

Sentiment Analysis of Threads Application Users Using Support Vector Machine Method

Nurul Alam¹, Mahdi^{2*}, Musta'inul Abdi³

^{1,2,3} Jurusan Tekniknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe Jln. B.Aceh Medan Km.280 Buketrata 24301
INDONESIA

*Penulis Korespondensi : mahdi@pnl.ac.id

INFORMASI ARTIKEL

Riwayat artikel:

Diajukan pada 18 Mei 25

Direvisi pada 02 Juni 25

Publikasi pada 20 Juni 25

Kata kunci:

Media Sosial

Threads

Analisis Sentimen

SVM

Keywords:

Social Media

Threads

Sentiment Analyst

SVM

ABSTRAK

Aplikasi Threads merupakan platform media sosial yang diluncurkan oleh Instagram, di mana pengguna dapat berkomunikasi, terhubung, dan berdiskusi seputar berbagai topik. Dengan pertumbuhan pengguna media sosial yang pesat di Indonesia, aplikasi Threads menarik perhatian sebagai platform baru dengan lebih dari 100 juta pengguna. Secara umum, Threads memiliki antarmuka dan algoritma yang mirip dengan Twitter, yang fungsinya mawadahi pengguna untuk membagikan unggahan berbasis teks, foto, dan video. Ulasan yang tersedia di Google Play Store sangat banyak untuk dianalisis secara manual, sehingga dilakukannya analisis sentimen untuk mengklasifikasi ulasan tersebut. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Threads menggunakan metode *Support Vector Machine* (SVM) terhadap ulasan pengguna di Google Play Store, dengan fokus pada ulasan berbahasa Indonesia untuk mengklasifikasikan opini menjadi positif, netral, dan negatif. Sebanyak 200 data ulasan dengan hasil 104 ulasan positif, 75 ulasan netral, dan 21 ulasan negatif. Hasil klasifikasi sentimen menggunakan metode SVM menunjukkan tingkat akurasi sebesar 80%. Akurasi ini menunjukkan bahwa metode SVM mampu secara efektif mengklasifikasikan sentimen pengguna terhadap aplikasi Threads. Hasil menunjukkan bahwa aplikasi Threads mendapat tanggapan positif dari mayoritas pengguna, yang menunjukkan bahwa aplikasi ini layak digunakan dan diterima dengan baik sebagai platform media sosial baru.

ABSTRACT

The Threads app is a social media platform launched by Instagram, where users can communicate, connect, and discuss various topics. With the rapid growth of social media users in Indonesia, Threads has gained attention as a new platform with over 100 million users. Generally, Threads features an interface and algorithm similar to Twitter, allowing users to share text-based posts, photos, and videos. Given the vast number of reviews available on the Google Play Store, manual analysis is impractical, hence sentiment analysis is conducted to classify these reviews. This study aims to analyze user sentiment toward the Threads app using the Support Vector Machine (SVM) method on user reviews in the Google Play Store, focusing on Indonesian-language reviews to classify opinions into positive, neutral, and negative. A total of 200 review samples were analyzed, yielding 104 positive reviews, 75 neutral reviews, and 21 negative reviews. The sentiment classification results using the SVM method achieved an accuracy of 80%. This accuracy indicates that the SVM method effectively classifies user sentiments toward the Threads app. The results show that Threads receives a positive response from the majority of users, suggesting that the app is well-received and acceptable as a new social media platform.

1. Pendahuluan

Media sosial telah menjadi platform penting bagi orang untuk berbagi pandangan, pengalaman, dan pendapat. Di Indonesia, jumlah pengguna media sosial terus meningkat pesat. Salah satu aplikasi terbaru yang diperkenalkan adalah Threads, yang dirilis oleh Instagram pada Juli 2023. Threads menawarkan fitur yang mirip dengan Twitter, di mana pengguna dapat berbagi teks, foto, dan video [1]. Meski popularitasnya melonjak dengan lebih dari 100 juta pengguna [2], ulasan pengguna yang beragam sulit dianalisis secara manual. Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen pengguna Threads menggunakan metode Support Vector Machine (SVM) [3].

Analisis sentimen adalah teknik untuk mengidentifikasi dan mengkategorikan opini atau emosi dalam teks menjadi sentimen positif, negatif, atau netral [4]. *Text mining* merupakan proses analisis dalam data yang berupa teks dimana sumber data didapatkan dari dokumen [5]. *Preprocessing* merupakan proses mengubah *dataset* kedalam bentuk yang mudah dipahami dengan tujuan untuk menghindari data yang rusak [3]. Tahapan dalam *preprocessing* berupa *cleansing*, *case folding*, *stopword removal*, dan *stemming* [6]. *Natural Language Processing* (NLP) merupakan *Artificial Intelligence* (AI) yang berfokus pada pengolahan bahasa alami atau bahasa yang digunakan oleh manusia. NLP digunakan untuk mencari kata dasar pada kalimat yang dimasukkan oleh user [7]. SVM yaitu salah satu metode klasifikasi dengan menggunakan *machine learning* yang memprediksi kelas berdasarkan model. Algoritma SVM bertujuan untuk mencari *Maximum Marginal Hyperplane* (MMH) menggunakan *support vector* dan *margin* [8]. Berikut rumus perhitungan *hyperplane* [9]:

$$f : w \cdot x + b = 0 \quad (1)$$

TF-IDF merupakan prosedur untuk memberikan nilai term pada dokumen dengan menggunakan pembobotan untuk mengklasifikasikan dokumen [10].

1. *Term Frequency* (TF) menghitung frekuensi jumlah kemunculan kata pada sebuah dokumen. Karena panjang dari setiap dokumen berbeda-beda, maka:

$$tf_{ij} = \frac{f_d(i)}{\max f_d(i)} \quad (2)$$

2. *Inverse Document Frequency* (IDF) berfungsi mengurangi bobot suatu term jika kemunculannya banyak tersebar di seluruh dokumen. Dengan kata lain IDF akan menilai kata yang sering muncul sebagai kata yang kurang penting berdasarkan kemunculan kata tersebut pada seluruh dokumen, semakin kecil nilai pada IDF maka semakin tidak penting kata tersebut.

$$idf(t, D) = \log \left(\frac{N}{df(t)+1} \right) \quad (3)$$

Salah satu metode yang banyak digunakan dalam analisis teks adalah SVM karena efektivitasnya dalam menangani volume data besar dan menghasilkan klasifikasi yang akurat. Penelitian sebelumnya menunjukkan bahwa SVM telah berhasil diterapkan dalam analisis ulasan aplikasi dan media sosial, dengan tingkat akurasi yang memuaskan dalam mengklasifikasikan opini pengguna.

Penelitian ini bertujuan untuk mengkaji sentimen pengguna aplikasi Threads melalui ulasan yang diambil dari Google Play Store. Metode SVM digunakan untuk mengelompokkan opini pengguna ke dalam kategori positif, negatif, atau netral. Selain itu, penelitian ini juga mengevaluasi tingkat akurasi metode SVM dalam mengklasifikasikan ulasan Threads. Diharapkan hasil penelitian ini dapat memberikan

wawasan yang bermanfaat bagi pengembang aplikasi dan calon pengguna mengenai persepsi terhadap aplikasi tersebut.

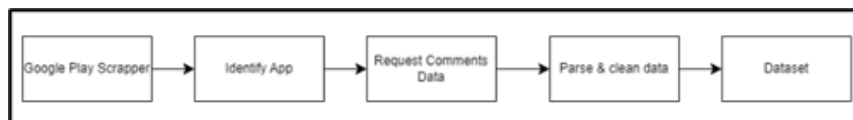
2. Metode

2.1 Sumber Data

Sumber data yang diperoleh untuk analisis sentimen didapat dari ulasan pengguna aplikasi Threads pada Google Play Store. Adapun pada penelitian ini hanya menggunakan satu cara yaitu dengan *scrapping data*. *Scrapping data* dilakukan dengan menggunakan *web scrapping*

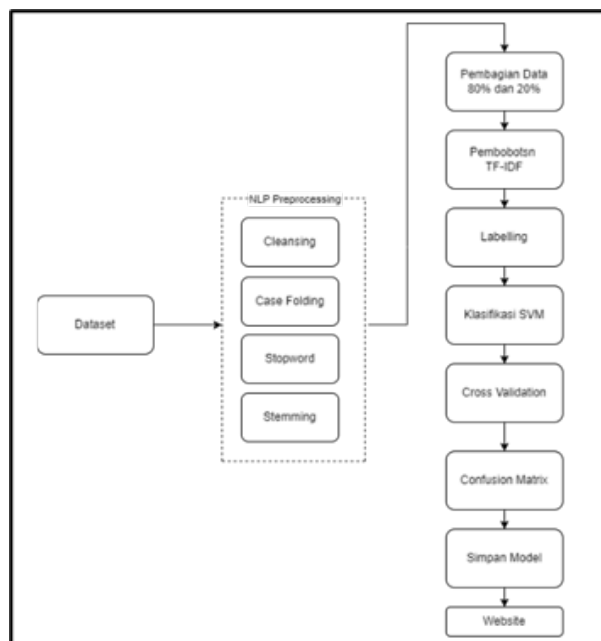
2.2 Tahapan Pengumpulan Data

Pengumpulan data didapat dari teknik *web scrapping* pada aplikasi Threads di Google Play Store berupa data ulasan pengguna yang nantinya akan dilakukan untuk membuat model prediksi. Gambar 1 memperlihatkan alur pengambilan data pada Google Play Store. Untuk proses *scrapping* diawali *setup libraries* menggunakan *google-play-scrapper*, untuk mengumpulkan data aplikasi dari Play Store. Tahap selanjutnya Identify App dengan menggunakan ID Aplikasi, kemudian menggunakan *library request comments data* untuk mengambil data yang diinginkan, selanjutnya *parse data* untuk menyusul ulang data agar lebih terstruktur. Setelah itu data yang sudah diproses disimpan dalam format .csv sebagai *dataset*.



Gambar 1. Tahapan *Scrapping*

2.3 Arsitektur Umum



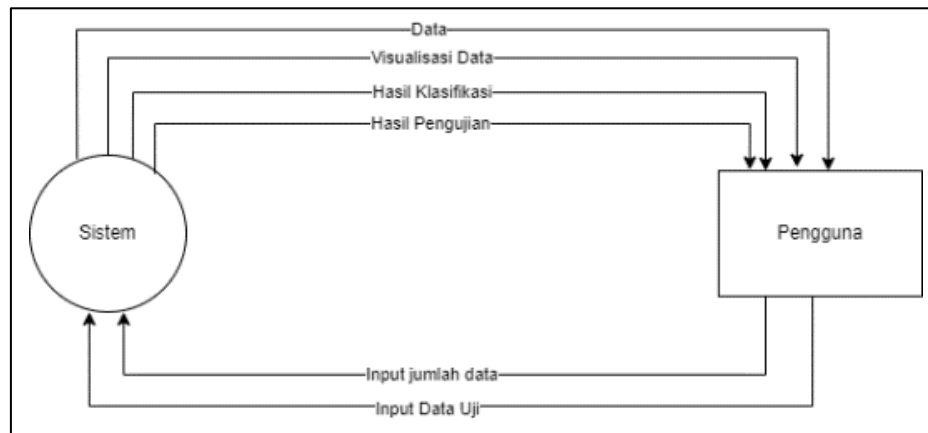
Gambar 2. Arsitektur Umum

Gambar 2 merupakan tahapan umum perancangan *machine learning*, dibutuhkan *dataset* untuk proses *training* dan *testing*. Tahapan selanjutnya adalah *preprocessing* dengan beberapa tahap yaitu *cleansing*, tokenizing, *stopword removal*, dan *stemming*. Proses selanjutnya pembobotan kata menggunakan metode TF-

IDF untuk melihat frekuensi kemunculan kata pada proses analisis. Tahapan analisis sentimen selanjutnya menggunakan metode SVM dengan menggunakan pengujian *K-Fold Cross Validation* dan *Confusion Matrix*. Kemudian model akan disimpan dan menampilkan hasil pada *website* yang sudah dibangun.

2.4 Diagram Konteks

Diagram Konteks memberikan gambaran umum tentang sistem secara keseluruhan. Diagram konteks menggambarkan aliran fungsional dalam sebuah proses pada sistem. Gambar 3 merupakan gambaran aliran yang menjelaskan ruang lingkup sistem secara garis besar. Pengguna dapat mengambil data review dengan jumlah yang diinginkan dan dapat meng-*input keyword*. Sistem dapat memberikan data dan hasil klasifikasi dari pengguna, Sistem juga memberikan pengguna terkait informasi berupa visualisasi data

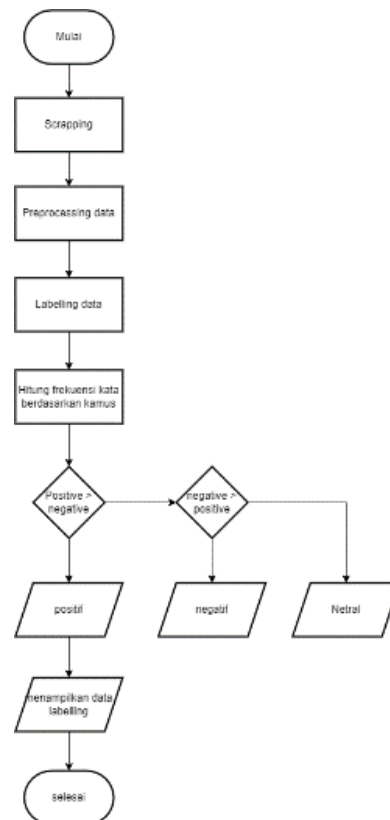


Gambar 3. Konteks Diagram

2.5 Rancangan Metode Penelitian

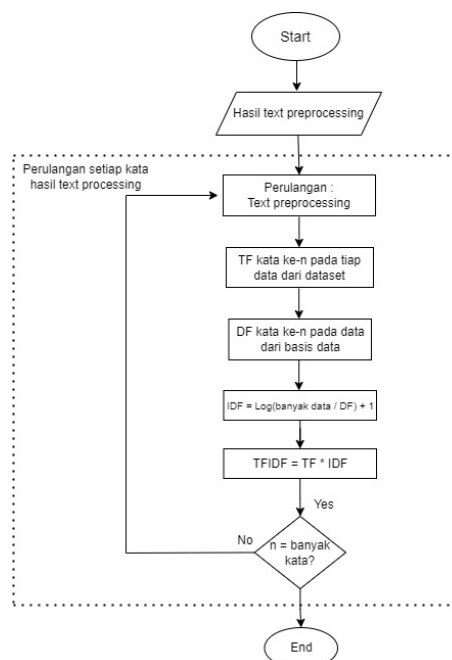
2.5.1 Labelling Data

Gambar 4 menjelaskan tahapan proses *labelling* data. Diawali dengan melakukan *scrapping* data dengan mengambil fitur *content*, kemudian dilanjutkan dengan proses *preprocessing* data. Pada *preprocessing* data melibatkan proses *cleansing*, *casefolding*, *stopword removal*, dan *stemming* data. Setelah dilakukan *preprocessing* data dilanjutkan dengan proses *labelling* data dengan menggunakan kamus yang disediakan pada <https://github.com/prasastoadi/ID-OpinionWords> yang berisikan data sentimen positif dan negatif. Data positif dan negatif akan dihitung jumlahnya apakah mengandung nilai positif, negatif, dan netral kemudian akan menampilkan hasil data setelah *labelling*.



Gambar 4. *Labeling Data*

2.5.2 Pembobotan TF-IDF



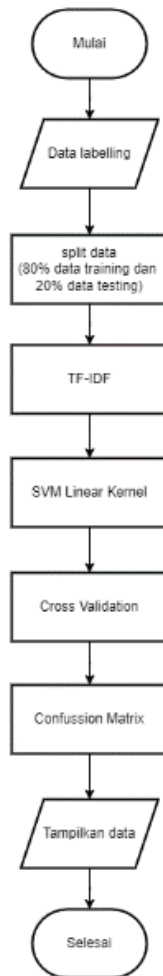
Gambar 5. *Pembobotan TF-IDF*

Gambar 5 menjelaskan tentang proses TF-IDF Setelah melalui tahap *preprocessing* dihasilkan *dataset* yang telah memiliki *token* pada masing-masing kalimat. Kemudian akan dihitung frekuensi kata terhadap

sebuah kalimat dari seluruh komentar. Berdasarkan proses tersebut menghasilkan nilai TF dan IDF. Kedua nilai tersebut kemudian dikalikan untuk memperoleh nilai bobot TF-IDF.

2.5.3 Klasifikasi SVM

Gambar 6 menjelaskan tahapan proses klasifikasi menggunakan metode SVM. Diawali dengan memasukkan data setelah *labelling*. Kemudian bagi data menjadi 80% data latih dan 20% data uji. Kemudian data akan dilakukan pembobotan TF-IDF. Setelah itu dilanjutkan dengan menginisialisasi model SVM menggunakan *kernel linear*. Setelah dilakukan pemodelan, model akan dilakukan pengujian akurasi menggunakan *Confussion Matrix*. Kemudian akan menampilkan hasil klasifikasi dan hasil pengujian.



Gambar 6. Klasifikasi SVM

3. Hasil Dan Pembahasan

3.1 Implementasi Proses

Implementasi proses menjabarkan prosedur dan tahapan yang terlibat berdasarkan perancangan model *machine learning* yang telah dijabarkan:

3.1.1 Dataset

Gambar merupakan *dataset* yang sudah dilakukan proses *scrapping* menggunakan *library web scrapping*.

at	score	content
2024-07-17 00:13:37	5	kgm yusak 😊
2024-07-16 21:46:37	5	pengguna baru
2024-07-16 19:37:33	5	Tambahkan kualitas ultra tinggi video dengan blu ray 4k 🤖🤖🤖🤖🤖
2024-07-16 19:18:49	4	masih belum mengerti jd baru mencoba nya
2024-07-16 18:42:23	3	Coba ditambahkan fitur bisa dikasih lagu gitu di postingan yang kita mau
2024-07-16 16:25:12	5	Oke juga
2024-07-16 15:26:46	3	assalamualaikum cukup puas dan merasa lebih nyaman
2024-07-16 12:00:07	1	Aplikasi sampah. Maksa buat instal dengan pancingan postingan Instagram
2024-07-16 11:03:21	5	Oke juga
2024-07-16 10:59:45	5	seru

Gambar 7. *Dataset*

3.1.2 *Cleansing*

content	cleansing
ok	ok
Baru saja gabung di Threads sementara masih bingung untuk n	baru saja gabung di threads sementara masih bingung
bagus banget sukaaa 🥰❤️	bagus banget sukaaa
overall good, gw sih minta adakan fitur list kyk sebelah agar kor	overall good gw sih minta adakan fitur list kyk sebelah
Mantap betul	mantap betul
terbaik Aplikasi ini Threads semoga kedepannya bisa lebih baik	terbaik aplikasi ini threads semoga kedepannya bisa lebih
lumayan baik hampir mirip dengan Twitter..	lumayan baik hampir mirip dengan twitter
top	top
Gabisa ganti foto profil	gabisa ganti foto profil
keren	keren

Gambar 8. *Cleansing*

Gambar 8 merupakan hasil *cleansing* data berfungsi untuk menghilangkan url, angka, simbol yang tidak diperlukan.

3.1.3 *Tokenizing / Case Folding*

cleansing	token_casefolding
ok	ok
baru saja gabung di threads sementara masih bingung un	baru saja gabung di threads sementara masih bingung
bagus banget sukaaa	bagus banget sukaaa
overall good gw sih minta adakan fitur list kyk sebelah aga	overall good gw sih minta adakan fitur list kyk sebelah
mantap betul	mantap betul
terbaik aplikasi ini threads semoga kedepannya bisa lebih	terbaik aplikasi ini threads semoga kedepannya bisa lebih
lumayan baik hampir mirip dengan twitter	lumayan baik hampir mirip dengan twitter
top	top
gabisa ganti foto profil	gabisa ganti foto profil
keren	keren

Gambar 9. *Tokenizing / Case Folding*

Gambar 9 merupakan hasil *tokenizing* and *case folding* berfungsi untuk menyamaratakan huruf menjadi huruf kecil dan memisahkan kata berupa spasi.

3.1.4 Stopword Removal

token_casfolding	stopword
ok	ok
baru saja gabung di threads sementara masih bingung un	gabung threads bingung postingan
bagus banget sukaaa	bagus banget sukaaa
overall good gw sih minta adakan fitur list kyk sebelah ag	overall good gw sih adakan fitur list kyk
mantap betul	mantap
terbaik aplikasi ini threads semoga kedepannya bisa lebih	terbaik aplikasi threads semoga kedepa
lumayan baik hampir mirip dengan twitter	lumayan twitter
top	top
gabisa ganti foto profil	gabisa ganti foto profil
keren	keren

Gambar 10. *Stopword Removal*

Gambar 10 merupakan hasil *Stopword Removal*, data yang dipakai yaitu bahasa Indonesia yang sesuai dengan KBBI.

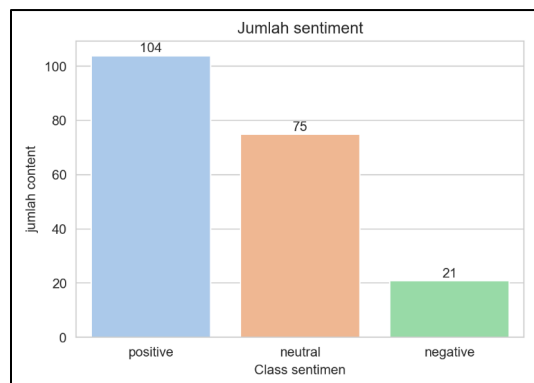
3.1.5 Stemming Data

stopword	stemming_data
ok	ok
gabung threads bingung postir	gabung threads bingung postingan
bagus banget sukaaa	bagus banget sukaaa
overall good gw sih adakan fitu	overall good gw sih adakan fitur list k
mantap	mantap
terbaik aplikasi threads semog	baik aplikasi threads moga depan
lumayan twitter	lumayan twitter
top	top
gabisa ganti foto profil	gabisa ganti foto profil
keren	keren

Gambar 11. *Stemming Data*

Gambar 11 merupakan hasil *stemming data* untuk mereduksi kata-kata yang terpisah menjadi utuh dalam satu kalimat.

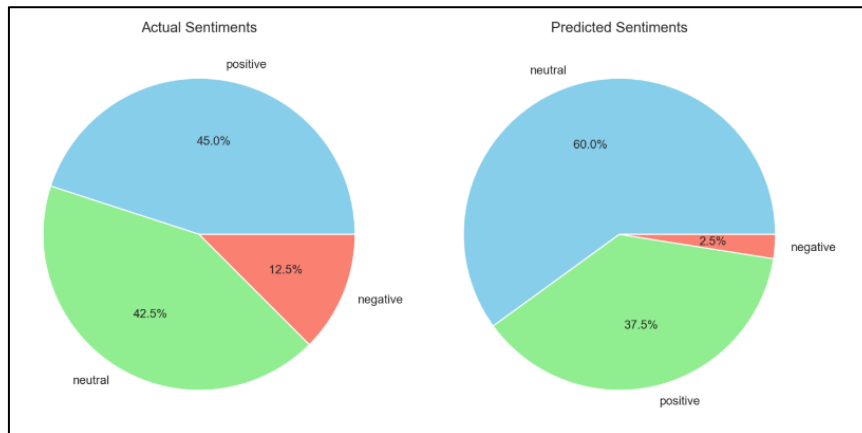
3.1.6 Hasil Labelling



Gambar 12. Hasil *Labelling*

Gambar 12 merupakan visualisasi data setelah proses *labelling*. Berdasarkan gambar 12 menunjukkan jumlah sentimen berlabel positif sebanyak 104 data, negatif 21 data dan netral 75 data.

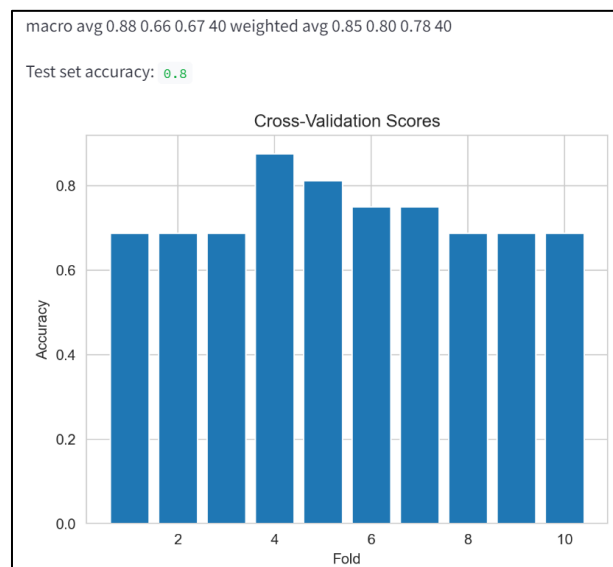
3.1.7 Hasil Implementasi Metode SVM



Gambar 13. Hasil Implementasi Metode SVM

Gambar 13 merupakan visualisasi data setelah dilakukan klasifikasi menggunakan metode SVM. Terdapat dua buah diagram lingkaran, yaitu *actual sentiment* dan *predicted sentiment*. Pada *actual sentiment* menjelaskan bahwa jumlah data positif memiliki persentase sebanyak 45%, jumlah data negatif memiliki persentase sebanyak 12,5%, dan jumlah persentase netral memiliki persentase sebanyak 42,5%. Sedangkan untuk *predicted sentiment* menjelaskan bahwa jumlah data positif memiliki persentase sebanyak 37,5%, jumlah data negatif memiliki persentase sebanyak 2,5%, dan jumlah persentase netral memiliki persentase sebanyak 60%.

3.1.8 Hasil Implementasi Metode SVM



Gambar 14. Hasil Pengujian *K-Fold Cross Validation*

Gambar 14 merupakan hasil hasil pengujian kinerja metode SVM menggunakan *cross validation*. Terdapat 10 iterasi pengujian yang mana setiap iterasi menghasilkan akurasi masing-masing berdasarkan baris data yang diuji, secara keseluruhan diperoleh akurasi rata-rata sebesar 80%.

3.1.9 Hasil Pengujian *Confussion Matrix*

Confusion matrix:

0	1	2
1	3	1
0	17	0
0	4	14

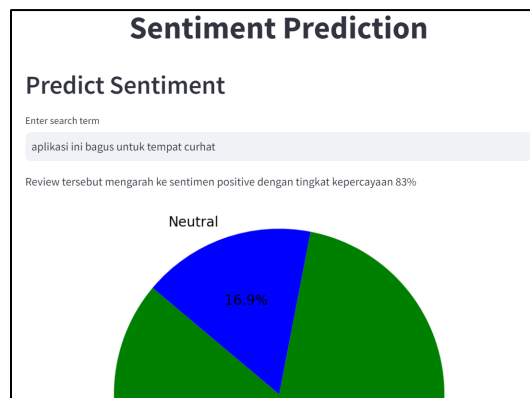
Classification report: precision recall f1-score support

negative	1.00	0.20	0.33	5
neutral	0.71	1.00	0.83	17
positive	0.93	0.78	0.85	18
accuracy			0.80	40

Gambar 15. Hasil Pengujian *Confussion Matrix*

Gambar 15 di atas, menunjukkan hasil evaluasi model SVM menggunakan *confussion matrix*. memperoleh akurasi klasifikasi sebesar 80% dari total data. Pengklasifikasi memiliki kinerja terbaik dalam memprediksi kelas positif dengan presisi 0,93 dan *recall* 0,78 sehingga menghasilkan skor F1 sebesar 0,85. Performa kelas netral tergolong sedang, dengan presisi 0,71 dan *recall* 1,0 sehingga menghasilkan skor F1 sebesar 0,83. Pengklasifikasi mengalami kesulitan besar pada kelas negatif, dengan presisi 1,0 dan *recall* 0,20 sehingga menghasilkan skor F1 sebesar 0,33. Matrix rata-rata makro dan rata-rata tertimbang menunjukkan bahwa meskipun pengklasifikasi memiliki keseimbangan kinerja yang wajar di semua kelas, masih ada ruang untuk perbaikan, terutama dalam menangani kelas negatif.

3.1.10. Hasil Pengujian Prediksi



Gambar 16. Hasil Pengujian Prediksi

Gambar 16 menunjukkan hasil prediksi sentimen, pengguna dapat meng-*input* kalimat apa pun dan sistem akan melakukan prediksi dengan melakukan proses klasifikasi SVM dan menampilkan visualisasi berupa diagram.

4. Kesimpulan

Penelitian ini berhasil mengklasifikasikan sentimen ulasan pengguna aplikasi Threads di Google Play Store ke dalam tiga kategori yaitu, positif, negatif, dan netral. Pada *actual sentiment* menjelaskan bahwa jumlah data positif memiliki persentase sebanyak 45%, jumlah data negatif memiliki persentase sebanyak

12,5%, dan jumlah persentase netral memiliki persentase sebanyak 42,5%. Sedangkan untuk *predicted sentiment* menjelaskan bahwa jumlah data positif memiliki persentase sebanyak 37,5%, jumlah data negatif memiliki persentase sebanyak 2,5%, dan jumlah persentase netral memiliki persentase sebanyak 60%.

Hasil penelitian menunjukkan bahwa metode SVM memberikan hasil akurasi yang baik dalam mengklasifikasikan sentimen. Dengan jumlah 200 data yang diambil memperoleh hasil pengujian menggunakan *K-Fold Cross Validation* memperoleh nilai akurasi sebesar 80% dan hasil evaluasi metode menggunakan *confusion matrix* memperoleh akurasi klasifikasi sebesar 80% dari total data. Pengklasifikasi memiliki kinerja terbaik dalam memprediksi kelas positif dengan presisi 0,93 dan *recall* 0,78 sehingga menghasilkan skor F1 sebesar 0,85. Performa kelas netral tergolong sedang, dengan presisi 0,71 dan *recall* 1,0 sehingga menghasilkan skor F1 sebesar 0,83. Pengklasifikasi mengalami kesulitan besar pada kelas negatif, dengan presisi 1,0 dan *recall* 0,20 sehingga menghasilkan skor F1 sebesar 0,33. Dapat disimpulkan bahwa metode Support Vector Machine cocok diterapkan untuk analisis sentimen pengguna aplikasi Threads di Google Play Store.

Referensi

- [1] W. Meliani and D. Gustian, "Analisis Sentimen Opini Publik Terhadap Aplikasi Threads Di Twitter Menggunakan Metode Naïve," 2023.
- [2] "Detik," Maret, 2024, [Online]. Available: <https://inet.detik.com/mobile-apps/d-7265063/jumlah-pengguna-tembus-130-juta-threads-fokus-garap-fitur-baru>
- [3] F. Bei and S. Sudin, "Analisis Sentimen Aplikasi Tiket Online Di Play Store Menggunakan Metode Support Vector Machine (Svm)," *Sismatik*, vol. 01, no. 01, pp. 91–97, 2021.
- [4] Z. Annisa and B. S. S. Ulama, "Analisis Sentimen Data Ulasan Pengguna Aplikasi 'PeduliLindungi' pada Google Play Store Menggunakan Metode Naïve Bayes Classifier Model Multinomial," *J. Sains dan Seni ITS*, vol. 11, no. 6, 2023, doi: 10.12962/j23373520.v11i6.94064.
- [5] "Ronen Feldman," 2007, [Online]. Available: Feldman, Ronen ,Sanger, dkk. (2007). The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data.Cambridge University Press, New York.
- [6] S. Wahyu Handani, D. Intan Surya Saputra, Hasirun, R. Mega Arino, and G. Fiza Asyrofi Ramadhan, "Sentiment analysis for go-jek on google play store," *J. Phys. Conf. Ser.*, vol. 1196, no. 1, 2019, doi: 10.1088/1742-6596/1196/1/012032.
- [7] A. L. Maitri and J. Sutopo, "Rancangan Bangun Chatbot Sebagai Pusat Informasi Lembaga Kursus Dan Pelatihan Menggunakan Pendekatan Natural Language Processing," *Eprints.Uty.Ac.Id*, pp. 1–9, 2019, [Online]. Available: <http://eprints.uty.ac.id/>
- [8] N. W. S. Agustini, D. Priadi, and R. V. Atika, "Profil Kimia dan Aktivitas Antibakteri Fraksi Aktif Nannochloropsis sp. sebagai Senyawa Penghambat Bakteri Penyebab Gangguan Kesehatan Mulut," *J. Pascapanen dan Bioteknologi Kelaut. dan Perikanan*, vol. 17, no. 1, p. 19, 2022, doi: 10.15578/jpbkp.v17i1.781.
- [9] N. W. Septiani, "Analisis Sentimen Menggunakan Support Vector Machine dan Maximum Entropy Sentiment Analysis Using Support Vector Machine and Maximum Entropy Method," vol. 4, no. 2, pp. 1–14, 2017.
- [10] M. Fitri, "Perancangan Sistem Temu Balik Informasi Dengan Metode Pembobotan Kombinasi TF-Idf Untuk Pencarian Dokumen Berbahasa Indonesia," *J. Sist. dan Teknol. Inf.*, vol. Vol. 1, no. 1, pp. 1–6, 2013.