

Analisis Sentimen Mahasiswa terhadap Fasilitas Kampus Berbasis *Lexicon*

Khanza Labita

Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

E-mail: [*khanzalabital@gmail.com](mailto:khanzalabital@gmail.com)

Abstrak – Fasilitas kampus memiliki peran penting dalam menunjang kenyamanan dan kualitas proses pembelajaran mahasiswa. Namun, pendapat mahasiswa terhadap fasilitas kampus sering kali belum dianalisis secara sistematis sehingga sulit digunakan sebagai bahan evaluasi. Penelitian ini bertujuan untuk menganalisis sentimen mahasiswa terhadap fasilitas Kampus 2 Universitas Nusantara PGRI Kediri berdasarkan ulasan teks yang diperoleh melalui survei Google Form. Data dikumpulkan dari mahasiswa aktif Angkatan 2022 dalam bentuk jawaban esai mengenai kondisi fasilitas kampus, meliputi parkir, kantin, toilet, musala, serta ruang kelas. Metode penelitian yang digunakan adalah analisis sentiment berbasis lexicon-based dengan memanfaatkan INSET Lexicon. Tahapan penelitian meliputi pengumpulan data, preprocessing teks, pelabelan sentimen, dan visualisasi hasil. Hasil penelitian menunjukkan bahwa sentimen mahasiswa terbagi ke dalam kategori positif, netral, dan negatif dengan dominasi sentiment positif dan netral terhadap fasilitas kampus. Temuan ini menunjukkan bahwa analisis sentiment berbasis teks dapat memberikan gambaran objektif mengenai persepsi mahasiswa dan berpotensi menjadi dasar evaluasi fasilitas kampus secara berkelanjutan.

Kata Kunci — Analisis Sentimen, Lexicon-Based, Survei Daring

1. PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat mendorong munculnya berbagai pendekatan baru dalam memahami dan menjelaskan opini publik, salah satunya melalui analisis sentimen. Analisis sentimen merupakan bagian dari bidang *Natural Language Processing* (NLP) yang berfungsi untuk mengidentifikasi dan mengklasifikasikan opini atau pandangan seseorang terhadap suatu objek ke dalam kategori tertentu, seperti positif, negatif, atau netral. Metode ini banyak dimanfaatkan untuk mengevaluasi persepsi Masyarakat terhadap produk, layanan, maupun kebijakan. Dalam konteks Pendidikan tinggi, pandangan mahasiswa terhadap fasilitas kampus menjadi aspek penting karena fasilitas yang memadai berperan langsung dalam menunjang proses pembelajaran dan menciptakan lingkungan akademik yang kondusif [1].

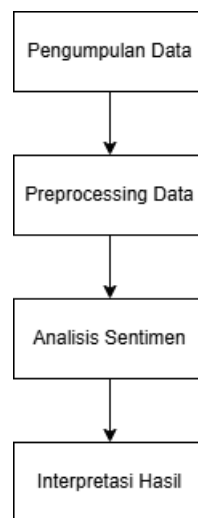
Meskipun fasilitas kampus memiliki peranan penting dalam mendukung kegiatan akademik maupun non-akademik, berbagai permasalahan terkait ketersediaan dan kualitas fasilitas masih sering ditemui. Mahasiswa kerap menyampaikan keluhan mengenai ketidaknyamanan fasilitas penunjang seperti kipas angin, kebersihan toilet, serta keterbatasan sarana pendukung lainnya. Namun demikian, diperlukan suatu pendekatan yang mampu mengakomodasi berbagai pandangan mahasiswa secara terstruktur sehingga dapat diperoleh gambaran yang lebih utuh mengenai kondisi fasilitas kampus [2].

Beberapa penelitian sebelumnya telah membahas pemanfaatan analisis sentiment dalam bidang Pendidikan untuk mengevaluasi kepuasan mahasiswa terhadap layanan akademik dan non-akademik [3]. Penelitian terbaru menunjukkan bahwa analisis sentimen berbasis teks efektif dalam memahami opini mahasiswa atau peserta pendidikan lebih mendalam daripada sekadar survei tertutup [4]. Pendekatan yang digunakan dalam penelitian-penelitian tersebut beragam, mulai dari penerapan algoritma klasifikasi sentimen hingga analisis statistik terhadap hasil survei. Meskipun demikian, sebagian penelitian masih berfokus pada data dari media sosial atau ulasan daring, sehingga kajian yang secara khusus memanfaatkan survei berbasis esai terstruktur untuk memetakan sentimen mahasiswa terhadap fasilitas kampus masih relatif terbatas [5].

Berdasarkan permasalahan dan penelitian sebelumnya, penelitian ini mengusulkan analisis sentimen mahasiswa terhadap fasilitas kampus dengan memanfaatkan survei berbasis *Google Form*. *Google Form* sendiri dipilih karena bersifat mudah diakses, fleksibel, dan memungkinkan pengumpulan data secara cepat serta terorganisir langsung dari responden. Penelitian ini diharapkan mampu memberikan pemetaan sentimen mahasiswa yang lebih sistematis dan berbasis data terhadap fasilitas kampus. Dengan demikian, hasil penelitian ini diharapkan dapat menjadi dasar pertimbangan dalam pengambilan keputusan serta perumusan strategi peningkatan kualitas kampus secara lebih tepat dan berkelanjutan.

2. METODE PENELITIAN

Terdapat beberapa tahapan dalam penelitian ini, diantaranya tahap pengumpulan data, preprocessing, analisis sentimen, dan terakhir interpretasi hasil.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan melalui survei daring menggunakan *Google Form*. Survei disebarkan kepada mahasiswa aktif fakultas Teknik Informatika Angkatan 2022 di kampus 2 Universitas Nusantara PGRI Kediri. Instrumen survei berupa pertanyaan esai yang membahas kondisi fasilitas kampus, meliputi area parkir, ruang kelas, toilet, musala, dan kantin. Jumlah responden yang berpartisipasi dalam penelitian ini kurang lebih 120 mahasiswa, dan data yang terkumpul digunakan sebagai bahan analisis untuk mengetahui sentimen mahasiswa terhadap fasilitas yang tersedia di Kampus 2 UNP Kediri.

2.2 Preprocessing Data

Data yang diperoleh dari hasil survei *Google Form* berupa teks bebas hasil jawaban esai responden. Oleh karena itu, sebelum dilakukan analisis sentimen, data perlu melalui tahap *preprocessing* untuk meningkatkan kualitas dan konsistensi data [6]. Tahap *preprocessing* bertujuan untuk mengurangi *noise* pada data teks serta memastikan bahwa informasi yang dianalisis benar-benar merepresentasikan opini responden [7]. Tahapan *preprocessing* yang dilakukan dalam penelitian ini meliputi beberapa proses sebagai berikut:

a. *Cleaning*

Proses pembersihan dilakukan dengan menghapus karakter khusus, tanda baca berlebih, serta symbol yang tidak memiliki makna sentimen [8]. Tujuan dari tahap ini adalah untuk mengurangi gangguan pada data teks yang dapat mempengaruhi hasil analisis.

b. *Case Folding*

Proses *case folding* dilakukan dengan mengubah seluruh teks dalam dataset menjadi huruf kecil guna menjaga konsistensi data. Tahapan ini bertujuan untuk menyeragamkan penulisan kata dengan menerapkan fungsi *lowercase* pada setiap teks [9].

c. *Stopword Removal*

Stopword adalah kata-kata umum dalam suatu bahasa yang sering muncul namun tidak memiliki pengaruh signifikan terhadap makna sentimen suatu teks, seperti kata penghubung atau kata keterangan tertentu. Proses ini dilakukan dengan menghapus kata-kata tersebut agar analisis sentimen dapat lebih fokus pada kata yang mengandung opini atau makna penting [10].

d. *Normalisasi*

Normalisasi merupakan proses penyesuaian bentuk kata dalam teks agar memiliki bentuk yang seragam dan baku [11].

2.3 Analisis Sentimen Berbasis *Lexicon*

Analisis sentimen pada penelitian ini menggunakan pendekatan *lexicon-based*, yaitu metode yang menentukan kecenderungan sentimen suatu teks berdasarkan kata-kata yang terkandung di dalamnya [12]. Setiap

kata pada jawaban esai responden dicocokkan dengan daftar kata (*lexicon*) yang telah memiliki polaritas sentimen, baik positif, negatif, maupun netral [13]. Pendekatan ini dipilih karena sesuai dengan karakteristik data berupa teks esai hasil survei dengan jumlah responden terbatas serta tidak memerlukan proses pelatihan data (*training data*).

Proses analisis sentimen diawali dengan menghitung skor sentimen dari setiap jawaban responden. Skor sentimen diperoleh dengan menjumlahkan nilai kata positif dan mengurangi nilai kata negatif yang muncul dalam satu teks [14]. Perhitungan skor sentimen dirumuskan sebagai berikut:

$$\text{Skor Sentimen} = \sum_{i=1}^n P_i - \sum_{j=1}^m N_j \dots\dots\dots (1)$$

Keterangan:

P_i = jumlah kata positif ke- i

N_j = jumlah kata negative ke- j

n = jumlah kata positif yang terdapat dalam teks

m = jumlah kata negatif yang terdapat dalam teks

Berdasarkan nilai skor sentimen yang diperoleh, setiap jawaban responden kemudian diklasifikasikan ke dalam tiga kategori sentimen, yaitu sentimen positif, negatif, dan netral. Klasifikasi sentimen dilakukan dengan ketentuan sebagai berikut: skor sentimen bernilai lebih dari nol dikategorikan sebagai sentimen positif, skor sentimen sama dengan nol dikategorikan sebagai sentimen netral, dan skor sentimen kurang dari nol dikategorikan sebagai sentimen negatif [15].

Setelah seluruh jawaban responden diklasifikasikan, dilakukan perhitungan distribusi sentimen untuk mengetahui kecenderungan opini mahasiswa secara keseluruhan terhadap fasilitas kampus. Distribusi sentimen dihitung dalam bentuk jumlah dan persentase masing-masing kategori sentimen. Persentase sentimen dihitung menggunakan rumus sebagai berikut:

$$\text{Presentase Sentimen} = \frac{f_s}{N} \times 100\% \dots\dots\dots (2)$$

F_s merupakan frekuensi sentimen pada kategori tertentu dan N merupakan jumlah keseluruhan responden. Hasil perhitungan persentase tersebut digunakan sebagai dasar dalam menganalisis kecenderungan sentimen mahasiswa terhadap fasilitas kampus.

3. HASIL DAN PEMBAHASAN

a. Dataset Ulasan

Dataset yang diperoleh berupa data teks tidak terstruktur yang berisi ulasan mahasiswa mengenai beberapa aspek fasilitas kampus, antara lain parkir, kantin, toilet, musala, serta ruang kelas. Seluruh data yang terkumpul kemudian diekspor ke dalam format file CSV untuk memudahkan proses pengolahan dan analisis lebih lanjut menggunakan bahasa pemrograman Python.

b. Preprocessing

Berdasarkan dataset ulasan yang telah dikumpulkan, selanjutnya dilakukan tahapan *preprocessing* data untuk menyiapkan teks sebelum dianalisis lebih lanjut. Tahap *preprocessing* ini bertujuan untuk membersihkan dan menormalkan data teks agar memiliki format yang konsisten serta mengurangi *noise* yang dapat memengaruhi hasil analisis sentimen. Proses *preprocessing* yang dilakukan meliputi *cleaning*, *case folding*, normalisasi, dan *stopword removal*.

Pada penelitian ini, contoh hasil *preprocessing* ditampilkan secara spesifik pada ulasan mahasiswa terkait ruang kelas. Pemilihan opini ruang kelas sebagai contoh bertujuan untuk memudahkan dalam memahami perubahan teks pada setiap tahapan *preprocessing*. Melalui tahapan tersebut, diperoleh dataset teks yang lebih bersih. Contoh hasil *preprocessing* dataset ditampilkan pada Tabel 1.

Tabel 1. *Preprocessing*

	Review Text	Cleaning	Case Folding	Normalisasi	Stopword Removal
0	Kurang dari segi manapun	Kurang dari segi manapun	kurang dari segi manapun	kurang dari segi manapun	['kurang', 'segi', 'manapun']
1	Nyaman apalagi kelas yg dibawah	Nyaman apalagi kelas yg dibawah	nyaman apalagi kelas yg dibawah	nyaman apalagi kelas yang di bawah	['nyaman', 'kelas', 'bawah']

2	Kipas di lantai atas tidak berfungsi dengan baik dan suasananya cukup bising, sehingga mengganggu kenyamanan belajar.	Kipas di lantai atas tidak berfungsi dengan baik dan suasananya cukup bising sehingga mengganggu kenyamanan belajar	kipas di lantai atas tidak berfungsi dengan baik dan suasananya cukup bising sehingga mengganggu kenyamanan belajar	kipas di lantai atas tidak berfungsi dengan baik dan suasananya cukup bising sehingga mengganggu kenyamanan belajar	['kipas', 'lantai', 'atas', 'berfungsi', 'baik', 'suasananya', 'cukup', 'bising', 'mengganggu', 'kenyamanan', 'belajar']
---	---	---	---	---	--

c. Pelabelan Data

Tahap selanjutnya adalah pelabelan data menggunakan metode *lexicon-based* dengan memanfaatkan *INSET Lexicon* untuk menghitung nilai polaritas dari setiap opini mahasiswa. Proses pelabelan dilakukan dengan menghitung nilai *sentiment score* berdasarkan jumlah kata positif dan negatif yang muncul dalam setiap ulasan. Apabila nilai skor sentimen bernilai 0, yang menunjukkan sentimen netral, maka nilai tersebut diubah secara acak menjadi 1 (positif) atau -1 (negatif) untuk menghindari dominasi kelas netral pada hasil analisis.

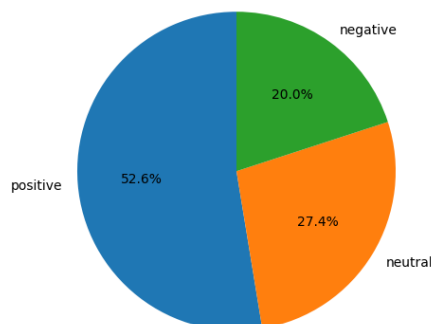
Selanjutnya, apabila nilai *sentiment score* lebih besar dari 0, maka ulasan diklasifikasikan ke dalam sentimen positif, sedangkan apabila nilai *sentiment score* kurang dari 0, maka ulasan diklasifikasikan sebagai sentimen negatif. Namun, apabila kondisi awal tidak terpenuhi atau skor polaritas tidak dapat ditentukan, maka sistem akan memberikan nilai *default* berupa skor 0 dengan label sentimen netral. Proses pelabelan ini diterapkan pada seluruh dataset yang berjumlah 46 data ulasan. Hasil pelabelan sentimen dari dataset tersebut selanjutnya digunakan sebagai dasar dalam proses analisis dan visualisasi sentimen. Contoh hasil pelabelan data ditampilkan pada Tabel 2.

Tabel 2. Hasil Pelabelan

Index	Preprocessing	Score	Sentimen
0	['kurang', 'segi', 'manapun']	-1	Negatif
1	['nyaman', 'kelas', 'bawah']	1	Positif
2	['kipas', 'lantai', 'atas', 'berfungsi', 'baik', 'suasananya', 'cukup', 'bising', 'mengganggu', 'kenyamanan', 'belajar']	1	Positif

d. Visualisasi

Distribusi Sentimen Mahasiswa terhadap Fasilitas Kampus



Gambar 2. Visualisasi Sentimen

Visualisasi sentimen menggunakan *diagram pie* dibuat berdasarkan keseluruhan data ulasan yang telah melalui proses pelabelan sentimen. Diagram ini digunakan untuk menunjukkan distribusi sentimen mahasiswa secara umum terhadap fasilitas kampus.

Berdasarkan Gambar 1, distribusi sentimen terdiri dari tiga kategori, yaitu sentimen positif, netral, dan negatif. Hasil visualisasi menunjukkan bahwa sentimen positif memiliki proporsi 52.6% dan sentimen netral memiliki proporsi sebesar 27.4% sedangkan sentimen negatif memiliki proporsi sebesar 20%. Temuan ini mengindikasikan bahwa sebagian besar mahasiswa memiliki persepsi yang cenderung netral terhadap fasilitas kampus, meskipun masih terdapat sejumlah ulasan yang menunjukkan sentimen positif dan negatif.

4. SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik beberapa simpulan sebagai berikut:

- Penelitian ini berhasil menerapkan analisis sentimen berbasis *lexicon-based* untuk mengidentifikasi kecenderungan opini mahasiswa terhadap fasilitas kampus berdasarkan data survei Google Form dalam bentuk teks esai.
- Tahapan preprocessing yang meliputi *cleaning*, *case folding*, *normalisasi*, dan *stopword removal* terbukti mampu menghasilkan data teks yang lebih bersih dan konsisten, sehingga meningkatkan kualitas proses pelabelan sentimen.
- Hasil visualisasi menunjukkan bahwa sentimen netral memiliki proporsi yang lebih dominan dibandingkan sentimen positif dan negatif, yang mengindikasikan bahwa persepsi mahasiswa terhadap fasilitas kampus cenderung cukup baik, meskipun masih terdapat aspek yang perlu mendapat perhatian.
- Kelebihan dari penelitian ini terletak pada penggunaan metode *lexicon-based* yang sederhana, tidak memerlukan data latih, serta mudah diimplementasikan pada dataset berukuran kecil hingga menengah. Namun, keterbatasan penelitian ini terdapat pada ketergantungan terhadap kamus *lexicon* dan penanganan sentimen netral yang masih bersifat acak.

5. SARAN

Berdasarkan keterbatasan penelitian yang telah dilakukan, beberapa saran untuk penelitian selanjutnya adalah sebagai berikut:

- Penelitian selanjutnya disarankan untuk menggunakan dataset dengan jumlah responden yang lebih besar dan cakupan fakultas yang lebih luas agar hasil analisis sentimen lebih representatif.
- Diperlukan pengembangan metode analisis sentimen dengan mengombinasikan pendekatan *lexicon-based* dan *machine learning* untuk meningkatkan akurasi serta mengurangi subjektivitas dalam penentuan sentimen netral.
- Penelitian lanjutan dapat menambahkan analisis sentimen berbasis aspek (*aspect-based sentiment analysis*) agar dapat mengidentifikasi secara spesifik kelebihan dan kekurangan pada setiap fasilitas kampus.

DAFTAR PUSTAKA

- [1] B. H. Hayadi and I. Maulita, "Sentiment Analysis of Public Discourse on Education in Indonesia Using Support Vector Machine (SVM) and Natural Language Processing," *Journal of Digital Society*, vol. 1, no. 1, pp. 68–90, 2025.
- [2] G. Priyadarshini, D. Karthika, M. Phil α , and D. Karthika σ , "Sentiment Analysis in Higher Education: A Systematic Mapping Review based Deep Neural Network," 2022.
- [3] A. R. Abidin and I. K. D. Nuryana, "Perbandingan metode klasifikasi data mining untuk mengukur tingkat kepuasan mahasiswa terhadap sistem informasi penilaian nonakademik unesa (sipena)," *Journal of Emerging Information System and Business Intelligence (JEISBI)*, vol. 4, no. 4, pp. 129–138, 2023.
- [4] G. hirzi Farizi, A. Voutama, and A. D. Hakim, "ANALISIS SENTIMEN MAHASISWA TERHADAP KETERSEDIAAN FASILITAS DI FAKULTAS ILMU KOMPUTER UNIVERSITAS SINGAPERBANGSA KARAWANG," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6544.
- [5] C. Grimalt-Álvaro and M. Usart, "Sentiment analysis for formative assessment in higher education: a systematic literature review," *J Comput High Educ*, vol. 36, no. 3, pp. 647–682, Dec. 2024, doi: 10.1007/s12528-023-09370-5.
- [6] W. Alfiyani, D. A. Fatah, and F. Irhamni, "Penerapan algoritma Naïve Bayes untuk analisis sentimen pada media sosial X terhadap performa tim nasional sepak bola Indonesia di era kepemimpinan Shin Tae-yong," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 3969–3977, 2025.

- [7] P. Punusingon, A. Yusupa, and V. Tarigan, “Analisis Sentimen Pemilihan Remaja Teladan Wilayah Ratahan Menggunakan Algoritma Naive Bayes,” *Journal Computer and Technology*, vol. 3, no. 1, pp. 65–73, 2025.
- [8] B. Zaman and E. Purwanti, *Mengurai Opini Digital: Panduan Praktis Analisis Sentimen berbasis Aspek*. PT. Pustaka Saga Jawadwipa, 2025.
- [9] A. Okta, K. Adi, F. Prayoganing Gusti, and F. Wijaya, “Analisis Sentimen Ulasan Pengguna Aplikasi Mobile Legends Pada Google Playstore Menggunakan Naïve Bayes”, *Stains*, vol. 4, no. 1, pp. 641–646, Jan. 2025, doi: 10.29407/avw8b190.
- [10] M. I. Raif, N. N. Hidayati, and T. Matulatan, “Otomatisasi Pendeteksi Kata Baku Dan Tidak Baku Pada Data Twitter Berbasis KBBI,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 337–348, 2024.
- [11] M. A. Nur and N. Wardhani, “Optimasi Normalisasi Kata Pada Data Twitter Untuk Meningkatkan Akurasi Analisis Sentimen (Studi Kasus Respon Masyarakat Terhadap Layanan Teman Bus),” *Jurnal Fokus Elektroda: Energi Listrik, Telekomunikasi, Komputer, Elektronika dan Kendali*, vol. 7, no. 4, pp. 237–243, 2022.
- [12] M. R. A. Yudianto, A. Rahim, P. Sukmasetya, and R. A. Hasani, “Perbandingan Metode Support Vector Machine dengan Metode Lexicon dalam Analisis Sentimen Bahasa Indonesia,” *Jurnal Teknologi Informasi*, vol. 6, no. 1, pp. 7–13, 2022.
- [13] S. A. Nugraha, “Penerapan Lexicon Based Untuk Analisis Sentimen Masyarakat Indonesia Terhadap Danantara,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 4949–4957, 2025.
- [14] A. R. Ismail and R. B. F. Hakim, “Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Menentukan Rekomendasi Pantai Di DI Yogyakarta Berdasarkan Data Twitter: Implementasi Lexicon Based,” *Emerging Statistics and Data Science Journal*, vol. 1, no. 1, pp. 37–46, 2023.
- [15] S. Rahayu, Y. Mz, J. E. Bororing, and R. Hadiyat, “Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP,” *Edumatic J. Pendidik. Inform*, vol. 6, no. 1, pp. 98–106, 2022.