

ANALISIS SENTIMEN ULASAN APLIKASI RUANGGURU PENERAPAN METODE SUPPORT VECTOR MACHINE (SVM)

Alit Damar Prabadaru¹, Ashifa Zahrawati², Salsabila Nurulita³, Muhammad Agmal Jibran⁴,
Nadya Cantika Apriani Dewana⁵, Fuad Nur Hasan⁶

Universitas Bina Sarana Informatika; Jl. Margonda Raya No.9, Depok; (021)8000063

Keywords:

*Ruang guru, Sentiment
Analysis, Support Vector
Machine, SVM, TF-IDF.*

Correspondent Email:

15230621@bsi.ac.id

Abstrak. Kemajuan pesat dalam teknologi pendidikan digital telah memicu pertumbuhan platform pembelajaran online, salah satunya adalah aplikasi Ruangguru. Banyaknya ulasan pengguna yang ada pada aplikasi ini menjadi sumber penting untuk menilai tingkat kepuasan pengguna terhadap layanan dan fitur-fiturnya. Penelitian ini memfokuskan pada analisis sentimen dari ulasan pengguna Ruang guru menggunakan pengklasifikasi Support Vector Machine (SVM). Data ulasan diperoleh dari Google Play Store dan dibagi menjadi dua kategori: positif dan negatif. Proses metodologis mencakup pengumpulan data, tahap-tahap praproses teks seperti pembersihan data, tokenisasi, penghapusan kata-kata umum (stopword), serta stemming. Selanjutnya, teks diproses menjadi vektor fitur numerik dengan metode pembobotan Term Frequency-Inverse Document Frequency (TF-IDF), diikuti dengan klasifikasi menggunakan algoritma SVM. Hasil eksperimen menunjukkan bahwa SVM mampu menghasilkan akurasi tinggi dalam mengklasifikasikan sentimen ulasan pengguna. Dengan demikian, hasil penelitian ini bertujuan memberikan informasi bagi tim pengembangan Ruangguru dengan mengungkap kebutuhan dan preferensi pengguna serta mengidentifikasi fitur aplikasi yang mungkin perlu perbaikan, guna mendukung upaya meningkatkan kualitas layanan dan pengalaman pengguna secara keseluruhan.

Abstract. Rapid advances in digital education technology have fueled the growth of online learning platforms, one of which is the Ruang guru app. The numerous user reviews on this app serve as an important source for assessing user satisfaction with its services and features. This study focuses on sentiment analysis of Ruang guru user reviews using a Support Vector Machine (SVM) classifier. Review data was obtained from the Google Play Store and divided into two categories: positive and negative. The methodological process included data collection, text preprocessing stages such as data cleaning, tokenization, stopwords removal, and stemming. The text was then processed into numeric feature vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting method, followed by classification using the SVM algorithm. Experimental results demonstrated that the SVM was capable of achieving high accuracy in classifying user review sentiment. Therefore, this study aims to inform the Ruangguru development team by uncovering user needs and preferences and identifying app features that may need improvement, supporting efforts to improve service quality and the overall user experience.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

1. PENDAHULUAN PENELITIAN

Transformasi teknologi informasi telah membawa perubahan besar pada berbagai

sektor, termasuk dalam dunia pendidikan. Maraknya berbagai platform pembelajaran daring memberikan pilihan baru bagi siswa dan

guru. Ulasan pengguna merupakan kunci untuk memahami tingkat kepuasan pengguna, pengalaman mereka, dan masalah yang mereka hadapi dengan suatu aplikasi. [1]

Namun, ketika terdapat begitu banyak ulasan, analisis manual menjadi tidak praktis. Pendekatan yang umum digunakan untuk analisis sentimen merupakan suatu metode yang menggunakan algoritma Support Vector Machine (SVM). SVM dikenal mampu menangani data teks dengan baik serta menghasilkan klasifikasi presisi.

Studi ini menerapkan analisis sentimen terhadap review pengguna aplikasi Ruangguru dengan memanfaatkan algoritma Support Vector Machine (SVM) untuk mengukur persepsi pengguna terhadap aplikasi tersebut. Temuan ini diharapkan dapat memberikan wawasan yang berharga.[2]

Analisis sentimen adalah cabang dari Pemrosesan Bahasa Alami (NLP) yang mengidentifikasi, mengukur, serta mengevaluasi opini publik. Menurut Liu, analisis ini menekankan pada identifikasi dan penafsiran sentimen, evaluasi, opini, sikap, dan emosi terkait suatu entitas—seperti produk, layanan, organisasi, individu, isu, peristiwa topik, atau atribut terkait.[3]

Support Vector Machine (SVM) merupakan metode klasifikasi yang memisahkan kelas-kelas data melalui pencarian hyperplane. Hyperplane optimal berada di tengah-tengah antara dua kelas dan memaksimalkan jarak margin dari support vector terdekat setiap kelas. [4]

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah cabang dari Pemrosesan Bahasa Alami (NLP) yang mengidentifikasi, mengukur, serta mengevaluasi opini publik. Menurut Liu, analisis ini menekankan pada identifikasi dan penafsiran sentimen, evaluasi, opini, sikap, dan emosi tentang suatu entitas—seperti produk, layanan, organisasi, orang, isu, acara, topik, atau atribut terkait.[3]

2.2 Metode Support Vector Machine

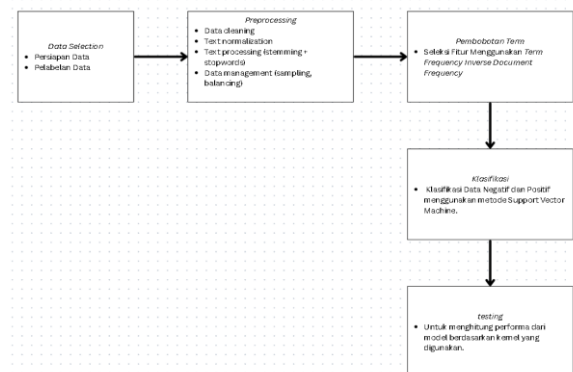
Support Vector Machine (SVM) merupakan metode klasifikasi yang memisahkan kelas data dengan menemukan hyperplane yang tepat. Hyperplane terbaik berada di posisi

tengah antara dua kelas dan berusaha mengoptimalkan jarak atau margin dari titik-titik terdekat dalam setiap kategori.[4] Pada tahun 1992, Vapnik memperkenalkan konsep dasar dari metode Support Vector Machine. Algoritma ini pada dasarnya bertujuan mengatasi masalah pengelompokan, khususnya dalam melakukan pemisahan data yang bersifat linear. Hyperplane bertindak sebagai garis terbaik yang memisahkan kelas -1 dan +1, dengan optimasi yang diukur berdasarkan jarak terjauh dari hyperplane ke titik data terdekat dikedua kategori. Keunggulan SVM terletak pada kemampuannya dalam memproses pemisahan data yang non-linear [5]

3. METODE DAN ALUR PENELITIAN

3.1 Alur Metode Penelitian

Alur Penelitian yang diterapkan mencakup beberapa Langkah, yakni pengumpulan informasi, pra-pemrosesan, klasifikasi dengan memanfaatkan algoritma Support Vector Machine dan Decision Tree, serta evaluasi sebagai tahap akhir. Untuk penjelasan yang lebih rinci, dapat dilihat pada gambar dibawah ini:



Gambar 1. Alur Metode Penelitian

3.2 Rancangan Penelitian

Dibawah ini adalah Langkah-langkah desain penelitian yang akan diterapkan dalam studi ini [6]

3.2.1 Pengumpulan Data

Tahap ini mencakup proses mengumpulkan informasi, yang dimulai dengan Tahap awal untuk memperoleh aplikasi Ruang Guru dengan

mempertimbangkan jumlah ulasan serta penilaian di Goggle dan PlayStore melalui parameter pencarian “Ruang Guru”. Setelah berhasil mendapatkan aplikasi Ruang Guru, Tahap berikutnya adalah mengumpulkan data opini berupa komentar ulasan yang diberikan oleh pengguna mengenai aplikasi Ruang Guru pada Goggle PlayStore. Data tersebut diperoleh dari Goggle PlayStore dengan memanfaatkan Python untuk mengekstrak ulasan, penilaian, waktu ulasan, dan jumlah suka.[7]

3.2.2 Preprocessing Data

Preprocessing melibatkan serangkaian tindakan untuk menyiapkan dan menyesuaikan data teks sebelum dianalisis lebih mendalam. tahap preprocessing bertujuan menata dan menyaring data agar data tersebut menjadi lebih bersih dan terstruktur sehingga lebih efisien dan efektif untuk dilakukan analisis lebih lanjut. [8]

1. Load dan filter data (tahun 2023 - 2025)

Tujuan: Memastikan semua baris memiliki teks yang valid sebelum diproses.

2. Hapus nilai kosong dan teks “nan”

Tujuan: Menghasilkan teks bersih dan seragam.

3. Bersihkan teks (Huruf kecil, Hapus simbol)

Tujuan: Mengurangi variasi kata agar model lebih fokus pada makna inti.

4. Lakukan stemming (Sastrawi)

Tujuan: Membuat fitur yang dipakai lebih relevan (kata bermakna kuat seperti “bagus”, “eror”, “bantu”).

5. Hapus stopwords

Tujuan: Hanya melatih model dengan data terbaru agar relevan dengan trend saat ini.

6. Sampling (6000 data)

Tujuan: Efisiensi waktu training tanpa mengubah distribusi data.

7. Labeling (rating -> sentimen)

Tujuan: Membuat target klasifikasi biner yang jelas.

8. Balancing data

Tujuan: Membuat jumlah data pada setiap kelas seimbang agar model tidak bias terhadap kelas yang dominan.

9. Vectorisasi TF-IDF

Tujuan: Mengukur dan memberi bobot pentingnya kata dalam dokumen.

3.2.3 Pembobotan Term

Setelah data melalui proses preprocessing, tahap selanjutnya adalah melakukan pembobotan untuk memperoleh nilai yang nantinya dapat digunakan dalam klasifikasi. Dalam fase ini, proses pembobotan kata dilakukan menggunakan metode Term Frequency – Inverse Document Frequency (TF – IDF) [9]

$$w_{if} = t f_{if} \times \log \left(\frac{D}{d f_i} \right) + 1$$

Simbol	Arti	Penjelasan
(w_{if})	Bobot kata i pada dokumen f	Nilai yang menunjukkan pentingnya kata tersebut dalam dokumen.
(TF_{if})	Term Frequency	Banyaknya kata i muncul dalam dokumen f .
(df_i)	Frequency Dokumen	Jumlah dokumen yang memuat kata i .
(D)	Jumlah dokumen total	Banyaknya dokumen dalam korpus.
$(\log(\frac{D}{df_i}))$	Inverse Document Frequency (IDF)	Menilai tingkat keunikan kata tersebut diseluruh kumpulan dokumen.
$(+1)$	Penyesuaian (smoothing)	Menjamin hasil log tidak nol dan menjaga skala tetap positif.

Tujuan dari tahap ini untuk mengubah representasi teks dalam dokumen menjadi format numerik, sehingga lebih mudah diproses oleh algoritma pembelajaran mesin. Lalu adapun teknik yang digunakan untuk pembobotan adalah *TF-IDF Vectorizer*, yang di implementasikan menggunakan kelas *TF-IDF Vectorizer* dari pustaka scikit-learn. Teknik ini memberikan nilai bobot untuk setiap istilah dalam dokumen ditentukan oleh frekuensi kemunculan kata tersebut, serta mempertimbangkan seberapa relevan kata tersebut dibandingkan dengan keseluruhan dokumen yang di maksud

3.2.4 Klasifikasi

Setelah melalui tahap Preprocessing dan Pembobotan, Proses berikutnya adalah mengklasifikasikan data menjadi negatif dan positif dengan metode Support Vector Machine. Di Bawah ini terdapat Rumus Dari Klasifikasi Svm: [10]

a. Rumus Klasifikasi SVM Linear

$$f(x) = \text{sign}(w \cdot x + b)$$

Keterangan:

$f(x)$ = hasil fungsi keputusan (decision function)

w = vektor bobot hasil pembelajaran

x = vektor fitur hasil TF-IDF

b = bias

Kriteria klasifikasi:

- Jika $f(x) > 0$ dan $f(x) > 0$, maka termasuk kelas positif
- Jika $f(x) < 0$ dan $f(x) < 0$, maka termasuk kelas negative

b. Rumus SVM *Non-Linear* dengan Kernel RBF

$$f(x) = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right)$$

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$$

Keterangan

- α_i = koefisien Lagrange hasil training
- y_i = label kelas (+1 untuk positif, -1 untuk negatif)
- $K(x_i, x)$ = fungsi kernel RBF
- γ = parameter kernel yang mengatur kompleksitas model

3.2.5 Testing

Pada proses pengujian dilakukan evaluasi terhadap hasil klasifikasi yang telah dikerjakan. Evaluasi ini dilakukan dalam beberapa percobaan. Untuk menilai kinerja dilakukan pengujian awal dengan menggunakan jenis kernel yang diterapkan. Pengujian tersebut memakai berbagai kombinasi pembagian antara data pelatihan dan data pengujian [11]

4. HASIL DAN PEMBAHASAN

Dalam bagian ini ditampilkan keseluruhan hasil penelitian disertai penjelasan dan pembahasan mendalam. Hasil dapat berbentuk gambar, grafik, tabel atau elemen pendukung lainnya yang memudahkan pemahaman pembaca serta dirujuk dalam teks. Untuk pembahasan yang cukup Panjang, penulisan dapat dibagi ke dalam beberapa subjudul, contohnya sebagai berikut:

4.1 Dataset

Dataset yang dipakai pada studi ini didapatkan melalui metode pengambilan data dari platform Ruang Guru, sebuah platform Pembelajaran Daring yang digunakan oleh berbagai institusi

pendidikan di seluruh Indonesia. Data yang dikumpulkan terdiri dari komentar pengguna asli yang memberikan umpan balik terkait pengalaman mereka menggunakan aplikasi tersebut untuk kegiatan pembelajaran daring.[12]

4.2 Hasil Pengujian Model

Pengujian dilakukan dengan dua jenis kernel pada algoritma *Support Vector Machine* (SVM), yakni *Linear Kernel* dan *Radial Basis Function* (RBF). [13]

Pembagian Data	Akurasi SVM Linear	Akurasi SVM RBF
50 : 50	89.40%	92.10%
60 : 40	89.91%	92.29%
70 : 30	90.47%	93.01%
80 : 20	91.39%	93.63%
90 : 10	91.44%	93.05%

Tabel 1. Hasil Pengujian dari Masing Masing Model

4.2.1 Evaluasi Model berdasarkan Jenis Kernel

Evaluasi model dilakukan menggunakan keseluruhan data latih dan uji, dengan menguji dua kernel SVM yang berbeda: Linear dan Radial Basis Function (RBF).[14]

Kernel	Sentimen	Recall	Precision	Accuracy
RBF	Negative	0.99	0.96	0.97
	Positive	0.96	0.99	0.97
Linear	Negative	0.98	0.96	0.97

	ve Positiv e	0.95	0.98	0.97
--	--------------------	------	------	------

Tabel 2 Hasil Pengujian Berdasarkan Kernel

4.2.2 Pengujian didasarkan pada kombinasi porsi data latih dan data uji

Pada uji coba kedua, data dibagi menggunakan Kernel *Radial Basis Function*

Data	Senti ment	Recall	Precis ion	Accur acy
50% - 50%	Negati ve	0.93	0.91	0.92
	Positi ve	0.91	0.93	0.92
60% - 40%	Negati ve	0.94	0.91	0.92
	Positi ve	0.91	0.93	0.92
70% - 30%	Negati ve	0.95	0.92	0.93
	Positi ve	0.91	0.93	0.93
80% - 20%	Negati ve	0.96	0.92	0.94
	Positi ve	0.92	0.95	0.94

90% - 10%	Negati ve	0.96	0.91	0.93
	Positi ve	0.90	0.96	0.93

Tabel 3 Hasil Klasifikasi Berdasarkan Pembagian Data

Berdasarkan Tabel 3, kombinasi 80% data latih dan 20% data uji menghasilkan akurasi tertinggi, yaitu 0.94 atau 94% [15]

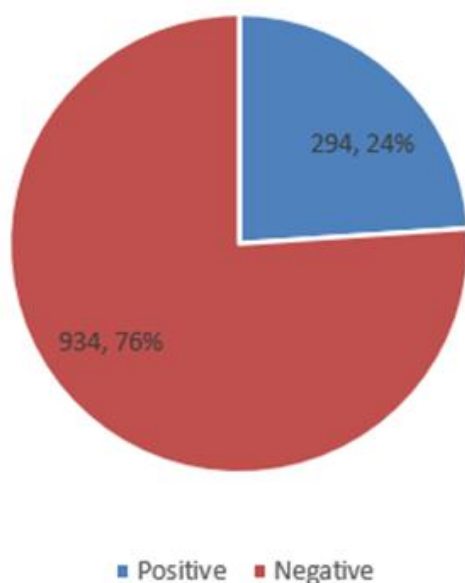
4.2.3 Hasil Evaluasi Penelitian

Sentiment	Persentase
Positive	294 23,94%
Negative	934 76,05%

Gambar 3. Hasil Evaluasi Penelitian

Gambar 3 di atas menunjukkan hasil analisis sentimen pada ulasan pengguna dikategorikan menjadi dua kelompok utama, yakni **sentimen positif** dan **sentimen negatif**. Hasil analisis menunjukkan bahwa mayoritas ulasan masuk dalam kategori Negatif, yang ditandai dengan kata-kata seperti “*lama*”, “*rusak*”, serta “*tidak sesuai*”.

Sementara itu, ulasan dengan sentiment Positif relatif lebih sedikit, umumnya memuat kata seperti “*bagus*”, “*puas*”, atau “*Cepat*”. Temuan ini menggambarkan bahwa mayoritas pengguna memiliki persepsi yang baik terhadap layanan atau produk yang dianalisis.



Gambar 4. Hasil Evaluasi Penelitian

Gambar 4 menyajikan data yang sama dalam bentuk *diagram pie chart* untuk memperjelas perbandingan proporsi antara ulasan positif dan negatif. Dari diagram tersebut, terlihat bahwa ulasan **Negatif** mendominasi dengan persentase yang lebih besar dibandingkan ulasan **Positif**. Visualisasi ini membantu memberikan gambaran yang lebih mudah dipahami mengenai tingkat kepuasan pelanggan secara keseluruhan, di mana proporsi ulasan positif menunjukkan citra produk atau layanan yang cukup baik di mata pengguna. [16]

5. KESIMPULAN AKHIR

Berdasarkan serangkaian tahapan penelitian yang telah dilaksanakan, menunjukan bahwa penerapan algoritma Support Vector Machine (SVM) pada analisis sentimen ulasan pengguna aplikasi Ruangguru berhasil menghasilkan performa klasifikasi yang sangat baik. Proses penelitian mencakup pengumpulan data ulasan, preprocessing teks, pembobotan TF-IDF, serta pengujian model, yang semuanya menunjukkan konsistensi hasil sesuai dengan tujuan penelitian.

Hasil eksperimen mengindikasikan bahwa kernel RBF mencapai akurasi tertinggi, yakni 94% pada skenario pembagian data latih 80% dan data uji 20%. Distribusi sentimen menonjolkan dominasi ulasan **Negatif**, yang mencerminkan bahwa sebagian besar

pengguna memberikan tanggapan **Negatif** terhadap layanan aplikasi Ruangguru

Penelitian ini berhasil mencapai ekspektasi yang dirumuskan di pendahuluan, yaitu mendapatkan gambaran objektif tentang persepsi pengguna melalui analisis sentimen. Temuan tersebut dapat menjadi fondasi untuk penyempurnaan fitur dan peningkatan kualitas layanan ke depan, serta membuka peluang penelitian lanjutan dengan model dan metode yang lebih variatif.

PENGHARGAAN DAN APRESIASI

Penulis menyampaikan rasa terima kasih kepada semua pihak yang telah memberikan dukungan untuk penelitian ini.

DAFTAR PUSTAKA

- [1] H. H. Mubaroroh, H. Yasin, and A. Rusgiyono, "ANALISIS SENTIMEN DATA ULASAN APLIKASI RUANGGURU PADA SITUS GOOGLE PLAY MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER DENGAN NORMALISASI KATA LEVENSHTAIN DISTANCE," vol. 11, no. 2, pp. 248–257, 2022, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [2] R. Mas, R. W. Panca, K. Atmaja1, and W. Yustanti2, "Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers)," *JEISBI*, vol. 02, p. 2021.
- [3] J. Perintis Kemerdekaan Km and K. STMIK Dipanegara Jl Perintis Kemerdekaan Km, "PROSIDING SEMINAR ILMIAH SISTEM INFORMASI DAN TEKNOLOGI INFORMASI Pusat Penelitian dan Pengabdian Pada Masyarakat (P4M) STMIK Dipanegara Makassar Perancangan Aplikasi Sistem Customer Relationship Managemen Pada Rapokalling Auto Service Berbasis Web," 2019.
- [4] I. Nurul Hassanah, S. Faisal, A. Mutoi Siregar, U. Buana Perjuangan Karawang Jl HSRonggo Waluyo, T. Timur, and J. Barat, "Kumpulan jurnaL Ilmu Komputer (KLIK)".
- [5] E. Bambang, R. A. Setia Priadi, and H. -, "PERANCANGAN DAN PEMBUATAN SISTEM INFORMASI PERSETUJUAN PERBAIKAN DAN PERGANTIAN ALAT KOMPUTER BERBASIS WEB (Studi Kasus pada PT. Lautan Teduh Interniaga)," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 2, no. 1, Jan. 2014, doi: 10.23960/jitet.

- [6] M. Romy Firdaus *et al.*, “Analisis Sentimen Dan Topic Modelling Dalam Aplikasi Ruangguru,” *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 4, p. 66, 2020, [Online]. Available: <http://tunasbangsa.ac.id/ejurnal/index.php/jsakti>
- [7] Y. Aprianti, A. Lia Hananto, S. Shofiah Hilabi, S. Informasi, and U. Buana Perjuangan Karawang, “Klasifikasi Sentimen Komentar Pengguna pada Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes,” vol. 9, p. 2025, doi: 10.47002/metik.v9i1.1023.
- [8] U. Khairani, V. Mutiawani, and H. Ahmadian, “Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 887–894, Aug. 2024, doi: 10.25126/jtiik.1148315.
- [9] Y. Aliya Rohim and V. Atina, “Analisis Sentimen terhadap Ulasan Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes.” [Online]. Available: <http://jiip.stkipyapisdampu.ac.id>
- [10] A. S. Simbolon, N. I. Pangaribuan, and N. M. Aruan, “ANALISIS SENTIMEN APLIKASI E-LEARNING SELAMA PANDEMI COVID-19 DENGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN CONVOLUTIONAL NEURAL NETWORK,” *SEMINASTIKA*, vol. 3, no. 1, pp. 16–25, Nov. 2021, doi: 10.47002/seminastika.v3i1.236.
- [11] Y. Caesar, I. Sabastian, A. Kindarto, and A. Fathurrohman, “Analisis Sentiment Masyarakat Terhadap Clash of Champions Ruang Guru Menggunakan Metode Support Vector Machine (SVM) Analysis of Public Sentiment Towards Ruang Guru’s Clash of Champions Using the Support Vector Machine (SVM) Method.”
- [12] A. Baita and N. Cahyono, “ANALISIS SENTIMEN MENGENAI VAKSIN SINOVAQ MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN K-NEAREST NEIGHBOR (KNN).”
- [13] Irfani Faizal Fakhir, Triyanto Mohamad, Hartanto Anggit Dwi, and Kusnawi, “Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine,” Feb. 2020.
- [14] E. Fitri, Y. Yuliani, S. Rosyida, and W. Gata, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” *TRANSFORMATIKA*, vol. 18, no. 1, pp. 71–80, 2020, [Online]. Available: www.nusamandiri.ac.id,
- [15] E. Fernando Baharsyah, T. Hasanah Bimastari Aviani, and C. Wulandari, “Copyright @ Analisis Sentimen Pengguna Terhadap Aplikasi Belajar Online Ruang Guru Pada Ulasan Google Play Store Menggunakan Algoritma Naïve Bayes dan Support Vector Machine”.
- [16] S. Helmiyah and A. Verdian, “JURNAL PENDIDIKAN STKIP ROSALIA LAMPUNG Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP dan Model BERT.”