

Metode Support Vector Machine Untuk Analisis Sentimen Aplikasi *Threads* di *Google Play Store*

Dimas Triully Prasetyo¹, Atiqah Meutia Hilda²

dprasetyo963@gmail.com¹, atiqahmeutihilda@uhamka.ac.id²

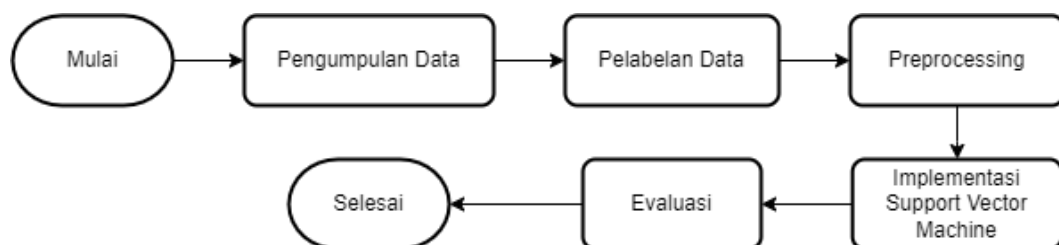
^{1,2} Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, DKI Jakarta, Indonesia

Informasi Artikel	Abstrak
Diterima : 16 Okt 2024 Direvisi : 23 Okt 2024 Disetujui : 30 Okt 2024	Penelitian ini berawal dari ketertarikan pengguna terhadap aplikasi <i>Twitter</i> terbaru dan aplikasi <i>Threads</i> di <i>Google Play Store</i> dengan tujuan utama menerapkan teknik <i>Support Vector Machine</i> (SVM) untuk menganalisa evaluasi pengguna terhadap sentimen, baik positif maupun negatif. Data yang terkumpul telah melalui proses preprocessing yang meliputi <i>cleaning</i> , <i>casefolding</i> , <i>tokenizing</i> , <i>filtering</i> , dan <i>stemming</i> . Setelah melalui <i>preprocessing</i> data sebanyak 1000 komentar diimplementasikan ke dalam <i>metode Support Vector Machine</i> menunjukkan 54,1% sentimen positif dan 45,9% sentimen negatif. Nilai akurasi sebesar 81.19% dan <i>confusion matrix</i> hasil dari nilai variabel TP (<i>True Positive</i>), TN (<i>True Negative</i>), FP (<i>False Positive</i>), dan FN (<i>False Negative</i>) juga ditampilkan oleh pengujian. Berdasarkan nilai akurasi yang diperoleh, dapat disimpulkan bahwa aplikasi <i>Threads</i> menerima umpan balik yang positif dan kemungkinan besar dapat ditingkatkan untuk memberikan hasil yang lebih baik lagi.
Kata Kunci	
Threads, Analisis Sentimen, Support Vector Machine, Google Play Store	
Keywords	Abstract
Threads, Sentiment Analysis, Support Vector Machine, Google Play Store	This research started with user interest in the latest Twitter app and the Google Play Store's Threads app. The primary goal was to apply the Support Vector Machine (SVM) technique to analyze user evaluations for sentiment, both positive and negative. The collected data went through a preprocessing process that included cleaning, casefolding, tokenizing, stop word removal, stemming, and filtering. After going through preprocessing data as many as 1000 comments were implemented into the Support Vector Machine method showing 54.1% positive sentiment and 45.9% negative sentiment. The accuracy value of 81.19% and the confusion matrix results from the values of the TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) variables are also displayed by the testing. Based on the accuracy value acquired, it can be inferred that the Threads application receives positive feedback and is likely to be improved to yield even better outcomes.

A. Pendahuluan

Media sosial sudah menjadi kebutuhan utama bagi seluruh kalangan di seluruh negara untuk mencari informasi dan berkomunikasi [1]. Salah satunya Indonesia, negara yang menempati posisi ketiga sebagai negara dengan penggunaan media sosial tertinggi pada tahun 2023 dengan rata-rata pengguna 8,4 jenis media sosial aktif per bulan [2]. Dari laporan “Digital2023Indonesia” oleh *Hoosuite (We are Social)*, pengguna media sosial aktif di Indonesia 60,4% dari total populasi yang berarti sekitar 167 juta pengguna media sosial dengan rata-rata setiap hari waktu pemakaian 3 jam. Media sosial akan terus berkembang dan tumbuh seiring dengan kemajuan teknologi digital, karena media sosial sudah memberikan dampak yang besar dalam mencari informasi dan komunikasi bagi seluruh dunia [3]. Belakangan ini ada salah satu media sosial terbaru yaitu *Threads* yang cukup populer bagi generasi muda saat ini. *Threads* adalah platform media sosial baru buatan Instagram yang diluncurkan oleh Mark Zuckerberg yaitu *CEO Meta* yang bertujuan untuk menyaingi *Twitter* [4]. Untuk mengakses *Threads* harus memiliki akun *Instagram* terlebih dahulu, secara tampilan dan fitur mirip dengan *Twitter* [5]. Dalam studi analisis sentimen yang menggunakan pendekatan *Naïve Bayes* dalam kaitannya dengan aplikasi *Threads* menghasilkan tingkat akurasi sebesar 76% cukup baik namun mungkin perlu dilakukan perbaikan [6]. Oleh karena itu, metode *Support Vector Machine* (SVM) dipilih untuk analisis sentimen dalam penelitian ini karena akurasinya yang lebih unggul dibandingkan dengan teknik-teknik alternatif. Dengan tingkat akurasi SVM sebesar 81,46% dan 75,41%, *Support Vector Machine* mengungguli *Naïve Bayes* untuk analisis sentimen pada data *Google Play Store* [7]. Kelebihan *Support Vector Machine* (SVM) adalah dapat menangani data besar untuk klasifikasi, terutama klasifikasi teks [8]. Diharapkan penelitian ini dapat membantu memenuhi pandangan para pengguna media sosial *Threads* dan pengembang untuk meningkatkan fitur dan fungsi pada aplikasi *Threads*. Karena tujuan penelitian berfokus pada menganalisis ulasan para pengguna *Threads* dan metode penelitian yang digunakan lebih baik daripada metode lainnya. Pembaca yang memiliki ketertarikan pada media sosial mungkin akan menemukan informasi dari penelitian ini berguna untuk berkontribusi dalam pengembangan aplikasi *Threads*.

B. Metode Penelitian



Gambar 1. Metodologi Penelitian

Untuk memproses data ke tahap file csv, urutan pertama dari metodologi penelitian ini adalah mengumpulkan data review *Google Play Store* pada aplikasi

Threads yang menggunakan metode web crawling [9]. Data ini selanjutnya diolah menggunakan *Google Collab*. Selanjutnya data yang berformat file csv diberi label positif atau negatif secara manual oleh peneliti, setelah dilakukan pelabelan sentimen lanjut melakukan *preprocessing*. Pada *preprocessing* akan dilakukan proses membenahi dan membersihkan data yang akan dianalisis sentimen [10]. Setelah semua proses di *preprocessing* lanjut, gunakan metode *Support Vector Machine* untuk menemukan *hyperline* terbaik [11]. Ini akan memisahkan tiap review ke dalam dua kelas, positif dan negatif, dan data ini akan digunakan untuk memprediksi data pengujian [12]. Data yang diimplementasikan akan dinilai dengan menghitung nilai *F1-Score*, akurasi, presisi, dan *recall* [13].

C. Hasil dan Pembahasan

1. Pengumpulan Data

Data ini dikumpulkan dari 1000 komentar pengguna pada aplikasi *Threads* di *Google Play Store*. Pengumpulan data dilakukan dengan *scrapping* menggunakan *tools Google Colab* dan bahasa pemrograman *Python* dari 27 Juni 2023 hingga 6 November 2023, dalam format CSV, seperti Gambar 2.



Gambar 2. Proses Pengumpulan Data Komentar Aplikasi Threads

2. Pelabelan Data

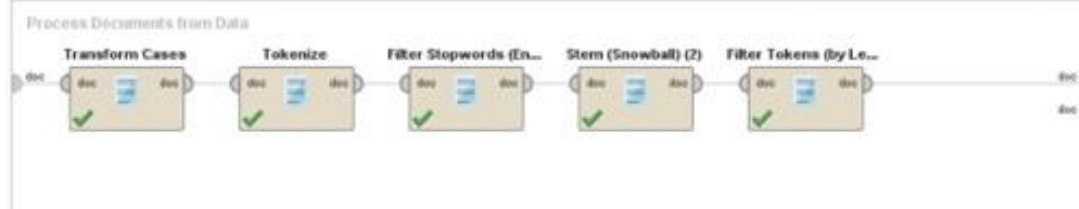
Setelah mengumpulkan data dan diubah ke dalam bentuk file csv, selanjutnya dilakukan pelabelan data secara manual di *Microsoft Excel* dengan memberi label sentimen positif negatif sesuai rating komentar dengan skala 1-2 sentimen negatif dan 3-4 setimen positif. Pada pelabelan didapat sebanyak 497 komentar negatif dan 503 komentar positif dari 1000 data komentar yang telah dikumpulkan seperti Tabel 1.

Tabel 1. Hasil Pelabelan Sesuai Rating Komentar

sentimen	comment	rate
Negatif	There's a "suggested Thread you might enjoy" in between Instagram posts. If I tap on it, it used to open Threads to that thread in the app. Now it takes me to the Play Store page for Threads instead, not the post in the app.	1
Negatif	it is just an extension of meta's monopoly on social media. and I haven't been able to change my bio in months. they didn't seem to put much thought into the app before releasing it.	1
Positif	4 me Threads is pretty cool. I don't know how 2 use it yet! or what's going on really per say, threads remains killer anywaysâ€Š.	4

3. Preprocessing

Kemudian dilakukan *preprocessing* pada data yang sudah dilabeli, *preprocessing* dilakukan menggunakan *RapidMiner Studio* [14]. Pada tahap ini menggunakan beberapa tahapan yang meliputi yaitu : *cleaning* untuk menghilangkan simbol-simbol, *case folding* memotong teks secara keseluruhan dalam dokumen, *tokenizing* membagi kumpulan karakter teks menjadi unit-unit kata; *filtering* mengekstrak kata-kata kunci, dan *stemming* mengurangi jumlah indeks yang berbeda yang ada dalam dokumen [15]. Tahapan ini menggunakan operator pada *RapidMiner Studio*, seperti Gambar 3 dibawah ini.



Gambar 3. Preprocessing

3.1. *Cleaning* untuk menghiangkan simbol-simbol seperti, mata uang, tanda baca, emoji, simbol matematika dan simbol lainnya.

Tabel 2. Hasil Proses *Cleaning*

Sebelum	Sesudah
Seriously? Is meta dumb, like, to the extent of its symbol? Why would I need insta to use this? Fond hope for increasing the number of insta users? Or overconfidence in the utility/usability of its other apps? Or a secret desire to mimick the junk of apple in creating a closed ecosystem in a software sphere? Despite the issues of copying another app that has gone to dogs recently? And what it is for? All I see on homepage is females doing stupid things onscreen while asking for a boyfriend!	Seriously Is meta dumb like to the extent of its symbol Why would I need insta to use this Fond hope for increasing the number of insta users Or overconfidence in the utilityusability of its other apps Or a secret desire to mimick the junk of apple in creating a closed ecosystem in a software sphere Despite the issues of copying another app that has gone to dogs recently And what it is for All I see on homepage is females doing stupid things onscreen while asking for a boyfriend

3.2. *Case Folding* merupakan proses mengubah kata menjadi semua huruf kecil atau huruf kapital.

Tabel 3. Hasil Proses *Case Folding*

Sebelum	Sesudah
As a neophyte More deliberate Help for Senior Citizens please	as a neophyte more deliberate help for senior citizens please

3.3. *Tokenizing* merupakan langkah yang memisahkan kata-kata dalam kalimat.

Tabel 4. Hasil Proses *Tokenizing*

Sebelum	Sesudah
---------	---------

<i>better than the alternative but still a cesspit of hate and anything meta is just dangerous software</i>	<i>['better', 'than', 'the', 'alternative', 'but', 'still', 'a', 'cesspit', 'of', 'and', 'hate', 'and', 'anything', 'meta', 'is', 'just', 'dangerous', 'software']</i>
---	--

3.4. *Filtering* dilakukan untuk menghilangkan kata atau singkatan dengan karakter minimal dan maksimal hurufnya sudah ditentukan.

Tabel 5. Hasil Proses *Filtering*

Sebelum	Sesudah
<i>['easy', 'to', 'and', 'fun']</i>	<i>['easy', 'to', 'and', 'fun']</i>

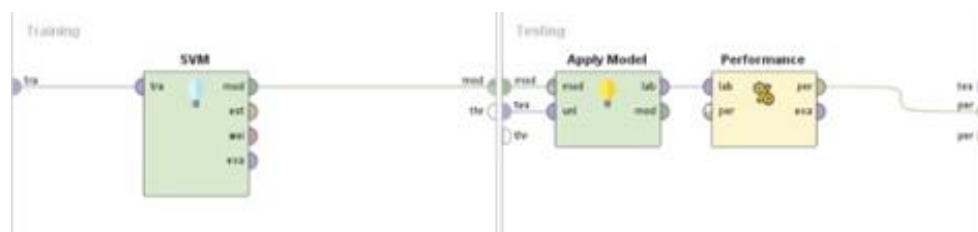
3.5. *Stemming* untuk mengubah kata-kata yang memiliki imbuhan menjadi kata dasarnya. Dalam melakukan analisis sentimen, proses ini membantu mengidentifikasi makna yang mendasari ulasan.

Tabel 6. Hasil Proses *Stemming*

Sebelum	Sesudah
<i>['stupid', 'popup', 'keeps', 'appearing', 'despite', 'having', 'the', 'app', 'installed', 'on', 'devices', 'when', 'using', 'through', 'ig']</i>	<i>['stupid', 'popup', 'keep', 'appear', 'despite', 'have', 'the', 'app', 'installed', 'on', 'device', 'when', 'use', 'through', 'ig']</i>

4. Implementasi *Support Vector Machine*

Tahap selanjutnya adalah implementasi menggunakan *Support Vector Machine*. Pada tahap ini akan melakukan percobaan data sebanyak 1000 data yang sudah melalui tahap preprocessing yang selanjutnya diimplementasikan dengan *Support Vector Machine* yang nanti mengeluarkan nilai *accuracy*, *precision*, dan *recall* [16]. Berikut proses penerapan *Support Vector Machine* pada gambar 4.

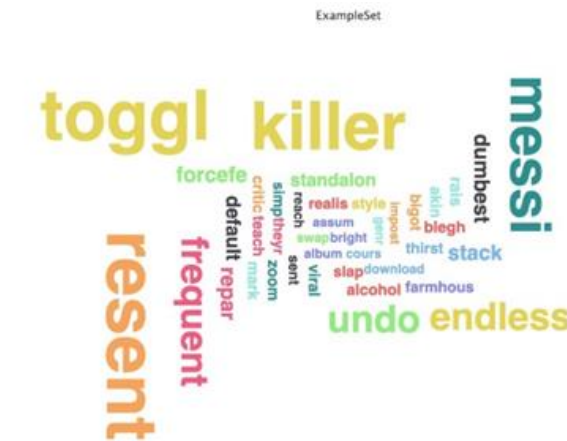


Gambar 4. Implementasi *Support Vector Machine*

Sesudah diimplementasi metode *Support Vector Machine* dengan membagi dua kelas menjadi positif dan negatif. Maka menghasilkan penelitian analisis sentimen dengan persentase sentimen positif 54,1% dan sentimen negatif 45,9% seperti pada gambar 5 dibawah ini.



Journal of Management Education 36(7) 809-824



Gambar 7. Worldcloud Negatif

Worldcloud Negatif pada Gambar 7 menampilkan kata-kata yang biasanya disebutkan secara negatif oleh pengguna *Threads* dalam evaluasi mereka. Pengguna cenderung lebih sering menggunakan sebuah kata sebagai titik fokus diskusi jika istilah tersebut muncul lebih besar di layar. Pengguna *Threads* biasanya memberikan ulasan positif tentang aplikasi seperti *visual*, *sleek*, *former*, *sensit*, dan sebagainya [14].

6. Evaluasi

Terakhir, melakukan evaluasi pada data *Support Vector Machine* yang sudah diimplementasikan, yang menghasilkan confusion matriks berisi nilai *accuracy*, *precision*, dan *recall*. Salah satu metode untuk menilai matriks kinerja model dengan membandingkan hasil kategorisasi adalah confusion matriks [20]. Nilai-nilai ini akan diterapkan pada aplikasi *Threads*, yang dapat dijangkau melalui *Google Play Store*, untuk analisis sentimen. Jumlah prediksi akurat yang dibuat oleh model ditunjukkan oleh nilai variabel TP (*True Positive*) dan TN (*True Negative*), sementara jumlah variabel FP (*False Positive*) dan FN (*False Negative*) menunjukkan jumlah prediksi yang tidak akurat yang dibuat oleh model. Ini adalah tanda-tanda *confusion matrix*. Seperti pada tabel 7.

Tabel 7. Hasil Confusion Matrix Support Vector Machine

	true Positif	true Negatif	class precision
pred. Positif	189	52	78.42%
pred. Negatif	21	126	85.71%
class recall	90.00%	70.79%	

D. Simpulan

Analisis sentimen dengan menggunakan metode *Support Vector Machine* terhadap aplikasi *Threads* di *Google Play Store* dengan dataset sebanyak 1000

komentar terdiri dari 497 sentimen negatif dan 503 sentimen positif dikumpulkan dari komentar pengguna aplikasi. Metode *Support Vector Machine* menemukan *hyperline* terbaik, yang membagi tiap review ke dalam dua kelas, positif dan negatif. Data ini akan digunakan untuk memprediksi data pengujian. Hasil pengujian dari *Support Vector Machine* menghasilkan nilai *accuracy* 81,19%, menunjukkan bahwa sentimen warga terhadap aplikasi *Threads* cenderung baik, dengan persentase sentimen positif 54.1% dan persentase sentimen negatif 45.9%. Hasil ini menunjukkan bahwa aplikasi *Threads* mendapat respon yang baik dari beberapa pengguna dan bagi para pengembang untuk meningkatkan demi kepuasan pengguna di masa depan. Penelitian ini diharapkan dapat membantu penelitian selanjutnya di platform media sosial lainnya supaya dapat memberikan analisis sentimen yang lebih akurat.

E. Referensi

- [1] W. Meliani and D. Gustian, "SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika) Universitas Nusa Putra, 12 Agustus 2023."
- [2] T. Hidayat and A. F. Nashirah, "Analisis Pengaruh Pemasaran Sosial Media pada Usaha Mikro , Kecil , dan Menengah di Kota Semarang," vol. 7, pp. 218–236, 2024.
- [3] M. Iqbal, A. Davy Wiranata, R. Suwito, and R. Faiz Ananda, "KLIK: Kajian Ilmiah Informatika dan Komputer Perbandingan Algoritma Naïve Bayes, KNN, dan Decision Tree terhadap Ulasan Aplikasi Threads dan Twitter," *Media Online*, vol. 4, no. 3, pp. 1799–1807, 2023, doi: 10.30865/klik.v4i3.1402.
- [4] F. Nufairi, N. Pratiwi, and F. Herlando, "Analisis Sentimen Pada Ulasan Aplikasi Threads Di Google Play Store Menggunakan Algoritma Support Vector Machine," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 1, pp. 339–348, Feb. 2024, doi: 10.29100/jipi.v9i1.4929.
- [5] A. Muhammad, F. Achmad, M. Basri, M. Reza, and A. Ansar, "Bidang: Teknik dan Manajemen Industri Topik: Manajemen Industri dan Kerekayasaan PENERAPAN TAM (TECHNOLOGY ACCEPTANCE MODEL) TERHADAP PENGGUNAAN APLIKASI THREADS PADA MAHASISWA TEKNIK INDUSTRI AGRO DI POLITEKNIK ATI MAKASSAR."
- [6] N. Samrin and M. N. Akbar, "Analisis Sentimen Komentar Pengguna Aplikasi Threads Pada Google Play Store Menggunakan Algoritma Multinomial Naïve Bayes," vol. 3, no. 2, p. 2023.
- [7] L. B. Ilmawan and M. A. Mude, "Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 154–161, Aug. 2020, doi: 10.33096/ilkom.v12i2.597.154-161.
- [8] Oryza Habibie Rahman, Gunawan Abdillah, and Agus Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 17–23, Feb. 2021, doi: 10.29207/resti.v5i1.2700.
- [9] F. Lim, A. Trista Ayunda Fredricklim, P. korespondensi, P. Studi Sistem Informasi, and F. Sains dan Teknologi, "Analisis Sentimen Masyarakat Indonesia Terhadap Aplikasi Threads di Twitter Menggunakan Naïve Baiyes," vol. 9, no. 2, pp. 55–64.

- [10] N. Suarna and W. Prihartono, "Analisis Sentimen Ulasan Aplikasi Threads Di Google Playstore Menggunakan Algoritma Naïve Bayes," 2024.
- [11] B. Sifa Amalia, Y. Umaidah, R. Mayasari, S. Karawang Jl HSRonggo Waluyo, K. Telukjambe Timur, and K. Karawang, "Analisis Sentimen Review Pelanggan Restoran Menggunakan Algoritma Support Vector Machine Dan K-Nearest Neighbor," vol. 19, no. 1, pp. 28–34, 2021.
- [12] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *Jurnal Informatika*, vol. 8, no. 2, pp. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.
- [13] F. T. Admojo and Y. I. Sulistya, "Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin," *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 1, 2022.
- [14] R. Nofitri and N. Irawati, "Analisis Data Hasil Keuntungan Menggunakan Software Rapidminer," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 5, no. 2, pp. 199–204, 2019, doi: 10.33330/jurteksi.v5i2.365.
- [15] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.
- [16] B. W. Sari and F. F. Haranto, "Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet," *Jurnal Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 171–176, Sep. 2019, doi: 10.33480/pilar.v15i2.699.
- [17] D. Amelia, A., U. Enri, and Y. U. Maidah, "Analisis Sentimen Ulasan Pengguna Klik Indomaret Pada Google Play Menggunakan Support Vector Machine," *INNOVATIVE: Journal Of Social Science Research*, vol. 3, pp. 2173–2185.
- [18] E. Suryati, A. Ari Aldino, N. Penulis Korespondensi, and E. Suryati Submitted, "Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM)," vol. 4, no. 1, pp. 96–106, 2023, doi: 10.33365/jtsi.v4i1.2445.
- [19] A. Meutia Hilda and M. Jafar Elly, "PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika Metode Klasifikasi Support Vector Machine (SVM) Untuk Analisis Sentimen Aplikasi Bing: Chat with AI & GPT-4 Di Google Play Store", doi: 10.33322/petir.v17i1.2283.
- [20] U. Brawijaya, F. T. Berton, D. E. Ratnawati, and A. Rahman, "Fakultas Ilmu Komputer Perbandingan Naïve Bayes Dan K-Nearest Neighbor Untuk Analisis Sentimen Terhadap Ulasan Aplikasi Threads," 2017. [Online]. Available: <http://j-ptiik.ub.ac.id>