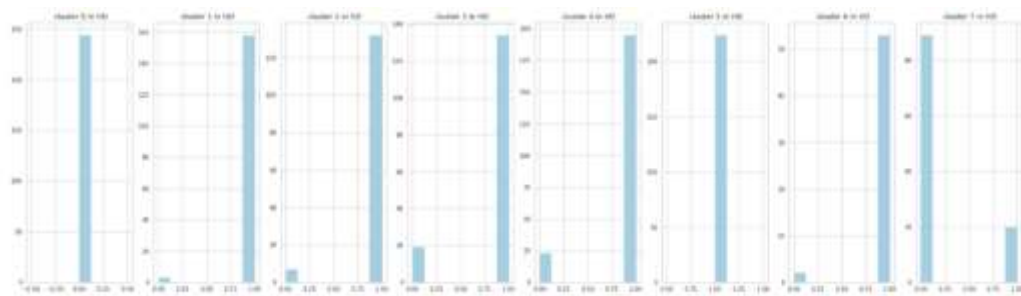
Gambar 6. Grafik Fitur inches



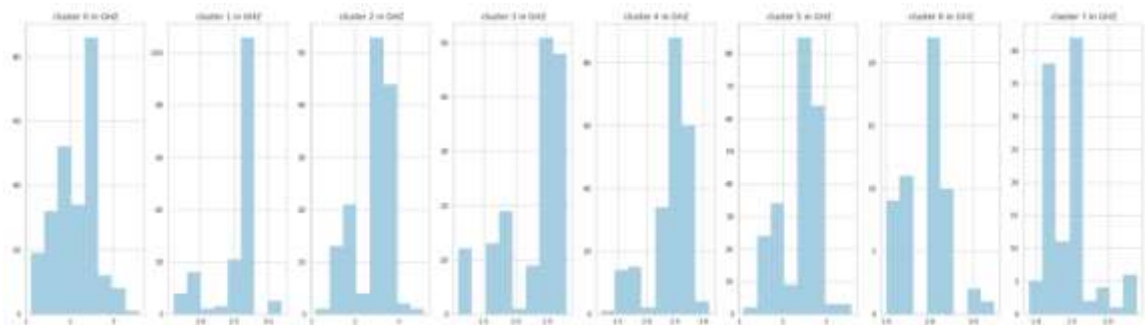
Gambar 7. Grafik Fitur Weight



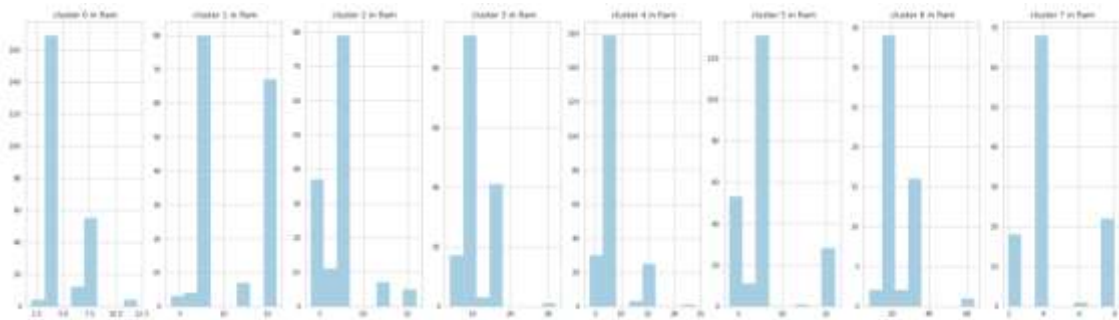
Gambar 8. Grafik Fitur Touchscreen



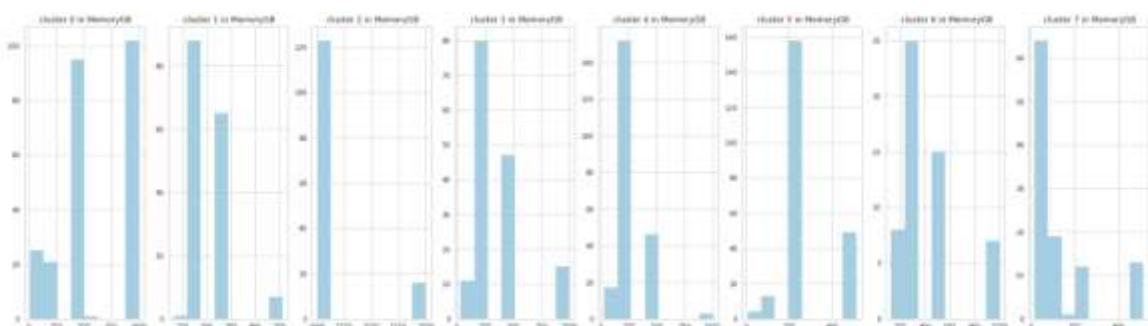
Gambar 9. Grafik Fitur HD



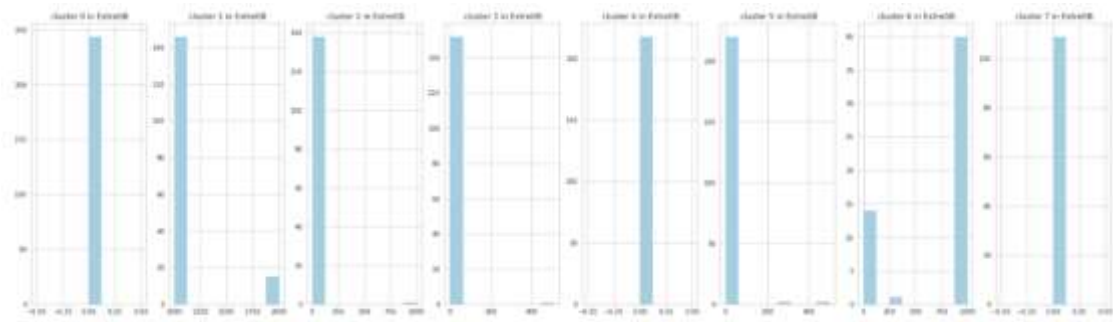
Gambar 10. Grafik Fitur GHZ



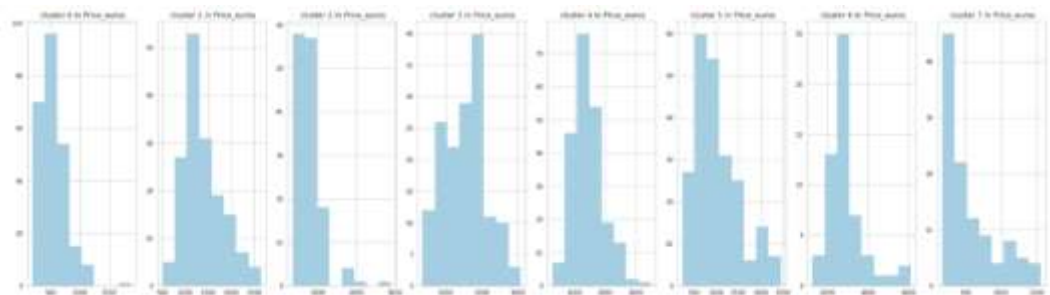
Gambar 11. Grafik Fitur RAM



Gambar 12. Grafik Fitur Memory



Gambar 13. Grafik Fitur ExtraGB



Gambar 14. Grafik Fitur Price

Berdasarkan grafik diatas, dapat disimpulkan bahwa:

1) *Klaster 0*

Terdiri dari laptop yang memiliki ukuran 15-16 inchi, berat 2-2,5 kg, touchscreen 0, HD pada angka 0. Spesifikasi GHZ pada angka 2-2,5 ghz, dengan RAM 4 dan 8, memory 500 GB dan 1 TB, ExtreGB pada angka 0, dan harga euro 500-800 euro.

2) *Klaster 1*

Terdiri dari laptop yang memiliki ukuran 15-16 dan 17 inchi, berat 2-3 kg, touchscreen 0, HD pada skala 0,75-1. Spesifikasi GHZ pada angka 2,5-2,75 ghz, dengan RAM 8 dan 16, memory 150 GB dan 250 GB, ExtreGB sebanyak 1000-1100, dan harga euro 1000-1500 euro.

3) *Klaster 2*

Terdiri dari laptop yang memiliki ukuran 15-16 inchi, berat 2-2,5 kg, touchscreen 0, HD pada skala 0,75-1. Spesifikasi GHZ pada angka 2,5-3 ghz, dengan RAM 4 sampai 8, memory 1 TB, ExtreGB sebanyak 0-125, dan harga euro 0-1500 euro.

4) *Klaster 3*

Terdiri dari laptop yang memiliki ukuran 12,5-14 inchi, berat 1-1,5 kg, touchscreen 1, HD pada skala 0,75-1. Spesifikasi GHZ pada angka 2-2,75 ghz, dengan RAM 8 dan 16, memory 250 GB, ExtreGB sebanyak 0-50, dan harga euro 1000-2000 euro.

5) *Klaster 4*

Terdiri dari laptop yang memiliki ukuran 13-14 inchi, berat 1-2 kg, touchscreen o, HD bernilai 1. Spesifikasi GHZ pada angka 2-2,75 ghz, dengan RAM 8, memory 250 GB, ExtreGB sebanyak o, dan harga euro 1000-1750 euro.

6) *Klaster 5*

Terdiri dari laptop yang memiliki ukuran 14 inchi, berat 2-2,125 kg, touchscreen o, HD bernilai 1. Spesifikasi GHZ pada angka 2,5-3 ghz, dengan RAM 4 dan 8, memory 250 GB, ExtreGB sebanyak 0-100, dan harga euro 500-1000 euro.

7) *Klaster 6*

Terdiri dari laptop yang memiliki ukuran 17 inchi, berat 4 kg, touchscreen o, HD pada skala 0,75-1. Spesifikasi GHZ pada angka 2,6-2,9 ghz, dengan RAM 16 dan 32, memory 150-400GB dan 500 GB, ExtreGB sebanyak 750-1000, dan harga euro 2000-3000 euro.

8) *Klaster 7*

Terdiri dari laptop yang memiliki ukuran 12-14 inchi, berat 1-1,5 kg, touchscreen o, HD bernilai o. Spesifikasi GHZ pada angka 1,25-1,5 ghz, dengan RAM 4, memory 150 GB, ExtreGB sebanyak o, dan harga euro 0-1000 euro.

Hasil dari klastering tersebut dapat digunakan sebagai dukungan keputusan untuk memberikan perbandingan antar laptop, agar konsumen dapat membeli laptop sesuai dengan spesifikasi yang dibutuhkan serta harga yang sesuai dengan dana yang tersedia.

Kesimpulan dan Saran

Dalam kasus ini dilakukan analisis untuk membandingkan jenis laptop yang bisa direkomendasikan kepada konsumen berdasarkan keperluan dan biaya yang tersedia. Hasil klastering memberikan perbandingan laptop berdasarkan spesifikasi dan harga. Klaster 2 cocok untuk yang butuh penyimpanan besar, performa baik, dan harga terjangkau. Klaster 3 cocok untuk portabilitas, layar touchscreen, dan performa baik dengan ukuran kecil. Klaster 4 cocok untuk laptop kompak, performa baik, dan harga terjangkau.

Daftar Pustaka

- Kamalia, A. Z., Pradana, E. A., & Surojudin, N. (2022). Penerapan algoritma K-Means dalam klasterisasi penjualan laptop. *Jurnal SIGMA*, 13(3), 133-138.
- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics (Switzerland)*, 9(8), 1–12. <https://doi.org/10.3390/electronics9081295>
- Das, D., Kayal, P., & Maiti, M. (2023). A K-means clustering model for analyzing the Bitcoin extreme value returns. *Decision Analytics Journal*, 6(December 2022), 100152. <https://doi.org/10.1016/j.dajour.2022.100152>
- Liu, G., Ji, F., Sun, W., & Sun, L. (2023). Optimization design of short-circuit test

- platform for the distribution network of integrated power system based on improved K-means clustering. *Energy Reports*, 9, 716–726.
<https://doi.org/10.1016/j.egyr.2023.04.319>
- Liu, W., Zou, P., Jiang, D., Quan, X., & Dai, H. (2022). Zoning of reservoir water temperature field based on K-means clustering algorithm. *Journal of Hydrology: Regional Studies*, 44(October), 101239. <https://doi.org/10.1016/j.ejrh.2022.101239>
- Niu, G., Ji, Y., Zhang, Z., Wang, W., Chen, J., & Yu, P. (2021). Clustering analysis of typical scenarios of island power supply system by using cohesive hierarchical clustering based K-Means clustering method. *Energy Reports*, 7, 250–256.
<https://doi.org/10.1016/j.egyr.2021.08.049>
- Setiawan, K. E., Kurniawan, A., Chowanda, A., & Suhartono, D. (2023). Clustering models for hospitals in Jakarta using fuzzy c-means and k-means. *Procedia Computer Science*, 216(2022), 356–363. <https://doi.org/10.1016/j.procs.2022.12.146>
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi algoritma K-means clustering dalam penentuan siswa kelas unggulan. *Jurnal Tekno Kompak*, 15(2), 25.
<https://doi.org/10.33365/jtk.v15i2.1162>
- Ulil, P., Aulia, F., & Saepudin, S. (2021). Penerapan data mining K-Means clustering untuk mengelompokkan berbagai jenis merk laptop. *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, 241–249.
- Wijaya, A. E., & Alfian, D. (2018). Sistem rekomendasi laptop menggunakan collaborative filtering dan content-based filtering. *Jurnal Computech & Bisnis*, 12(1), 11–27.