

Metode dan Algoritma Dalam Sentimen Analisis: Systematic Literature Review

Erba Lutfina¹, Wiwin Andriana², Sanina Quamila Putri Wiratmaja³, Ervina Febrianti⁴

^{1,2,3,4} Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Jl. Penanggungan No.41a, Bandar Lor, Kec. Kota Kediri, 64129, Indonesia

DOI: <http://dx.doi.org/10.53416/stmj.v4i2>

Info Artikel

Sejarah Artikel:

Disubmit 22-07-2024

Direvisi 04-08-2024

Disetujui 05-08-2024

Keywords:

Analisis data; analisis sentimen; sentiment analysis; Systematic Review; Metode SLR

Abstrak

Penelitian ini menyoroti keunggulan analisis sentimen menggunakan algoritma dalam memahami opini publik, terutama dalam konteks meningkatnya kompleksitas konten digital. Dengan tujuan untuk menyelidiki dan menyajikan perkembangan terkini dalam analisis sentimen, penelitian ini menggunakan metode Systematic Literature Review (SLR) untuk mengidentifikasi dan mengevaluasi metode-metode analisis sentimen yang telah dikembangkan sebelumnya. Langkah-langkah penelitian melibatkan identifikasi metode analisis sentimen yang signifikan, evaluasi kelebihan dan kekurangannya, serta penelaahan kritis terhadap kemajuan terbaru. Dengan menerapkan algoritma dalam proses ini, diharapkan dapat menganalisis dan menggambarkan perkembangan penelitian analisis sentimen secara komprehensif. Melalui penerapan metode SLR, penelitian ini diharapkan memberikan wawasan mendalam tentang tren, tantangan, dan peluang penelitian masa depan dalam analisis sentimen, menciptakan pemahaman yang lebih baik tentang metode-metode analisis sentimen yang efektif, dan merinci ekspektasi hasil yang dapat diharapkan dalam pengembangan analisis sentimen.

Abstract

This study highlights the advantages of sentiment analysis using algorithms in understanding public opinion, especially in the context of the increasing complexity of digital content. To investigate and present the latest developments in sentiment analysis, this study uses the Systematic Literature Review (SLR) method to identify and evaluate previously developed sentiment analysis methods. The research steps involve identifying significant sentiment analysis methods, assessing advantages and disadvantages, and critically reviewing recent advances. By applying algorithms in this process, it is expected to be able to analyze and describe the development of sentiment analysis research comprehensively. Through the application of the SLR method, this study is expected to provide in-depth insights into trends, challenges, and opportunities for future research in sentiment analysis, create a better understanding of effective sentiment analysis methods, and detail the expected results that can be expected in the development of sentiment analysis.

□ Alamat Korespondensi:
E-mail: erba.lutfina@dsn.dinus.ac.id

1. Pendahuluan

Di era digitalisasi saat ini tidak bisa dipungkiri bahwa masyarakat tidak bisa terlepas dari informasi-informasi yang bertebaran di internet, seperti informasi dari website, e-commerce, terutama media sosial seperti Twitter, Facebook, Instagram, Tik Tok dan lainnya (Hartmann et al., 2023). Media sosial merupakan wadah untuk berkomunikasi, saling berkomunikasi, saling berinteraksi dan mengemukakan pendapat. Oleh karena itu sentiment publik atau pandangan masyarakat menjadi hal yang sangat penting karena sifatnya yang dapat mempengaruhi perilaku, keputusan, dan tindakan individu maupun kelompok di dalam masyarakat selain itu, sentimen publik juga dapat berdampak pada citra dan reputasi sebuah organisasi, perusahaan, atau pemerintahan (Saputra et al., 2021).

Analisa sentimen publik merupakan studi komputasi mengenai opini atau pendapat yang disampaikan oleh masyarakat atau perorangan yang dapat memuat sentimen positif, netral, atau negatif. Adanya hal tersebut membuat perusahaan dan organisasi kini mulai memperhatikan sentimen atau pendapat publik terhadap produk atau jasa yang mereka tawarkan, maupun isu-isu tertentu yang tengah viral di media sosial (Pramesti & Pratiwi, 2023).

Saat ini banyak sekali metode-metode analisis sentimen yang digunakan di berbagai penelitian. Hal ini terasa sulit bagi peneliti untuk mendapatkan metode mana yang paling cocok untuk digunakan dalam penelitian sentiment analysis. Dikarenakan keberagaman metode-metode yang digunakan, peneliti juga sangat kesulitan dalam mengevaluasi kualitas metodologi penelitian yang cocok sehingga menghasilkan akurasi yang tepat (Dussoye & Cadarsaib, 2018).

Salah satu metode yang digunakan untuk mengidentifikasi isu-isu dan mengkaji informasi dari berbagai sumber daya, seperti jurnal dan buku adalah metode SLR (Systematic Literature Review) atau disebut juga tinjauan pustaka sistematis merupakan sebuah metode dengan pendekatan identifikasi secara terstruktur terhadap aspek-aspek penting dari sejumlah penelitian yang tersedia (Lutfina et al., 2023). Metode SLR ini mempunyai beberapa kelebihan yaitu lebih mudah dan efisien dalam menghasilkan literature review yang lebih detail, akurat, dan kompleks, serta dapat memberikan informasi yang baik dalam menggambarkan sesuatu yang lebih dalam (Sanjaya et al., 2023). Oleh karena itu, dalam penelitian ini penulis menggunakan Sistem Literature Review (SLR) sebagai metode untuk meninjau kembali analisis sentimen yang telah dilakukan oleh penelitian sebelumnya pada

1. Pertanyaan Penelitian:

Tabel 1. Pertanyaan Penelitian

ID	Pertanyaan Penelitian	Motivasi
RQ1	Metode apa saja yang digunakan dalam sentimen analisis?	Mengetahui macam metode yang ada dalam sentimen analisis

beberapa tahun terakhir dan untuk membantu peneliti untuk melihat tren penggunaan metode analisis sentimen saat ini.

2. Metodologi Penelitian

2.1. Objek Penelitian

Objek pada penelitian yang dilakukan yaitu mengenai sentimen analisis. Pengambilan topik sentimen analisis sebagai objek penelitian dilandasi dengan alasan sebagai berikut:

1. Tantangan dalam pemilihan metode analisis yang tepat

Beragamnya metode analisis sentiment yang tersedia. Peneliti menghadapi tantangan dalam memilih metode yang paling sesuai dan efektif untuk tujuan penelitian mereka. Penelitian ini bertujuan untuk memberikan panduan yang lebih jelas dalam pemilihan metode.

2. Ruang lingkup metode analisis yang beragam

Dalam literatur saat ini, terdapat berbagai metode analisis sentimen yang digunakan. Penelitian ini bertujuan untuk mengevaluasi metode-metode tersebut agar dapat memberikan panduan tentang kecocokan dan efektivitas masing-masing dalam berbagai konteks.

3. Pengaruh media sosial

Dengan berkembangnya media sosial dan perkembangan teknologi digital, masyarakat semakin aktif berkomunikasi dan berbagi pendapat mereka di platform seperti Twitter atau X. Analisis sentiment menjadi kritis untuk memahami dinamika opini public dalam konteks digital ini.

2.2. Metode Penelitian

Systematic Literature Review (SLR) merupakan metode yang digunakan dalam penelitian ini untuk mengidentifikasi dan menginterpretasi berbagai penelitian terdahulu yang sesuai dengan pertanyaan penelitian mengenai topik yang menjadi perhatian (Lutfina et al., 2022). Metode SLR yang digunakan bersifat kualitatif dengan merangkum hasil dari penelitian dalam bentuk deskriptif.

RQ2	Algoritma apa yang populer dalam analisis sentimen?	Mengetahui algoritma yang sering digunakan dalam sentimen analisis.
RQ3	Algoritma apa yang paling efektif dalam sentimen analisis?	Mengetahui algoritma yang paling efektif dalam sentimen analisis.
RQ4	Apakah jumlah dataset mempengaruhi tingkat keakuratan dalam sentiment analisis?	Mengetahui jumlah dataset paling optimal dalam analisis sentimen.
RQ5	Bagaimana cara menilai keakuratan hasil sentimen analisis?	Mengetahui bagaimana cara menilai keakuratan hasil dari sentimen analisis.
RQ6	Apa faktor-faktor yang dapat mempengaruhi hasil dari sentimen analisis?	Mengetahui faktor-faktor yang mempengaruhi hasil dari sentimen.
RQ7	Bidang apa saja yang menjadi objek dalam sentimen analisis?	Mengetahui berbagai macam bidang yang menjadi objek dalam sentimen analisis.

Terdapat sebanyak 7 pertanyaan dalam penelitian tentang metode yang digunakan dalam sentimen analisis, yang mana akan dicari jawabannya dalam penelitian literature review ini.

2. Strategi Pencarian

Systematic literature review yang dilakukan dalam melakukan review jurnal, dengan menerapkan strategi pencarian dengan langkah-langkah sebagai berikut (Lutfina et al., 2023):

- Mengumpulkan berbagai jurnal dari google scholar
- Dari berbagai jurnal yang diperoleh kemudian di filter berdasarkan jurnal dengan tema sentimen analisis.
- Kemudian diambil jurnal sentimen analisis yang memakai berbagai metode klasifikasi, dan didapat 30 jurnal.
- Mereview 30 jurnal yang didapat, dengan menentukan setiap tujuan, metode dari setiap jurnal.
- Menentukan Research Question, dan menjawabnya berdasarkan 30 jurnal yang telah direview.

3. Seleksi Studi

Setelah melakukan tahap *selection of the articles*, dengan menggunakan filter untuk Batasan dalam tahun terbit jurnal/artikel yaitu 5 tahun terakhir. Dan Batasan pada kata kunci yang dimasukkan yaitu sentiment analisis. Maka diperoleh sebanyak 30 artikel yang relevan dengan penelitian terkait. Selanjutnya akan dilakukan tahap *categorization of the articles* yaitu pengkategorian isi artikel yang disesuaikan dengan penelitian. Berikut hasil analisis:

- Penelitian yang melakukan analisis sentimen pada platform Instagram mengenai layanan BPJS dengan algoritma Naïve Bayes Classifier, dengan dataset data training sebanyak 250 komentar dengan data testing 50 komentar dilakukan pengujian menggunakan metode frekuensi dan menghasilkan hasil akurasi sebesar 73% (Karim, 2020).
- Penelitian selanjutnya dengan judul Analisis Sentimen Berbasis Aspek Dengan Deep Learning Ditinjau Dari Sudut Pandang Filsafat Ilmu yang diterbitkan pada tahun 2021 dengan metode dan dataset yang tidak disebutkan pada jurnal. Pengujian dilakukan dengan metode confusion matrix yang menghasilkan Tingkat akurasi sebesar 88% (Wahyuni et al., 2023).
- Pada penelitian lain dengan judul Bidirectional Gru Dengan Attention Mechanis m Pada Analisis Sentimen Pln Mobile, dengan metode eksperimen: Data Gathering, Data Pre-Processing, Proposed Method, Model Test and Experiment, Result Evaluation and Validation. Menggunakan dataset yang diambil/scrapping dari ulasan PLN Mobile pada google playstore dengan jumlah data sebanyak 236905 dan diuji menggunakan Confusion Matrix menghasilkan Tingkat akurasi sebesar 97,48% (Rohman et al., 2023).

- Penelitian selanjutnya dengan judul Analisis Sentimen Pengguna Marketplace Bukalapak dan Tokopedia di Twitter Menggunakan Machine Learning dengan dataset data pencarian tentang Tokopedia dan Bukalapak masing-masing sebanyak 50 catatan dan diuji menggunakan metode frekuensi menghasilkan Tingkat akurasi sebesar 82% (Saputra et al., 2021).
- Pada penelitian lain dengan judul Analisis Sentimen Film di Twitter Menggunakan Metode Support Vector Machine dengan dataset data training sebanyak 160 (80%) data dan data testing sebanyak 40 (20%) data. Kemudian diuji menggunakan Confusion Matrix dan menghasilkan Tingkat akurasi sebesar 87,5% (Khairudin et al., 2023).
- Dengan judul Analisis Sentimen Politik Berdasarkan Big Data Dari Media Sosial Youtube: Sebuah Tinjauan Literature yang menggunakan metode klasifikasi SVM, NBC, KNN, Decision Tree dan Random Forest lalu diuji dengan metode Confusion Matrix menghasilkan Tingkat akurasi paling tinggi yaitu KNN sebesar 87% (Ainul Wildan et al., 2021).
- Dalam penelitian lain dengan judul Analisis Cyberbullying Pada Komentar Instagram Dengan Metode Klasifikasi Support Vector Machine menggunakan dataset yang diambil secara luring sebanyak 400 data dan dengan total fitur 1799, lalu dokumen komentar dibagi menjadi data training sebesar 70% dan data testing sebesar 30%. Kemudian diuji menggunakan metode frekuensi dan menghasilkan nilai akurasi sebesar 90% (Athira Luqyana et al., 2018).
- Pada penelitian yang berjudul Analisis Sentimen Twitter Terhadap Uu Omnibus Law Menggunakan Metode Support Vector Machine (SVM) dan Naïve Bayes Classifier (NBC), menggunakan dataset informasi Twitter dari bulan Januari 2022 hingga bulan November 2022 dengan hasil sebanyak 3018 data. Kemudian dilakukan pengujian menggunakan Confusion Matrix dan menghasilkan Tingkat akurasi sebesar 79,4% (Cholifah et al., 2023).
- Penelitian lain dengan judul Klasifikasi Sentimen Masyarakat Terhadap Kebijakan Vaksin Covid-19 Pada Twitter Dengan Imbalance Classes Menggunakan Naïve Bayes yang dipublikasikan tahun 2022 dengan dataset 12000 raw data Tweet, dengan rincian data yang digunakan 8000 Tweet data-training, 778 Tweet data-dev, dan 400 Tweet data-test. Diuji menggunakan metode Confusion Matrix dan menghasilkan Tingkat akurasi sebesar 61% (Yohana et al., 2022).
- Pada penelitian dengan judul Lstm (Long Short Term Memory) For Sentiment Covid-19 Vaccine Classification on Twitter. Menggunakan metode Bidirectional LSTM ini memproses data dari dua arah dengan hidden layer yang terpisah, dengan dataset 13.115 raw data Tweet berbahasa Indonesia dari 23 Maret 2021 hingga 20 April 2021 dengan rincian data yang digunakan adalah 8000 tweet data training, 778 tweet data development, dan 400 tweet data testing. Kemudian dilakukan pengujian dengan metode Confusion Matrix yang menghasilkan Tingkat akurasi sebesar 66% (Ihsan et al., 2022).
- Pada penelitian dengan judul Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Metode Logistic Regression dengan pembobotan kata TF-IDF dengan dataset 9178 data Tweet, dengan rincian 8000

- Tweet sebagai data training, 778 Tweet sebagai Data Dev dan 20 Tweet sebagai data test. Kemudian diuji menggunakan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 67% (Ash & Surya, 2022).
- Pada penelitian dengan judul Analisis Sentimen Terhadap Pelayanan Krl Commuterline Berdasarkan Data Twiter Menggunakan Algoritma Bernoulli Naïve Bayes dengan dataset raw data tweet berbahasa Indonesia dengan total 9100 data, dan dibagi menjadi 6.504 tweet data training, 1.626 tweet data dev, dan 20 tweet data testing. Kemudian diuji dengan Confusion Matrix menghasilkan nilai akurasi sebesar 85% (Saraswati & Rimirasih, 2020).
 - Selanjutnya pada penelitian dengan judul Sentiment Analysis of Social Media Twitter with Case Of Anti-Lgbt Campaign In Indonesia Using Naïve Bayes, Decision Tree, And Random Forest Algorithm dengan dataset 3744 data tweet dengan 2808 tweet data training, dan 936 tweet data testing. Kemudian diuji menggunakan metode Classification of Performance Evaluation menghasilkan tingkat akurasi sebesar 83,43% (Fitri et al., 2019).
 - Pada penelitian lain dengan judul Eksperimen Sistem Klasifikasi Analisis Sentimen Twitter Pada Akan Resmi Pemerintahan Kota Surabaya berbasis Pembelajaran Mesin dengan menggunakan metode Naïve Bayes dan kombinasi penggunaan Stopword. Penelitian ini menggunakan dataset 5836 data tweet dari 1 September 2015 hingga 13 Oktober 2015 yang kemudian diuji kembali menggunakan metode Confusion Matrix dan menghasilkan tingkat akurasi sebesar 67,05% (Faradhillah et al., 2016).
 - Pada penelitian dengan judul Penerapan Deep Sentiment Analysis Pada Angket PenilaianTerbuka Menggunakan K-Nearest Neighbor dengan dataset angket penelitian perkuliahan dan wawancara dan diuji menggunakan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 95,6% (et al., 2016).
 - Dengan judul penelitian lain yaitu Sentimen Analysis to Determine Accommodation, Shopping and Culinary Location On Foursquare In Kupang City. Penelitian ini menggunakan metode Naïve Bayes dengan pembobotan kata TF-IDF dengan dataset sebanyak 671 data test yang berasal dari tips pertama yang dituliskan di masing-masing lokasi pada 5 Juni 2014. Kemudian dilakukan pengujian menggunakan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 66,22% (Aliandu, 2015).
 - Pada penelitian dengan judul Sentimen Analisis Pandangan Masyarakat Terhadap Vaksinasi Covid 19 Menggunakan K-Nearest Neighbors dengan dataset sebanyak 2241 data Twitter, kemudian diuji menggunakan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 79,25% (Apriliani et al., 2023).
 - Dengan judul penelitian Sentimen Analisis Aplikasi E-Commerce Berdasarkan Ulasan Pengguna Menggunakan Algoritma Stochastic Gradient Descent dengan dataset sebanyak 10000 data ulasan Tokopedia dan 10000 data ulasan Shopee yang kemudian setelah diuji menggunakan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 84% .
 - Pada penelitian lain dengan judul Analisis Sentimen dan Sistem Pendukung Keputusan Menginap di Hotel Menggunakan Metode CRISP-DM dan SAW. Penelitian ini menggunakan metode Support Vector Machine

- dengan dataset sebanyak 1680 data klasifikasi ulasan, kemudian diuji kembali dengan metode Confusion Matrix menghasilkan tingkat akurasi sebesar 99,74% (Singgalen, 2023a).
- Penelitian selanjutnya dengan judul Penerapan Metode K-Nearest Neighbor Pada Sentimen Analisis Pengguna Twitter Terhadap KTT G20 Di Indonesia. Dataset yang digunakan pada penelitian ini adalah data tweet dengan total 25000 data, kemudian diuji menggunakan metode Frekuensi menghasilkan tingkat akurasi sebesar 99% (Andriana et al., 2023).
 - Pada penelitian lain dengan judul Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naïve Bayes Classifier Di Twitter dengan dataset sebanyak 301 data tweet, kemudian dilakukan pengujian menggunakan metode K-Fold Cross Validation dan menghasilkan tingkat akurasi sebesar 60% (Febriyani & Februariyanti, 2023).
 - Pada penelitian lain dengan judul Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (Svm) dengan dataset data ulasan sebanyak 3000 yang kemudian diuji kembali menggunakan metode frekuensi dan menghasilkan tingkat akurasi sebesar 98% (Idris et al., 2023).
 - Penelitian lain dengan judul Analisis Sentimen Terhadap Kenaikan Harga Bahan Pokok Menggunakan Metode Naïve Bayes Classifier dengan jumlah dataset sebanyak 2070 data tweet. Kemudian dilakukan pengujian menggunakan metode Confusion Matrix dan menghasilkan nilai akurasi sebesar 94,38% (Muslimin & Lusiana, 2023).
 - Penelitian dengan judul Perbandingan Algoritma Naïve Bayes dan Svm Dalam Sentimen Analisis Marketplace Pada Twitter. Penelitian ini menggunakan dataset 1000 Data dari Tokopedia, Shopee, Lazada. Kemudian dilakukan pengujian untuk mendapatkan hasil akurasi dengan metode Frekuensi. Setelah dilakukan pengujian penelitian ini menghasilkan tingkat akurasi sebesar 88,93% (Kurniawan et al., 2023).
 - Penelitian lain dengan judul Analisis Sentimen Terhadap Aplikasi Chatgpt Pada Twitter Menggunakan Algoritma Naïve Bayes. Penelitian ini menggunakan dataset berupa data tweet pada tanggal 20 Mei 2023 dengan query yang digunakan yaitu Chatgpt sebanyak 301 data. Kemudian dilakukan pengujian menggunakan metode Rapidminer dan menghasilkan tingkat akurasi sebesar 80% (Rifaldi et al., 2023).
 - Pada penelitian lain dengan judul Analisis Sentimen Pengunjung Pulau Komodo Dan Pulau Rinca Di Website Tripadvisor Berbasis Crisp-Dm. penelitian ini menggunakan metode K-Nearest Neighbor, Naïve Bayes Classifier, Support Vector Machine, dan Decision Tree. Dengan dataset data wisatawan Pulau Komodo sebanyak 564 data dan data ulasan wisatawan Pulau Rinca sebanyak 364 data. Kemudian dilakukan pengujian untuk mendapatkan tingkat akurasi dengan metode frekuensi dan menghasilkan nilai akurasi sebesar 99,39% (Singgalen, 2023b).
 - Penelitian dengan judul Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine dengan dataset data Tweet sebanyak 849 data yang kemudian dilakukan pengujian menggunakan metode Confusion Matrix dan menghasilkan tingkat akurasi sebesar 84,76% (Pramesti & Pratiwi, 2023)
 - Dengan judul penelitian Analisis Sentimen

- Terhadap Masyarakat Adanya Uang Kertas Baru Menggunakan Algoritma K-Nearest Neighbor. Penelitian ini menggunakan dataset berupa data komentar sebanyak 510 data. Kemudian penelitian ini di uji menggunakan metode Rapidminer dan menghasilkan tingkat akurasi sebesar 75,06% (Hasanah et al., 2023).
- Selanjutnya pada penelitian lain dengan judul Penerapan Algoritma K-Nearest Neighbor(K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring. Penelitian ini menggunakan dataset berupa Data Tweet Bahasa Indonesia dari tanggal 1 Februari 2020- 30 September 2020 sebanyak 1825 data. Kemudian dilakukan pengujian menggunakan metode Confusion Matrix dan menghasilkan tingkat akurasi sebesar 87%(Supriyanto et al., 2023).
 - Pada penelitian dengan judul Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. Penelitian ini menggunakan dataset data tweet berbahasa Indonesia yang dikumpulkan selama bulan Januari 2017 sebanyak 2000 data. Kemudian penelitian ini diuji menggunakan metode Confusion Matrix dan menghasilkan tingkat akurasi paling tinggi dengan nilai 67,2% saat nilai K adalah 5 (Deviyanto & Wahyudi, 2018).

3. Hasil dan Pembahasan

3.1. RQ1 Metode apa saja yang digunakan dalam sentimen analisis?

Untuk mengetahui macam-macam metode yang ada dalam sentimen analisis, maka munculah research question mengenai metode apa saja yang digunakan dalam sentiment analisis. Terdapat 30 jurnal yang kemudian dikelompokkan berdasarkan usulan dari paper yang di review berdasarkan metode penelitian yang digunakan untuk menjawab research question.

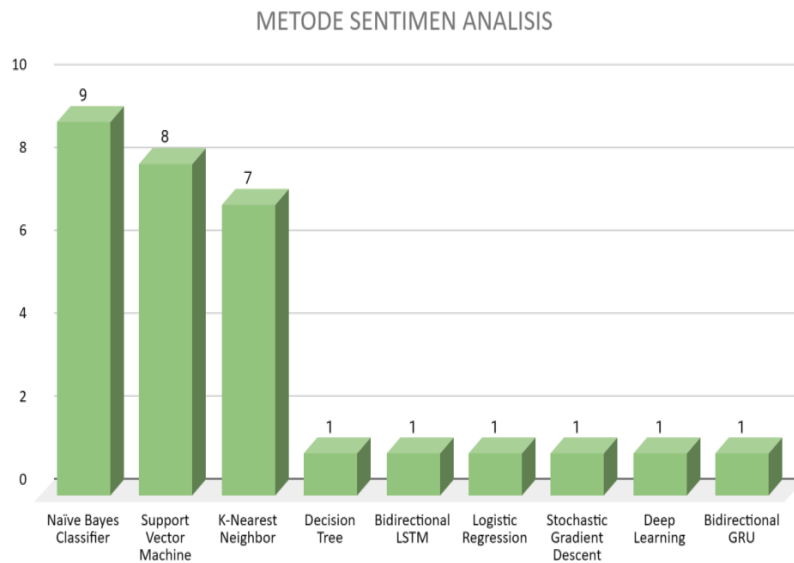
Hasil ini menjawab RQ1 yang menunjukkan bahwa metode yang digunakan dalam sentimen analisis yaitu Naïve Bayes Classifier, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Bidirectional LSTM, Logistic Regression, Stochastic Gradient Descent, Deep Learning, Bidirectional GRU.

3.2. RQ2 Algoritma apa yang populer dalam analisis sentimen?

Untuk mengetahui algoritma yang sering digunakan dalam sentimen analisis, maka munculah research question mengenai algoritma apa yang populer dalam analisis sentimen. Dari 30 jurnal yang sudah di review, menunjukkan bahwa algoritma yang populer dalam analisis sentimen yaitu Naïve Bayes Classifier dengan jumlah jurnal 9 jurnal.

Prinsip kerja metode Naïve Bayes yaitu algoritma klasifikasi probabilistik berdasarkan teorema Bayes. Dalam konteks analisis sentimen, algoritma ini menghitung probabilitas bahwa suatu dokumen (teks) termasuk dalam kategori sentiment tertentu.

Algoritma Naïve Bayes Classifier menjadi algoritma paling populer karena Naïve Bayes memiliki Tingkat akurasi yang tinggi ketika diterapkan pada database yang berjumlah banyak dan beragam. Selain itu, algoritma Naïve Bayes Classifier juga optimal dalam meminimalkan Tingkat error dibandingkan dengan algoritma lainnya. Hal ini dibuktikan dengan penelitian yang telah dilakukan sebelumnya pada media yang sama yaitu twitter, algoritma ini menghasilkan akurasi paling tinggi saat dibandingkan dengan algoritma lainnya seperti Decision Tree dan Random Forest.



Gambar 1. Grafik Metode Sentimen Analisis

Dari grafik diatas dapat diketahui bahwa metode sentiment yang paling sering digunakan yaitu metode Naïve Bayes Classifier dengan total 9 jurnal. Kemudian diikuti oleh metode K-Nearest Neighbor dengan total 7 jurnal. Metode masing-masing satu kali dalam jurnal adalah metode Decision Tree, Bidirectional LSTM, Logistic Regression, Stochastic Gradient Descent, Deep Learning, dan Bidirectional GRU.

Dalam menentukan algoritma yang efektif, faktor akurasi, yaitu presentase prediksi sentimen yang benar dari semua prediksi menjadi perhatian penting dalam melakukan sentimen analisis. Dalam melakukan analisis sentimen memerlukan klasifikasi yang akurat sebagai kelas positif, netral, atau negatif. Ketika terlalu banyak teks diklasifikasikan ke kategori yang salah, maka kemungkinan terjadi kesalahan dalam menafsirkan maksud suatu hal secara substansif.

3.3. RQ3 Algoritma apa yang paling efektif dalam sentimen analisis?

Tabel 2. Hasil akurasi metode

Jurnal ke -	Metode	Akurasi	Diatas 90%
3	Bidirectional GRU	97,48%	1
15	KNN	95,60%	2
20	KNN	99%	
30	KNN	84,93%	
12	NBC	85%	1
23	NBC	94,38%	4
5	SVM	87,50%	
7	SVM	90%	
19	SVM	99,74%	
22	SVM	98%	
26	SVM	99,39%	

Untuk mengetahui algoritma yang paling efektif dalam sentimen analisis, maka munculah research question mengenai algoritma apa yang paling efektif dalam sentimen analisis. Dari 30 jurnal yang sudah di review, menunjukkan bahwa algoritma yang paling efektif dalam analisis sentiment yaitu Support Vector Machine dengan terdapat 4 penelitian yang memiliki hasil akurasi diatas 90%.

3.4. RQ4 Apakah jumlah dataset mempengaruhi Tingkat keakuratan dalam sentiment analisis?

Untuk mengetahui pengaruh jumlah dataset dalam analisis sentimen, maka munculah research question mengenai apakah jumlah dataset mempengaruhi Tingkat keakuratan dalam sentimen analisis.

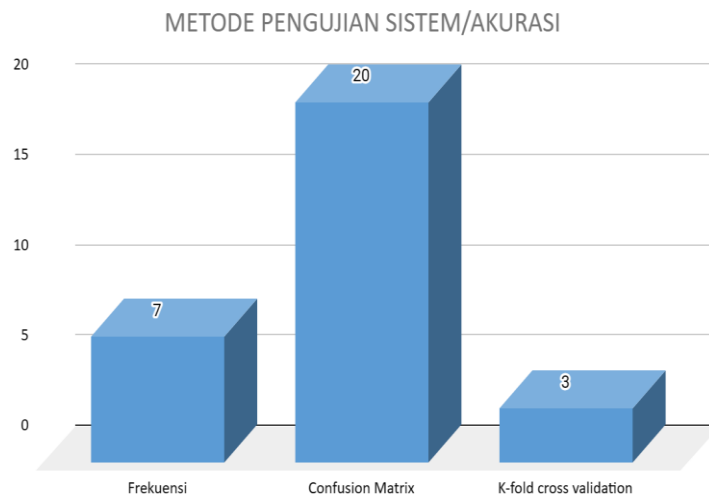
Tabel 3. Dataset yang digunakan

Jurnal ke-	Metode	Akurasi	Jumlah Dataset
3	Bidirectional GRU	97,48%	236905
20	KNN	99%	25000
22	SVM	98%	3000
23	NBC	94,38%	2070
30	KNN	84,93%	1825
19	SVM	99,74%	1680
24	SVM	88,93%	1000
26	SVM	99,39%	982
27	SVM	84,76%	849
7	SVM	90%	400
5	SVM	87,50%	200
2	Deep Learning	88%	30
6	KNN	87%	30
15	KNN	95,60%	29

Dari sampel diatas, dapat diketahui bahwa metode KNN dalam jurnal ke-20 yang digunakan menghasilkan akurasi sebesar 90% dengan jumlah dataset sebanyak 25.000. Di sisi lain, dari jurnal ke- 15, dengan penggunaan metode yang sama yaitu KNN menghasilkan akurasi sebesar 95,60% dengan jumlah dataset hanya 29. Dari perbandingan diatas, maka dapat disimpulkan bahwa jumlah dataset tidak mempengaruhi Tingkat keakuratan dalam sentimen analisis.

3.5. RQ5 Metode apa saja yang digunakan dalam pengujian keakuratan hasil setimen analisis?

Untuk mengetahui metode yang sering digunakan dalam pengujian keakuratan hasil sentimen analisis, maka munculah research question mengenai metode apa saja yang digunakan dalam pengujian keakuratan hasil sentimen analisis. Dari 30 jurnal yang sudah di review, menunjukkan bahwa terdapat beberapa cara dalam menilai keakuratan hasil sentimen analisis yaitu dengan Confusion Matrix, Frekuensi, dan K-Fold Cross Validation.

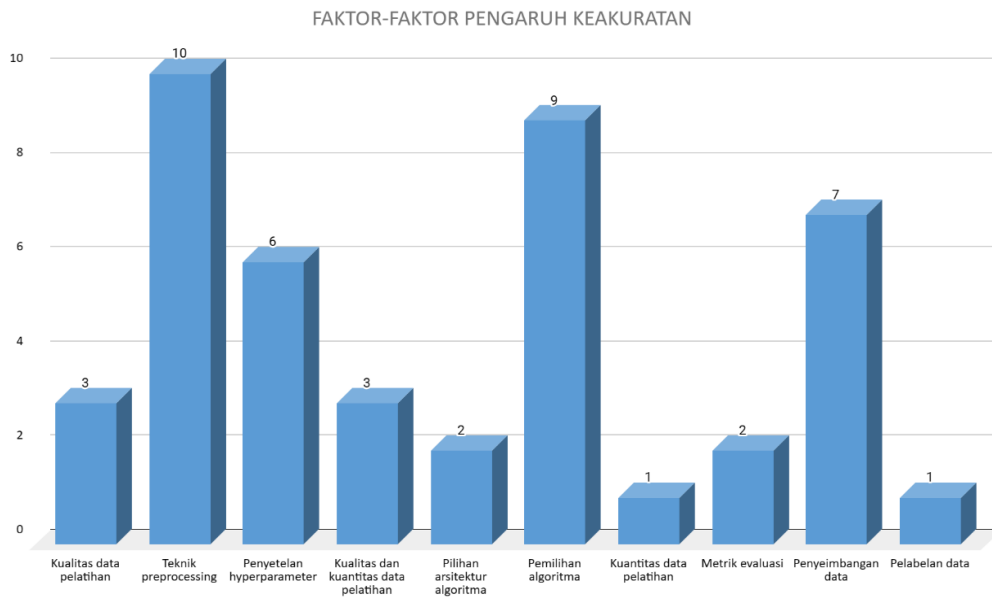


Gambar 2. Grafik Pengujian Sistem

Dari grafik diatas, dapat diketahui bahwa metode pengujian sistem/akurasi yang paling sering digunakan yaitu metode Confusion Matrix dengan total 18 jurnal. Selanjutnya terdapat metode Frekuensi dengan total 6

jurnal. Sedangkan metode K-Fold Cross Validation dan Classification of Performance Evaluation dengan masing-masing hanya 1 jurnal.

3.6. RQ6 Apa faktor-faktor yang dapat mempengaruhi hasil dari sentimen analisis?



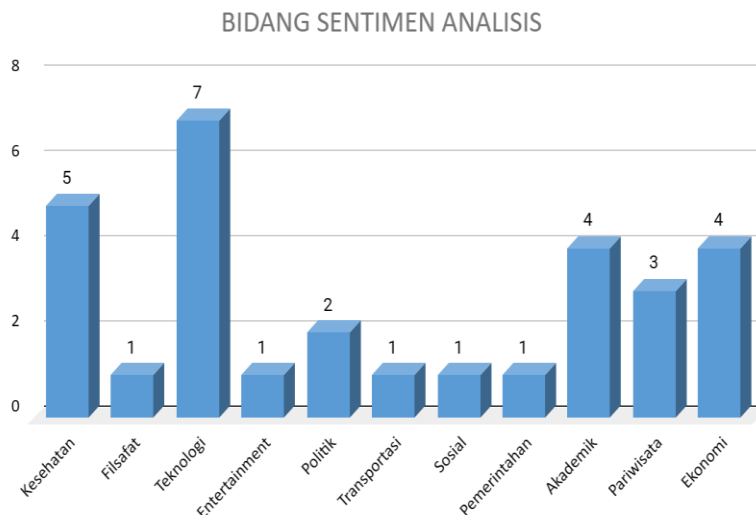
Gambar 3. Grafik Faktor-Faktor Yang Mempengaruhi Keakuratan

Dari grafik diatas, dapat diketahui bahwa faktor yang paling berpengaruh terhadap keakuratan (Akurasi) yaitu teknik preprocessing dengan total 10 jurnal. Selanjutnya terdapat faktor pemilihan algoritma dengan total 9 jurnal. Kemudian diikuti oleh faktor penyeimbangan data dengan total 7 jurnal. Selanjutnya yaitu faktor penyetelan hyperparameter dengan total 6 jurnal. Sedangkan faktor lain seperti kualitas dan kuantitas data pelatihan, pemilihan arsitektur algoritma, metrik evaluasi, dan

pelabelan data jarang sekali berpengaruh terhadap keakuratan dengan masing-masing kemunculan kurang dari 4 jurnal.

3.7. RQ7 Bidang apa saja yang menjadi objek dalam sentimen analisis?

Untuk mengetahui macam-macam bidang yang menjadi objek dalam sentimen analisis, maka munculah research question mengenai bidang apa saja yang menjadi objek dalam sentimen analisis.



Gambar 4. Grafik Bidang Sentimen Analisis

Dari grafik diatas, dapat diketahui bahwa bidang yang paling sering menjadi objek penelitian yaitu teknologi dengan total 7 jurnal. Selanjutnya terdapat bidang kesehatan dengan

total 5 jurnal. Kemudian diikuti oleh bidang akademik dan ekonomi dengan total 4 jurnal. Selanjutnya yaitu bidang pariwisata dengan total 3 jurnal. Kemudian terdapat bidang

politik dengan total 2 jurnal. Sedangkan bidang lain seperti filsafat, entertainment, transportasi, social dan pemerintahan masing-masing hanya 1 jurnal.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil beberapa kesimpulan sebagai berikut:

1. Terdapat banyak metode yang dapat digunakan dalam melakukan penelitian mengenai sentimen analisis, di antaranya yaitu Naïve Bayes Classifier, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Bidirectional LSTM, Logistic Regression, Stochastic Gradient Descent, Deep Learning, dan Bidirectional GRU.
2. Berdasarkan hasil literatur review, metode Naïve Bayes Classifier merupakan metode yang paling sering digunakan dalam melakukan sentimen analisis. Kemunculan metode Naïve Bayes Classifier yaitu sebanyak 9 kali dari 30 literatur.
3. Metode yang paling efektif dalam analisis sentimen yaitu Support Vector Machine. Pemilihan metode paling efektif dihitung berdasarkan total jurnal yang menghasilkan akurasi diatas 90%, dimana dari total 8 jurnal yang menggunakan metode Support Vector Machine, 4 jurnal diantaranya memiliki hasil akurasi diatas 90%.
4. Jumlah dataset tidak berpengaruh terhadap tingkat keakuratan dalam sentimen analisis. Hal ini dibuktikan dengan hasil dari Jurnal ke-20 yang menggunakan metode KNN menghasilkan akurasi sebesar 99% dengan jumlah dataset sebanyak 25.000. Di sisi lain, dari Jurnal ke-15, dengan penggunaan metode yang sama yaitu KNN menghasilkan akurasi sebesar 95,60% dengan jumlah dataset hanya sebanyak 29 data.
5. Metode pengujian sistem/akurasi yang paling sering digunakan yaitu metode Confusion Matrix dengan total 18 jurnal.
6. Faktor yang paling berpengaruh terhadap keakuratan (Akurasi) yaitu teknik preprocessing.
7. Bidang yang paling sering menjadi objek penelitian yaitu pada bidang teknologi.

5. Daftar Pustaka

- Ainul Wildan, R. S., Adam Rajagede, R., & Rahmadi, R. (2021). Analisis Sentimen Politik Berdasarkan Big Data dari Media Sosial Youtube: Sebuah Tinjauan Literatur. *Prosiding Automata*, 2(1).
- Aliandu, P. (2015). Sentiment Analysis to Determine Accommodation, Shopping and Culinary Location on Foursquare in Kupang City. *Procedia Computer Science*, 72, 300–305. <https://doi.org/10.1016/j.procs.2015.12.144>
- Andriana, H., Shofia Hilabi, S., & Hananto, A. (2023). Penerapan Metode K-Nearest Neighbor pada Sentimen Analisis Pengguna Twitter terhadap KTT G20 di Indonesia. *JURIKOM (Jurnal Riset Komputer)*, 10(1), 60–67. <https://doi.org/10.30865/jurikom.v10i1.5427>
- Apriliansi, D., Susanto, A., Hidayattullah, M. F., & Sasmito, G. W. (2023). Sentimen Analisis Pandangan Masyarakat Terhadap Vaksinasi Covid 19 Menggunakan K-Nearest Neighbors. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(1), 34–37. <https://doi.org/10.30591/jpit.v8i1.4759>
- Ash, S., & Surya, A. (2022). Analisis Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Metode Logistic Regression. *Jurnal CoSciTech (Computer Science and Information Technology)*, 99–106.
- Athira Luqyana, W., Cholissodin, I., & Perdana, R. S. (2018). Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 4704–4713. <http://j-ptiik.ub.ac.id>
- Cholifah, C., Hanny Hikmayanti Handayani, & Ayu Ratna Juwita. (2023). Analisis Sentimen Twitter Terhadap Uu Omnibus Law Menggunakan Metode Support Vector Machine (Svm) Dan Naïve Bayes Classifier (Nbc). *Jurnal Informatika Teknologi Dan Sains*, 4(4), 483–488. <https://doi.org/10.51401/jinteks.v4i4.2191>
- Deviyanto, A., & Wahyudi, M. D. R. (2018). Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 3(1), 1. <https://doi.org/10.14421/jiska.2018.31-01>
- Dusoye, H., & Cadersaib, Z. (2018). Sentiment analytics framework integrating Twitter and Odo ERP. *2017 International Conference on Infocom Technologies and Unmanned Systems: Trends and Future Directions, ICTUS 2017, 2018-Janua*, 145–151. <https://doi.org/10.1109/ICTUS.2017.8285994>
- Faradhillah, N. Y. A., Kusumawardani, R. P., & Hafidz, I. (2016). Eksperimen Sistem Klasifikasi Analisa Sentimen Twitter pada Akun Resmi Pemerintah Kota Surabaya Berbasis Pembelajaran Mesin (Experiments on Sentiment Classification System for Tweets of the Official Account of the City Government of Surabaya based on Mach. *Prosiding Seminar Nasional Sistem Informasi Indonesia 2016*, 15–24. <http://is.its.ac.id/pubs/oajis/index.php/home/detail/1645/EKSPERIMEN-SISTEM-KLASIFIKASI-ANALISA-SENTIMEN-TWITTER-PADA-AKUN-RESMI-PEMERINTAH-KOTA-SURABAYA-BERBASIS-PEMBELAJARAN-MESIN>
- Febriyani, E., & Februariyanti, H. (2023). Analisis

- Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter. *Jurnal Tekno Kompak*, 17(1), 25. <https://doi.org/10.33365/jtk.v17i1.2061>
- Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naive Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
- Hartmann, J., Heitmann, M., Siebert, C., & Schamp, C. (2023). More than a Feeling: Accuracy and Application of Sentiment Analysis. *International Journal of Research in Marketing*, 40(1), 75–87. <https://doi.org/10.1016/j.ijresmar.2022.05.005>
- Hasanah, S., Purwasih, I., & Santoso, I. (2023). Analisis Sentimen Terhadap Masyarakat Adanya Uang Kertas Baru Menggunakan Algoritma K-Nearest Neighbor(Knn). *IKRA-ITH Informatika ...*, 7(2), 105–114. <http://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/2813%0Ahttps://journal.s.upi-yai.ac.id/index.php/ikraith-informatika/article/download/2813/2065>
- Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, 5(1), 32–35. <https://doi.org/10.37905/jjee.v5i1.16830>
- Ihsan, M., Benny Sukma Negara, & Surya Agustian. (2022). LSTM (Long Short Term Memory) for Sentiment COVID-19 Vaccine Classification on Twitter. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 13(1), 79–89. <https://doi.org/10.31849/digitalzone.v13i1.9950>
- Karim, A. (2020). Analisis Sentimen Pada Komentar Sosial Media Instagram Layanan Kesehatan BPJS Menggunakan Naive Bayes Classifier. *Repository Unisula*, 5(3), 248–253.
- Khairudin, M., Sukendar, A., & Somantri, A. (2023). Analisis Sentimen Film Di Twitter Menggunakan Metode Support Vector Machine. *Jurnal Sains Dan Sistem Teknologi Informasi*, 5(1), 97–102. <https://doi.org/10.59811/sandi.v5i1.47>
- Kurniawan, I., Lia Hananto, A., Shofia Hilabi, S., Hananto, A., Priyatna, B., & Yuniar Rahman, A. (2023). Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter. *Jurnal Teknik Informatika Dan Sistem Informasi*, 10(1), 731–740. <http://jurnal.mdp.ac.id>
- Lutfina, E., Nugroho, A., Abdillah, M. Z., Informasi, S., Sains, F., Karangturi, U. N., & Patah, J. R. (2022). SYSTEMATIC LITERATURE REVIEW UNTUK IDENTIFIKASI METODE FRAGMENTASI PADA PERANCANGAN SISTEM DATABASE TERDISTRIBUSI. *Science, Technology and Management Journal* <http://Journal.Unkartur.Ac.Id/Index.Php/Stmj>
- SYSTEMATIC, 2(2), 56–61.
- Lutfina, E., Setiawan, R. O. C., Nugroho, A., & Abdillah, M. Z. (2023). PERANCANGAN APLIKASI PEMBELAJARAN DENGAN KONSEP GAMIFIKASI Systematic Literature Review. *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi*, 7(1), 78–87. <https://doi.org/10.46880/jmika.vol7no1.pp78-87>
- Muslimin, M., & Lusiana, V. (2023). Analisis Sentimen Terhadap Kenaikan Harga Bahan Pokok Menggunakan Metode Naive Bayes Classifier. *Jurnal Media Informatika Budidarma*, 7(3), 1200–1209. <https://doi.org/10.30865/mib.v7i3.6418>
- Pramesti, L. A., & Pratiwi, N. (2023). Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine. *Journal of Information System Research (JOSH)*, 4(4), 1145–1154. <https://doi.org/10.47065/josh.v4i4.3807>
- Riany, J., Fajar, M., & Lukman, M. P. (2016). Penerapan Deep Sentiment Analysis pada Angket Penilaian Terbuka Menggunakan K-Nearest Neighbor. *Sisfo*, 06(01), 147–156. <https://doi.org/10.24089/j.sisfo.2016.09.011>
- Rifaldi, M. I., Ramadhan, Y. R., & Jaelani, I. (2023). Analisis Sentimen Terhadap Aplikasi Chatgpt Pada Twitter Menggunakan Algoritma Naive Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 7(2), 802–814. <http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/viewFile/687/662>
- Rohman, M. A., Suhartono, & Chamidy, T. (2023). Bidirectional GRU dengan Attention Mechanism pada Analisis Sentimen PLN Mobile Bidirectional GRU with Attention Mechanism on Sentiment Analysis of PLN Mobile. *Techno.Com*, 22(2), 358–372.
- Sanjaya, F., Lutfina, E., Nugroho, A., & Abdillah, M. (2023). Systematic Literature Review Untuk Perancangan Sistem Informasi Stok Opname Gudang Berbasis Web. *Science, Technology and Management Journal*, 3(1), 182–192. <http://journal.unkartur.ac.id/index.php/stmj>
- Saputra, I., AJI PAMBUDI, R. S., DARONO, H. E., AMSURY, F., FAHDIA, M. R., RAMADHAN, B., & ARDIANSYAH, A. (2021). Analisis Sentimen Pengguna Marketplace Bukalapak dan Tokopedia di Twitter Menggunakan Machine Learning. *Faktor Exacta*, 13(4), 200. <https://doi.org/10.30998/faktorexacta.v13i4.7074>
- Saraswati, M., & Rimirasih, D. (2020). Analisis Sentimen Terhadap Pelayanan Krl Commuterline Berdasarkan Data Twitter Menggunakan Algoritma Bernoulli Naive Bayes. *Jurnal Ilmiah Informatika Komputer*, 25(3), 225–238. <https://doi.org/10.35760/ik.2020.v25i3.3256>
- Singgalen, Y. A. (2023a). Analisis Sentimen dan Sistem Pendukung Keputusan Menginap di Hotel

- Menggunakan Metode CRISP-DM dan SAW. *Journal of Information System Research (JOSH)*, 4(4), 1343–1353. <https://doi.org/10.47065/josh.v4i4.3917>
- Singgalean, Y. A. (2023b). Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM. *Journal of Information System Research (JOSH)*, 4(2), 614–625. <https://doi.org/10.47065/josh.v4i2.2999>
- Supriyanto, J., Alita, D., & Isnain, A. R. (2023). Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 4(1), 74–80. <https://doi.org/10.33365/jatika.v4i1.2468>
- Wahyuni, L., Triandi, B., Zarlis, M., & Nasution, Z. (2023). Pendekatan Filsafat Ilmu Terhadap Perkembangan Deep Learning dalam Perspektif Aksiologi. *Computer Science Research and Its Development Journal*, 15(1), 62–72. <http://csridjournal.potensi-utama.ac.id/index.php/CSRIDjournal/article/view/39%0Ahttps://csridjournal.potensi-utama.ac.id/index.php/CSRIDjournal/article/download/39/16>
- Yohana, P., Agustian, S., & Gusti, S. K. (2022). Klasifikasi Sentimen Masyarakat terhadap Kebijakan Vaksin Covid-19 pada Twitter dengan Imbalance Classes Menggunakan Naive Bayes. *Seminar Nasional Teknologi ...*, 69–80. <http://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/19012%0Ahttp://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/viewFile/19012/8336>