



## Application of *K-Means Clustering* Method to Analyze Consumer Perceptions of Local Products (Case Study: Aerostreet on Shopee)

Amanda Fauza<sup>1</sup>, Mahdi<sup>2\*</sup>, Musta'inul Abdi<sup>3</sup>

<sup>1,2,3</sup>Jurusan Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe Jl. Banda Aceh-Medan Km.280 Buketrata 24301  
INDONESIA

\*Penulis Korespondensi : mahdi@pnl.ac.id

### INFORMASI ARTIKEL

*Riwayat artikel:*

Diajukan pada 21 Mei 25  
Direvisi pada 05 Juni 25  
Publikasi pada 20 Juni 25

*Kata kunci:*

Analisis Ulasan  
*K-Means Clustering*  
Persepsi Konsumen  
*Clustering*

*Keywords:*

Review Analysis  
*K-Means Clustering*  
Consumer Perception  
*Clustering*

### ABSTRAK

Penelitian ini bertujuan untuk implementasi algoritma *K-Means Clustering* dalam menganalisis persepsi konsumen terhadap produk lokal brand Aerostreet di marketplace Shopee. Metode ini diterapkan untuk memahami bagaimana persepsi konsumen tercermin dari ulasan produk yang tersedia di platform tersebut. Setelah merancang dan menguji sistem dengan menggunakan metode *K-Means Clustering*, hasil penelitian menunjukkan bahwa data ulasan konsumen menghasilkan dua *cluster* dengan karakteristik berbeda di setiap *cluster*. Dari 750 komentar yang dianalisis dari 15 produk Aerostreet di Shopee, ditemukan bahwa kelompok kata positif lebih dominan dibandingkan kelompok kata negatif dalam setiap *cluster*. Hal ini mengindikasikan bahwa persepsi masyarakat terhadap produk lokal Aerostreet cenderung positif. Penelitian ini diharapkan dapat memberikan informasi yang berguna bagi pemangku kepentingan dan pihak yang membutuhkan insight mengenai persepsi konsumen terhadap produk lokal di pasar online.

### ABSTRACT

*This study aims to implement the K-Means Clustering algorithm to analyze consumer perceptions of the local brand Aerostreet on the Shopee marketplace. This method is applied to understand how consumer perceptions are reflected in product reviews available on the platform. After designing and testing the system using the K-Means Clustering method, the results show that consumer review data produced two clusters with distinct characteristics in each cluster. Out of 750 comments analyzed from 15 Aerostreet products on Shopee, it was found that positive word groups were more dominant than negative word groups in each cluster. This indicates that public perception of the local Aerostreet brand is generally positive. The study is expected to provide valuable information for stakeholders and those seeking insights into consumer perceptions of local products in the online market.*

## 1. Pendahuluan

Perkembangan di industri *fashion* saat ini berkembang dengan pesat. Banyak negara yang berlomba-lomba untuk memproduksi berbagai macam produk untuk dijual belikan, baik di dalam negeri maupun di luar negeri. Tidak kalah saing di Indonesia saat ini banyak memiliki berbagai macam produk *fashion*. Produk *fashion* yang saat ini banyak diminati oleh masyarakat Indonesia adalah produk lokal. Produk lokal adalah

produk yang diproduksi di dalam negeri. Aerostreet merupakan salah satu *brand* produk lokal yang banyak diminati oleh masyarakat Indonesia.

Aerostreet juga merupakan salah satu produk sepatu lokal Indonesia yang memulai penjualannya di pasar *offline* dan saat ini sudah ada di pasar *online* yaitu Shopee[1]. Saat ini Aerostreet tidak hanya memproduksi sepatu saja, namun Aerostreet juga memproduksi produk *fashion* lainnya seperti *t-shirt*, sandal, hingga jaket. Namun industri *fashion* sangat dinamis, dan tren konsumen bisa berubah dengan cepat. Aerostreet mungkin akan mengalami kesulitan dalam beradaptasi dengan perubahan tren dan preferensi konsumen, yang mengakibatkan ketidakrelevanan produk di pasar saat ini [2].

Salah satu ukuran keberhasilan dari sebuah produk adalah bagaimana persepsi konsumen dapat meningkatkan kepercayaan terhadap suatu produk sehingga mereka mempunyai keinginan membeli yang sangat besar terhadap produk tersebut dan Aerostreet dapat mengembangkan produknya [3]. Hal ini tentu saja dapat diketahui dengan melakukan penilaian persepsi konsumen melalui ulasan konsumen pada produk.

*K-Means Clustering* adalah teknik pengelompokan data non-hierarki yang memisahkan data ke dalam *cluster*, mengelompokkan data dengan fitur yang sama bersama-sama dan mengelompokkan data dengan karakteristik yang berbeda ke dalam kelompok yang berbeda[4][5]. Penerapan dari metode *K-Means Clustering* adalah untuk memudahkan melakukan penilaian persepsi konsumen melalui ulasan dengan mengelompokkan ulasan konsumen ke dalam *cluster* yang akan ditentukan.

Persepsi adalah suatu proses yang dilewati seseorang untuk menyeleksi, mengorganisasikan, dan menginterpretasikan informasi-informasi tertentu dalam rangka membentuk makna tertentu mengenai produk atau merek tertentu[6]. Persepsi konsumen terhadap suatu perusahaan semakin berperan penting di era bisnis digital saat ini, karena konsumen dapat dengan mudah berbagi pengalaman negatif konsumen kepada penjual melalui media digital seperti media sosial[7].

## 2. Metode

### 2.1 Data dan Pengumpulan Data

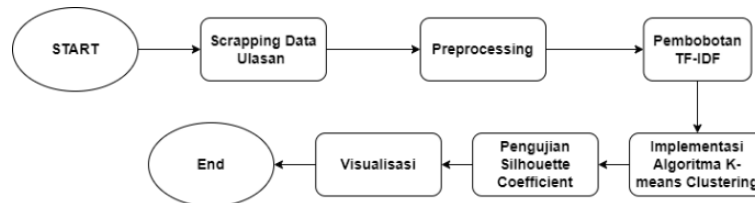
Sumber data untuk penelitian ini diperoleh dari ulasan atau *review* konsumen produk *brand* lokal Aerostreet pada *marketplace* Shopee. Adapun ulasan yang diambil sebanyak 5 varian setiap produk Sepatu, T-Shirt, dan Sandal. Pengumpulan data akan dilakukan melalui *scraping* dan API Shopee pada toko Aerostreet Official di Shopee. Data yang diambil sebanyak minimal 10 maksimal 200 data komentar pada setiap produk yang diinginkan. Namun pada penelitian ini akan mengambil 50 data komentar setiap produk.

### 2.1 Perancangan Model *Machine Learning*

Pada bagian ini akan membahas tentang perancangan model *machine learning*. Perancangan model ini dilakukan dengan beberapa tahapan yang dimulai dari arsitektur umum, data yang digunakan, *preprocessing data*, TF-IDF, dan perhitungan akurasi menggunakan *Silhouette Score*. Perancangan model *machine learning* ini akan dilakukan sepenuhnya oleh *developer*.

### 2.1.1 Rancangan Alur Sistem

Gambar 1 menjelaskan proses sistem yang akan dibuat. Dimulai dari *scraping* data ulasan produk Aerostreet di Shopee, kemudian data ulasan akan melalui proses *preprocessing* yaitu pembersihan data. Selanjutnya data yang sudah bersih maka akan dibobotkan dengan metode TF-IDF, selanjutnya data yang sudah diberi bobot akan melakukan proses *K-Means Clustering*. Setelah melakukan proses *K-Means Clustering* data yang sudah masuk *cluster* akan di hitung akurasi dengan menggunakan metode *Silhouette Score*. Terakhir kita dapat menampilkan data *cluster* yang ada.



**Gambar 1.** Alur Sistem

### 2.1.2 Preprocessing Data

*Preprocessing data* dilakukan setelah mendapatkan data melalui proses *scraping*. Data didapatkan pada proses *scraping* masih data mentah sehingga perlu dibersihkan. Tahapan *preprocessing* terdiri dari *clean text*, *tokenize*, *remove stopwords*, dan *stemming*.

### 2.1.3 TF-IDF

Proses pembobotan kata menggunakan metode TF-IDF. Pada proses ini akan melakukan pembobotan setiap kata yang ada pada komentar. Pada Python terdapat beberapa *library* yang akan digunakan untuk pembobotan kata menggunakan metode TF-IDF. *Library scikit-learn* adalah salah satu *library machine learning* yang akan digunakan pada penelitian ini. *Scikit-learn* menyediakan modul “*TfidfVectorizer*” yang dapat digunakan untuk menghitung skor TF-IDF dari kumpulan dokumen.

### 2.1.4 Silhouette Coefficient

*Silhouette Coefficient* adalah sebuah metrik yang digunakan untuk mengukur seberapa baik sebuah *clustering* (pengelompokan) yang telah dilakukan. Pada penelitian ini untuk melakukan perhitungan *Silhouette Coefficient* menggunakan *library python scikit-learn* dengan modul “*metrics*”.

### 2.1.5 Word Cloud

*Word Cloud* adalah teknik visualisasi yang digunakan untuk menampilkan kata-kata yang paling sering muncul dalam sebuah teks atau data. Pada penelitian ini setelah proses *clustering* data komentar, selanjutnya setiap *cluster* akan dilakukan *Word Cloud* untuk melihat kata-kata apa saja yang paling banyak muncul di setiap *cluster*. Teknik *Word Cloud* ini menggunakan *library* pada Python untuk menghasilkan visualisasi kata pada setiap *cluster*.

## 3. Hasil Dan Pembahasan

### 3.1 Hasil Scraping Data

Data yang berhasil di *scraping* adalah 15 produk yang masing-masing terdiri dari 5 produk sandal, 5 produk sepatu, dan 5 produk *t-shirt*. Data yang diambil setiap produk sebanyak 50 data komentar, sehingga total keseluruhan data berjumlah 750. Berikut hasil data yang telah di-*scraping*.

#	Bintang	Komentar
1	5	Tampilan:bagus. Warna:hitam seperti yang di harapkan Terimakasih aero street. Sepatunya sampai dengan selamat. Untuk sepatunya nyaman. Belum coba dipakai seharian ya. Semoga saja cocok. Pengiriman juga cepat. Sekali lagi terimakasih
2	5	Warna:bagus Tampilan:sangat kren dan elegan Barang sesuai dengan pesanan saya, dan sepatu nya sangat sangat wah kren abis elegan mewah nampak seperti sepatu mahal.. Sangat memuaskan beli sepatu di toko ini.. Dari bulan Desember mau beli eh habis terus dan alhamdulillah dapat stelah ready stok full
3	5	Gak nyangka pas datang sepatunya bagus banget.. barang mulus, gak ada yg cacat.. Terus juga sepatunya udah di sol, jadi bakal awet.. Rekomended bagi yang ingin beli
4	5	Kualitassangat bagus dgn harga segini Styling:cakep banget, lebih dewasa Sepadan dengan Harga:sangat sepadan Kualitas sangat bgus, dgn harga segini, tampilan sangat keren, cocok yg gk suka warna rame
5	5	Tampilan:sangat pas cocok sekali nyaman di pakainya Warna:black dengan harga terjangkau
6	5	Warnahitam Tampilansesuai dgn foto Enggak nyangka bakal datang secepat ini, aku pesan tgl 5 Desember dan udh datang aja tgl 7. Woilahhh gaskeun kaliahlah beli disini selagi promo besar-besaran. Barangnya juga bagus dan berkualitas. Dijamin puas.
7	5	Kenyamanantukurannya pas dan soldnya tinggi Sepadan dengan Harga:Sangat worth it Styling:Casual Sudah ke tiga kalinya beli sepatu di Aerostreet dan gak pernah mengecewakan. Pengiriman juga sangat cepat
8	5	Warna:baguss bagus bgttt, nyaman, cepet bgt dikirim ke paling timur indonesia👍
9	5	Tampilan:Sepatu sesuai dengan gambar Warna:Hitam Baru pertama beli sepatu di sini untuk hadiah.. Sepatu nya kokoh, ukuran sesuai, model

Gambar 2. Hasil Scraping

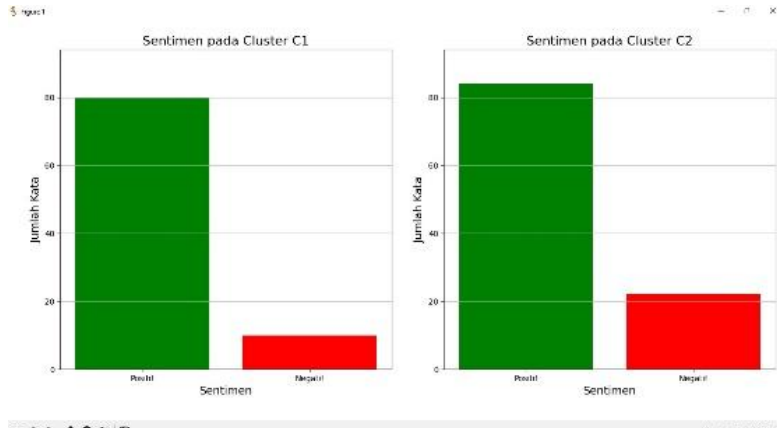
### 3.2 Implementasi Analisis Cluster Menggunakan Metode K-Means Clustering dan Silhouette Score

Metode *K-Means Clustering* adalah merupakan metode *clustering* berbasis *centroid* dimana kumpulan data dibagi menjadi *K-Cluster*. *K-Means Clustering* adalah salah satu algoritma *clustering* paling sederhana dan paling umum digunakan. Data yang diuji sebanyak 750 komentar dari 15 produk Aerostreet. Berikut program penerapan metode *K-Means Clustering* untuk menganalisis jumlah *cluster* pada komentar produk Aerostreet di Shopee dan metode pendukung *Silhouette Score*.

Tahapan analisis *cluster* menggunakan metode *K-Means Clustering*. Diawali dengan pengambilan data dari hasil *preprocessing*, selanjutnya dilakukan pembobotan setiap kata. Proses pembobotan ini dilakukan untuk melihat frekuensi kemunculan suatu kata dan mengubah kata ke bentuk biner agar dapat diolah dengan sistem. Kemudian dilakukan tahap analisis *clustering* menggunakan *library* dan rumus *K-Means Clustering* dengan memasukkan jumlah *cluster* yaitu 2, kemudian masukkan nilai *centroid* nya, lalu masukkan data yang ingin dikelompokkan ke dalam *cluster* dan hitung jarak *centroid* setiap data, selanjutnya mengelompokkan data berdasarkan kedekatan *centroid* ke *cluster*. Terakhir melakukan perulangan hingga *cluster* yang diikuti *centroid* tidak berubah. Dilanjutkan dengan mengukur akurasi setiap *cluster* menggunakan *Silhouette Score*.

### 3.3 Hasil Uji K-Means Clustering

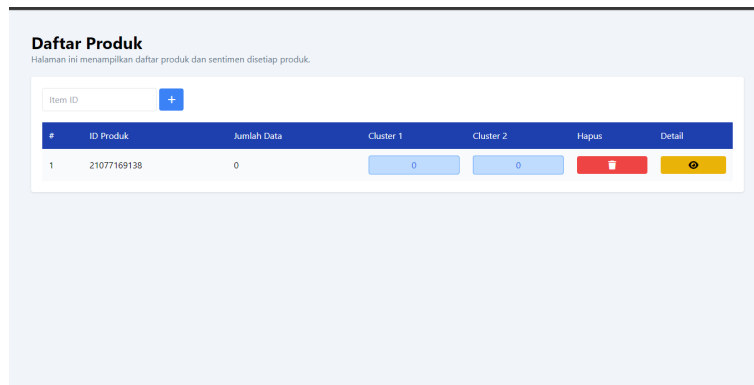
Grafik pada gambar 3 menampilkan analisis sentimen pada dua cluster berbeda, C1 dan C2. Pada Cluster C1, terlihat dominasi sentimen positif dengan jumlah yang jauh lebih tinggi daripada sentimen negatif, menunjukkan bahwa mayoritas responden atau data dalam *cluster* ini memiliki pandangan yang baik. Sebaliknya, Cluster C2 menunjukkan proporsi yang lebih seimbang, meskipun sentimen positif masih mendominasi, terdapat peningkatan ketimbang *cluster* pertama, tetapi tetap ada sejumlah sentimen negatif yang terdeteksi. Ini mengindikasikan bahwa terdapat perbedaan dalam persepsi atau pengalaman antara kedua *cluster* tersebut.



Gambar 3. Hasil Uji *K-Means Clustering*

### 3.4 Halaman Daftar Produk

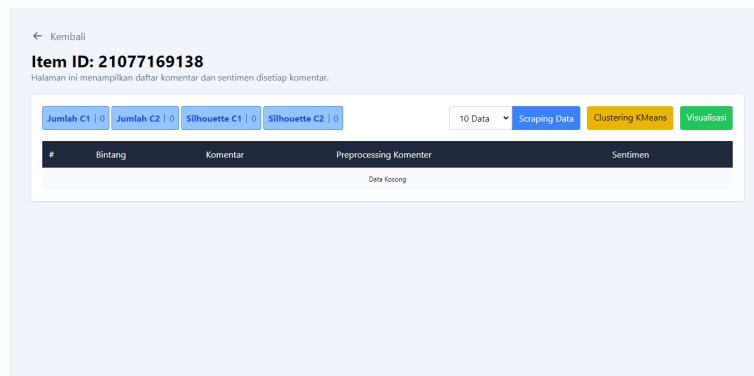
Gambar 4 halaman daftar produk adalah halaman utama yang dapat diakses oleh pengguna. Pada halaman ini pengguna memasukkan Item ID produk Aerostreet yang diinginkan.



Gambar 4. Halaman Daftar Produk

### 3.5 Halaman Proses Item ID

Gambar 5 halaman item id adalah halaman yang akan tampil setelah *user* menekan tombol detail dari halaman daftar produk. Halaman ini akan melakukan proses *scraping* komentar, melakukan proses *clustering K-Means*, dan menampilkan visualisasi *Word Cloud* setiap *cluster*.



Gambar 5. Halaman Proses Item ID

### 3.6 Pengujian Sistem

Berdasarkan pengujian *Black Box* yang dilakukan pada sistem, didapatkan lulus pengujian sebanyak 6 *test case* dari 7 *test case*. Dapat disimpulkan pengujian *Black Box* pada sistem penerapan metode *K-Means Clustering* untuk menganalisis persepsi konsumen produk lokal studi kasus: Aerostreet di Shopee memiliki persentase kelulusan sebesar 85% dan persentase ketidakkulusan sebesar 15%.

Sedangkan pengujian *White Box* yang telah dilakukan didapatkan hasil pengujian sebesar 100% dari 1 *test case*. Secara keseluruhan hasil pengujian menunjukkan bahwa kode program pada metode *K-Means Clustering* ini dinyatakan berhasil dan benar.

## 4. Kesimpulan

Hasil penelitian penerapan metode *K-Means Clustering* untuk menganalisis persepsi konsumen terhadap produk lokal dengan studi kasus Aerostreet di Shopee menghasilkan dua *cluster* dengan anggota yang berbeda pada setiap *cluster*. Penelitian ini berhasil menunjukkan bahwa dari 750 data komentar yang diambil dari 15 produk Aerostreet di Shopee, diperoleh kelompok kata yang mencerminkan persepsi positif dan negatif. Setiap *cluster* pada masing-masing produk menghasilkan kelompok kata positif sebanyak 80% dan kata negatif sebanyak 20%, sehingga dapat disimpulkan bahwa persepsi masyarakat terhadap produk lokal merek Aerostreet adalah persepsi yang baik.

## Referensi

- [1] Y. Dianti, “*濟無*No Title No Title No Title,” *Angew. Chemie Int. Ed.* 6(11), 951–952., pp. 5–24, 2017, [Online]. Available: <http://repo.iain-tulungagung.ac.id/5510/5/BAB 2.pdf>
- [2] R. K. Ahmad Zakiy Ramadhan, “Meningkatkan Brand Awareness melalui Strategi Kolaborasi,” *Pros. SNADES 2022 – Desain Kolaborasi Interdisip. di Era Digit.*, pp. 177–183, 2022.
- [3] A. Shiratina and P. Afiatun, “Pengaruh Persepsi Konsumen Dan Motivasi Konsumen Terhadap Keputusan Pembelian Produk Kain Rajut,” *JURISMA J. Ris. Bisnis Manaj.*, vol. 7, no. 1, 2017, doi: 10.34010/jurisma.v7i1.472.
- [4] T. Amalina, D. Bima, A. Pramana, and B. N. Sari, “Metode *K-Means Clustering* Dalam Pengelompokan Penjualan Produk Frozen Food,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. 15, pp. 574–583, 2022, [Online]. Available: <https://doi.org/10.5281/zenodo.7052276>
- [5] S. I. Safitri, C. Suhery, and S. Bahri, “Implementasi Algoritma K–Means Untuk Clustering Sentimen Pada Opini Kualitas Pelayanan Jasa Penerbangan,” *Coding J. Komput. dan Apl.*, vol. 09, no. 02, pp. 186–197, 2021, [Online]. Available: <https://jurnal.untan.ac.id/index.php/jcskommipa/article/view/47377>
- [6] A. Fahmi, R. Wandu, and Z. Husna, “Implementasi Metode *K-Means Clustering* Dalam Menentukan Jenis Uang Terlaris Pada Ud. Tegar Rizqi,” *TECHSI -J. Tek. Inform.*, vol. 13, no. 1, p. 92, 2021, doi: 10.29103/techsi.v13i1.3148.
- [7] Nurvaedah, “Analisis Persepsi Konsumen Terhadap Minat Beli Online Era New Normal Covid-19 (Studi Kasus Pada Pengguna Aplikasi Grab Food Di Kota Palopo),” *J. Perilaku Konsum.*, vol. 2020, pp. 1–76, 2021.
- [8] Seputar Birokrasi.com, “Belanja Produk Lokal: Solusi untuk Meningkatkan Daya Saing dan Kualitas Produk Indonesia.” [Online]. Available: <https://seputarbirokrasi.com/belanja-produk-lokal-solusi-untuk-meningkatkan-daya-saing-dan-kualitas-produk-indonesia/#:~:text=Produk lokal adalah produk yang,dasar yang dihasilkan dari Indonesia.B>