



SYSTEMATIC LITERATURE REVIEW (SLR): ANALISIS SENTIMEN PEMILIHAN CALON PRESIDEN 2024 MENGGUNAKAN METODE NAIVE BAYES

Destri Komalasari¹⁾, Nopita Amelia²⁾, Novita Damayanti³⁾, dan Habi Baturohmah⁴⁾

^{1, 2, 3, 4)} Program Studi Sistem Informasi, Universitas Nusa Putra

Jl. Raya Cibatu Cisaat No.21, Cibolang Kaler, Kec. Cisaat, Kab. Sukabumi, Jawa Barat 43155

e-mail: destri.komalasari_si20@nusaputra.ac.id¹⁾, nopita.amelia_si20@nusaputra.ac.id²⁾,

novita.damayanti_si20@nusaputra.ac.id³⁾, habi.baturohmah@nusaputra.ac.id⁴⁾

*Korespondensi: e-mail: destri.komalasari_si20@nusaputra.ac.id

ABSTRAK

Indonesia akan menyelenggarakan pemilihan presiden secara demokratis pada tahun 2024. Setiap tokoh politik yang dicalonkan sebagai presiden akan mempertimbangkan seberapa populer mereka di mata masyarakat. Menjelang pemilihan umum, hal itu secara signifikan mempengaruhi pembentukan opini publik, pandangan, sentimen, dan preferensi politik terutama di Facebook dan platform media sosial lainnya. Beberapa kegunaan media sosial antara lain untuk mencari teman baru, belajar tentang olahraga, ekonomi, pariwisata, dan politik. Karena salah satunya adalah pimpinan capres 2024, maka penulis ingin mengetahui informasi apa saja tentang capres capres 2024 yang terdapat pada opini publik di situs media sosial seperti facebook. Masalah ini dapat diatasi dengan bantuan penelitian di bidang analisis sentimen, yang berfokus pada studi komputasi opini, perilaku, dan perasaan tentang entitas yang diekspresikan dalam teks. Berdasarkan hasil analisis sentimen, penelitian ini bertujuan untuk menggunakan metode Naive Bayes untuk membagi reaksi publik terhadap berita tentang capres 2024 menjadi tiga kategori: positif, negatif, dan netral.

Kata Kunci: *Calon Presiden, Sosial Media, Analisis Sentimen, Naive Bayes*

ABSTRACT

Indonesia will hold democratic presidential elections in 2024. Each political figure nominated for president will consider how popular they are in the eyes of the public. In the run-up to general elections, it significantly influences the formation of public opinion, views, sentiments and political preferences especially on Facebook and other social media platforms. Some of the uses of social media include making new friends, learning about sports, the economy, tourism, and politics. Because one of them is the leader of the 2024 presidential candidate, the author wants to know any information about the 2024 presidential candidate that is found in public opinion on social media sites such as Facebook. This problem can be overcome with the help of research in the field of sentiment analysis, which focuses on the computational study of opinions, behaviors, and feelings about entities expressed in texts. Based on the results of sentiment analysis, this study aims to use the Naive Bayes method to divide public reactions to news about the 2024 presidential candidates into three categories: positive, negative, and neutral.

Keywords: *Presidential Candidates, Social Media, Sentiment Analysis, Naive Bayes*



I. PENDAHULUAN

Dalam hal ini, Indonesia adalah negara demokrasi yang menyelenggarakan pemilihan umum untuk memilih kepala negara. Setiap lima tahun, pemilu biasanya diadakan serentak di Indonesia. Pada tahun 2024 akan diadakan perlombaan politik resmi Indonesia. Sebagai seorang anggota parlemen yang ingin mencalonkan diri, tentunya akan mempertimbangkan tingkat ketidakpercayaan yang dilihat dari reaksi publik. Orang menggunakan platform media sosial salah satunya Facebook, untuk menyuarakan pendapat mereka.[1]

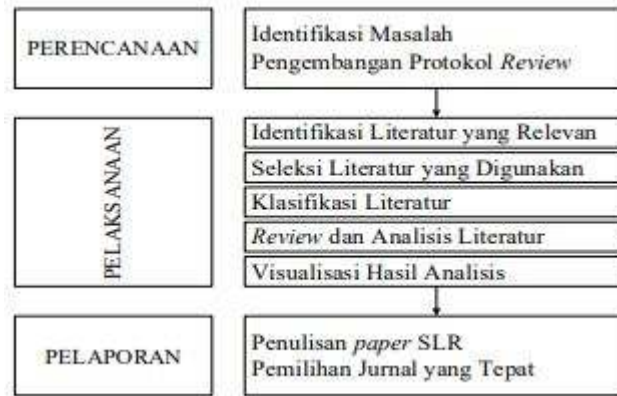
Sebagai aturan umum, postingan di Facebook digunakan untuk mentransfer data tentang klien dan menawarkan data, serta menyampaikan berita. Perasaan pengguna juga bisa diungkapkan dalam konten postingan. Untuk memproses data teks dan mengekstrak informasi yang berguna tentang pendapat seseorang tentang tokoh politik atau publik, diperlukan analisis yang tepat. Analisis sentimen yang mengkaji pendapat atau opini yang menyatakan atau mengungkapkan persepsi yang mengandung sentimen positif atau negatif, dapat digunakan untuk mengolah kumpulan komentar berupa teks. Sentimen yang diharapkan melengkapi pemeriksaan tersebut datang dari komentar di halaman fanpage calon presiden Indonesia 2024 di Facebook.(Santoso & Nugroho, 2019)

Beberapa pengujian sebelumnya telah dilakukan dengan memanfaatkan metode Naive Bayes pada data media social untuk menganalisa sentiment terhadap calon presiden Indonesia. Pada penelitian ini, data akan diklasifikasikan dengan menggunakan metode klasifikasi Naive Bayes, yang telah dikumpulkan. Teknik klasifikasi ini mengandalkan nilai probabilitas statistik langsung dan mengasumsikan tingkat independensi yang tinggi dari aturan Bayesian. Informasi awal dan data sampel digunakan dalam teknik analisis Bayesian ini. Metode Naive Bayes Classifier sendiri digunakan dengan menguji data testing dengan memanfaatkan data training. Metode Naive Bayes Classifier ini memberikan presisi tinggi saat diterapkan pada kumpulan data yang sangat besar dan berbagai informasi. Hal ini ditunjukkan oleh penelitian yang menunjukkan bahwa pendekatan ini memiliki sejumlah kemampuan tambahan, termasuk kecepatan, kesederhanaan, dan akurasi yang tinggi.

II. METODE PENELITIAN

A. *Systematic Literature Review*

Systematic Literature Review (SLR) adalah strategi eksplorasi yang diharapkan dapat mengenali dan menilai hasil penelitian dengan metode terbaik mengingat teknik eksplisit dari hasil pemeriksaan. Penulis telah menelaah analisis sentimen berbasis metode Naive Bayes data mining pemilihan calon presiden dalam literatur yang diulasnya. Gambar 1 dibawah ini menggambarkan alur penelitian ini.



Gambar 1. Alur penelitian

a. Tahap Perencanaan

Dengan menggunakan Algoritma Naive Bayes dan mencari artikel mengenai subjek tersebut, proses identifikasi isu dalam penelitian analisis sentimen akan dilakukan pada tahap pertama yang disebut dengan tahap perencanaan. Kata kunci yang digunakan dalam mencari artikel adalah:

1. Analisis Sentimen dengan Algoritma Naive Bayes
2. Optimalisasi Algoritma Naive Bayes
3. Implementasi Algoritma Naive Bayes pada Analisis Sentimen dari Media Sosial

b. Tahap Pelaksanaan

Pada tahap ini terdiri dari 5 tahapan, yaitu diantaranya :

1. Mengidentifikasi literatur analisis sentimen dengan menggunakan algoritma Naive Bayes yang relevan dengan pokok bahasan pada penelitian ini.
2. Berdasarkan relevansi topik dan tahun penerbitan makalah, pilih salah satu literatur yang tersedia.
3. Literatur yang terpilih kemudian akan dikategorikan sesuai dengan topik yang dipilih dan tahun publikasi penelitian.
4. Lakukan prosedur review artikel dan menganalisis literatur pengetahuan yang ada.
5. Buat representasi visual dari temuan analisis.

c. Tahap Pelaporan

Tahap terakhir adalah perincian dimana pada tahap ini paper SLR akan siap dari proses literature review yang telah dilakukan dan selanjutnya memilih jurnal yang tepat untuk men publikasikan paper yang telah dibuat.[7]

B. Preprocessing Data

1. Cleaning, yaitu untuk mengurangi kebisingan proses selama proses klasifikasi, pembersihan mengacu pada proses meninjau kata yang tidak diperlukan. Karakter adalah kata yang dihilangkan.
2. Case Folding, yaitu proses mengubah semua huruf dalam dokumen menjadi huruf kecil, yang diubah hanya huruf a sampai z.



3. Tokenizing, yaitu proses memotong kalimat menjadi kata-kata dengan memisahkan kata-kata dalam kumpulan data dan mencari tahu bagaimana strukturnya.
4. Filtering, yaitu tahap mengambil kata-kata penting dari konsekuensi tokenizing. Algoritme stoplist, yang menghapus kata-kata yang tidak penting, atau daftar kata, yang menyimpan kata-kata penting, dapat digunakan dalam proses pemfilteran.

C. Metode Naïve Bayes Classifier

Metode Naive Bayes adalah metode yang digunakan untuk menghitung probabilitas kemunculan kata-kata dalam dokumen dan menggunakan nilai tersebut untuk menentukan apakah dokumen tersebut termasuk pada kategori positif, negatif, atau netral. Rumus dasar metode Naive Bayes ini adalah sebagai berikut:

$$P(C|W) = P(W|C) * P(C) / P(W)$$

Dimana:

$P(C|W)$ adalah probabilitas kelas C (positive/negative/neutral) pada saat diketahui adanya fitur W dalam dokumen

$P(W|C)$ adalah probabilitas kemunculan fitur W pada data training set dengan label C

$P(C)$ adalah probabilitas prior dari kelas C yang dihitung dari jumlah data training set dengan label C dibagi total data training set

$P(W)$ adalah probabilitas kemunculan fitur W secara keseluruhan

Pada penelitian ini, metode yang akan digunakan metode klasifikasi Naive Bayes untuk melakukan analisis sentimen terhadap teks-teks yang berkaitan dengan pemilihan calon presiden 2024 di media sosial Facebook. (Ardiansyah, 2019)

III. HASIL DAN PEMBAHASAN

Berdasarkan hasil pencarian untuk artikel melalui platform Google Scholar sesuai dengan kata kunci yang digunakan, informasi diperoleh sebagai artikel ilmiah, prosiding, dan skripsi. Topik penelitian dari artikel ilmiah ini adalah analisis sentimen dengan menggunakan algoritma Naive Bayes. Setelah itu, dilakukan proses klasifikasi, dan tahun terbit artikel ditunjukkan pada Tabel 1 di bawah ini.



Tabel 1. *Literature Review*

No.	Referensi	Judul	Tools	Metode	Hasil	Media
1.	(Santoso & Nugroho, 2019)	Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook	R Program	Naïve Bayes	Telah disimpulkan bahwa informasi yang baru ditangani dapat digunakan untuk hapus kata-kata seperti huruf dan tanda baca berlebihan, dan hapus kata-kata sering digunakan sebagai komentar.	Facebook
2.	(Juanita, 2020)	Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes	WEKA	Naïve Bayes	Melakukan pelabelan untuk menentukan sentimen positif, negatif dan netral. Lalu, lakukan proses preprocessing dan perubahan teks selesai. Dengan metode klasifikasi Naive Bayes ini untuk analisis sentimen, prosedur ini menghasilkan kumpulan data pola persepsi.	Twitter
3.	(Mahawardana et al., 2022)	Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap Figure Pemimpin Menggunakan Python	Python	Naïve Bayes	Memperoleh data dari tweet berbahasa Indonesia. Lalu, gunakan fitur yang disediakan Python. Gunakan Library tweepy, untuk pengumpulan data, yang memiliki akses langsung ke API Twitter melalui console atau script.	Twitter

Pada penelitian ini, peneliti akan mengumpulkan data pada komentar media sosial Facebook untuk tujuan penerapan algoritme Naive Bayes. Dengan jumlah komentar acak yang diambil dari akun Anies Baswedan, Prabowo Subianto, dan Ganjar Pranowo. Selanjutnya, preprocessing data akan dilakukan dengan tahapan cleaning, case folding, tokenizing, dan filtering, dimana data yang sudah diolah dapat digunakan untuk menghilangkan kata-kata seperti tanda baca dan huruf yang tidak diperlukan serta kata-kata umum yang digunakan sebagai komentar.

Setelah melakukan penghapusan beberapa data akan ada data yang terpilih. Selanjutnya dilakukan proses pelabelan pada data dan proses klasifikasi dalam penelitian. Eksplorasi ini dilakukan dengan memanfaatkan RapidMiner dengan perbandingan data 80% untuk data penyiapan dan sebesar 20% untuk data pengujian dari data komentar tersebut.



IV. KESIMPULAN

Berdasarkan penemuan Systematic Literature Review (SLR) menunjukkan bahwa penelitian, metodologi, dan teori analisis sentiment masih kurang, sehingga memerlukan penyelidikan tambahan. Selain itu, terdapat kesenjangan dalam tools atau metode yang digunakan, sehingga diperlukan peningkatan jumlah penelitian analisis sentiment menggunakan metode Naive Bayes berdasarkan studi kasus dan sumber data.

REFERENCES

- [1] Alfiah Zulqornain, J., & Pandu Adikara, P. (2021). *Analisis Sentimen Tanggapan Masyarakat Aplikasi Tiktok Menggunakan Metode Naive Bayes dan Categorical Proportional Difference (CPD)*. 5(7), 2886–2890. <http://j-ptiik.ub.ac.id>
- [2] Ardiansyah, R. (2019). Analisis Sentimen Calon Presiden Dan Wakil Presiden Periode 2019-2024 Pasca Debat Pilpres Di Twitter. *ScientiCO : Computer Science and Informatics Journal*, 2(1), 21. <https://doi.org/10.22487/j26204118.2019.v2.i1.13068>
- [3] Kusumawati, R. (2022). *Analisis Sentimen Pengguna Twitter mengenai pemilihan Presiden (Pilpres) Tahun 2024 Dengan Metode Naive Bayes Classifier*.
- [4] Mahawardana, P. P. O., Imawati, I. A. P. F., & Dika, I. W. (2022). Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap “Figure Pemimpin” Menggunakan Python. *Jurnal Manajemen Dan Teknologi Informasi*, 12(2), 50–56.
- [5] Nardilasari, A. P., Hananto, A. L., Hilabi, S. S., & Priyatna, B. (2024). *Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM*. 7(1), 11–18.
- [6] Santoso, E. B., & Nugroho, A. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook. *Eksplora Informatika*, 9(1), 60–69. <https://doi.org/10.30864/eksplora.v9i1.254>
- [7] Septianingrum, F., & Irawan, A. S. Y. (2021). Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma Naive Bayes: Sebuah Literature Review. *Jurnal Media Informatika Budidarma*, 5(3), 799. <https://doi.org/10.30865/mib.v5i3.2983>
- [8] Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naive Bayes : Systematic Review. *Faktor Exacta*, 13(1), 35. <https://doi.org/10.30998/faktorexacta.v13i1.5548>
- [9] Nugroho, B. I., Ma’arif, Z., & Arif, Z. (2022). Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media. *Jurnal Sistem Informasi Dan ...*, 3(2), 46–51. <http://journal.peradaban.ac.id/index.php/jsitp/article/download/1265/860>



SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)
Universitas Nusa Putra, 12 Agustus 2023