E-ISSN: 3025-0994
Halaman 1298-1311
Volume 4 Nomor 4 Tahun 2026

Analisis Sentimen Ulasan Pengguna Aplikasi DSW *(Depok Single Window)* Di Google *Play Store* Menggunakan Algoritma *K-Means*

Aris Pramuja[1], Nanda Anggraini Maulana[2], Sherliani Putri Saharani[3], Ade Fitria Lestari[4]
[1,2,3,4] Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika Jakarta
Email: [1]19221403@bsi.ac.id, [2]19221349@bsi.ac.id, [3]19221398@bsi.ac.id, [4]ade.afr@bsi.ac.id

Abstrak

Aplikasi Depok Single Window (DSW) merupakan platform layanan publik terpadu yang dikembangkan Pemerintah Kota Depok, namun tingkat pengetahuan masyarakat terhadapnya masih rendah dan ribuan ulasan pengguna di Google Play Store masih menunjukkan berbagai kendala teknis seperti kesulitan login, loading lambat, notifikasi tidak real-time, hingga kegagalan pendaftaran online. Ulasan yang tidak terstruktur ini menyulitkan pengelola dalam mengidentifikasi pola permasalahan secara cepat dan akurat. Penelitian ini bertujuan mengimplementasikan algoritma K-Means Clustering untuk menganalisis sentimen dan mengelompokkan topik ulasan pengguna, serta mengidentifikasi pola keluhan guna menghasilkan rekomendasi perbaikan teknis bagi pengembang aplikasi. Metode yang digunakan meliputi pengumpulan 1.190 data ulasan, text preprocessing, pembobotan kata dengan TF-IDF, penentuan jumlah klaster optimal menggunakan Elbow Method dan Silhouette Score, serta klasterisasi menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa jumlah klaster optimal adalah tiga (k=3). Klaster 0 merepresentasikan ulasan positif berisi apresiasi pengguna terhadap aplikasi, Klaster 1 berisi ulasan netral atau informatif terkait pengunaan layanan, sedangkan Klaster 2 didominasi keluhan teknis seperti kegagalan pendaftaran online, kesulitan login, dan masalah akses CCTV dengan rata-rata rating terendah. Hasil klasterisasi ini diharapkan menjadi masukan bagi pengembang aplikasi dan Pemerintah Kota Depok dalam meningkatkan kualitas layanan publik digital.

Kata kunci: *Analisis Sentimen, K-Means Clustering, Text Mining, Tf-Idf, Depok Single Window*

Abstract

Depok Single Window (DSW) is an integrated public service application developed by the Depok City Government, yet public awareness of it remains low, and thousands of user reviews on the Google Play Store still reveal technical issues such as login

difficulties, slow loading, non-real-time notifications, and online registration failures. These unstructured reviews make it difficult for administrators to quickly and accurately identify recurring problem patterns. This study aims to implement the K-Means Clustering algorithm to perform sentiment analysis and topic grouping of user reviews, and to identify complaint patterns for measurable technical improvement recommendations. The method applied includes collecting 1.190 review data, text preprocessing, term weighting using TF-IDF, determining the optimal number of clusters using the Elbow Method and Silhouette Score, and clustering with K-Means. The results show that the optimal number of clusters is three (k=3). Cluster 0 represents positive reviews containing user appreciation, Cluster 1 contains neutral or informative reviews on general service usage, while Cluster 2 is dominated by technical complaints such as online registration failures, login difficulties, and CCTV access problems, with the lowest average rating. These findings are expected to help developers and the Depok City Government improve digital public service quality.

Keywords : *sentiment analysis, k-means clustering, text mining, tf-idf, depok single window*

LATAR BELAKANG

Pergeseran pelayanan publik ke arah sistem berbasis elektronik telah mendorong pemerintah daerah untuk menghadirkan berbagai platform layanan terpadu. Upaya ini bertujuan untuk mengintegrasikan berbagai kebutuhan administrasi dalam satu sistem sehingga proses birokrasi menjadi lebih efisien dan memudahkan masyarakat. Namun, pada kenyataannya, tingkat pengetahuan masyarakat terhadap keberadaan layanan tersebut masih tergolong rendah, yaitu hanya mencapai 22,5% (Novitasari et al., 2024). Rendahnya tingkat pengetahuan tersebut menunjukkan adanya kesenjangan antara ketersediaan teknologi dan pemanfaatannya oleh masyarakat, yang sering disebabkan oleh kurangnya sosialisasi serta kendala teknis yang dialami pengguna saat menggunakan aplikasi. Kondisi tersebut sejalan dengan penelitian Munawar et al. (2025) yang menyatakan bahwa kualitas layanan aplikasi pelayanan publik berbasis digital berpengaruh positif dan signifikan terhadap kepuasan pengguna. Penelitian tersebut menjelaskan bahwa aspek keandalan sistem (*reliability*), daya tanggap (*responsiveness*), serta kemudahan penggunaan (*ease of use*) menjadi faktor penting yang memengaruhi pengalaman pengguna dalam memanfaatkan aplikasi layanan publik digital.

Kualitas pengalaman pengguna dapat terlihat dengan jelas melalui ribuan ulasan yang ditinggalkan pada platform distribusi aplikasi digital, di mana hingga Maret 2026, jumlah ulasan untuk aplikasi Depok Single Window (DSW) telah mencapai sebanyak 1.190 data dengan rata-rata rating 4,1. Ulasan pengguna pada platform digital tidak hanya mencerminkan tingkat kepuasan pengguna, tetapi juga dapat dimanfaatkan sebagai sumber informasi dalam mengevaluasi kualitas pelayanan publik digital. Menurut Karomah et al. (2025), analisis terhadap ulasan dan persepsi

pengguna mampu memberikan gambaran mengenai kelebihan dan kelemahan layanan digital sehingga dapat menjadi dasar dalam upaya peningkatan kualitas pelayanan publik. Meskipun angka tersebut menunjukkan persepsi yang cukup positif, hasil observasi terhadap kolom komentar menunjukkan bahwa pengguna masih sering menghadapi berbagai kendala teknis yang menghambat fungsi utama aplikasi. Beberapa permasalahan yang ditemukan secara langsung dalam ulasan tersebut meliputi kesulitan saat melakukan proses *login* ke dalam sistem, waktu pemuatan data (*loading*) yang memakan waktu lama, hingga fitur notifikasi yang tidak berjalan secara *real-time*. Selain itu, masyarakat juga banyak mengeluhkan kegagalan sistem pada fitur pendaftaran *online* yang merupakan layanan krusial, sehingga kondisi tersebut menunjukkan bahwa DSW masih memerlukan evaluasi mendalam agar dapat memenuhi kebutuhan pengguna secara optimal.

Mengingat keluhan-keluhan tersebut disampaikan dalam bentuk teks yang tidak terstruktur, pengelola sering mengalami kesulitan dalam mengidentifikasi serta memetakan kategori permasalahan utama secara cepat dan akurat. Kondisi ini menyebabkan ribuan masukan dari masyarakat Kota Depok tersebut hanya akan menjadi tumpukan informasi digital yang tidak memberikan manfaat langsung bagi pengembangan sistem jika tidak dilakukan pengolahan data yang sistematis dan tepat. Teknik pengolahan teks digunakan sebagai solusi untuk mengubah ribuan ulasan tersebut menjadi informasi yang lebih terstruktur dan mudah dipahami. Pengelompokan opini secara otomatis sangat diperlukan untuk mengetahui pola ketidakpuasan serta preferensi pengguna tanpa harus melakukan pengecekan manual satu per satu (Amalia et al., 2023). Melalui pendekatan pembelajaran mesin tanpa pengawasan (*unsupervised learning*), sistem dapat menemukan pola dalam data ulasan untuk mengidentifikasi kelompok permasalahan secara lebih akurat (Alfarizi et al., 2026).

Penerapan algoritma *K-Means* menjadi pilihan yang tepat karena kemampuannya dalam mengolah data teks yang kompleks serta kemudahannya dalam melakukan pengelompokan data secara efektif (Alfarizi et al., 2026). Algoritma ini bekerja dengan mengelompokkan ulasan berdasarkan kemiripan kata yang digunakan, sehingga dapat memberikan gambaran umum mengenai pola permasalahan teknis yang paling sering dikeluhkan oleh pengguna (Sunarko et al., 2023). Hasil dari pengelompokan ini tidak hanya akan menyajikan data statistik, tetapi juga menjadi landasan strategis bagi pengelola dalam menentukan prioritas perbaikan layanan publik digital. Oleh karena itu, penelitian ini mengimplementasikan metode *K-Means Clustering* untuk menganalisis dan mengelompokkan ulasan pada aplikasi Depok Single Window (DSW), guna mengidentifikasi kategori sentimen dan kendala utama yang dirasakan oleh masyarakat secara lebih terstruktur.

Kebaruan (*state of the art*) penelitian ini terletak pada pemilihan metode serta data yang digunakan. Studi ini merupakan implementasi pertama yang menggunakan algoritma *K-Means* secara langsung sebagai teknik *unsupervised learning* pada data ulasan platform layanan satu pintu tersebut. Berbeda dengan penelitian sebelumnya yang menggunakan metode terbimbing seperti *Support Vector Machine* (SVM) dan

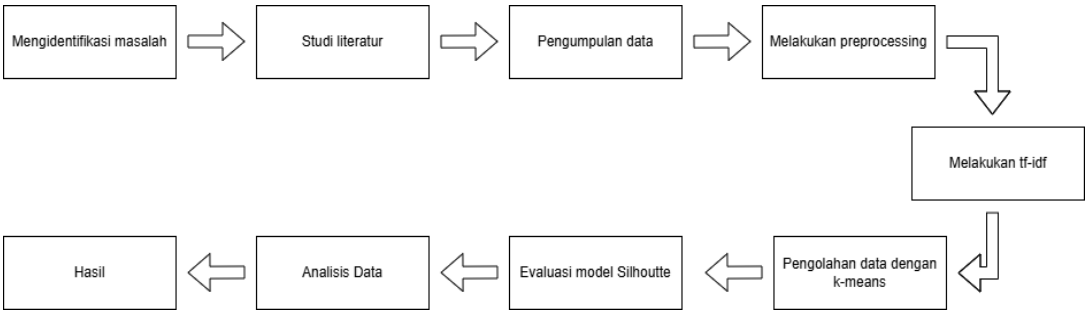
memerlukan proses pelabelan data secara manual, penggunaan *K-Means* dalam penelitian ini memungkinkan pola-pola permasalahan muncul secara objektif dari data asli tanpa adanya campur tangan peneliti. Sisi kebaruan lainnya ditunjukkan melalui aspek waktu penelitian dan hasil temuan pada rating aplikasi. Data yang dianalisis mencakup periode hingga Maret 2026, sehingga memberikan gambaran yang lebih terbaru dibandingkan penelitian sebelumnya. Fokus pada pengelompokan pola keluhan menggunakan algoritma diharapkan mampu menyajikan rincian permasalahan teknis secara lebih jelas guna mendukung perbaikan sistem layanan digital secara lebih tepat.

METODE

Berisi Penelitian ini menggunakan pendekatan *data mining* dengan metode *K-Means Clustering* untuk menganalisis sentimen dan mengelompokkan ulasan pengguna aplikasi Depok Single Window (DSW) di Google Play Store. *Data mining* merupakan proses penggalian informasi dari kumpulan data menggunakan teknik statistik, matematika, dan *machine learning* untuk memperoleh pola atau pengetahuan yang bermanfaat (Alfarizi et al., 2026).

Data penelitian diperoleh dari platform Google Play Store yang menyediakan informasi mengenai ulasan pengguna aplikasi Depok Single Window (DSW). Pengumpulan data dilakukan menggunakan teknik *web scraping* secara otomatis menggunakan bahasa pemrograman Python dengan pustaka *google-play-scraper*. Sebagaimana diterapkan dalam penelitian Sunarko et al. (2023), pengumpulan data ulasan dari Google Play Store melalui teknik *scraping* terbukti mampu menghasilkan dataset yang representatif untuk keperluan analisis sentimen. Data yang berhasil dikumpulkan dikelola menggunakan pustaka *pandas* untuk mengorganisir ribuan ulasan mentah ke dalam struktur data tabular (DataFrame), sehingga diperoleh sebanyak 1.190 data ulasan pengguna yang mencakup informasi nama pengguna, rating, teks ulasan, dan tanggal publikasi dalam rentang waktu Agustus 2018 hingga Maret 2026.

Tahapan analisis data terdiri dari proses *preprocessing*, pembobotan kata *TF-IDF*, penentuan jumlah klaster optimal, implementasi algoritma *K-Means*, evaluasi hasil *clustering*, hingga interpretasi hasil pengelompokan. Secara keseluruhan, alur dan tahapan penelitian ini digambarkan melalui kerangka penelitian yang ditunjukkan pada Gambar 1 berikut.



Gambar 1. Kerangka Penelitian
Sumber: (Hasil Penelitian, 2026)

Proses *text preprocessing* merupakan tahapan penting untuk membersihkan data teks dari unsur-unsur yang tidak diperlukan agar proses analisis dapat berjalan lebih optimal (Daely et al., 2026). Berdasarkan tahapan teknis yang dipaparkan oleh Ardiawan et al. (2022) dan Daely et al. (2026), pemrosesan teks dalam penelitian ini meliputi:

1. Data Cleaning: Menangani *missing value* dan menghapus data duplikat menggunakan fungsi isnull().sum() dan duplicated().sum() pada pustaka *pandas*. Selanjutnya, tipe data kolom ulasan dikonversi secara konsisten menjadi *string* menggunakan fungsi astype(str)
2. Case Folding: Menyeragamkan seluruh karakter huruf dalam teks menjadi huruf kecil (*lowercase*) agar sistem tidak menganggap kata yang sama sebagai entitas berbeda (Ardiawan et al., 2022).
3. Cleansing: Menghapus elemen non-alfabet seperti angka, simbol, tanda baca, URL, *mention*, *hashtag*, serta spasi berlebih menggunakan *Regular Expression* (re) untuk mengurangi *noise* pada data (Daely et al., 2026).
4. Normalization: Mengubah kata-kata tidak baku, singkatan, atau istilah gaul menjadi bentuk kata baku sesuai kaidah Bahasa Indonesia menggunakan kamus pemetaan kata (*dictionary*).
5. Tokenizing: Memecah rangkaian kalimat ulasan menjadi potongan kata tunggal (*token*) sebagai unit analisis dasar (Daely et al., 2026).
6. Filtering (Stopword Removal): Menghapus kata-kata umum yang memiliki frekuensi tinggi namun tidak membawa makna penting (seperti kata hubung "dan", "yang", "di") menggunakan daftar *stopword* dari pustaka Sastrawi yang dikombinasikan dengan *stopword* khusus domain (seperti "nya", "yg", "aja", "min").
7. Stemming: Mengubah kata berimbuhan menjadi bentuk kata dasar menggunakan pustaka Sastrawi agar kata-kata yang memiliki akar makna sama dapat direpresentasikan dalam bentuk seragam (Sunarko et al., 2023).

8. Handling Empty Data: Menghapus baris ulasan yang menjadi kosong akibat eliminasi *stopword* atau *stemming*, sehingga menghasilkan dataset bersih akhir sebanyak 1.074 baris yang siap diproses.

Data teks yang telah bersih kemudian ditransformasikan ke dalam bentuk representasi numerik menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF). Metode ini mengukur tingkat kepentingan suatu kata berdasarkan frekuensi kemunculannya dalam sebuah ulasan (*Term Frequency*) dan tingkat kelangkaannya pada seluruh dokumen (*Inverse Document Frequency*) (Sunarko et al., 2023; Ramadhan et al., 2021). Perhitungan bobot kata secara matematis dilakukan berdasarkan persamaan:

$$tf - idf = \left(1 + log\ log\ \left(tf_{\{t,d\}}\right)\right) \times log\ log\ \left(\frac{N}{df_t}\right)$$

Keterangan:
$TF_{t,d}$ = *Term Frequency*, yaitu jumlah kemunculan kata t dalam dokumen d
N = total jumlah seluruh dokumen yang menjadi korpus data penelitian ($N = 1.074$)
DF_t = *Document Frequency*, yaitu jumlah dokumen yang mengandung kata t

Pembobotan TF-IDF diimplementasikan menggunakan TfidfVectorizer dari pustaka *scikit-learn* dengan konfigurasi max_features=1000 dan min_df=2. Penentuan jumlah klaster optimal dilakukan dengan mengombinasikan *Elbow Method* dan *Silhouette Score*. Berdasarkan evaluasi tersebut, ditetapkan jumlah klaster optimal sebanyak tiga klaster ($k = 3$).

Proses pengelompokan dilakukan menggunakan algoritma *K-Means Clustering*, yaitu metode *unsupervised learning* yang difungsikan untuk membentuk klaster berdasarkan kedekatan jarak antarvektor teks berbobot TF-IDF. Menurut Alfarizi et al. (2026), algoritma ini bekerja dengan menghitung jarak antara data ulasan dengan titik pusat klaster (*centroid*) secara iteratif hingga posisi pusat *centroid* menjadi stabil. Perhitungan jarak geometris antara objek data ulasan x dengan titik pusat klaster c menggunakan metode *Euclidean Distance* dengan persamaan (Alfarizi et al., 2026):

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

Keterangan:
$d(x,c)$ = jarak geometris antara objek data ulasan x dengan titik pusat klaster c
x_i = nilai koordinat data ulasan pada dimensi ke-i

c_i = nilai koordinat titik pusat klaster pada dimensi ke-i
n = jumlah dimensi fitur hasil pembobotan TF-IDF ($n = 1000$)

Setelah jarak setiap data terhadap *centroid* diperoleh, data dikelompokkan berdasarkan jarak terdekat. Selanjutnya dilakukan pembaruan nilai *centroid* menggunakan nilai rata-rata seluruh data dalam setiap klaster. Proses iterasi dilakukan secara berulang hingga nilai *centroid* konvergen.

Evaluasi kualitas hasil pengelompokan dilakukan menggunakan *Silhouette Score* untuk mengukur tingkat kesamaan data dalam klaster yang sama (*cohesion*) serta perbedaannya dengan data pada klaster lain (*separation*). Menurut Syahkur et al. (2024), *Silhouette Score* mengevaluasi kualitas klasterisasi melalui perbandingan antara rata-rata jarak data terhadap anggota dalam klaster yang sama dan rata-rata jarak data terhadap anggota dalam klaster lain yang paling dekat. Semakin tinggi nilai *Silhouette*, maka semakin baik kualitas klasterisasi yang dihasilkan karena menunjukkan bahwa data dalam suatu klaster memiliki tingkat kesamaan yang tinggi serta pemisahan yang jelas dengan klaster lain. Setelah terbentuk tiga klaster optimal, pelabelan sentimen (positif, netral, dan negatif) pada masing-masing klaster dilakukan melalui analisis manual (*post-processing*) berdasarkan interpretasi visualisasi *word cloud* serta analisis kualitatif kata-kata berbobot tertinggi.

HASIL DAN PEMBAHASAN

hotel. Pada bagian ini disampaikan hasil pengolahan data penelitian beserta pembahasan yang komprehensif, dengan menekankan nilai kebaruan, inovasi, serta implikasi dari penelitian yang dilakukan. Data dan fakta yang disajikan merupakan data matang yang dituangkan dalam bentuk tabel dan visualisasi untuk memudahkan pemahaman pembaca.

Hasil Pengumpulan Data *(Scraping*) dan Preprocessing

Objek penelitian ini berfokus pada ulasan pengguna aplikasi Depok Single Window (DSW) di Google Play Store yang dikumpulkan menggunakan teknik *web scraping* berbasis pustaka *google-play-scraper*, menghasilkan sebanyak 1.190 data ulasan dalam rentang waktu Agustus 2018 hingga Maret 2026. Contoh cuplikan dataset ulasan mentah hasil *scraping* yang merepresentasikan ulasan pengguna dari Google Play Store dapat dilihat pada Gambar 1 berikut:

Nama_user	Rating	Ulasan	Tanggal
Guntur Za	4	ga bisa unduh form F-1.06	2/26/2026 05:58
Wulan Maryanti	1	daftar peroses pembuatan akte g jdi jdi pdhal udh ngirim dari 2 minggu lalu, dan katanya di proses 3 hari kerja eh ini sampe udh 2 minggu g jdi jdi, katanya nanti diinfokan trus pdhal klo kita g follow up g akan diinfokan kurangny apa dan hrus nunggu berapa lama lgii	2/18/2026 09:55
Novi hermansyah Herman	5	baik	2/16/2026 00:21
Yulia anisa Fitri	1	buat apa apk di ciptainnnnnn, tp gabisa DAFTAR BUAT KE PUSKESMAS, NYUSAHIN AJAAA SIHHHHHHHHHH	2/15/2026 22:29
Ade Chintya	1	pemdaftaran tdk bisa ... lemot ... padahal sinyal stabil ,sudah coba ajakan 3 akun kawan dgn data berbeda, perlu di perbaiki sistem / pekerja nya ganti aja pemeliharaan gk beres	2/12/2026 14:21
Ciprut Aja	1	mau daftar online buat lusa kok g bisa2...malahan yg muncul riwayat online 4th lalu g bisa dirubah	2/6/2026 20:21
Dark Style	2	kaga bisa di akses CCTV nya	1/22/2026 05:19
Paguyuban Trashmouth	1	Tolong untuk akses CCTV lebih diperhatikan lagi perbaikkannya. Jangan tiap hari lost data terus.	1/21/2026 12:02
Muhammad rizky Rizky	1	CCTV rusak mulu gak bisa liat debit air ciliwung	1/21/2026 11:27
YUYU YIYI	1	cctv-nya susah diakses	1/21/2026 10:37
Halim Himawan	1	mati semua cctv nya	1/16/2026 18:30
Musa Indra	5	good	1/13/2026 09:04
Uchie Cahysdie	4	CCTV tidak bisa dibuka min, tolong perbaiki...	1/12/2026 16:00
luthfy Daniel	5	Sangat mudah digunakan	12/24/2025 13:57

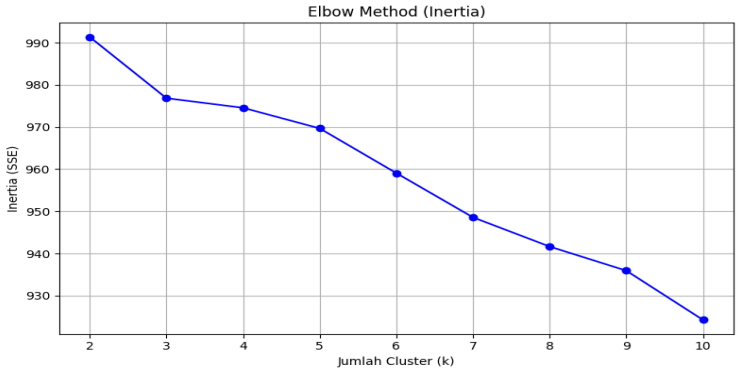
Gambar 1. Hasil Dataset Ulasan
Sumber: (Hasil Penelitian, 2026)

Data mentah tersebut kemudian melalui tahapan *text preprocessing* (pembersihan *missing value* dan duplikat, *case folding*, *cleansing*, *normalization*, *tokenizing*, *stopword removal*, *stemming*, dan *handling empty data*). Berdasarkan penanganan data kosong setelah proses penyaringan, diperoleh sebanyak 1.074 baris data bersih yang siap diproses ke tahap pembobotan.

Pembobotan TF-IDF dan Evaluasi Optimal

Data teks yang telah bersih selanjutnya ditransformasikan ke dalam representasi numerik menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF) dengan konfigurasi max_features=1000 dan min_df=2. Sebelum proses pembentukan klaster, penentuan jumlah klaster optimal (k) dianalisis menggunakan kombinasi *Elbow Method* dan *Silhouette Score*.

Berdasarkan evaluasi kurva penurunan nilai *inertia* pada *Elbow Method*, titik tekukan (*elbow point*) diidentifikasi pada rentang $k = 3$, sehingga ditetapkan $k = 3$ sebagai jumlah klaster optimal. Grafik pengujian *Elbow Method* ditunjukkan pada Gambar 2 berikut:



Gambar 2. Grafik Elbow Method
Sumber: (hasil penelitian. 2026)

Berdasarkan Gambar 2 di atas, grafik *Elbow Method* digunakan untuk mengukur nilai *inertia* atau *Within-Cluster Sum of Squares* (WCSS), yaitu total kuadrat jarak antara setiap data dengan titik pusat (*centroid*) pada klaster yang ditempatinya. Berdasarkan pengujian variasi nilai k dari 2 hingga 10, nilai *inertia* menunjukkan kecenderungan penurunan secara konsisten mulai dari 991,33 pada $k = 2$ hingga mencapai 924,28 pada $k = 10$ yang menandakan bahwa jarak data dengan *centroid*-nya menjadi semakin rapat.

Meskipun demikian, kurva penurunan nilai *inertia* tersebut mulai melandai secara bertahap pada rentang $k = 3$ hingga $k = 5$. Titik pelandaian ini menunjukkan bahwa penambahan jumlah klaster pada kisaran tersebut tidak lagi memberikan penurunan nilai *inertia* atau peningkatan kualitas pengelompokan yang signifikan. Berdasarkan pola kurva tersebut, ditetapkan $k = 3$ sebagai jumlah klaster optimal yang digunakan dalam proses klasterisasi ulasan pengguna aplikasi DSW.

Hasil Klasterisasi *K-Means* dan Analisis Karakteristik

Berdasarkan penerapan algoritma *K-Means Clustering* dengan $k = 3$, sebanyak 1.074 data ulasan berhasil dikelompokkan ke dalam tiga klaster. Sebaran spasial klaster beserta titik pusat (*centroid*) yang direduksi menggunakan *Principal Component Analysis* (PCA 2D) divisualisasikan pada Gambar 3 berikut:

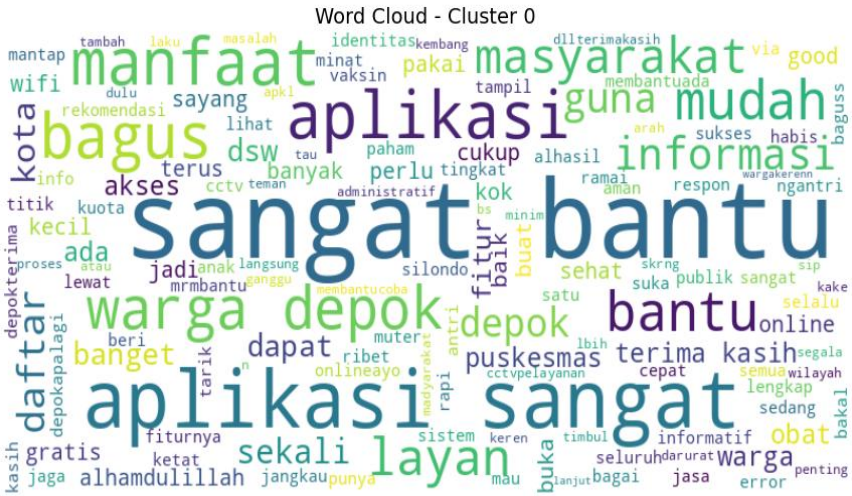
	cluster	jumlah_ulasan	rata_rata_rating
0	0	103	4.883495
1	1	661	3.199697
2	2	310	1.896774

Gambar 3. Distribusi Jumlah Ulasan tiap Klaster
Sumber: (Hasil Penelitian, 2026)

Pelabelan sentimen pada masing-masing klaster dilakukan melalui analisis manual (*post-processing*) berdasarkan interpretasi visualisasi *word cloud* serta analisis kualitatif terhadap kata-kata berbobot tertinggi dan nilai rata-rata *rating* pada setiap kelompok. Distribusi jumlah ulasan dan rata-rata *rating* pada masing-masing klaster dirangkum sebagai berikut:

1. **Klaster 0 (Sentimen Positif):** Memiliki 103 ulasan dengan rata-rata *rating* tertinggi sebesar 4,88. Kata-kata dominan seperti *"sangat bantu"* dan *"terima kasih"* mencerminkan apresiasi nyata pengguna terhadap manfaat fungsi layanan aplikasi DSW.
2. **Klaster 1 (Sentimen Netral / Informatif):** Merupakan kelompok terbesar yang mencakup 661 ulasan dengan rata-rata *rating* 3,20. Dominasi kata *"aplikasi"* dan *"layanan"* menunjukkan bahwa sebagian besar ulasan berfokus pada informasi umum seputar pemanfaatan layanan digital Pemkot Depok.
3. **Klaster 2 (Sentimen Negatif / Keluhan Teknis):** Memiliki 310 ulasan dengan rata-rata *rating* terendah yaitu 1,90. Klaster ini didominasi oleh keluhan teknis krusial seperti kegagalan pendaftaran *online*, kendala *login*, waktu muat (*loading*) yang lambat, hingga masalah akses CCTV (*"sulit"*, *"daftar"*, *"online"*, *"puskesmas"*, *"cctv"*, *"lambat"*).

Visualisasi *Word Cloud* yang menampilkan kata-kata dominan pada masing-masing klaster (seperti contoh untuk Klaster 0) ditunjukkan pada Gambar 4 berikut:





Gambar 4. Hasil Visualisasi Word Cloud per Klaster
Sumber: (Hasil Penelitian, 2026)

Pelabelan sentimen dan analisis karakteristik pada masing-masing klaster dilakukan melalui analisis manual (*post-processing*) berdasarkan interpretasi visualisasi *word cloud* serta analisis kualitatif terhadap kata-kata berbobot tertinggi dan nilai rata-rata *rating* pada setiap kelompok. Visualisasi *Word Cloud* yang menampilkan kata-kata dominan pada masing-masing klaster selengkapnya ditunjukkan pada Gambar berikut:

1. **Word Cloud - Cluster 0 (Sentimen Positif / Apresiasi Pengguna):** Visualisasi *Word Cloud* pada Gambar 4 (Cluster 0) didominasi oleh kata-kata dengan frekuensi tinggi seperti *"sangat"*, *"bantu"*, *"manfaat"*, *"mudah"*, dan *"bagus"*. Ukuran teks yang besar pada kata-kata tersebut menunjukkan bahwa kelompok ulasan ini berisikan ekspresi kepuasan, apresiasi, serta ulasan positif

dari pengguna yang merasakan kemudahan dan manfaat nyata dari fitur-fitur yang disediakan oleh aplikasi DSW.

2. **Word Cloud - Cluster 1 (Sentimen Netral / Informatif):** Visualisasi *Word Cloud* pada Cluster 1 memperlihatkan dominasi kata-kata seperti *"aplikasi"*, *"baik"*, *"dsw"*, *"depok"*, *"update"*, *"lancar"*, dan *"mudah"*. Klaster ini mencerminkan ulasan yang bersifat umum, informatif, atau netral, di mana pengguna banyak membahas mengenai pemanfaatan aplikasi secara fungsional sehari-hari serta pembaruan sistem yang dirasakan.
3. **Word Cloud - Cluster 2 (Sentimen Negatif / Keluhan Teknis):** Visualisasi *Word Cloud* pada Cluster 2 didominasi oleh kata-kata kunci keluhan yang sangat menonjol seperti *"daftar"*, *"online"*, *"puskesmas"*, *"mau"*, *"sulit"*, *"terus"*, *"padahal"*, dan *"klau"*. Ukuran teks yang besar pada kata-kata tersebut secara jelas menggambarkan bahwa kelompok ulasan ini berisi keluhan teknis yang berulang, terutama terkait kendala kegagalan saat melakukan pendaftaran *online* (seperti pendaftaran ke puskesmas), proses yang bertele-tele, serta masalah akses sistem yang menyulitkan pengguna.

SIMPULAN

Penelitian ini berhasil menganalisis dan mengelompokkan ulasan pengguna aplikasi Depok Single Window (DSW) pada Google Play Store menggunakan algoritma *K-Means Clustering* yang mengelompokkan 1.074 data ulasan bersih ke dalam tiga klaster optimal ($k = 3$). Hasil pengelompokan menunjukkan bahwa sentimen pengguna terbagi menjadi Klaster 0 sebagai kelompok sentimen positif (103 ulasan, rata-rata *rating* 4,88) yang memuat apresiasi pengguna, Klaster 1 sebagai kelompok sentimen netral (661 ulasan, rata-rata *rating* 3,20) mengenai penggunaan layanan umum, serta Klaster 2 sebagai kelompok sentimen negatif (310 ulasan, rata-rata *rating* 1,90) yang terfokus pada keluhan teknis. Temuan ini mengonfirmasi bahwa permasalahan utama pengguna aplikasi DSW terpusat pada kendala kegagalan pendaftaran *online* (khususnya pendaftaran puskesmas), proses *login*, waktu pemuatan data yang lambat, serta stabilitas akses CCTV. Implikasi dari temuan ini memberikan masukan strategis bagi pengembang aplikasi dan Pemkot Depok untuk memprioritaskan perbaikan teknis pada modul pendaftaran dan keandalan peladen (*server*) guna meningkatkan kualitas pelayanan publik digital ke depan.

DAFTAR PUSTAKA

Alfarizi, M. F., Putra, A. H., & Fathurramdan, M. D. (2026). Analisis Segmentasi Tingkat Kepuasan Pengguna Layanan Sistem Pemerintahan Berbasis Elektronik Pemerintah Kota Bandung Menggunakan Algoritma K-Means Clustering. *Jurnal Multidisiplin Saintek*, *10*. https://doi.org/10.8734/Kohesi.v1i2.365

Amalia, N., Royanti, N. I., Indrayanti, I., & Ismanto, B. (2023). Analisa Sentimen Pelanggan pada Review Belanja Online Berbasis Text Mining Menggunakan Metode K-Means. *Journal of Information System Research (JOSH)*, *4*(4), 1441–1447. https://doi.org/10.47065/josh.v4i4.3781

Ardiawan, R., Sari, Y. A., & Rahayudi, B. (2022). *Pengelompokan Ulasan Produk HP pada Marketplace Tokopedia menggunakan Metode Semi Supervised K-Means* (Vol. 6, Number 1). http://j-ptiik.ub.ac.id

Chandra, R., Djumiarti, T., & Afrizal, T. (2024). *Evaluasi Aplikasi Depok Single Window Dalam Penerapan Konsep Smart City Di Kota Depok*.

Daely, G., Sanjaya, A. E., & Wijaya, A. (2026). Analisis Pola Kepuasan Pengunjung Amanzi Waterpark Palembang Menggunakan Algoritma K-Means Clustering. *Jurnal Ilmu Komputer Dan Informatika*, *2*(3). https://jurnal.globalscients.com/index.php/jiki.930

Haidar, R. M., & Sulastri. (2024). *Klusterisasi Data Review Pengguna Aplikasi Marketplace Blibli.Com dengan Algoritma K-Means dan K-Medoids*.

Karomah, R. W., Sahid, A., Husein, A., Radhitya, F. R., Fathanah, N. N., & Gitasafira, I. I. (2025). Analisis Kualitas Pelayanan Publik Digital Pada Jogja Smart Service (Jss) Menggunakan Metode E-Govqual: Studi Literatur Berbasis Review Play Store. In *Jurnal of Public Administration and Governance Insights* (Vol. 1, Number 2).

Khairuna, R., Nurdin, & Ar Razi. (2025). Analisis Perbandingan Algoritma K-Means Dan K-Medoids Untuk Klusterisasi Teks Ulasan Pada Aplikasi Cookpad. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, *10*(2), 445–458. https://doi.org/10.36341/rabit.v10i2.6131

Munawar al, H. rahmat alief, Hidayat Rian Rahmat, & Nopriadi. (2025). Jurnal Ilmiah Manajemen Informasi dan Komunikasi Pengaruh Kualitas Layanan Aplikasi Jogja Smart Service terhadap Kepuasan Pengguna The Effect of The Quality of The Jogja Smart Service Application on User Satisfaction. *Jurnal Ilmiah Manajemen Informasi Dan Komunikasi*, *9*(2).

Novitasari, S., Razaq, F. M., Kamilah, Z., & Samudra, A. A. (2024). Evaluasi Efektivitas Depok Single Window (Dsw) Dalam Meningkatkan Kualitas Pelayanan Publik Di Kota Depok. *Jurnal Ilmiah Ilmu Administrasi Negara Dan Ilmu Komunikasi*, *10*, 72–81.

Radiatussiva, R., Nurazizah, D. P., Dwisukma, B. P., & Tinggi Teknologi Cipasung, S. (2026). *Klasterisasi Sentimen Pada Ulasan Aplikasi Line Bank Menggunakan Algoritma K-Means Clustering* (Vol. 3, Number 1).

Ramadhan, R., Sari, Y. A., & Adikara, P. P. (2021). *Perbandingan Pembobotan Term Frequency-Inverse Document Frequency dan Term Frequency-Relevance Frequency terhadap Fitur N-Gram pada Analisis Sentimen* (Vol. 5, Number 11). http://j-ptiik.ub.ac.id

santoso, lukman, & priyadi. (2024). *Mengoptimalkan Proses Pembersihan Data dalam Analisis Big Data Menggunakan Pipeline Berbasis AI*. *17*(2). https://doi.org/10.51903/elkom.v17i2.2311

Sunarko, G. S., Wasino,), & Sutrisno, T. (2023). *Klasterisasi Sentimen Ulasan Pengguna Aplikasi Bca Mobile pada Platform Google Play Store Dengan Algoritma K-Means Clustering*.

Syahkur, R. M., Hartama, D., & Solikhun, S. (2024). Evaluasi Jumlah Cluster pada Algoritma K-Means++ Menggunakan Silhouette dan Elbow dengan Validasi Nilai DBI dalam Mengelompokkan Gizi Balita. *JST (Jurnal Sains Dan Teknologi)*, *13*(3), 487–496. https://doi.org/10.23887/jstundiksha.v13i3.86419