SNESTIK

Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika

https://ejurnal.itats.ac.id/snestik dan https://snestik.itats.ac.id

Informasi Pelaksanaan :
SNESTIK IV - Surabaya, 27 April 2024
Ruang Seminar Gedung A, Kampus Institut Teknologi Adhi Tama Surabaya

Informasi Artikel:
DOI : 10.31284/p.snestik.2024.5877

Prosiding ISSN 2775-5126

Fakultas Teknik Elektro dan Teknologi Informasi-Institut Teknologi Adhi Tama Surabaya
Gedung A-ITATS, Jl. Arief Rachman Hakim 100 Surabaya 60117 Telp. (031) 5945043
Email : snestik@itats.ac.id

Analisis Sentimen terhadap Kecurangan Pemilu dan SIREKAP di Twitter menggunakan Metode Vader Lexicon dan Naïve Bayes

Mukhamad Angga Gumilang, Achmad Sirojudin, Fauzan Abdilllah, Muhammad Yusril Amin, Winda Budi Lestari

Politeknik Negeri Jember
angga.gumilang@polije.ac.id

ABSTRACT

Elections are crucial to democracy, serving as a means to elect representatives for executive and legislative bodies. In Indonesia, traditional voting methods are being supplemented by electronic systems like the Vote Recapitulation System (SIREKAP). However, issues of election fraud persist, drawing attention on social media platforms like Twitter. This study employs sentiment analysis techniques, including VADER lexicon and Naive Bayes algorithm, to gauge public sentiment towards election integrity and SIREKAP. Through Twitter data analysis, insights into public perceptions are gained. The results, comprising 456 negative, 302 neutral, and 553 positive comments regarding the 2024 election process, reflect the complex spectrum of public opinion. Similarly, sentiments regarding election fraud, with 1079 negative and 181 positive comments, highlight the multifaceted views of the populace. Visual representations such as word clouds aid in identifying key themes. This research offers valuable insights for stakeholders, aiding in understanding public opinion dynamics and informing decision-making processes in enhancing democratic integrity.

Keywords: *General elections; SIREKAP; sentiment analysis; twitter; election fraud.*

ABSTRAK

Pemilihan umum adalah krusial dalam demokrasi sebagai sarana pemilihan wakil untuk badan eksekutif dan legislatif. Di Indonesia, metode pemilihan tradisional kini disokong oleh sistem elektronik seperti Sistem Rekapitulasi Suara (SIREKAP). Namun, masalah kecurangan pemilu tetap menjadi sorotan, terutama di platform media sosial seperti Twitter. Penelitian ini menggunakan teknik analisis sentimen, termasuk leksikon VADER dan algoritma Naive Bayes, untuk menilai sentimen publik terhadap integritas pemilu dan SIREKAP. Melalui

analisis data Twitter, diperoleh wawasan tentang persepsi publik. Hasilnya, yang terdiri dari 456 komentar negatif, 302 netral, dan 553 positif mengenai proses pemilu 2024, mencerminkan spektrum kompleks dari pendapat publik. Demikian juga, sentimen terkait kecurangan pemilu, dengan 1079 komentar negatif dan 181 komentar positif, menyoroti pandangan yang beragam dari masyarakat. Representasi visual seperti wordcloud membantu mengidentifikasi tema-tema utama. Penelitian ini menawarkan wawasan berharga bagi para pemangku kepentingan, membantu memahami dinamika opini publik dan menginformasikan proses pengambilan keputusan dalam meningkatkan integritas demokratis.

Kata kunci: Pemilihan umum; SIREKAP; analisis sentimen; twitter; kecurangan pemilu;

PENDAHULUAN

Pemilihan umum adalah pilar demokrasi penting yang menentukan wakil rakyat di pemerintahan, dilakukan melalui mekanisme pemungutan suara di Tempat Pemungutan Suara (TPS) dengan kertas suara konvensional [1]. Namun, perkembangan teknologi membawa adopsi sistem elektronik seperti Sistem Rekapitulasi Suara (SIREKAP) untuk memfasilitasi proses pemilu [2]. Meskipun demikian, kecurangan dalam pemilu tetap menjadi isu, dengan media sosial menjadi saluran utama ekspresi dan penyebaran informasi masyarakat [3].

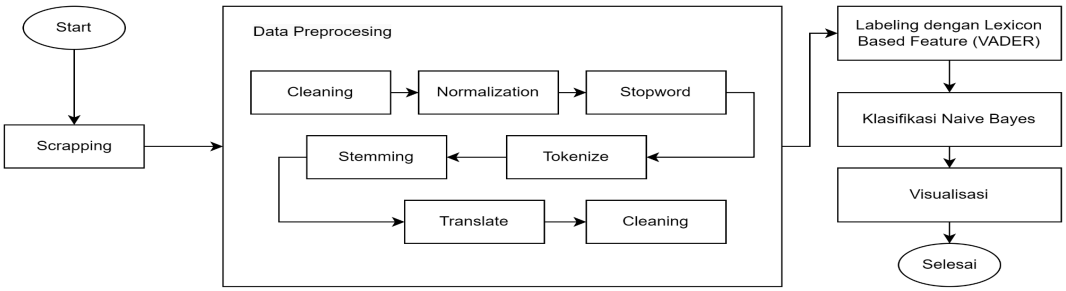
Media sosial, khususnya Twitter, telah menjadi platform utama bagi masyarakat untuk berbagi pandangan, pemikiran, dan respons terhadap berbagai isu, termasuk pemilu. Twitter sendiri merupakan sebuah platform media sosial yang memungkinkan penggunanya untuk menyampaikan pemikiran atau gagasan mengenai topik tertentu melalui tulisan atau pesan singkat [4]. Twitter juga merupakan media sosial yang sedang populer saat ini, seperti yang diungkapkan oleh Khusnul Khotimah Sirajuddin dan rekannya. Dalam penelitiannya, dari 100 pengguna Generasi Z yang aktif di media sosial, sebanyak 69 di antaranya lebih memilih untuk menggunakan Instagram sebagai platform yang paling sering mereka kunjungi [5]. Dengan demikian, Twitter menjadi salah satu wadah bagi masyarakat untuk mengekspresikan pandangan serta mengungkapkan keresahan terhadap isu atau permasalahan yang timbul.

Dalam konteks ini, Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kecurangan pemilu dan SIREKAP melalui data yang dikumpulkan dari media sosial Twitter. Analisis sentimen adalah jenis penelitian yang menggunakan pendapat atau opini mengenai suatu topik untuk menilai apakah opini tersebut bersifat positif, negatif, atau netral [6], [7]. Analisis sentimen juga dapat digunakan sebagai indikator untuk menilai seberapa populer atau berapa banyak dukungan yang diterima oleh seorang figur publik atau tokoh masyarakat [8].

Penelitian sebelumnya yang dilakukan oleh Solagratia Saron Tandiapa dan Gladly Caren Rorimpandey melaksanakan sebuah penelitian yang bertajuk ANALISIS SENTIMEN ULASAN PENGGUNA PADA APLIKASI THREADS MENGGUNAKAN METODE LEXICON BASED DAN NAIVE BAYES CLASSIFIER. Penelitian ini mencatat perbandingan hasil akurasi klasifikasi antara metode lexicon based dan naive bayes classifier, yaitu 55% dan 51% secara berturut-turut. Dengan perbedaan akurasi tersebut, disimpulkan bahwa analisis sentimen dengan metode lexicon based memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan metode naive bayes classifier [9]. Penelitian lain yang dilakukan oleh Putri Agung Permatasari dan rekan-rekannya dalam karya berjudul Survei Tentang Analisis Sentimen Pada Media Sosial, juga mengulas literatur tentang analisis media sosial termasuk metode yang umum digunakan dan platform media sosial yang sering dijadikan subjek penelitian. Dalam penelitian tersebut menyebutkan bahwa Algoritma Naive Bayes terbukti efektif dalam menganalisis teks [10].

Melalui Penelitian ini, diharapkan dapat terungkap pandangan dan perasaan masyarakat secara menyeluruh terhadap integritas pemilu dan proses SIREKAP yang berkembang di media sosial Twitter dan mengetahui hasil akurasi kepuasan masyarakat dengan metode metode VADER (Valence Aware Dictionary and sentimen Reasoner) lexicon dan algoritma naïve bayes. Hasil analisis sentimen ini dapat memberikan wawasan yang berharga bagi pihak-pihak terkait, seperti penyelenggara pemilu, partai politik, dan masyarakat umum, dalam memahami dinamika opini publik dan perasaan masyarakat terkait dengan proses pemilu yang sedang berlangsung.

METODE



Gambar 1 Alur Penelitian

Dalam Penelitian ini terdapat beberapa lakukan yang dilakukan yang mana dapat dilihat pada **Gambar 1**. Langkah Pertama yang dilakukan adalah Scraping yaitu kegiatan mengambil data atau informasi dari sebuah situs web secara otomatis. Langkah kedua yang dilakukan adalah Data preprocessing yaitu proses pengolahan data dalam data mining yang bertujuan untuk menyiapkan dan mengubah data agar sesuai dengan tahap-tahap dalam proses data mining. Langkah ketiga dalam proses ini adalah melakukan proses labeling dengan menerapkan metode vader lexicon. VADER Lexicon adalah salah satu metode analisis dalam kategori algoritma berbasis lexicon. Dalam algoritma ini, data dinilai dan dievaluasi berdasarkan kamus kata-kata. Outputnya mencakup kelas polaritas seperti positif, netral, dan negatif, serta skor komposit atau total[9]. Langkah keempat adalah melakukan proses klasifikasi yaitu tindakan menempatkan (mengklasifikasikan) data ke dalam satu atau lebih kelas yang telah ditentukan sebelumnya [11] dan metode yang digunakan untuk proses klasifikasi dalam penelitian ini adalah Naive Bayes. Metode Naive Bayes adalah suatu teknik klasifikasi yang sederhana yang memperkirakan semua probabilitas berdasarkan pada teorema Bayes. Metode ini menggabungkan nilai frekuensi dari basis data dan merupakan algoritma untuk menghitung peluang bersyarat (posterior), yakni peluang suatu kejadian X terjadi ketika kejadian H telah diketahui, yang disimbolkan dengan P(X|H)[12], [13]. Kemudian, langkah terakhir adalah visualisasi data, di mana proses ini bertujuan untuk menyajikan gambaran yang jelas terhadap data yang telah dikelompokkan ke dalam berbagai kelas atau kategori yang telah ditentukan sebelumnya.

HASIL DAN PEMBAHASAN

Data Preprocessing

Proses kategorisasi komentar-komentar terkait Pemilu 2024 dimulai dengan mengidentifikasi tiga kategori sentimen utama: negatif, netral, dan positif. Komentar-komentar yang mengekspresikan ketidakpuasan, keraguan, atau kritik terhadap proses pemilu atau kandidat diklasifikasikan sebagai negatif, sementara yang bersifat berita atau fakta tanpa ekspresi sentimen jelas dianggap netral. Komentar-komentar yang menyatakan dukungan, kepuasan, atau harapan terhadap pemilu atau kandidat masuk dalam kategori positif.Contoh sentimen pada 3 kategori, seperti yang ditunjukkan pada **Tabel 1** dibawah ini.

Tabel 1. Jenis-jenis Komentar

Kategori	Komentar
Komentar Positif	Pada akhirnya pemilu ini bukan siapa lagi yang menang tapi siapa yang setuju dan berjuang menegakan kebenaran dan kejujuran diatas kecurangan-kecurangan. Yang menang belum tentu mulia. Yang Kalah belum tentu hina. Dicaci tidak tumbang Dipuja tidak jumawa. Salam akal sehat

Komentar Netral	Hasto PDIP Sebut Data Sirekap Alami 753 Kali Perubahan Usai KPU Tetapkan Hasil Pemilu 2024 https://t.co/Lq4RAoSosT #Pemilu
Komentar Negatif	percuma ga si? kalo emang iya ada kecurangan pemilu oleh paslon 2 besar kemungkinan juga paslon 2 mencurangi gugatan tersebut? kalo kalah yaudah kalaj aja yg legowo besok ikut pemilu lagi kayak paslon 2 beberapa kali kalah dan skarang bisa menang

Tahap preprocessing analisis sentimen melibatkan membersihkan teks komentar dari karakter tidak diinginkan, normalisasi untuk konsistensi istilah, dan menghilangkan kata-kata umum yang tidak relevan. Proses tokenisasi membagi teks menjadi token untuk analisis, sementara stemming mengubah kata menjadi bentuk dasar. Komentar bahasa asing mungkin diterjemahkan untuk konsistensi, dan teks diperiksa untuk memastikan keakuratan preprocessing, sehingga siap untuk analisis sentimen lebih lanjut terkait Pemilu 2024.

Labeling dengan Lexicon Based VADER

Setelah tahapan preprocessing, komentar-komentar yang telah disiapkan kemudian dianalisis menggunakan metode Lexicon Based VADER untuk menentukan sentimen mereka secara lebih terperinci. Metode ini memanfaatkan sebuah lexicon yang berisi daftar kata-kata bersama dengan skor sentimen yang terkait. Setiap kata dalam komentar dievaluasi berdasarkan lexicon ini, dan skor sentimen keseluruhan dari sebuah komentar dihitung berdasarkan skor-skor kata-kata individual serta struktur kalimatnya. Hasil analisis ini menyediakan pemahaman yang lebih mendalam tentang nuansa sentimen dalam komentar-komentar tersebut, memungkinkan untuk mengidentifikasi tidak hanya apakah sebuah komentar bersifat positif, negatif, atau netral, tetapi juga seberapa kuatnya sentimen tersebut. Dengan demikian, proses labeling menggunakan Lexicon Based VADER memberikan wawasan yang lebih detail dan kompleks terhadap opini publik terkait Pemilu 2024, yang dapat digunakan untuk informasi dan pengambilan keputusan yang lebih baik dalam konteks politik dan sosial.

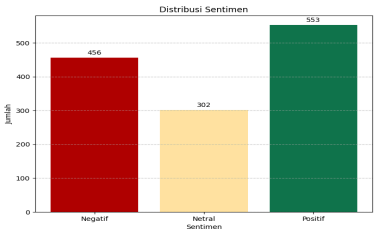
Klasifikasi Naive Bayes

Setelah proses labeling, langkah selanjutnya adalah melakukan training model klasifikasi. Dalam konteks ini, model klasifikasi Naive Bayes dapat digunakan. Pertama, data yang telah dilabeli dengan sentimen dari Lexicon Based VADER digunakan sebagai data latih untuk model Naive Bayes. Model ini akan dilatih untuk mengidentifikasi pola dan hubungan antara fitur-fitur (misalnya, kata-kata atau token) dalam komentar dan sentimen yang terkait.

Selanjutnya, setelah pelatihan selesai, model Naive Bayes dapat digunakan untuk mengklasifikasikan komentar-komentar berdasarkan pada fitur-fitur yang diambil dari komentar tersebut. Dengan memanfaatkan probabilitas dari kemunculan fitur-fitur tersebut dalam kategori sentimen yang berbeda, model Naive Bayes dapat memberikan prediksi tentang sentimen mana yang paling mungkin terkait dengan suatu komentar. Hasil dari klasifikasi ini dapat memberikan pemahaman yang lebih mendalam tentang sentimen yang tersirat dalam komentar-komentar terkait Pemilu 2024, serta memungkinkan untuk mengklasifikasikan komentar-komentar tersebut ke dalam kategori sentimen yang tepat (negatif, netral, atau positif) secara otomatis.

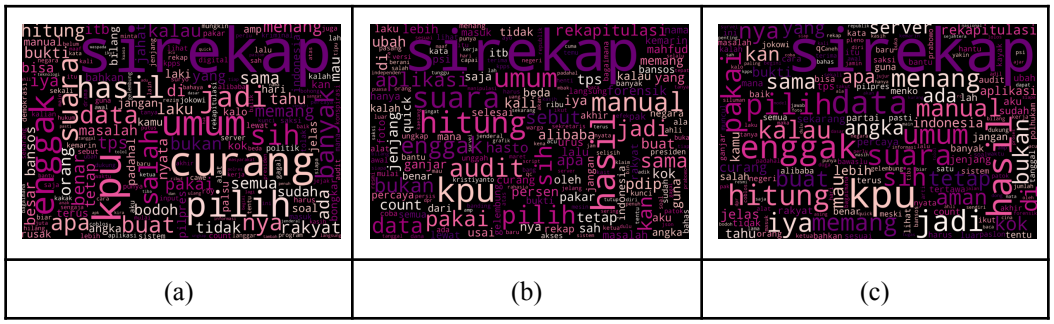
1. SIREKAP

Distribusi sentimen pada **Gambar 2** dibawah yang bervariasi ini mencerminkan kompleksitas opini publik yang perlu dipahami secara mendalam untuk meningkatkan proses demokratisasi di masa mendatang.



Gambar 2. Distribusi Sentimen SIREKAP

Wordcloud yang didasarkan pada komentar-komentar tersebut disajikan dalam **Gambar 3a-3c** di bawah ini.

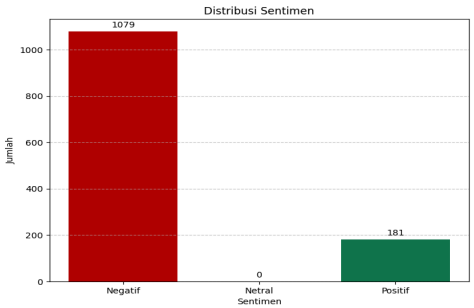


(a) (b) (c)

Gambar 3. a) Wordcloud Negatif, b) Wordcloud Netral, c) Wordcloud Positif

2. Kecurangan Pemilu

Distribusi sentimen pada **Gambar 4** dibawah ini menggambarkan kompleksitas pandangan masyarakat terhadap isu penting ini, yang perlu dipahami secara mendalam untuk memperbaiki integritas dan transparansi dalam proses demokrasi di masa depan.



Gambar 4. Distribusi Sentimen Kecurangan Pemilu

Gambar 5a-5b di bawah menunjukkan Word Cloud yang didasarkan pada komentar-komentar tersebut.



(a) (b)

Gambar 5. a) Wordcloud Negatif, b) Wordcloud Positif

KESIMPULAN

Analisis sentimen terhadap komentar-komentar di platform SIREKAP terkait Pemilu 2024 dan isu kecurangan menunjukkan spektrum opini yang luas dari masyarakat. Ada ketidakpuasan dan kritik yang signifikan terhadap proses pemilihan, namun juga ada dukungan dan kepuasan terhadapnya. Analisis menggunakan metode Lexicon Based VADER dan klasifikasi Naive Bayes membantu memahami nuansa sentimen tersebut. Sebagian besar komentar terkait Pemilu mengekspresikan ketidakpuasan, sementara dalam konteks isu kecurangan, mayoritas komentar bersifat negatif, yaitu 85.6%.Pemahaman yang mendalam diperlukan untuk meningkatkan proses demokratisasi ke depan, dengan fokus pada integritas dan transparansi. Penyajian visual menggunakan word cloud memperkuat representasi kata-kata kunci dalam setiap kategori sentimen, mempermudah identifikasi tema-tema utama dalam diskusi publik. Analisis ini memberikan wawasan berharga bagi pengambil keputusan dalam merancang strategi untuk meningkatkan integritas dan transparansi dalam proses demokrasi.

DAFTAR PUSTAKA

[1] Y. Yusrin and S. Salpina, “Partisipasi Generasi Millenial dalam Mengawasi Tahapan Pemilu 2024,” *J. Educ.*, vol. 5, no. 3, pp. 9646–9653, Feb. 2023, doi: 10.31004/joe.v5i3.1842.

[2] I. A. Pradesa, “ANALISIS PENGGUNAAN SISTEM REKAPITULASI SUARA (SIREKAP) DALAM MENGHADAPI PROBLEMATIKA PEMILU 2024,” *Open Access*, vol. 03, no. 04, 2024.

[3] M. A. Lubis, M. Y. A. Gea, and N. Muniifah, “Penerapan Asas Pemilu Terhadap Electronic Voting (E-Voting) Pada Pemilu Tahun 2024,” *J. Ilm. Penegakan Huk.*, vol. 9, no. 1, pp. 44–56, Jun. 2022, doi: 10.31289/jiph.v9i1.6491.

[4] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” *J. Tek. Inform. Dan Sist. Inf.*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3216.

[5] Khusnul Khotimah Sirajuddin, Ahmad, and Dian Novita Siswanti, “Hubungan Harga Diri Dengan Kepuasan Hidup Generasi Z Pengguna Media Sosial Instagram,” *PESHUM J. Pendidik. Sos. Dan Hum.*, vol. 2, no. 2, pp. 255–265, Feb. 2023, doi: 10.56799/peshum.v2i2.1390.

[6] D. W. Ardras and A. Voutama, “ANALISIS SENTIMEN ANTI LGBT DI INDONESIA MELALUI MEDIA SOSIAL TWITTER,” *J. Tek.*, vol. 15, no. 1, pp. 23–28, Mar. 2023, doi: 10.30736/jt.v15i1.926.

[7] D. Oktavia and Y. R. Ramadahan, “Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM)”.

[8] F. P. Rachman, "Perbandingan Model Deep Learning untuk Klasifikasi Sentiment Analysis dengan Teknik Natural Languange Processing," *J. Teknol. Dan Manaj. Inform.*, vol. 7, no. 2, pp. 113–121, Dec. 2021, doi: 10.26905/jtmi.v7i2.6506.

[9] S. S. Tandiapa and G. C. Rorimpandey, "ANALISIS SENTIMEN ULASAN PENGGUNA PADA APLIKASI THREADS DENGAN METODE LEXICON BASED DAN NAIVE BAYES CLASSIFIER".

[10] P. A. Permatasari, L. Linawati, and L. Jasa, "Survei Tentang Analisis Sentimen Pada Media Sosial," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 177, Dec. 2021, doi: 10.24843/MITE.2021.v20i02.P01.

[11] D. Darwis, N. Siskawati, and Z. Abidin, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, Feb. 2021, doi: 10.33365/jtk.v15i1.744.

[12] A. D. Cahyo, "METODE NAIVE BAYES UNTUK KLASIFIKASI MASA STUDI SARJANA," vol. 3, 2023.

[13] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.