

**PENGELOMPOKAN KEPUASAN PUBLIK
TERHADAP KEBIJAKAN DIGITALISASI
PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA *K-MEANS***

SKRIPSI SARJANA INFORMASI

Oleh

Arscelandy Daffa Ananta

227006516037



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2026

**PENGELOMPOKAN KEPUASAN PUBLIK
TERHADAP KEBIJAKAN DIGITALISASI
PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA *K-MEANS***

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Arscelandy Daffa Ananta

227006516037



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL**

2026

HALAMAN PENGESAHAN

TUGAS AKHIR

PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP KEBIJAKAN
DIGITALISASI PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA K-MEANS



Arscelandy Daffa Ananta

227006516037

Dosen Pembimbing 1

(Moch. Firmansyah, S. Kom., M.Kom)

Dosen Pembimbing 2

(Bayu Yasa Wedha S.T., M.Kom)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

**PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP KEBIJAKAN
DIGITALISASI PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA K-MEANS**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 27 Februari 2026



Arscelandy Daffa Ananta

NPM 227006516037

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

**PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP
KEBIJAKAN DIGITALISASI PELAYANAN PUBLIK DI
MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA
K-MEANS**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2025-2026 pada tanggal 24 Februari Tahun 2026

Dosen Pembimbing 1

Dosen Pembimbing 2

(Moch. Firmansyah, S. Kom., M.Kom)

(Bayu Yasa Wedha S.T., M.Kom)

NIDN: 0317128602

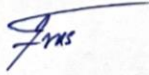
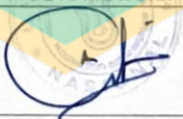
NIDN: 050023014

Ketua Program Studi

Dr. Ir. Andrianingsih, S.Kom., MMSI.

NIDN 10314106802

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama	: Arscelandy Daffa Ananta	
NPM	: 227006516037	
Fakultas/Akademi	: Fakultas Teknologi Komunikasi dan Informatika	
Program Studi	: Sistem Informasi	
Tanggal Sidang	: 24 Februari 2026	
JUDUL DALAM BAHASA INDONESIA :		
PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP KEBIJAKAN DIGITALISASI PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA K-MEANS		
JUDUL DALAM BAHASA INGGRIS :		
CLASSIFICATION OF PUBLIC SATISFACTION WITH PUBLIC SERVICE DIGITALIZATION POLICIES ON SOCIAL MEDIA TWITTER WITH ALGORITHM K-MEANS		
TANDA TANGAN DAN TANGGAL		
Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 2 Maret 2026	TGL : 2 Maret 2026	TGL : 2 Maret 2026
		

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Arscelandy Daffa Ananta

NPM : 227006516037

Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika

Program Studi : Sistem Informasi

Tanggal Sidang : 24 Februari 2026

JUDUL DALAM BAHASA INDONESIA :

PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP
KEBIJAKAN DIGITALISASI PELAYANAN PUBLIK DI
MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA
K-MEANS

JUDUL DALAM BAHASA INGGRIS :

CLASSIFICATION OF PUBLIC SATISFACTION WITH
PUBLIC SERVICE DIGITALIZATION POLICIES ON
SOCIAL MEDIA TWITTER WITH ALGORITHM K-MEANS

TANDA TANGAN DAN TANGGAL

Pembimbing 2	Ka. Prodi	Mahasiswa
TGL : 2 Maret 2026	TGL : 2 Maret 2026	TGL : 2 Maret 2026
		

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi dengan judul **“PENGELOMPOKAN KEPUASAN PUBLIK TERHADAP KEBIJAKAN DIGITALISASI PELAYANAN PUBLIK DI MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA *K-MEANS*”** sebagai salah satu syarat kelulusan Program Studi Sarjana Fakultas Teknologi Komunikasi dan Informatika. Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir, Bapak Moch. Firmansyah, S. Kom., M.Kom, dan Bapak Bayu Yasa Wedha S.T., M.Kom. yang telah meluangkan banyak waktu, tenaga, pikiran, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi.

Penulis juga mengucapkan banyak terima kasih kepada :

1. Bapak Dr. Agung Triayudi, S.Kom., M.Kom. selaku Dekan Fakultas Teknologi Komunikasi dan Informatika (FTKI) Universitas Nasional.
2. Ibu Dr. Ir. Andrianingsih, S.Kom., MMSI. selaku Ketua Program Studi Sistem Informasi Universitas Nasional.
3. Bapak Moch. Firmansyah, S. Kom., M.Kom. Selaku Dosen Pembimbing Penulis yang selalu memotivasi dan memberi arahan penulis.
4. Bapak Bayu Yasa Wedha S.T., M.Kom. Selaku Dosen Pembimbing Penulis yang selalu memotivasi dan memberi arahan penulis.
5. Kepada Bapak Andy Haryanto selaku ayah dan Didit Arsanti selaku ibu, atas doa, kasih sayang, serta dukungan moral dan motivasi yang tiada henti.
6. Teman-teman Segardu yang telah membantu dan mendukung, dan memotivasi penulis tiada henti.
7. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang memberikan banyak ilmu.

Akhir kata, semoga Tuhan Yang Maha Esa membalas kebaikan dan bantuan

yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga skripsi ini dapat memberikan manfaat di bidang Teknologi Informatika.

Jakarta, 12 Desember 2025

A handwritten signature in black ink, appearing to be "R. R. R.", is written above the name "Penulis".

Penulis

ABSTRAK

Transformasi digital dalam pelayanan publik melalui kebijakan digitalisasi dan penerapan Sistem Pemerintahan Berbasis Elektronik (SPBE) menjadi upaya pemerintah untuk meningkatkan efisiensi dan kualitas layanan kepada masyarakat. Namun, keberhasilan kebijakan tersebut juga perlu dievaluasi berdasarkan tingkat kepuasan publik sebagai pengguna layanan. Penelitian ini bertujuan untuk mengelompokkan opini masyarakat terhadap kebijakan digitalisasi pelayanan publik berdasarkan data media sosial Twitter. Data dikumpulkan melalui proses web scraping menggunakan beberapa kata kunci terkait digitalisasi pelayanan publik pada periode 1 April 2024 hingga 1 April 2025. Sebanyak 2.403 tweet diperoleh, kemudian dilakukan preprocessing dan filtering duplikasi sehingga dataset akhir berjumlah 2.007 tweet. Data teks selanjutnya direpresentasikan dalam bentuk numerik menggunakan metode pembobotan TF-IDF dan dianalisis menggunakan algoritma *K-Means* clustering. Penentuan jumlah cluster optimal dilakukan dengan metode Elbow, yang menunjukkan bahwa jumlah cluster terbaik adalah $k = 4$. Hasil clustering menghasilkan empat kelompok opini publik dengan karakteristik tema yang berbeda, mulai dari dukungan positif terhadap transformasi digital, pembahasan administratif terkait sistem pemerintahan elektronik, pengalaman teknis penggunaan layanan digital, hingga kritik terhadap proses layanan yang masih dianggap berbelit. Penelitian ini menunjukkan bahwa algoritma *K-Means* mampu digunakan untuk memetakan kepuasan publik secara otomatis dan terstruktur, serta dapat menjadi masukan bagi pemerintah dalam evaluasi dan peningkatan kualitas pelayanan publik digital.

Kata Kunci: Digitalisasi Pelayanan Publik, Kepuasan Publik, Twitter, Text Mining, TF-IDF, *K-Means* Clustering

ABSTRACT

Digital transformation in public services through digitalization policies and the implementation of the Electronic-Based Government System (SPBE) has become a major governmental effort to improve service efficiency and quality. However, the success of this policy must also be assessed based on public satisfaction as service users. This study aims to cluster public opinions regarding public service digitalization policies using Twitter social media data. The dataset was collected through web scraping using relevant keywords during the period from April 1, 2024, to April 1, 2025. A total of 2,403 tweets were obtained, and after preprocessing and duplicate filtering, the final dataset consisted of 2,007 tweets. The textual data were transformed into numerical representations using the TF-IDF weighting method and analyzed using the K-Means clustering algorithm. The optimal number of clusters was determined using the Elbow Method, indicating that the best value was $k = 4$. The clustering results produced four main groups of public opinions with different thematic characteristics, including positive support for digital transformation, administrative discussions on electronic government systems, technical experiences in accessing digital services, and criticism regarding complicated service procedures. This research demonstrates that the K-Means algorithm can effectively provide an automatic and structured mapping of public satisfaction, which may serve as valuable input for the government in evaluating and improving digital public services.

Keywords: Public Service Digitalization, Public Satisfaction, Twitter, Text Mining, TF-IDF, K-Means Clustering

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
ABSTRAK.....	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian	4
1.5.1 Manfaat Teoritis	4
1.5.2 Manfaat Praktis	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu.....	6
2.2 Landasan Teori	15
2.2.1 Digitalisasi Pelayan Publik	15
2.2.2 Kepuasan Public Terhadap Layanan Digital	15
2.2.3 Media Sosial Twitter	16
2.2.4 Text Mining.....	16

2.2.5	Text Preprocessing dan TF-IDF	16
2.2.6	Algoritma <i>K-Means</i> Clustering.....	17
2.2.7	Hubungan antara TF-IDF dan <i>K-Means</i>	18
2.2.8	Elbow Method.....	19
2.2.9	Web Scrapping	19
2.2.10	<i>Python</i>	19
BAB III METODOLOGI PENELITIAN.....		20
3.1	Lokasi Penelitian.....	20
3.2	Waktu Penelitian	20
3.3	Jenis Penelitian.....	21
3.4	Sumber Data.....	21
3.5	Jenis Data	22
3.6	Tekhnik Pengumpulan Data	22
3.7	Perangkat dan Tools Penelitian.....	22
3.8	Tahapan Penelitian	23
3.8.1	Proses Pengumpulan Data.....	24
3.8.2	Processing Data.....	25
3.8.3	Pembobotan TF-IDF Dan <i>K-Means Clustering</i>	26
3.8.4	Evaluasi Model.....	27
BAB IV HASIL DAN DISKUSI		28
4.1	Hasil Pengumpulan Data Scrapping Twitter	28
4.2	Hasil Text Preprocessing.....	28
4.2.1	<i>Case Folding</i>	29
4.2.2	Cleansing.....	31
4.2.3	Tokenizing	32

4.2.4	Stopword Removal.....	33
4.2.5	Stemming	34
4.2.6	Filtering Duplikasi	36
4.3	Pembobotan TF-IDF	37
4.4	Penentuan Jumlah Cluster (Metode Elbow).....	37
4.5	Proses Clustering dengan <i>K-Means</i>	39
4.6	<i>Interpretasi Hasil Clustering</i>	43
4.6.1	<i>Interpretasi Cluster 0</i>	44
4.6.2	<i>Interpretasi Cluster 1</i>	45
4.6.3	<i>Interpretasi Cluster 2</i>	46
4.6.4	<i>Interpretasi Cluster 3</i>	47
4.7	Pembahasan.....	50
4.7.1	Hubungan Hasil Clustering dengan Pengelompokan Kepuasan Publik	50
4.7.2	Interpretasi Pengelompokan Kepuasan Publik Berdasarkan Cluster	51
4.7.3	Pembahasan Pola Visualisasi yang Mendukung Hasil Cluster	52
4.7.4	Implikasi Hasil terhadap Evaluasi Kebijakan Digitalisasi	53
4.8	IMPLEMENTASI SISTEM	53
4.8.1	Use Case Diagram Sistem.....	54
4.8.2	Activity Diagram Sistem.....	54
4.8.3	Tampilan Halaman Utama Sistem.....	56
4.8.4	Pengaturan Parameter Sistem.....	56
4.8.5	Proses Preprocessing Data	57
4.8.6	TF-IDF Vectorization	58

4.8.7	Penentuan Jumlah Cluster dengan Elbow Method.....	58
4.8.8	Proses Clustering Menggunakan Algoritma <i>K-Means</i>	59
4.8.9	Ringkasan Implementasi Sistem	64
BAB V PENUTUP.....		66
5.1	Kesimpulan	66
5.2	Saran.....	67
Daftar PUSTAKA		69



DAFTAR GAMBAR

Gambar 3. 1 Diagram Alur Penelitian.....	23
Gambar 3. 2Diagram Alur Analisi Data	23
Gambar 3. 3gambar proses scraping data	24
Gambar 3. 4Proses Pembobotan TF-IDF dan Implementasi <i>K-Means</i>	25
Gambar 4. 1 Hasil Scraping	27
Gambar 4. 2 Hasil Proses <i>Case Folding</i> Data Tweet	28
Gambar 4. 3 Hasil Proses Cleansing Data Tweet.....	30
Gambar 4. 4 Hasil Proses Tokenizing Data Tweet.....	32
Gambar 4. 5 Hasil Proses stopwords Removal.....	33
Gambar 4. 6 Hasil Proses Stemming Data Tweet	34
Gambar 4. 7 Hasil Filtering Duplikasi Data Tweet.....	35
Gambar 4. 8 Hasil Transformasi TF-IDF.....	36
Gambar 4. 9Grafik Metode Elbow untuk Penentuan Jumlah Cluster Optimal.....	37
Gambar 4. 10 Hasil Pemberian Label Cluster pada Tweet	38
Gambar 4. 11 Distribusi Jumlah Tweet pada Setiap Cluster Hasil <i>K-Means</i>	38
Gambar 4. 12 Distribusi Jumlah Tweet pada Setiap Cluster.....	40
Gambar 4. 13 Diagram Pie Distribusi Tweet pada Setiap Cluster	41
Gambar 4. 14 Visualisasi 2D Hasil Clustering Menggunakan PCA	41
Gambar 4. 15 Visualisasi Kata Dominan pada Cluster 0	43
Gambar 4. 16 Visualisasi Kata Dominan pada Cluster 1	44
Gambar 4. 17 Visualisasi Kata Dominan pada Cluster 2	45
Gambar 4. 18 Visualisasi Kata Dominan pada Cluster 3	46
Gambar 4. 19 Use Case diagram sistem.....	51
Gambar 4. 20 Activity diagram sistem.....	53
Gambar 4. 21 Tampilan Halaman Utama Sistem.....	54
Gambar 4. 22 Pengaturan Parameter Sistem.....	54
Gambar 4. 23 Contoh Hasil Preprocessing sistem.....	55
Gambar 4. 24 Hasil TF-IDF Vectorization Sistem	56
Gambar 4. 25 Hasil Elbow Method untuk Menentukan Jumlah Cluster	57

Gambar 4. 26 Hasil Proses Clustering Menggunakan Algoritma <i>K-Means</i>	58
Gambar 4. 27 Distribusi Data pada Setiap Cluster	59
Gambar 4. 28 Visualisasi Cluster Menggunakan PCA	59
Gambar 4. 29 Kata Dominan dan Wordcloud pada Setiap Cluster.....	60
Gambar 4. 30 wordcloud cluster 0	60
Gambar 4. 31 wordcloud cluster 1	61
Gambar 4. 32 wordcloud cluster 2	61
Gambar 4. 33 wordcloud cluster 3	62
Gambar 4. 34 Fitur Unduh Hasil Pengelompokan.....	62



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	6
Tabel 3. 1 Tabel Waktu Penelitian.....	20
Tabel 4. 1 Perbandingan Teks Sebelum dan Sesudah <i>Case Folding</i>	28
Tabel 4. 2 Perbandingan Teks Sebelum dan Sesudah Cleansing	30
Tabel 4. 3 Contoh Hasil Tokenizing.....	32
Tabel 4. 4 Contoh Perbandingan Teks Sebelum dan Sesudah Stopword Removal.....	33
Tabel 4. 5 Perbandingan Teks Sebelum dan Sesudah Stemming	34
Tabel 4. 6 Jumlah Data Sebelum dan Sesudah Filtering Duplikasi	35
Tabel 4. 7 Distribusi Jumlah Tweet pada Setiap Cluster.....	39
Tabel 4. 8 Kategori Kepuasan Berdasarkan Interpretasi Cluster <i>K-Means</i>	47

