**JUTIS (Jurnal Teknik Informatika Unis)**
**Vol. 14 No. 1 Bulan April 2026** **p-ISSN: 2252-5351**
Program Studi Teknik Informatika – Fakultas Teknik **e-ISSN: 2656-0860**
Universitas Islam Syekh-Yusuf Tangerang

# Implementasi Algoritma K-Means Clustering pada Sistem Rekomendasi Produk Berbasis Python

**Muh. Ali Zainal Abidin[1,*], Suprianto[2]**
[1,2] Fakultas Sains dan Teknologi, prodi Informatika , Universitas Muhammdiyah Sidoarjo, Indonesia
[*1]aliabidin982@gmail.com, [2] suprianto@umsida.ac.id

*E-mail koresponden : aliabidin982@gmailcom

**Abstrak**

*Pertumbuhan pesat platform perdagangan elektronik menyebabkan meningkatnya jumlah produk yang tersedia, sehingga menimbulkan permasalahan information overload bagi konsumen dalam menentukan produk yang sesuai dengan kebutuhan dan preferensi mereka. Kondisi ini mendorong perlunya sistem rekomendasi yang mampu membantu pengguna dalam proses pengambilan keputusan pembelian secara efektif. Namun, sebagian besar sistem rekomendasi masih menghadapi kendala cold-start, yaitu keterbatasan data histori interaksi pengguna. Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means Clustering pada sistem rekomendasi produk berbasis Python sebagai solusi rekomendasi tanpa bergantung pada data pengguna. Metode yang digunakan dalam penelitian ini adalah pendekatan kuantitatif dengan memanfaatkan data produk yang terdiri dari atribut harga, penilaian, dan persediaan. Data terlebih dahulu melalui tahap pre-processing menggunakan metode Min-Max Scaling untuk menyamakan skala nilai setiap atribut. Selanjutnya, algoritma K-Means Clustering diterapkan dengan jumlah cluster sebanyak tiga kelompok untuk mengelompokkan produk berdasarkan kemiripan karakteristik menggunakan jarak Euclidean. Hasil clustering kemudian dimanfaatkan sebagai dasar sistem rekomendasi berbasis item-based similarity, di mana produk yang berada dalam cluster yang sama direkomendasikan kepada pengguna. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu menghasilkan pengelompokan produk yang jelas dan mudah diinterpretasikan, yaitu cluster produk premium, standar, dan ekonomis. Sistem rekomendasi yang dikembangkan mampu menampilkan rekomendasi produk yang relevan berdasarkan hasil clustering tanpa memerlukan data histori pengguna. Dengan demikian, sistem rekomendasi berbasis clustering ini dapat menjadi solusi yang sederhana, efisien, dan praktis untuk diterapkan pada sistem perdagangan elektronik berskala kecil hingga menengah.*

***Kata kunci:*** *K-Means Clustering, Python, rekomendasi produk, sistem rekomendasi, unsupervised learning.*

## A. Pendahuluan

Perkembangan teknologi informasi dan pertumbuhan pesat platform digital telah mendorong meningkatnya jumlah produk yang tersedia di berbagai sistem perdagangan elektronik. Kondisi ini memberikan keuntungan bagi konsumen dalam hal variasi pilihan, namun di sisi lain menimbulkan permasalahan baru berupa kesulitan dalam menentukan produk yang paling sesuai dengan kebutuhan dan preferensi pengguna. Fenomena information overload ini menyebabkan konsumen membutuhkan bantuan sistem cerdas yang mampu menyaring dan merekomendasikan produk secara efektif. Oleh karena itu, sistem rekomendasi menjadi salah satu komponen penting dalam mendukung pengambilan keputusan pembelian pada platform digital modern. Beberapa penelitian menunjukkan bahwa penerapan sistem rekomendasi yang tepat dapat meningkatkan

kepuasan pengguna, mempercepat proses pencarian produk, serta mendorong peningkatan penjualan secara signifikan Dalam praktiknya, penerapan sistem rekomendasi berbasis clustering tidak hanya berperan sebagai alat bantu pencarian produk, tetapi juga sebagai mekanisme pendukung strategi bisnis berbasis data. Dengan mengelompokkan produk ke dalam segmen-segmen tertentu, pelaku usaha dapat memperoleh pemahaman yang lebih baik mengenai pola karakteristik produk yang diminati oleh konsumen. Informasi ini dapat dimanfaatkan untuk pengambilan keputusan strategis, seperti pengelolaan stok, penentuan harga, serta penyusunan promosi yang lebih terarah. Oleh karena itu, sistem rekomendasi tidak hanya memberikan manfaat dari sisi pengguna akhir, tetapi juga memberikan nilai tambah dari sisi pengelolaan sistem dan manajemen produk secara keseluruhan. (Al-Hagery, 2020)

Sistem rekomendasi secara umum dikembangkan menggunakan dua pendekatan utama, yaitu collaborative filtering dan content-based filtering. Namun, kedua pendekatan tersebut memiliki keterbatasan, terutama ketika sistem dihadapkan pada kondisi cold-start, yaitu ketika data histori pengguna atau interaksi masih terbatas. Untuk mengatasi permasalahan tersebut, pendekatan berbasis unsupervised learning seperti clustering menjadi alternatif yang relevan karena tidak bergantung pada data pelatihan berlabel maupun riwayat interaksi pengguna. Algoritma clustering memungkinkan sistem untuk mengelompokkan item berdasarkan kemiripan karakteristik intrinsik, sehingga rekomendasi dapat diberikan meskipun tanpa informasi perilaku pengguna sebelumnya. Penelitian oleh Prakash et al. (2025) menunjukkan bahwa sistem rekomendasi berbasis clustering mampu memberikan solusi yang efektif dalam kondisi cold-start dengan mengandalkan karakteristik produk sebagai dasar pengelompokan. (Khatwal & Hasan, 2022)

Pendekatan rekomendasi berbasis item-based clustering memiliki keunggulan dalam hal kesederhanaan alur kerja dan interpretasi hasil. Berbeda dengan sistem rekomendasi yang bergantung pada profil pengguna, pendekatan ini memfokuskan rekomendasi pada kemiripan antar produk berdasarkan atribut yang terukur. Dengan demikian, sistem dapat langsung memberikan rekomendasi yang relevan ketika pengguna memilih sebuah produk tertentu. Pendekatan ini dinilai efektif untuk sistem yang belum memiliki basis pengguna yang besar atau belum menyimpan data histori interaksi yang memadai. Sejumlah penelitian menunjukkan bahwa item-based recommendation berbasis clustering mampu menghasilkan rekomendasi yang stabil dan mudah diadaptasi pada berbagai skenario implementasi (Jeon & Lee, 2023)

Pemilihan atribut produk seperti harga, penilaian, dan persediaan sebagai dasar pengelompokan didasarkan pada pertimbangan bahwa ketiga atribut tersebut merupakan indikator utama dalam pengambilan keputusan pembelian konsumen. Harga mencerminkan tingkat keterjangkauan produk, penilaian merepresentasikan persepsi kualitas berdasarkan pengalaman pengguna lain, sedangkan persediaan menggambarkan ketersediaan produk di sistem. Kombinasi ketiga atribut ini memungkinkan sistem untuk membentuk cluster yang tidak hanya mencerminkan kesamaan teknis antar produk, tetapi juga kesamaan nilai dari sudut pandang konsumen. Pendekatan serupa telah diterapkan

pada beberapa penelitian sebelumnya dan terbukti mampu menghasilkan segmentasi produk yang relevan serta mudah dipahami (Xu, 2025)

Selain aspek algoritmik, implementasi sistem rekomendasi juga perlu mempertimbangkan kemudahan penggunaan dan transparansi hasil. Penyajian hasil rekomendasi dalam bentuk tabel yang menampilkan produk-produk dengan karakteristik serupa memungkinkan pengguna untuk memahami alasan di balik rekomendasi yang diberikan oleh sistem. Transparansi ini penting untuk meningkatkan kepercayaan pengguna terhadap sistem rekomendasi, khususnya pada sistem yang tidak menggunakan data personal pengguna. Dengan menampilkan rekomendasi berdasarkan cluster yang sama, sistem memberikan gambaran bahwa rekomendasi didasarkan pada kesamaan atribut produk, bukan pada asumsi subjektif atau preferensi tersembunyi. (Legito et al., 2023)

Salah satu algoritma unsupervised learning yang paling banyak digunakan dalam sistem rekomendasi adalah K-Means Clustering. Algoritma ini bekerja dengan mengelompokkan data ke dalam sejumlah cluster berdasarkan kedekatan jarak antar data terhadap pusat cluster (centroid), yang umumnya dihitung menggunakan jarak Euclidean. K-Means dikenal memiliki keunggulan dalam hal kesederhanaan, efisiensi komputasi, serta kemudahan implementasi, sehingga cocok diterapkan pada dataset kecil hingga menengah. Berbagai penelitian terdahulu (Putri & Leu, 2019) telah membuktikan efektivitas K-Means dalam sistem rekomendasi di berbagai domain, seperti e-commerce, musik, film, dan marketplace digital (Ye & Rong, 2022; Kumar et al., 2024; Rachman & Voutama, 2024). Selain itu, studi komparatif yang dilakukan oleh Pily dan Rio (2025) menunjukkan bahwa K-Means dengan jarak Euclidean memberikan hasil segmentasi yang cukup stabil dan mudah diinterpretasikan dalam konteks rekomendasi produk.

Meskipun K-Means memiliki keunggulan, proses pre-processing data menjadi tahapan krusial untuk memastikan kualitas hasil pengelompokan. Dalam konteks data produk, perbedaan skala antar variabel seperti harga, penilaian, dan persediaan dapat menyebabkan bias dalam perhitungan jarak jika tidak dilakukan normalisasi. Oleh karena itu, metode Min-Max Scaling (Muhammed, 2023) sering digunakan untuk menyamakan rentang nilai setiap variabel ke dalam interval yang seragam, umumnya antara 0 hingga 1. Pendekatan ini bertujuan agar setiap fitur memiliki kontribusi yang seimbang dalam proses clustering. Beberapa penelitian terkini menegaskan bahwa normalisasi data berpengaruh signifikan terhadap akurasi dan konsistensi hasil K-Means, terutama ketika fitur yang digunakan memiliki skala yang sangat berbeda (Zarzour et al., 2022; Miranda & Sriani, 2025). Dengan pre-processing yang tepat, hasil cluster yang terbentuk dapat merepresentasikan karakteristik produk secara lebih akurat.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means Clustering pada sistem rekomendasi produk berbasis Python dengan memanfaatkan fitur harga, penilaian, dan persediaan sebagai dasar pengelompokan. Sistem yang dikembangkan dirancang untuk memberikan rekomendasi produk secara item-based, di mana produk yang dipilih oleh pengguna akan menjadi acuan untuk menampilkan produk lain yang berada dalam cluster yang sama. Pendekatan ini

memungkinkan (Andra & Baizal, 2022) sistem memberikan rekomendasi tanpa memerlukan data histori pengguna, sehingga cocok diterapkan pada sistem dengan keterbatasan data interaksi. Kontribusi penelitian ini terletak pada implementasi sistem rekomendasi berbasis clustering yang sederhana namun fungsional, serta penyajian hasil rekomendasi dalam bentuk tabel yang mudah dipahami oleh pengguna. Diharapkan hasil penelitian ini dapat menjadi referensi bagi pengembangan sistem rekomendasi produk berbasis unsupervised learning, khususnya pada skala usaha kecil hingga menengah yang membutuhkan solusi praktis dan efisien.

## B. Metode

### Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian terapan dengan pendekatan kuantitatif yang bertujuan untuk mengimplementasikan algoritma K-Means Clustering pada sistem rekomendasi produk. Pendekatan kuantitatif digunakan karena penelitian ini melibatkan pengolahan data numerik berupa atribut produk yang dianalisis menggunakan metode statistik dan algoritma pembelajaran mesin. Penelitian ini berfokus pada pemanfaatan data produk sebagai dasar pengelompokan dan pemberian rekomendasi tanpa melibatkan data histori interaksi pengguna.

### Dataset dan Variabel Penelitian

Dataset yang digunakan dalam penelitian ini berupa data produk yang terdiri dari sejumlah item dengan beberapa atribut utama. Data yang digunakan bersifat data sekunder/simulasi yang disusun untuk merepresentasikan karakteristik produk pada sistem perdagangan elektronik. Jumlah data produk yang digunakan sebanyak *n* data, sebagaimana ditunjukkan pada Tabel 1.

| Produk | Harga | Penilaian | Persediaan |
|---|---|---|---|
| Produk A | 15000 | 4.3 | 120 |
| Produk B | 25000 | 4.8 | 80 |
| Produk C | 9000 | 3.9 | 200 |
| Produk D | 35000 | 4.7 | 60 |
| Produk E | 6000 | 3.5 | 300 |
| Produk F | 45000 | 4.9 | 50 |
| Produk G | 12000 | 4.1 | 140 |
| Produk H | 20000 | 4.6 | 90 |
| Produk I | 10000 | 3.8 | 220 |
| Produk J | 30000 | 4.5 | 70 |
| Produk K | 8000 | 3.6 | 250 |
| Produk L | 40000 | 4.9 | 55 |

**Tabel 1.** Data Set Produk

Variabel yang digunakan dalam penelitian ini meliputi tiga atribut utama, yaitu harga produk, penilaian produk, dan persediaan. Variabel harga merepresentasikan tingkat keterjangkauan produk dalam satuan rupiah, variabel penilaian merepresentasikan tingkat kualitas produk berdasarkan skala penilaian pengguna, sedangkan variabel

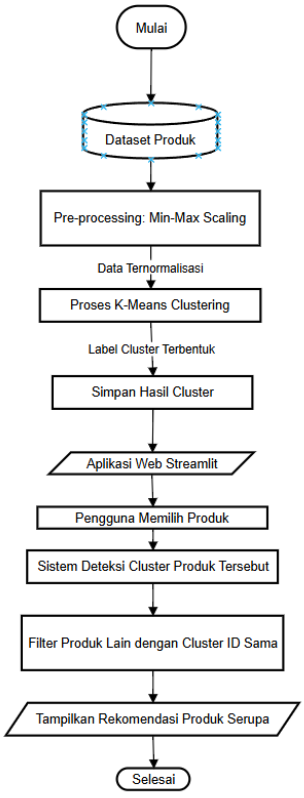
**p-ISSN: 2252-5351**
**e-ISSN: 2656-0860**

persediaan menunjukkan jumlah stok produk yang tersedia. Ketiga variabel tersebut digunakan sebagai dasar dalam proses pengelompokan produk menggunakan algoritma K-Means.

**Tahapan Penelitian**

Tahapan penelitian dalam pengembangan sistem rekomendasi produk ini terdiri dari beberapa langkah utama, yaitu pengumpulan data, *pre-processing* data, proses clustering menggunakan algoritma K-Means, pemberian label cluster, serta penyajian rekomendasi produk berbasis cluster. Alur tahapan penelitian secara keseluruhan ditunjukkan pada Gambar 1.

1. Pengumpulan data
2. Pre-processing (Min-Max Scaling)
3. Clustering dengan K-Means
4. Pemberian label cluster
5. Sistem rekomendasi berbasis cluster




**Gambar 1.** Flowchart Alur Sistem

Pada tahap *pre-processing*, data produk dinormalisasi menggunakan metode Min-Max Scaling untuk menyamakan rentang nilai setiap variabel ke dalam interval 0 hingga 1. Tahapan ini bertujuan untuk menghindari bias perhitungan jarak akibat perbedaan skala antar variabel. Selanjutnya, data yang telah dinormalisasi diproses menggunakan algoritma K-Means untuk membentuk kelompok produk berdasarkan tingkat kemiripan

karakteristik.

**Metode Clustering Menggunakan K-Means**

Algoritma K-Means Clustering digunakan sebagai metode utama dalam pengelompokan produk pada penelitian ini. K-Means merupakan metode *unsupervised learning* yang mengelompokkan data berdasarkan kedekatan jarak antar data terhadap pusat cluster (*centroid*). Jarak kemiripan antar produk dihitung menggunakan jarak Euclidean.

Dalam penelitian ini, jumlah cluster ditentukan sebanyak tiga cluster (k = 3), yang merepresentasikan kelompok produk dengan karakteristik harga dan penilaian yang berbeda. Proses clustering dimulai dengan inisialisasi centroid secara acak, dilanjutkan dengan perhitungan jarak setiap data terhadap centroid, pengelompokan data ke cluster terdekat, serta pembaruan posisi centroid hingga proses mencapai kondisi konvergen. Hasil clustering berupa label cluster untuk setiap produk yang selanjutnya digunakan sebagai dasar sistem rekomendasi.

**Metode Rekomendasi Produk**

Metode rekomendasi produk pada penelitian ini menggunakan pendekatan *item-based similarity* berbasis hasil clustering. Produk yang berada dalam cluster yang sama dianggap memiliki karakteristik yang serupa. Ketika pengguna memilih satu produk, sistem akan menampilkan produk lain yang memiliki label cluster yang sama sebagai rekomendasi.

Pendekatan ini memungkinkan sistem memberikan rekomendasi tanpa memerlukan data histori pengguna, sehingga sesuai untuk sistem dengan keterbatasan data interaksi. Rekomendasi disajikan dalam bentuk tabel yang menampilkan daftar produk dengan karakteristik serupa berdasarkan hasil pengelompokan.

**Tools dan Lingkungan Implementasi**

Sistem rekomendasi produk dikembangkan menggunakan bahasa pemrograman Python. Library Pandas digunakan untuk pengolahan dan manipulasi data, sedangkan Scikit-learn digunakan untuk implementasi algoritma K-Means Clustering dan Min-Max Scaling. Antarmuka sistem dibangun menggunakan framework Streamlit sehingga sistem dapat diakses melalui web dan menampilkan hasil rekomendasi secara interaktif.

## C. Hasil dan Pembahasan

**Hasil Pengelompokan Produk**

Hasil penerapan algoritma K-Means Clustering terhadap dataset produk menghasilkan tiga kelompok (*cluster*) utama yang terbentuk berdasarkan atribut harga, penilaian, dan persediaan. Sebelum proses clustering dilakukan, seluruh data terlebih dahulu melalui tahap *pre-processing* menggunakan metode Min-Max Scaling untuk menyamakan rentang nilai setiap variabel ke dalam interval 0 hingga 1. Tahapan ini

**p-ISSN: 2252-5351**
**e-ISSN: 2656-0860**

bertujuan untuk memastikan bahwa perbedaan skala antar variabel tidak memengaruhi perhitungan jarak Euclidean pada proses clustering.

**Data Produk**

| | Produk | Harga | Penilaian | Persediaan | Cluster |
|---|---|---|---|---|---|
| 0 | Produk A | 15000 | 4.300000 | 120 | 2 |
| 1 | Produk B | 25000 | 4.800000 | 80 | 1 |
| 2 | Produk C | 9000 | 3.900000 | 200 | 0 |
| 3 | Produk D | 35000 | 4.700000 | 60 | 1 |
| 4 | Produk E | 6000 | 3.500000 | 300 | 0 |
| 5 | Produk F | 45000 | 4.900000 | 50 | 1 |
| 6 | Produk G | 12000 | 4.100000 | 140 | 2 |
| 7 | Produk H | 20000 | 4.600000 | 90 | 2 |
| 8 | Produk I | 10000 | 3.800000 | 220 | 0 |
| 9 | Produk J | 30000 | 4.500000 | 70 | 1 |

**Gambar 2. Hasil Proses Clustering**

Berdasarkan hasil pengolahan data, setiap produk berhasil dikelompokkan ke dalam salah satu dari tiga cluster yang terbentuk. Visualisasi hasil clustering ditunjukkan pada **Gambar 2**, yang memperlihatkan sebaran produk berdasarkan atribut harga dan penilaian dengan warna yang merepresentasikan masing-masing cluster. Visualisasi tersebut menunjukkan bahwa produk-produk dalam satu cluster memiliki kedekatan karakteristik yang relatif tinggi dibandingkan dengan produk pada cluster lainnya. Hal ini mengindikasikan bahwa algoritma K-Means mampu mengelompokkan produk secara efektif berdasarkan kemiripan atribut numerik yang digunakan.

## Karakteristik Setiap Cluster Produk

Hasil clustering selanjutnya dianalisis untuk mengidentifikasi karakteristik masing-masing cluster. Analisis ini bertujuan untuk memahami pola segmentasi produk yang terbentuk sebagai dasar sistem rekomendasi. Ringkasan karakteristik setiap cluster ditunjukkan pada **Tabel 2**.

**Tabel 2** Karakteristik Cluster

| Produk | Harga | Penilaian | Persediaan | Cluster |
|---|---|---|---|---|
| Produk A | 15000 | 4.3 | 120 | 2 |
| Produk B | 25000 | 4.8 | 80 | 1 |
| Produk C | 9000 | 3.9 | 200 | 0 |
| Produk D | 35000 | 4.7 | 60 | 1 |
| Produk E | 6000 | 3.5 | 300 | 0 |
| Produk F | 45000 | 4.9 | 50 | 1 |
| Produk G | 12000 | 4.1 | 140 | 2 |
| Produk H | 20000 | 4.6 | 90 | 2 |
| Produk I | 10000 | 3.8 | 220 | 0 |
| Produk J | 30000 | 4.5 | 70 | 1 |

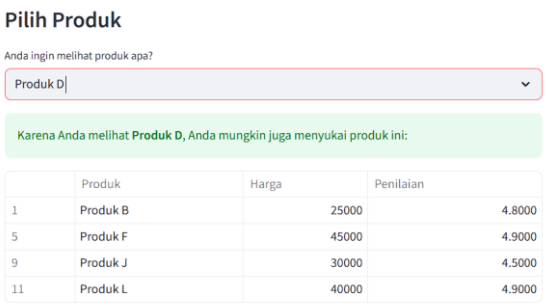
| | | | | |
|---|---|---|---|---|
| Produk K | 8000 | 3.6 | 250 | 0 |
| Produk L | 40000 | 4.9 | 55 | 1 |

Cluster pertama (Cluster 1) merepresentasikan kelompok produk dengan harga relatif tinggi dan penilaian pengguna yang sangat baik. Produk-produk pada cluster ini memiliki rentang harga antara Rp30.000 hingga Rp45.000 dengan nilai penilaian berkisar antara 4,5 hingga 4,9. Karakteristik tersebut menunjukkan bahwa cluster ini dapat dikategorikan sebagai **produk premium**, yang umumnya ditujukan bagi konsumen yang mengutamakan kualitas dibandingkan harga. Cluster kedua (Cluster 0) terdiri dari produk-produk dengan harga paling terjangkau dan tingkat persediaan yang relatif tinggi. Produk dalam cluster ini memiliki harga di bawah Rp10.000 dengan jumlah stok antara 200 hingga 300 unit. Berdasarkan karakteristik tersebut, cluster ini dapat diklasifikasikan sebagai **produk ekonomis**, yang cenderung menarik bagi konsumen dengan sensitivitas harga yang tinggi. Cluster ketiga (Cluster 2) berada pada posisi tengah antara cluster premium dan ekonomis. Produk-produk pada cluster ini memiliki harga berkisar antara Rp12.000 hingga Rp20.000 dengan penilaian yang cukup baik, yaitu antara 4,1 hingga 4,6. Cluster ini dikategorikan sebagai **produk standar**, yang menawarkan keseimbangan antara harga dan kualitas.

Hasil analisis ini menunjukkan bahwa pengelompokan produk menggunakan K-Means mampu menghasilkan segmentasi yang logis dan mudah diinterpretasikan, sehingga dapat dimanfaatkan secara langsung sebagai dasar sistem rekomendasi produk.

### Hasil Sistem Rekomendasi Produk Berbasis Cluster

Berdasarkan hasil pengelompokan produk, sistem rekomendasi dikembangkan menggunakan pendekatan *item-based similarity*, di mana produk yang berada dalam cluster yang sama dianggap memiliki karakteristik serupa. Ketika pengguna memilih satu produk, sistem secara otomatis memfilter data untuk menampilkan produk lain yang memiliki label cluster yang sama.




**Gambar 2.** Sistem Rekomendasi

Sebagai ilustrasi, ketika pengguna memilih Produk D yang termasuk dalam Cluster 1, sistem menampilkan daftar produk lain dalam cluster yang sama, seperti Produk B, F, J, dan L. Hasil rekomendasi ini ditampilkan dalam bentuk tabel rekomendasi sebagaimana ditunjukkan pada **Tabel 3**. Tabel tersebut memuat informasi produk yang memiliki karakteristik harga dan penilaian yang serupa dengan produk yang dipilih pengguna. Dengan demikian, sistem mampu memberikan rekomendasi produk premium lainnya secara cepat dan relevan tanpa memerlukan data histori interaksi pengguna. Pendekatan ini menunjukkan bahwa sistem rekomendasi berbasis clustering dapat berfungsi secara efektif meskipun tanpa memanfaatkan preferensi pengguna secara eksplisit. Hal ini menjadi keunggulan utama sistem dalam kondisi keterbatasan data interaksi atau pada sistem yang masih berada pada tahap awal pengembangan.

**Pembahasan**

Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means Clustering pada sistem rekomendasi produk mampu menghasilkan pengelompokan produk yang konsisten dan mudah diinterpretasikan. Segmentasi produk ke dalam kategori premium, standar, dan ekonomis memberikan gambaran yang jelas mengenai posisi setiap produk berdasarkan atribut harga, penilaian, dan persediaan. Temuan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa K-Means efektif digunakan untuk segmentasi produk pada sistem rekomendasi berbasis atribut numerik. Selain itu, pemanfaatan hasil clustering sebagai dasar rekomendasi produk berbasis *item-based similarity* terbukti mampu memberikan rekomendasi yang relevan bagi pengguna. Sistem tidak hanya membantu pengguna dalam menemukan produk dengan karakteristik serupa, tetapi juga meningkatkan efisiensi proses pencarian produk. Pendekatan ini juga memiliki keunggulan dalam hal kesederhanaan implementasi dan transparansi hasil, karena rekomendasi diberikan berdasarkan kesamaan atribut produk yang dapat dipahami oleh pengguna. Namun demikian, penelitian ini masih memiliki keterbatasan, terutama pada jumlah data produk yang digunakan dan belum diterapkannya metode evaluasi kuantitatif untuk mengukur kualitas cluster secara numerik. Meskipun demikian, hasil penelitian ini menunjukkan bahwa sistem rekomendasi berbasis clustering dapat menjadi solusi praktis dan efisien untuk sistem perdagangan elektronik berskala kecil hingga menengah yang memiliki keterbatasan data pengguna.

## D. Kesimpulan

Penelitian ini berhasil mengimplementasikan algoritma K-Means Clustering pada sistem rekomendasi produk berbasis Python dengan memanfaatkan atribut harga, penilaian, dan persediaan sebagai dasar pengelompokan. Berdasarkan hasil pengolahan data, produk-produk dapat dikelompokkan ke dalam beberapa cluster yang memiliki karakteristik yang jelas dan mudah diinterpretasikan. Hal ini menunjukkan bahwa algoritma K-Means mampu melakukan segmentasi produk secara efektif meskipun

tanpa menggunakan data histori interaksi pengguna. Hasil penelitian juga menunjukkan bahwa pemanfaatan cluster sebagai dasar sistem rekomendasi berbasis *item-based similarity* dapat memberikan rekomendasi produk yang relevan. Ketika pengguna memilih suatu produk, sistem mampu menampilkan produk lain yang berada dalam cluster yang sama sehingga memiliki karakteristik yang serupa. Pendekatan ini membuktikan bahwa sistem rekomendasi dapat dibangun menggunakan metode unsupervised learning dan tetap berfungsi secara optimal pada kondisi keterbatasan data pengguna, khususnya dalam mengatasi permasalahan *cold-start*.

Dengan demikian, tujuan penelitian telah tercapai, yaitu mengembangkan sistem rekomendasi produk berbasis clustering yang sederhana, efisien, dan dapat diimplementasikan pada skala usaha kecil hingga menengah. Sistem yang dihasilkan tidak hanya membantu pengguna dalam proses pengambilan keputusan, tetapi juga memberikan solusi praktis bagi pengelola sistem perdagangan elektronik yang memiliki keterbatasan data dan sumber daya.

**Saran**

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat menjadi arah pengembangan penelitian selanjutnya. Pertama, penelitian berikutnya disarankan untuk menggunakan jumlah data produk yang lebih besar dan lebih beragam agar hasil clustering dapat merepresentasikan kondisi pasar secara lebih komprehensif. Penambahan variabel lain, seperti kategori produk, jumlah penjualan, atau ulasan pengguna, juga dapat meningkatkan kualitas segmentasi produk. Kedua, evaluasi kualitas clustering secara kuantitatif dapat dikembangkan dengan menerapkan metode pengukuran seperti *Silhouette Score*, *Davies-Bouldin Index*, atau *Elbow Method* untuk menentukan jumlah cluster yang optimal. Selain itu, sistem rekomendasi dapat dikombinasikan dengan metode lain, seperti collaborative filtering atau pendekatan hybrid, guna meningkatkan tingkat personalisasi rekomendasi sesuai dengan preferensi pengguna.

Terakhir, pengembangan antarmuka sistem dan pengujian langsung terhadap pengguna (*user testing*) dapat dilakukan untuk menilai tingkat kepuasan dan kegunaan sistem secara lebih objektif. Dengan pengembangan tersebut, diharapkan sistem rekomendasi berbasis clustering dapat memberikan kontribusi yang lebih luas dalam pengembangan aplikasi e-commerce dan sistem pendukung keputusan berbasis data.

**Daftar Pustaka**

Al-Hagery, M. A. (2020). A novel Based-Approach Composed of Clustering Algorithm & Cosine Similarity for Products Recommendation. *International Journal of Education*, *14*, 133–141. https://doi.org/10.46300/9109.2020.14.16

Khatwal, R., & Hasan, S. N. (2022). *Cold Start Problem in Recommendation System: A Solution Model Based on Clustering and Association Rule Techniques*. 1–8. https://doi.org/10.1109/IMPACT55510.2022.10029293

Jeon, I.-J., & Lee, H.-S. (2023). Clustering-based Product Recommendation Model Reflecting Customer Response History and Product Information. *Journal of the Korean Operations Research and Management Science Society*. https://doi.org/10.7737/jkorms.2023.48.3.029

Xu, L. (2025). A Clustering Method for Product Cannibalization Detection Using Price Effect. *Electronics*, *14*(15), 3120–3120. https://doi.org/10.3390/electronics14153120

Legito, Yoanti, F., Wattimena, & Rofi', Y. U. (2023). E-Commerce Product Recommendation System Using Case-Based Reasoning (CBR) and K-Means Clustering. *International Journal Software Engineering and Computer Science*. https://doi.org/10.35870/ijsecs.v3i2.1527

Putri, D. C. G., & Leu, J.-S. (2019). *Towards the Implementation of Movie Recommender System by Using Unsupervised Machine Learning Schemes*. 99–113. https://doi.org/10.1007/978-3-030-52988-8_9

Muhammed, L. A. (2023). Role of data normalization in k-means algorithm results. *Nucleation and Atmospheric Aerosols*. https://doi.org/10.1063/5.0119267

Andra, D., & Baizal, A. (2022). E-commerce Recommender System Using PCA and K-Means Clustering. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *6*(1), 57–63. https://doi.org/10.29207/resti.v6i1.3782

Prakash, K. L., Pavithra, E., Nagalakshmi, N., Reddy, V. S. N., Annapoorna, S., & Chitty, A. (2025, February). Cluster based Personalized Recommender System to address Cold-Start Challenge. In *2025 1st International Conference on Secure IoT, Assured and Trusted Computing (SATC)* (pp. 1-6). IEEE.
https://ieeexplore.ieee.org/abstract/document/11137121

Zhang, W., & Wu, Z. (2024). E-commerce recommender system based on improved K-means commodity information management model. *Heliyon*, *10*(9).
https://www.cell.com/heliyon/fulltext/S2405-8440(24)05076-X

Zuo, F., Siniauski, U., Yang, H., & Wang, G. (2022). OP-K-Means: Optimized algorithm for recommendation system based on user preferences. In *Journal of Physics: Conference Series* (Vol. 2171, No. 1, p. 012002). IOP Publishing.
https://iopscience.iop.org/article/10.1088/1742-6596/2171/1/012002/meta

Pily, A. K. E., & Rio, U. (2025). Komparasi K-Means Clustering dengan Euclidean dan Cosine Similarity untuk Segmentasi dan Rekomendasi Produk pada Data E-Commerce. *The Indonesian Journal of Computer Science*, *14*(2).
http://ai-soc.org/ijcs/index.php/ijcs/article/view/4713

MARDIANA, A., & Budiman, B. (2022, December). Implementasi Metode K-Means Untuk

Clustering Produk Yang Kurang Diminati Berdasarkan Data Penjualan. In *Seminar Teknologi Majalengka (STIMA)* (Vol. 6, pp. 198-210).

Zarzour, H., Maazouzi, F., Al-Zinati, M., Nusayr, A., Alsmirat, M., Al-Ayyoub, M., & Jararweh, Y. (2022, September). Using k-means clustering ensemble to improve the performance in recommender systems. In *2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)* (pp. 176-180). IEEE. https://ieeexplore.ieee.org/abstract/document/9923070

Miranda, M., & Sriani, S. (2025). Implementation of K-Means Clustering in Grouping Sales Data at Zura Mart. *Journal of Applied Informatics and Computing*, *9*(2), 547-555.

Al Rasyid, H., Soebari, B. F. K., & Kartika, D. S. Y. (2022, September). Implementasi Algoritma K-Means Clustering Untuk Pengelompokan Penjualan Produk Pada Online Shop Toko Gizi. In *Prosiding Seminar Nasional Teknologi dan Sistem Informasi* (Vol. 2, No. 1, pp. 242-248). https://sitasi.upnjatim.ac.id/index.php/sitasi/article/view/304

Ye, X., & Rong, T. (2022, April). An improved k-means clustering collaborative filtering recommendation algorithm. In *2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)* (pp. 545-548). IEEE. https://ieeexplore.ieee.org/abstract/document/9948405

Rahman, N. F., Priyatna, B., Nurapriani, F., & Paryono, T. (2025). PENERAPAN ALGORITMA K-MEANS UNTUK KLASTERISASI FILM PADA PLATFORM AMAZON PRIME VIDEO. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(5), 8200-8206. https://www.ejournal.itn.ac.id/jati/article/view/15065

Kumar, M. R., Vishnu, S., Roshen, G., Kumar, D. N., Revathi, P., & Baster, D. R. L. (2024, February). Product recommendation using collaborative filtering and k-means clustering. In *2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)* (Vol. 5, pp. 1722-1728). IEEE. https://ieeexplore.ieee.org/abstract/document/10486625

Rachman, M. D., & Voutama, A. (2024). Implementasi Algoritma K-Means Dalam Sistem Rekomendasi Musik Menggunakan Python. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 3857-3862.

Bagdja, M. N., Subagio, R. T., & Kartika, V. D. (2024). PENERAPAN METODE K-MEANS UNTUK MENGKLASIFIKASI MINAT KONSUMEN TERHADAP PRODUK TOKO ONLINE. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(5), 10690-10695.

Kundiman, I. M., Tampunabale, M. M., Kondoj, M. A., Langi, H. S., & Walukow, S. B. (2025). Product and Store Recommendation System Using K-Means Clustering and Hybrid Filtering on Marketplace. *Journal La Multiapp*, *6*(4), 837-848. https://mail.newinera.com/index.php/JournalLaMultiapp/article/view/2378