

Analisis Sentimen pada Aplikasi Translate Google Menggunakan Metode SVM (Studi Kasus: Komentar Pada Playstore)

¹Sri Ayu Ashari, ²Muhammad Wahyu Ade Saputra, ³Esta Larosa, ⁴Bait Syaiful Rijal

^{1),4)} Program Studi Pendidikan Teknologi Informasi, Fakultas Teknik, Universitas Negeri Gorontalo
Gorontalo, Indonesia

²⁾ Magister Teknik Informatika, Universitas AMIKOM Yogyakarta, Sleman, Indonesia

³⁾ Program Studi Pendidikan Teknik Mesin, Fakultas Teknik, Universitas Negeri Gorontalo
Gorontalo, Indonesia

e-mail: sriayu@ung.ac.id

Abstrak

Penelitian ini bertujuan menganalisis *review* dalam memahami opini dan emosi yang diungkapkan oleh pengguna terhadap aplikasi *Translate Google* pada *Google Play Store* menggunakan analisis sentimen. Dengan menggunakan metode *Support Vector Machine (SVM)* dalam analisis sentimen aplikasi *Translate Google*, untuk mendapatkan pemahaman yang lebih baik tentang bagaimana pengguna merespon aplikasi tersebut. Hal ini dapat membantu pengembang dalam meningkatkan pengalaman aplikasi, merespon kebutuhan dan preferensi pengguna dengan lebih baik. Analisis *review* pengguna ini menggunakan metode SVM. Alat pengukur dalam penelitian ini menggunakan, pertama *Lexicon* Indonesia sebagai alat untuk mendapatkan hasil positif dan negatif, kedua, term *frequency-inverse document frequency (tf-idf)* sebagai pendukung untuk hasil dari evaluasi. Aplikasi *Translate Google* memiliki 1000 dataset ulasan pengguna yang dikumpulkan dari *Google play store*. Hasil analisis menggunakan *Support Vector Machine* menghasilkan akurasi 95%, dengan "tidak" sebagai hasil *review* positif dan negatif terbanyak dari 1580 *review*.

Kata kunci: Analisis Sentimen, *Google Playstore*, *Translate Google*, *Support Vector Machine (SVM)*.

Abstract

This research aims to analyze reviews in understanding the opinions and emotions expressed by users regarding the Google Translate application on the Google Play Store using sentiment analysis. By using the Support Vector Machine (SVM) method in sentiment analysis of the Google Translate application, to get a better understanding of how users respond to the application. This can help developers improve the application experience, responding better to user needs and preferences. This user review analysis uses the SVM method. The measuring tool in this research uses, firstly, the Indonesian Lexicon as a tool to obtain positive and negative results, secondly, term frequency-inverse document frequency (tf-idf) as a support for the results of the evaluation. Google Translate app has a dataset of 1000 user reviews collected from Google play store. The results of analysis using Support Vector Machine produced 95% accuracy, with "no" as the result of the most positive and negative reviews out of 1580 reviews.

Keywords: *Google Playstore*, *Sentiment Analysis*, *Support Vector Machine (SVM)*, *Translate Google*.

Diterima: 20 Oktober 2023

Disetujui: 13 Desember 2023

Dipublikasi: 31 Desember 2023

©2023 Sri Ayu Ashari, Muhammad Wahyu Ade Saputra, Esta Larosa, Bait Syaiful Rijal
Under the license CC BY-SA 4.0

Pendahuluan

Translate google sebuah aplikasi untuk melakukan terjemahan bahasa, sehingga memudahkan kita untuk mengerti atau memahami bahasa tersebut (Alam, 2020). Penelitian ini bertujuan untuk melakukan analisis sentimen pada aplikasi *translate google* dengan menggunakan metode *Support Vector Machine (SVM)* berdasarkan komentar-komentar pengguna di platform *Playstore*. Analisis sentimen adalah metode yang digunakan untuk mengevaluasi dan mengklasifikasikan opini atau perasaan dari teks, baik itu positif, negatif, atau netral (Wulandari et al., 2021). Penelitian ini melibatkan penggabungan antara latar belakang penelitian dan literatur review yang tajam, eksplisit, dan didukung oleh sitasi dari sumber-sumber yang relevan.

Dalam era digital, aplikasi ponsel pintar semakin populer dan menjadi bagian penting dalam kehidupan sehari-hari (Daeng et al., 2017). Pengguna aplikasi *Translate Google* juga aktif memberikan komentar dan ulasan pada halaman *Playstore* untuk berbagi pengalaman mereka menggunakan aplikasi tersebut. Namun, pengelolaan volume besar data komentar tersebut menjadi tugas yang rumit, dan memahami sentimen dari pengguna secara manual sangat tidak efisien. Oleh karena itu, diperlukan sebuah metode analisis sentimen otomatis yang dapat memberikan wawasan berharga bagi pengembang dalam meningkatkan kualitas dan pengalaman pengguna aplikasi.

Metode *Support Vector Machine (SVM)* telah diakui sebagai salah satu teknik klasifikasi yang efektif sebagai analisis sentimen pada data teks (Hermawan et al., 2023). SVM mampu mengatasi masalah klasifikasi dalam ruang fitur berdimensi tinggi dan memiliki kinerja yang baik bahkan pada dataset yang kompleks. Beberapa penelitian sebelumnya telah menggunakan SVM untuk analisis sentimen pada berbagai domain, termasuk aplikasi mobile, dengan hasil yang memuaskan.

Pendekatan dalam penelitian ini didasarkan pada tahap pra-pemrosesan data yang meliputi pembersihan, normalisasi, dan penghapusan kata-kata tidak relevan. Selanjutnya, fitur-fitur ekstraksi digunakan untuk mengubah teks komentar menjadi representasi vektor numerik. SVM akan digunakan sebagai model klasifikasi untuk melatih data latih yang sudah dilabeli sentimen positif dan negative (Sodik & Kharisudin, 2021). Adapun pendukungnya adalah *lexicon* Indonesia untuk mengambil hasil positif dan negatif dan *term frequency-inverse document frequency (tf-idf)* sebagai pendukung untuk hasil evaluasi.

Penelitian ini memiliki implikasi yang signifikan dalam meningkatkan pemahaman tentang opini pengguna terhadap aplikasi *Translate Google*, sehingga pengembang dapat merespons dengan cepat dan memperbaiki aspek yang perlu ditingkatkan. Dengan menggunakan SVM sebagai metode analisis sentimen, diharapkan pada

penelitian ini dapat memberikan kontribusi berharga bagi pengembang aplikasi *mobile* di masa depan dan mendorong penelitian lebih lanjut dalam bidang analisis sentimen teks.

Metode

Metode Sastrawi (Stemming)

Metode Sastrawi digunakan untuk melakukan proses *stemming* pada teks komentar (Najiyah & Hariyanti, 2021). *Stemming* adalah proses mengubah kalimat menjadi bentuk dasar atau kalimat dasar dengan menghilangkan imbuhan atau awalan tertentu. Hal ini dilakukan untuk mengurangi variasi kata dan menggabungkan kalimat dengan akar kalimat yang sama. Dengan menggunakan metode Sastrawi, komentar-komentar akan diolah menjadi bentuk dasar, sehingga perbedaan bentuk kata tidak memengaruhi analisis sentimen.

Analysis Sentiment

Analisis sentimen adalah proses untuk mengekstraksi, memahami, dan menganalisis opini, sikap, dan emosi yang terkandung dalam teks, seperti ulasan, komentar, atau posting media sosial (Darwis et al., 2020). Tujuannya adalah untuk memahami sentimen yang diungkapkan oleh penulis teks terhadap suatu topik, entitas, atau peristiwa tertentu (Purbaya et al., 2023).

Langkah-langkah dalam analisis sentiment meliputi:

- 1) Pra-pemrosesan teks: Data teks pra-diproses untuk menghilangkan kebisingan dan mempersiapkannya untuk analisis lebih lanjut. Tahapan ini mencakup langkah-langkah seperti penghilangan *stopwords* (kata-kata umum yang tidak membawa makna sentimen), *stemming* (mengubah kata-kata menjadi bentuk dasarnya), dan normalisasi teks.
- 2) Klasifikasi sentimen: Setelah pra-pemrosesan, teks diklasifikasikan menjadi kategori sentimen yang relevan, seperti positif, negatif, atau netral. Ada beberapa pendekatan yang dapat digunakan, termasuk metode berbasis aturan, metode berbasis pembelajaran mesin seperti *Naive Bayes*, *Support Vector Machine (SVM)*, atau algoritma pembelajaran mendalam seperti jaringan saraf.
- 3) Analisis aspek dan entitas: Selain mengklasifikasikan sentimen secara keseluruhan, analisis sentimen juga dapat melibatkan identifikasi aspek atau entitas spesifik yang menjadi fokus teks. Misalnya, dalam analisis sentimen tentang sebuah produk, kita dapat mengidentifikasi sentimen terkait fitur-fitur produk tertentu.

- 4) Evaluasi dan interpretasi: Setelah sentimen diklasifikasikan, langkah selanjutnya adalah mengevaluasi hasilnya dan memberikan interpretasi. Ini dapat melibatkan penghitungan statistik seperti akurasi, presisi, dan *recall*, serta mengidentifikasi pola, tren, atau temuan menarik dari data sentimen.
- 5) Umpan balik dan pengambilan keputusan: Hasil analisis sentimen dapat digunakan untuk memberikan umpan balik kepada pengembang produk atau layanan yang ditinjau, membantu mereka memahami umpan balik pengguna dan meningkatkan produk. Selain itu, analisis sentimen juga dapat digunakan sebagai sumber informasi untuk pengambilan keputusan bisnis.

Lexicon Indonesia

Metode *Lexicon* Indonesia digunakan untuk mengidentifikasi dan menilai sentimen pada teks berbahasa Indonesia (Syakur, 2021). Dalam penelitian ini, *lexicon* bahasa Indonesia yang telah dikurasi akan digunakan sebagai referensi untuk menentukan polaritas kata dalam komentar pengguna. Setiap kata dalam komentar akan dicocokkan dengan *entri* di dalam *lexicon*, dan polaritasnya akan digunakan untuk menghitung skor sentimen keseluruhan dari komentar tersebut (Ruslim et al., 2019). Skor sentimen dapat berupa positif, negatif, atau netral, tergantung pada jumlah kata-kata yang mengandung polaritas tertentu dalam komentar (Aung & Myo, 2017) .

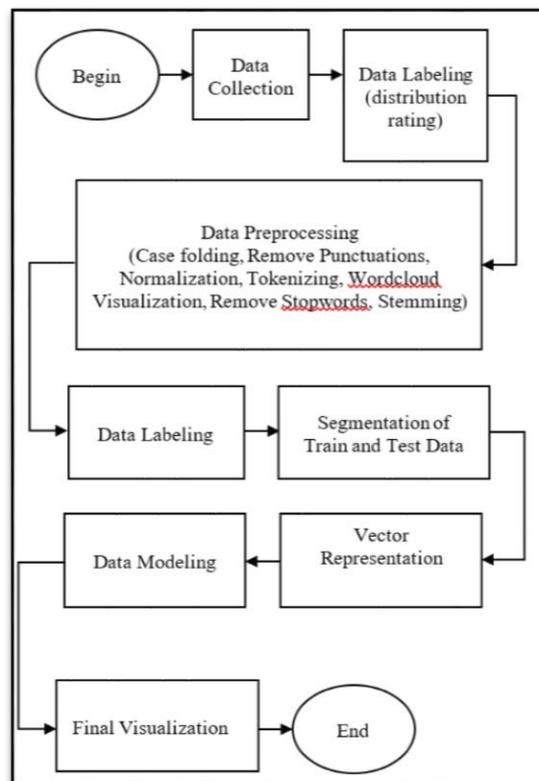
TF-IDF (*Term Frequency-Inverse Document Frequency*)

Metode TF-IDF digunakan untuk ekstraksi fitur dari teks komentar (Gifari et al., 2022). Setelah proses *stemming*, teks akan diubah menjadi representasi vektor numerik menggunakan metode TF-IDF. TF-IDF menghitung frekuensi kata dalam suatu komentar (*term frequency*) dan mengimbanginya dengan frekuensi kata tersebut dalam seluruh dataset komentar (*inverse document frequency*) (Sari et al., 2021). Dengan demikian, kata-kata yang sering muncul dalam komentar tertentu, tetapi jarang muncul dalam seluruh dataset, akan memiliki bobot yang tinggi dalam representasi vektor.

Metode SVM (*Support Vector Machine*)

Metode SVM digunakan sebagai model klasifikasi untuk mengklasifikasikan sentimen dalam komentar, yaitu positif, negatif, atau netral (Fikri & Sarno, 2019). Setelah komentar-komentar diubah menjadi representasi vektor menggunakan TF-IDF, model SVM akan dilatih menggunakan data latih yang sudah dilabeli sentimen (Praghakusma & Charibaldi, 2021). SVM bertujuan untuk menemukan batas keputusan yang optimal yang memisahkan kelas sentimen, sehingga dapat mengklasifikasikan sentimen komentar-komentar yang belum diketahui dengan akurat.

Dengan menguji efektivitas metode yang digunakan, penelitian ini menggunakan pendekatan validasi silang (*cross-validation*) untuk membagi data menjadi kumpulan data pelatihan dan pengujian. Data pelatihan digunakan untuk melatih model SVM, sedangkan data pengujian digunakan untuk menguji kinerja model. Performa model akan dievaluasi menggunakan metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-score. Selain itu, perbandingan hasil analisis sentimen dengan sentimen yang sebenarnya (*ground truth*) pada komentar juga dilakukan untuk mengukur tingkat kesesuaian (*accuracy*) dari metode yang diimplementasikan.



Gambar 1. Tahapan penelitian.

Google Colab

Google Colab, singkatan dari *Google Colaboratory*, adalah *platform cloud* berbasis web yang disediakan oleh *Google* yang diluncurkan pada bulan Maret 2017. *Google Colab* memberikan lingkungan pengembangan berbasis *Jupyter Notebook* yang memungkinkan pengguna untuk menulis dan menjalankan kode *Python* secara interaktif (Gelar Guntara, 2023).

Google Colab dirancang untuk menyediakan lingkungan pengembangan *Python* yang interaktif, mudah digunakan, dan bersifat kolaboratif. Tujuan utamanya adalah mempermudah pengembangan dan eksperimen dalam bidang data *science*, *machine*

learning, dan *deep learning*. Hal ini memudahkan pengguna untuk mengembangkan, menjalankan, dan berbagi kode *Python* secara interaktif.

Integrasi dengan *Google Colab* terintegrasi dengan *Google Drive*, memungkinkan pengguna untuk menyimpan, memuat, dan berbagi *notebook*, *dataset*, dan file lainnya secara langsung. Integrasi ini memfasilitasi kolaborasi dan aksesibilitas file di antara anggota tim yang bekerja secara bersama-sama.

Dalam beberapa tahun terakhir, *Google Colab* telah menjadi salah satu alat yang populer di kalangan praktisi *data science* dan pengembang *machine learning*. Dengan ketersediaan sumber daya gratis, aksesibilitas yang mudah, dan fitur-fitur kolaboratifnya, *Google Colab* telah mempermudah pengembangan dan eksperimen dalam bidang ilmu data dan *machine learning*.

Hasil dan Pembahasan

Hasil

Penelitian ini meliputi beberapa langkah yang dimulai dengan pengumpulan data, memberikan label pada data (mengelompokkan berdasarkan peringkat), pra pemrosesan data, pemberian label pada data, membagi data menjadi bagian latih dan uji, menghasilkan representasi vektor, memodelkan data, dan akhirnya, melakukan visualisasi akhir. Tahapan penelitian diilustrasikan pada gambar 1:

Pengumpulan Data

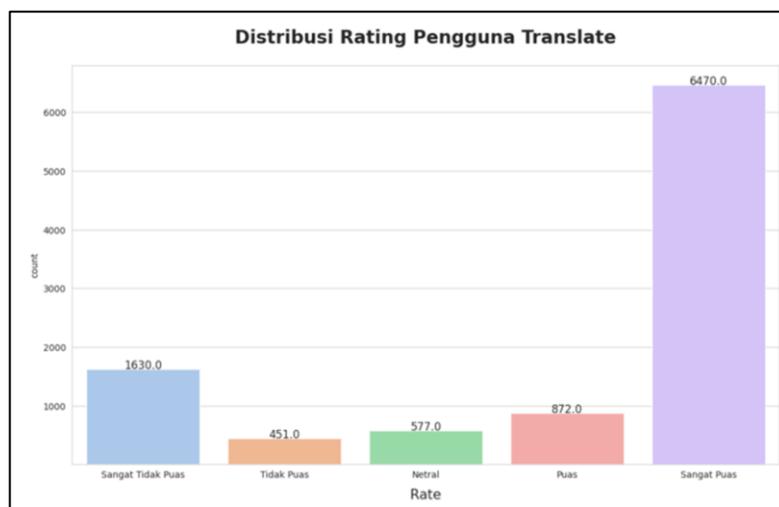
Dalam mengambil ulasan aplikasi *Translate Google* di *Google Play Store*, dengan pemodelan scraping menggunakan *Google Colab* yaitu *Google-Play-Scraper*. Rating dan Ulasan dalam *Google Play Store* dapat mengindikasikan seberapa baik kinerja suatu aplikasi. Untuk mendapatkan semua informasi yang diperlukan dalam analisis, seperti ID ulasan, nama pengguna, isi ulasan, skor, tanggal, dan lainnya. Setelah berhasil mengumpulkan data ulasan tersebut, informasi tersebut kemudian disimpan dalam format dokumen *Comma Separated Values (CSV)*.

Pelabelan Data (Distribusi Peringkat)

Ada 2 kolom yang mempunyai banyak *missing values* berjumlah 9894, yaitu *replyContent* dan *repliedAt*. Selain itu, terdapat kolom yang berisi *missing values* yang berjumlah 3209, yaitu *review Created Version*. Kolom *reply Content* dan *replied At* tidak dilakukan pengisian nilai karena berisi balasan ulasan yang diberikan oleh pihak *Translate Google* terhadap ulasan pengguna dan tanggal balasannya. Sedangkan

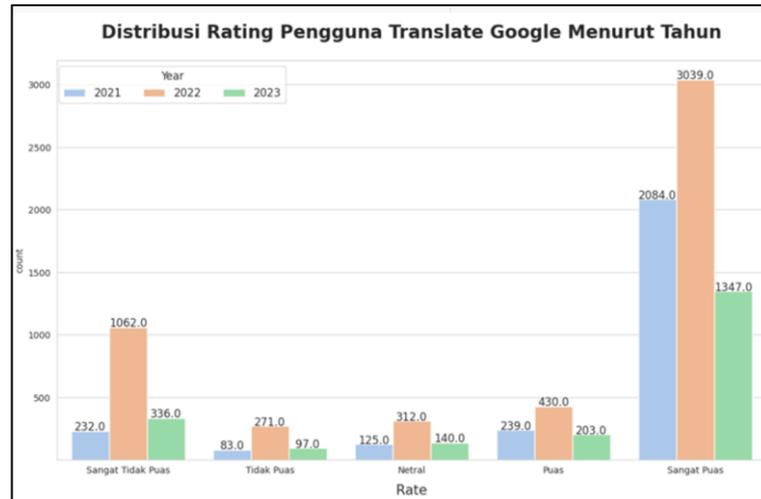
missing values yang terdapat di kolom *review Created Version* tidak diisi karena berisi versi dari aplikasi *Translate Google*.

Pelabelan data (Distribusi Peringkat) ulasan pengguna aplikasi *Translate Google* menggunakan metode manual. Distribusi peringkat menghapus nilai yang hilang pada kolom (*reply Content, replied At, app Version, review Created Version*) dan memisahkan serta mengganti nama kolom waktu lebih spesifik lagi menjadi bulan dan tahun. Ulasan pengguna dilabeli “Sangat Tidak Puas”, “Tidak Puas”, “Netral”, “Puas”, dan “Sangat Puas”, dengan nilai distribusi peringkat pengguna mendapatkan hasil dari “Sangat Tidak Puas” 1630, “Tidak Puas” 451, “Netral” 577, “Puas” 872, dan “Sangat Puas” 6470.



Gambar 2. Distribusi Pengguna *Translate Google* dengan 5 label.

Terdapat juga distribusi peringkat pengguna *Translate Google* dari tahun ke tahun yang di mana perbandingan dari tahun-tahun sebelumnya terdiri dari Tahun 2021 sampai dengan 2023. Di antara tahun tersebut mulai dari “Sangat Tidak Puas” sampai dengan “Sangat Puas”, Tahun 2022 lebih tinggi nilai distribusinya dari tahun 2021 dan 2023. Label “Sangat Tidak Puas” dengan angka 1062. Label “Tidak Puas” dengan angka 271. Label “Netral” dengan angka 312. Label “Puas” dengan angka 430. Dan dengan hasil 3039 yang berlabelkan “Sangat Puas”.



Gambar 3. Distribusi peringkat pengguna *Translate Google* berdasarkan Tahun.

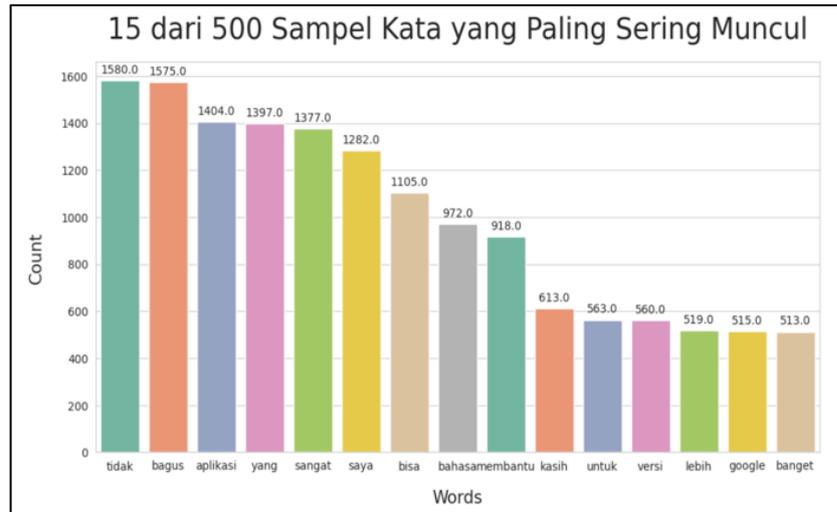
Prapemrosesan Data

Prapemrosesan data adalah proses pertama pada klasifikasi sebuah teks. Prapemrosesan data memiliki beberapa proses sebagai berikut:

- 1) *Case Folding*: Dalam proses ini, karakter teks dijadikan huruf kecil dan *lowercase*
- 2) *Remove Punctuations*: Langkah ini mencabut semua simbol tanda baca, angka, *Uniform Resource Locator (URL)*, kode ASCII, dan emoji.
- 3) *Normalization*: Proses ini mengonversi kata-kata yang singkat dan tidak standar menjadi bentuk kata baku sesuai dengan Kamus Besar Bahasa Indonesia (KBBI).
- 4) *Tokenizing*: Proses ini membagi teks menjadi bagian-bagian kata yang terstruktur sehingga dapat dihitung jumlah frekuensi kata yang muncul.
- 5) *Wordcloud Visualization*: Proses ini memvisualisasikan kumpulan kata yang diambil dari teks. Hasil dari sampel distribusi frekuensi dari 500 ulasan *Translate Google*, kata “tidak” muncul 1580 kali dalam ulasan.



Gambar 4. Visualisasi *Wordcloud*



Gambar 5. 15 dari 500 sampel kata yang paling sering muncul

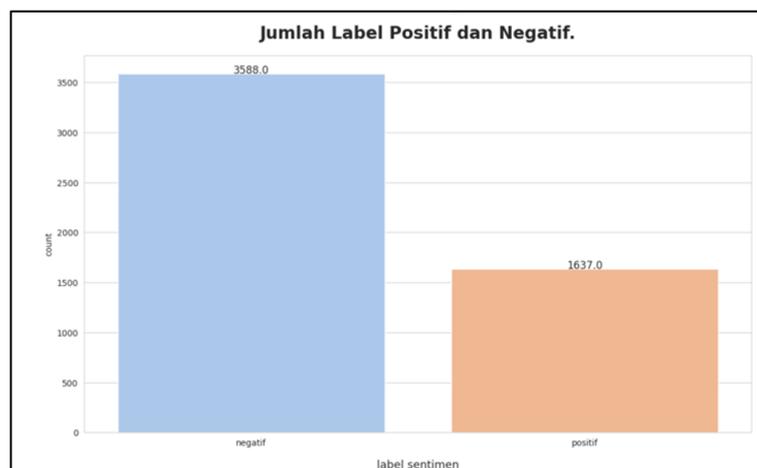
- 6) *Remove Stopwords*: Proses ini menghilangkan kalimat yang berulang kali muncul dalam teks yang dianggap tidak memiliki arti dengan 758 basis data *stopwords* penyajian bentuk bahasa Indonesia.
- 7) *Stemming*: Proses stemming mencari kalimat dasar dengan proses meleburkan awalan dan akhiran suatu kalimat. Proses *stemming* dilakukan menggunakan pemrograman *Python* bernama Sastrawi.

Pelabelan Data

Ulasan masih belum mempunyai sentimen sehingga sulit untuk mencari tahu apakah pengguna memberikan ulasan positif atau negatif. Proses pemberian sentimen tidak mungkin dilakukan secara manual dengan melihat ulasan secara satu per satu karena membutuhkan waktu yang lama dan memerlukan seorang ahli di bidang bahasa yang dapat menafsirkan ulasan kemudian mengelompokkan ke sentimen positif dan negatif. Maka dari itu, perlunya proses pemberian label sentimen dengan menerapkan metode yang berbasis *lexicon* atau biasa dikenal dengan *lexicon-based method*. Kamus yang digunakan adalah *InSet Lexicon* yang terdiri dari kamus positif dan negatif.

Membuat fungsi sentiment *analysis lexicon* indonesia yang berisi *for loop* dan kondisi *if* untuk melabeli setiap ulasan. Ulasan diberi *score* terlebih dahulu dengan menyesuaikan dengan bobot kata yang terdapat di kamus. +5 untuk kata positif dan -5 untuk negatif. Setelah itu, bobot ulasan dijumlahkan kemudian di kelompokkan menjadi sentimen positif, negatif, dan netral. Netral ini diartikan sebagai hasil kalkulasi ulasan bernilai 0. Karena fokus analisis ini hanya sentimen positif dan negatif, maka ulasan yang mengandung sentimen netral dihilangkan.

Mengonversikan *whitespaces* menjadi *np.nan* kemudian menghapusnya menggunakan fungsi *dropna*, setelah itu *reset index* tabel agar urutan menjadi penomoran menjadi lebih rapih. Sehingga menghasilkan negatif 3588, positif 1637, dan netral 1242. Selanjutnya ke tahap menghapus ulasan yang bersentimen netral karena tujuan dari penelitian ini hanya menggunakan 2 sentimen yaitu positif dan negatif.



Gambar 6. Jumlah label positif 1637 dan negatif 3588.

Segmentasi Data

Segmentasi data *training* dan data testing dengan mencari parameter untuk data *splitting* terbaik yang menghasilkan nilai akurasi paling tinggi dan meminimalisir model melakukan kesalahan prediksi, sehingga *confusion matrix* difungsikan melalui librari SVM, dengan model *training* dan *testing* yang telah diklasifikasikan sebelumnya, sehingga mendapatkan matriks dengan ukuran 2x2 untuk representatif kelas aktual dan prediksi. Hasil model *training* dengan menggunakan data baru yang belum terproses.

Berdasarkan hasil penelitian terkait parameter *test_size* dan *random_state* terbaik, bahwa ukuran *test_size* adalah 0,1 dan *random_state* adalah 25 memperoleh nilai akurasi, *recall*, dan *precision* lebih tinggi dibandingkan ukuran yang lain. Sehingga ukuran tersebut akan digunakan untuk melakukan *splitting dataset*.

Representasi Vektor

Dengan menggunakan *Term Frequency–Inverse Document Frequency (TF–IDF)*, algoritma yang umum digunakan untuk menghitung bobot kalimat pada suatu dokumen adalah *Term frequency (TF)*. Frekuensi kata (TF) menunjukkan banyaknya frekuensi kata yang sering muncul dalam suatu dokumen. *Inverse document frequency (IDF)* berfungsi untuk mengurangi bobot kata yang sering muncul dan menentukan pentingnya

makna suatu kata dalam dokumen. Ini juga sering mengganggu proses pencarian kata yang unik.

Persamaan (1) menunjukkan perhitungan $TF:TF(t,d)$ menunjukkan frekuensi kalimat dari kalimat t dalam dokumen d , $N(t,d)$ menunjukkan banyaknya kalimat t yang muncul dalam dokumen d , dan T adalah jumlah kalimat dalam dokumen. Persamaan (2) menunjukkan perhitungan IDF, dengan $IDF(t)$ mewakili jarangness frekuensi kalimat dalam dokumen, N mewakili jumlah dokumen, dan $N(t)$ mewakili jumlah dokumen dengan kalimat t . Maka, untuk menghitung TF-IDF digunakan (3), merupakan perkalian dari TF dan IDF.

Pemodelan Data

Setelah melalui tahapan representasi vektor, tahapan selanjutnya menerapkan pemodelan data dengan library SVM menggunakan pemrograman *Python* yang dinamakan *scikit-learn*. Tahapan pemodelan data ini dilakukan pemodelan prediksi yang dimana -1 adalah negatif dan 1 adalah positif. Nilai b mendapatkan hasil -0.012367731. Nilai akurasi rata-rata pada data training dan label 99% dan akurasi rata-rata pada data test dan label 95% dengan *mean_squared_error (y_test, predict)* 22%.

Pembahasan

Visualisasi Akhir

Dengan Visualisasi akhir ini dapat mewakili hasil prediksi yang diperoleh positif 317 dan negatif 728. Jumlah sentimen setelah pemodelan mendapatkan hasil persentase sentimen positif: 30.33 % dan persentase sentimen negatif: 69.67 %, diantara keduanya mempunyai persentase perbedaan antara positif dan negatif sebesar 39.33 %. Model berhasil memprediksi data uji ulasan sebagai sentimen negatif sebesar 69.67% atau sebanyak 728 ulasan dan sentimen positif sebesar 30.33% atau sebanyak 317 ulasan. Dapat disimpulkan bahwa model lebih banyak memprediksi ulasan baru sebagai sentimen negatif dibandingkan positif dengan tingkat akurasi 95%, *Recall* 90% dan *Precision* 92%.

	precision	recall	f1-score	support
Negative	0.96	0.97	0.96	721
Positive	0.92	0.90	0.91	324
accuracy			0.95	1045
macro avg	0.94	0.93	0.94	1045
weighted avg	0.95	0.95	0.95	1045

Gambar 7. Visualisasi

Hasil dari penelitian ini adalah implementasi analisis sentimen pada *aplikasi Translate Google* menggunakan metode SVM berdasarkan komentar-komentar yang ditinggalkan oleh pengguna di *platform Playstore*. Dalam proses eksperimen, data komentar yang telah diolah dan dipreproses digunakan sebagai dataset latih dan pengujian model SVM. Proses validasi silang digunakan untuk mengevaluasi kinerja model. Hasil analisis sentimen menunjukkan akurasi yang tinggi dalam mengklasifikasikan sentimen komentar menjadi positif atau negatif. Hal ini memiliki hasil penelitian yang sama dengan penelitian yang telah dilakukan oleh (Mustakim & Priyanta, 2022) bahwa klasifikasi menggunakan algoritma SVM mampu menghasilkan akurasi yang tinggi dan unggul.

Hasil penelitian ini menunjukkan bahwa metode SVM yang diimplementasikan untuk analisis sentimen pada aplikasi *Translate Google* berhasil mencapai tingkat akurasi yang memuaskan. Proses penggunaan *Lexicon Indonesia* dan Sastrawi membantu dalam mengidentifikasi polaritas kata dalam teks komentar sehingga model SVM dapat mengambil keputusan klasifikasi yang lebih tepat. Hasil ini sesuai dengan tujuan penelitian untuk memberikan wawasan lebih dalam tentang sentimen pengguna terhadap aplikasi.

Selain itu, dengan menerapkan metode TF-IDF dalam ekstraksi fitur, representasi vektor *numerik* yang dihasilkan mampu mempertahankan informasi penting dari teks komentar. Penggunaan metode SVM sebagai model klasifikasi memungkinkan untuk menemukan batas keputusan yang optimal dalam mengklasifikasikan sentimen. Hal ini menyiratkan bahwa metode yang diimplementasikan dalam penelitian ini dapat diandalkan untuk analisis sentimen pada aplikasi *Translate Google*.

Dalam diskusi selanjutnya, ditemukan bahwa sebagian besar komentar pengguna pada aplikasi *Translate Google* cenderung bersifat positif, dengan apresiasi terhadap kemudahan akses informasi akademik, kualitas layanan, dan kecanggihan fitur yang ditawarkan. Namun, beberapa komentar negatif juga disoroti, khususnya terkait dengan

masalah teknis dan keterbatasan fungsionalitas tertentu. Temuan ini dapat menjadi masukan berharga bagi pengembangan selanjutnya dengan tujuan untuk meningkatkan kualitas aplikasi dan memahami kebutuhan pengguna secara lebih baik.

Secara keseluruhan, penelitian ini berhasil mengimplementasikan metode SVM dengan dukungan *Lexicon* Indonesia, Sastrawi, dan TF-IDF dalam analisis sentimen pada aplikasi *Translate Google*. Hasil penelitian mencerminkan efektivitas metode yang diusulkan dalam memberikan wawasan dan pemahaman yang mendalam terhadap opini dan sentimen pengguna aplikasi. Penelitian ini diharapkan dapat memberikan kontribusi bagi perkembangan aplikasi *mobile* di bidang pendidikan tinggi dan mendorong penelitian lebih lanjut dalam analisis sentimen teks.

Kesimpulan

Berdasarkan hasil riset yang dilakukan mengenai analisis sentimen pada aplikasi *Translate Google* menggunakan metode SVM dengan studi kasus komentar pada *Playstore*, dapat disimpulkan bahwa metode SVM yang diimplementasikan bersama dengan *Lexicon* Indonesia, Sastrawi, dan TF-IDF berhasil menghasilkan klasifikasi sentimen yang akurat untuk komentar-komentar pengguna. Metode ini memberikan wawasan berharga bagi pengembangan aplikasi dan layanan yang lebih baik serta pemahaman yang mendalam tentang opini dan sentimen pengguna.

Hasil eksperimen menunjukkan bahwa metode SVM memiliki kinerja yang baik dalam mengklasifikasikan sentimen komentar sebagai positif, negatif, atau netral. Proses penggunaan *Lexicon* Indonesia membantu dalam mengidentifikasi polaritas kata dalam teks, sedangkan metode Sastrawi mempermudah proses stemming untuk mengubah kata-kata menjadi bentuk dasar sehingga variasi kata tidak mempengaruhi analisis sentimen. Penggunaan metode TF-IDF dalam ekstraksi fitur juga berhasil menghasilkan representasi vektor numerik yang mempertahankan informasi penting dari teks komentar.

Penelitian ini memiliki potensi untuk diaplikasikan pada domain lain dalam analisis sentimen pada aplikasi *mobile* dan *platform* berbahasa Indonesia. Hasil riset ini dapat menjadi landasan bagi penelitian lebih lanjut dalam pengembangan metode analisis sentimen teks yang lebih canggih dan efisien. Sebagai kesimpulan, metode SVM dengan dukungan *Lexicon* Indonesia, Sastrawi, dan TF-IDF merupakan pendekatan yang efektif dan dapat diandalkan dalam analisis sentimen pada aplikasi *Translate Google*.

Daftar Pustaka

- Alam, A. (2020). Google Translate Sebagai Alternatif Media Penerjemahan Teks Bahasa Asing Ke Dalam Bahasa Indonesia. *Instruksional*, 1(2), 159. <https://doi.org/10.24853/instruksional.1.2.159-163>
- Aung, K. Z., & Myo, N. N. (2017). Sentiment analysis of students' comment using lexicon based approach. *Proceedings - 16th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2017*, 149–154. <https://doi.org/10.1109/ICIS.2017.7959985>
- Daeng, I. T. M., Mewengkang, N. ., & Kalesaran, E. R. (2017). Jurnal Kemudahan Smartphone. *Acta Diurna*, 6(1), 1–15.
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Edutic - Scientific Journal of Informatics Education*, 7(1), 1–11. <https://doi.org/10.21107/edutic.v7i1.8779>
- Fikri, M., & Sarno, R. (2019). A comparative study of sentiment analysis using SVM and Senti Word Net. *Indonesian Journal of Electrical Engineering and Computer Science*, 13(3), 902–909. <https://doi.org/10.11591/ijeecs.v13.i3.pp902-909>
- Gelar Guntara, R. (2023). Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(1), 55–60. <https://doi.org/10.47233/jteksis.v5i1.750>
- Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, 2(1), 36–40. <https://doi.org/10.46229/jifotech.v2i1.330>
- Hermawan, A., Jowensen, I., Junaedi, J., & Edy. (2023). Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine. *JST (Jurnal Sains Dan Teknologi)*, 12(1), 129–137. <https://doi.org/10.23887/jstundiksha.v12i1.52358>
- Mustakim, H., & Priyanta, S. (2022). Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 16(2), 113. <https://doi.org/10.22146/ijccs.68903>
- Najiyah, I., & Hariyanti, I. (2021). Sentimen Analisis Covid-19 Dengan Metode Probabilistic Neural Network Dan Tf-Idf. *Jurnal Responsif : Riset Sains Dan Informatika*, 3(1), 100–111. <https://doi.org/10.51977/jti.v3i1.488>

- Praghakusma, A. Z., & Charibaldi, N. (2021). Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi). *JSTIE (Jurnal Sarjana Teknik Informatika) (E-Journal)*, 9(2), 88. <https://doi.org/10.12928/jstie.v9i2.20181>
- Purbaya, M. E., Rakhmadani, D. P., Maliana Puspa Arum, & Luthfi Zian Nasifah. (2023). Implementation of n-gram Methodology to Analyze Sentiment Reviews for Indonesian Chips Purchases in Shopee E-Marketplace. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(3), 609–617. <https://doi.org/10.29207/resti.v7i3.4726>
- Ruslim, K. I., Adikara, P. P., & Indriati. (2019). Analisis Sentimen Pada Ulasan Aplikasi Mobile Banking Menggunakan Metode Support Vector Machine dan Lexicon Based Features. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(7), 6694–6702.
- Sari, H., Ginting, G. L., & Zebua, T. (2021). Penerapan Algoritma Text Mining dan TF-IDF Untuk Pengelompokan Topik Skripsi Pada Aplikasi Repository STMIK Budi Darma. *Terapan Informatika Nusantara*, 2(7), 414–432.
- Sodik, F., & Kharisudin, I. (2021). Analisis Sentimen dengan SVM , NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. *Prisma*, 4, 628–634.
- Syakur, A. (2021). Implementasi Metode Lexicon Base Untuk Analisis Sentimen Kebijakan Pemerintah Dalam Pencegahan Penyebaran Virus Corona Covid-19 Pada Twitter. *Jurnal Ilmiah Informatika Komputer*, 26(3), 247–260. <https://doi.org/10.35760/ik.2021.v26i3.4720>
- Wulandari, D. A., Rohmat Saedudin, R., & Andreswari, R. (2021). Analisis Sentimen Media Sosial Twitter Terhadap Reaksi Masyarakat Pada Ruu Cipta Kerja Menggunakan Metode Klasifikasi Algoritma Naive Bayes. *E-Proceeding of Engineering*, 8(5), 9007–9016.