

Analisis Sentimen Mahasiswa terhadap Kualitas Jaringan Internet UINSU Tuntungan Menggunakan Algoritma SVM

Nabila Intan Zahrani¹, Sriani²

^{1,2} Ilmu Komputer, Universitas Islam Negeri Sumatera Utara
Jl. Lap. Golf No.120, Kp. Tengah, Kec. Pancur Batu,
Kota Medan, Sumatera Utara 20353, Indonesia

nabilazahrani38@gmail.com

Abstrak

Kualitas jaringan internet yang tidak stabil di UINSU Kampus IV Tuntungan menjadi salah satu kendala dalam menunjang aktivitas akademik mahasiswa, khususnya pada proses pembelajaran digital, akses referensi, serta pengunggahan tugas. Penelitian ini dilakukan untuk mengkaji persepsi mahasiswa terhadap kualitas jaringan internet di lingkungan kampus menggunakan algoritma *Support Vector Machine* (SVM). Pengumpulan data dilakukan melalui wawancara langsung dan penyebaran kuesioner berbasis *Google Form*. Data teks yang diperoleh selanjutnya melalui tahapan pra-proses, pelabelan sentimen menggunakan pendekatan *lexicon-based*, pembobotan kata menggunakan metode *TF-IDF*, serta proses klasifikasi menggunakan *SVM*. Hasil analisis menunjukkan bahwa opini mahasiswa didominasi oleh sentimen negatif, khususnya terkait kecepatan akses dan kestabilan jaringan yang dinilai belum optimal. Model klasifikasi yang dibangun mencapai akurasi sebesar 88,78% dengan nilai *precision*, *recall*, dan *F1-score* yang relatif seimbang. Temuan tersebut membuktikan bahwa algoritma SVM efektif dalam mengidentifikasi pola sentimen mahasiswa dan dapat dimanfaatkan sebagai bahan evaluasi bagi pihak kampus dalam upaya peningkatan kualitas layanan jaringan internet.

Kata kunci: Analisis Sentimen, Jaringan Internet, SVM, TF-IDF

Abstract

The unstable quality of the internet network at UINSU Campus IV Tuntungan has become one of the obstacles in supporting students' academic activities, particularly in digital learning processes, access to references, and the submission of assignments. This study was conducted to examine students' perceptions of the quality of the campus internet network using the Support Vector Machine (SVM) algorithm. Data were collected through direct interviews and the distribution of Google Form-based questionnaires. The textual data obtained then underwent preprocessing stages, sentiment labeling using a lexicon-based approach, term weighting using the TF-IDF method, and classification using SVM. The results of the analysis indicate that students' opinions are dominated by negative sentiment, particularly related to access speed and network stability, which are considered not yet optimal. The developed classification model achieved an accuracy of 88.78%, with relatively balanced precision, recall, and F1-score values. These findings demonstrate that the SVM algorithm is effective in identifying patterns in student sentiment and can be utilized as an evaluation tool for the university in efforts to improve the quality of internet network services.

Keywords: Sentiment Analysis, Internet Network, SVM, TF-IDF

I. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah menjadikan internet sebagai kebutuhan peran penting dalam berbagai bidang kehidupan, khususnya pada sektor pendidikan. Perguruan tinggi

sebagai lembaga pendidikan tinggi sangat bergantung pada kualitas jaringan internet untuk mendukung kegiatan akademik, seperti akses *e-learning*, riset, komunikasi, dan administrasi [1]. Sarana teknologi informasi di lingkungan perguruan tinggi, baik berupa perangkat keras maupun

infrastruktur jaringan, seharusnya mampu mendukung akses internet yang optimal, baik di dalam kelas saat perkuliahan berlangsung, maupun di luar. Ketersediaan koneksi yang merata dan stabil sangat penting untuk menunjang pembelajaran digital yang fleksibel [2]. Kualitas jaringan yang buruk dapat berdampak pada efektivitas pembelajaran serta produktivitas mahasiswa dan dosen [3].

Salah satu perguruan tinggi yang mengalami tantangan dalam kualitas jaringan internet adalah Kampus IV Tuntungan (UINSU) Universitas Islam Negeri Sumatera Utara, Medan. Beberapa mahasiswa mengeluhkan layanan jaringan internet yang dinilai kurang optimal, meskipun secara infrastruktur kampus telah cukup menunjang aktivitas akademik berbasis digital. Permasalahan ini diduga berkaitan dengan pengelolaan jaringan terpusat yang berada di bawah koordinasi Pusat Teknologi Informasi dan Pangkalan Data (Pustipada), sehingga dapat memengaruhi stabilitas dan kecepatan koneksi pada layanan daring kampus. Dengan demikian, tujuan penelitian ini adalah untuk menganalisis dan memahami sentimen mahasiswa terhadap kualitas jaringan internet di UINSU Tuntungan, Medan [4].

Karena kemampuan untuk mengolah data yang sangat besar, algoritma Support Vector Machine (SVM) adalah pilihan yang tepat serta mampu menghasilkan klasifikasi yang lebih akurat dibandingkan beberapa algoritma lain [5]. Dalam ranah analisis sentimen, Algoritma SVM digunakan untuk mengelompokkan opini mahasiswa ke dalam kategori sentimen positif, negatif, dan netral. [6]. Melalui metode ini, memberikan gambaran yang lebih komprehensif terkait pengalaman dan keluhan mahasiswa terhadap jaringan internet di UINSU Tuntungan.

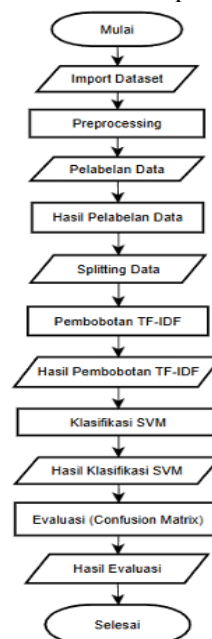
Sejumlah penelitian terdahulu mengindikasikan bahwa algoritma *Support Vector Machine* (SVM) menunjukkan kinerja yang baik dalam penerapan analisis sentimen. Fachrudin et al. (2024) melaporkan bahwa SVM mampu menghasilkan tingkat akurasi sebesar 84,33% dalam menganalisis sentimen masyarakat terhadap layanan internet Telkomsel di Twitter [7]. Hermanto et al. (2020) juga membuktikan keunggulan SVM dibandingkan *Naïve Bayes* dalam klasifikasi komplain mahasiswa dengan akurasi 84.45% [8]. Penelitian lain oleh Indrainsi dan Ernawati (2022) menerapkan SVM pada sentimen pembelajaran daring dan memperoleh akurasi 72% [9], sedangkan Sari dan Kurniawan (2024) sukses menerapkan SVM pada analisis opini publik terhadap program Walikota Medan dengan akurasi 81% [10]. Selain itu, Lubis dan Putri (2025) menggunakan data primer mahasiswa UINSU untuk

mengevaluasi sentimen terhadap *e-learning* dan memperoleh akurasi 82.5% [11]. Meskipun seluruh penelitian tersebut membuktikan efektivitas SVM, fokusnya masih terbatas pada layanan publik, pembelajaran daring, *e-learning*, atau opini masyarakat luas di media sosial. Hingga saat ini, belum ada penelitian khusus yang ditemukan mengkaji sentimen mahasiswa UINSU tentang kualitas jaringan internet kampus dengan menggunakan data primer hasil wawancara dan *Google Form*, sehingga menghadirkan celah penelitian yang menjadi dasar penelitian ini.

Adapun keunggulan penelitian ini terletak pada penggunaan data primer yang lebih valid dan kontekstual, berfokus langsung pada pengalaman mahasiswa di lingkungan kampus, serta memberikan penilaian spesifik terhadap kualitas jaringan internet kampus yang belum pernah diteliti sebelumnya. Selain itu, penerapan algoritma SVM pada data yang lebih terstruktur menjadikan hasil klasifikasi lebih stabil dan relevan, sehingga penelitian ini mampu memberikan rekomendasi nyata bagi pihak kampus dalam meningkatkan kualitas layanan jaringan internet yang mendukung aktivitas akademik mahasiswa [12].

II. METODE PENELITIAN

Penelitian ini dilaksanakan melalui serangkaian langkah yang sistematis, mengikuti tahapan umum dalam analisis sentimen, mulai dari pengumpulan data, proses pra-pengolahan, transformasi teks, hingga evaluasi dan penyajian hasil. Gambar 1 dibawah menjelaskan alur tahapannya.



Gambar 1. Tahapan Penelitian

A. Import Dataset

Dataset yang berisi komentar mahasiswa terkait jaringan internet dikumpulkan melalui kuesioner, media sosial, ataupun sistem *feedback* kampus dan diimpor ke lingkungan analisis menggunakan *library* seperti *pandas*. Pada tahap ini dilakukan pengecekan struktur data, jumlah record, nilai kosong, dan distribusi kelas sentimen jika tersedia.

B. Preprocessing

Tahap *preprocessing* bertujuan membersihkan data teks sehingga siap digunakan pada proses pembobotan dan klasifikasi. Beberapa proses utama meliputi *case folding*, *tokenization*, *stopword removal*, *normalization*, serta proses *stemming* untuk mengubah kata menjadi bentuk dasarnya. *Preprocessing* penting dilakukan karena kualitas input berpengaruh signifikan terhadap performa model klasifikasi [13].

C. Pelabelan Data

Untuk pelabelan data, pendekatan berbasis *lexicon* digunakan untuk membagi komentar siswa ke dalam kategori sentimen positif, negatif, dan netral. Metode ini memanfaatkan kamus kata yang telah diberi polaritas sentimen, sehingga proses penentuan sentimen dilakukan secara otomatis melalui proses pencocokan kata dalam teks terhadap daftar kosakata yang tersedia pada kamus. Teknik *Lexicon-Based* banyak digunakan dalam analisis sentimen karena mampu mengidentifikasi sikap atau emosi dalam teks dengan cepat dan konsisten. Hasil pelabelan kemudian dikodekan secara numerik agar dapat diproses pada tahap klasifikasi [14].

D. Splitting Data

Dataset dibagi menjadi 20% dan 80% masing-masing untuk pelatihan dan pelatihan model SVM. Data pengujian digunakan sebagai dasar penilaian kinerja model. Pembagian data dilakukan secara acak dengan menerapkan stratifikasi agar distribusi setiap kelas tetap seimbang [15].

E. Pembobotan TF-IDF

Pada tahap ini teks diubah menjadi fitur numerik dengan metode *TF-IDF* sehingga dapat diolah oleh algoritma *SVM*. Metode ini meningkatkan nilai kata-kata yang memiliki tingkat kepentingan tinggi dalam suatu dokumen namun frekuensinya rendah pada keseluruhan korpus. [16]. Adapun beberapa tahapan dan rumus proses pembobotan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) adalah:

Term Frequency (TF)

$$tf = \frac{\text{jumlah frekuensi kata terpilih}}{\text{jumlah kata}} \quad (1)$$

Invers Document Frequency (IDF)

$$idf = \log \frac{N}{df} \quad (2)$$

Term Frequency - Inverse Document Frequency (TF-IDF)

$$W_{t,d} = W_{tft,d} \times idf \quad (3)$$

F. Klasifikasi SVM

Support Vector Machine berfungsi untuk menggunakan *hyperplane* terbaik dengan margin maksimum untuk membedakan kelas sentimen. SVM efektif pada data berteks karena memiliki kemampuan guna menentukan batas pemisah terbaik pada ruang fitur berdimensi tinggi [17].

Rumus *Hyperplane*

$$f(x) = w^T x + b \quad (4)$$

Tujuan Optimasi/Margin

$$\text{maximize } \frac{2}{\|w\|} \quad (5)$$

Decision Function

$$y = \text{sign}(w^T x + b) \quad (6)$$

G. Evaluasi

Kinerja model dievaluasi menggunakan *Confusion Matrix* untuk menghitung akurasi, presisi, recall, dan nilai *F1*. *Matrix* ini menunjukkan performa klasifikasi dalam mengenali masing-masing kelas [18].

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Test Data}} \times 100\% \quad (7)$$

$$\text{Precision} = \frac{TP}{TP+FP}, \text{Recall} = \frac{TP}{TP+FN}, F1 - \text{score} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada pengumpulan data digunakan delapan sampel data kuesioner yang berisi tanggapan mahasiswa terkait kualitas jaringan internet kampus. Data tersebut menggambarkan berbagai persepsi pengguna, mulai dari yang menilai jaringan sudah cukup baik hingga yang merasa bahwa jaringan masih kurang mendukung kegiatan akademik. Sebagian besar respon menunjukkan bahwa jaringan perlu ditingkatkan agar lebih stabil dan mampu mendukung aktivitas perkuliahan, terutama dalam proses pengunggahan tugas serta akses internet harian. Sampel data ini kemudian dijadikan dasar untuk analisis lebih lanjut mengenai kondisi jaringan dan kebutuhan perbaikan. Sebagai contoh, sebagaimana disajikan pada Tabel 1.

Tabel 1. Sampel Kuesioner

No	Kuesioner
1	Masih kurang, karna jaringan tidak baik
2	Sudah cukup baik namun perlu ditingkatkan lagi
3	tidak terlalu
4	Masih kurang mendukung perlu di tingkatkan lagi jaringan di kampus agar lebih stabil
5	untuk sekedar mengupload tugas masih terbilang bagus
6	Menurut saya belum
7	Menurut saya jaringan internet kampus masih kurang mendukung aktivitas akademik mahasiswa sehari-hari
8	saya rasa kurang mendukung, harus memakai paket internet sendiri

A.1 Deskripsi Dataset

Data yang digunakan dalam penelitian ini berasal dari mahasiswa UINSU Kampus IV Tuntungan. melalui pengisian kuesioner berbasis Google Form serta wawancara singkat terkait kualitas jaringan internet kampus. Pengambilan data dilakukan pada periode Juni–Agustus 2025 dengan *Purposive sampling* adalah metode sampling *non-probability*, dengan pemilihan responden didasarkan pada kriteria tertentu sebagai mahasiswa aktif UINSU Kampus IV Tuntungan yang pernah menggunakan jaringan internet kampus dalam kegiatan perkuliahan, pengunggahan tugas, dan akses pembelajaran daring.

Secara keseluruhan, data yang berhasil dikumpulkan berjumlah 533 komentar mahasiswa, yang digunakan sebagai dataset utama dalam proses analisis sentimen dan merepresentasikan opini mahasiswa berdasarkan pengalaman akademik sehari-hari. Pada Tabel 1 ditampilkan 8 data kuesioner sebagai sampel, yang disajikan untuk memberikan gambaran bentuk dan karakteristik data teks yang digunakan dalam penelitian, bukan sebagai keseluruhan dataset, serta hanya berfungsi sebagai ilustrasi proses analisis. Setelah melalui proses preprocessing dan data dikategorikan ke dalam kategori sentimen positif dan negatif menggunakan pendekatan *lexicon-based* labeling, data dikategorikan ke dalam dua kategori ini.

Dari total 533 data, didapatkan distribusi sebanyak 287 data sentimen negatif dan 246 data sentimen positif, yang menunjukkan bahwa sentimen negatif lebih dominan. Dominasi sentimen negatif tersebut mencerminkan kecenderungan ketidakpuasan mahasiswa terhadap kualitas jaringan internet kampus, namun dataset ini tetap dapat digunakan secara efektif pada tahap klasifikasi

dengan menerapkan algoritma *Support Vector Machine* (SVM).

B. Preprocessing

Pendekatan *preprocessing* teks merupakan langkah besar menuju analisis sentimen untuk memvalidasi bahwa data yang diperoleh memiliki kualitas yang layak untuk di proses oleh algoritma klasifikasi. Pada tahap awal dilakukan proses pembersihan teks (*cleaning*) untuk menyingkirkan komponen yang tidak penting seperti tanda baca, angka, dan simbol yang tidak memberikan kontribusi terhadap analisis opini. Kemudian dilakukan *case folding* untuk menyeragamkan seluruh huruf menjadi huruf kecil agar sistem tidak membedakan variasi penulisan kata yang sama. Tahap berikutnya adalah normalisasi guna mengonversi kata tidak baku ke dalam bentuk yang sesuai dengan kaidah Bahasa Indonesia, diikuti dengan proses *tokenizing* yang memecah kalimat menjadi potongan kata. Kata-kata umum yang tidak berpengaruh terhadap makna sentimen kemudian dihapus melalui *stopword removal*. Tahap akhir yaitu *stemming*, merubah kata ke bentuk dasar sehingga makna inti tetap dipertahankan dan variasi kata dapat diminimalkan.

Hasil *preprocessing* tersebut menghasilkan bentuk teks jauh lebih bersih, sederhana, dan fokus terhadap kata-kata penting yang merepresentasikan inti dari setiap komentar kuesioner. Data yang telah melalui proses ini selanjutnya siap digunakan dalam tahap pelabelan sentimen dan pembobotan fitur menggunakan metode *TF-IDF* sebagai dasar sebelum dilakukan klasifikasi yang dilakukan menggunakan algoritma *Support Vector Machine* (SVM).

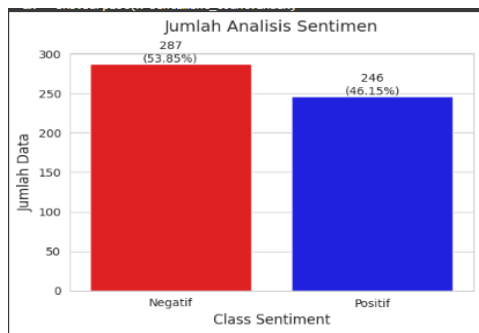
Tabel 2. Hasil Preprocessing Data

No	Komentar	Stemming
1	Masih kurang, karna jaringan tidak baik	[kurang, karena, jaringan, baik]
2	Sudah cukup baik namun perlu ditingkatkan lagi	[cukup, baik, tingkat]
3	tidak terlalu	[tidak, terlalu]
4	Masih kurang mendukung perlu di tingkatkan lagi jaringan di kampus agar lebih stabil	[kurang, dukung, tingkat, jaringan, kampus, stabil]

5	untuk sekadar mengupload tugas masih terbilang bagus	[sekadar, upload, tugas, bilang, bagus]
6	Menurut saya belum	[belum]
7	Menurut saya jaringan internet kampus masih kurang mendukung aktivitas akademik mahasiswa sehari-hari	[jaringan, internet, kampus, kurang, dukung, aktivitas, akademik, mahasiswa, hari, hari]
8	saya rasa kurang mendukung, harus memakai paket internet sendiri	[kurang, dukung, pakai, paket, internet, sendiri]

C. Pelabelan Data

Pada tahap ini, data hasil *preprocessing* diberi label sentimen menggunakan pendekatan *lexicon-based* dengan mencocokkan teks terhadap kamus sentimen positif dan negatif. Kategori sentimen ditentukan berdasarkan dominasi kata yang muncul, sehingga dihasilkan dataset berlabel yang siap digunakan pada proses pelatihan dan pengujian model SVM. Hasil pelabelan ditampilkan pada Gambar 2.



Gambar 2. Hasil *Labeling Lexicon Based*

Sumber dan Penyusunan Kamus Sentimen

Kamus sentimen yang digunakan dalam penelitian ini merupakan kamus buatan (*custom lexicon*) yang disusun secara manual oleh peneliti. Penyusunan kamus dilakukan karena konteks penelitian yang spesifik, yaitu opini mahasiswa terhadap kualitas jaringan internet kampus, sehingga diperlukan daftar kata yang relevan dengan domain akademik dan layanan jaringan.

Proses penyusunan kamus sentimen diawali dengan mengidentifikasi kosakata yang memiliki frekuensi kemunculan tinggi dalam data komentar mahasiswa setelah tahap *preprocessing*. Kosakata tersebut selanjutnya dibagi menjadi dua kategori, yaitu kata berpolaritas positif dan kata berpolaritas negatif, berdasarkan makna semantik dan konteks penggunaannya dalam kalimat. Contoh kata positif antara lain baik, cukup, stabil, dan bagus, sedangkan contoh kata negatif meliputi lambat, kurang, gangguan, dan tidak stabil.

Penggunaan kamus buatan dipilih karena kamus sentimen umum seperti *InSet* atau *SentiWordNet-ID* belum sepenuhnya mampu merepresentasikan istilah dan ekspresi yang sering digunakan mahasiswa dalam menilai kualitas jaringan internet kampus. Dengan demikian, kamus buatan ini diharapkan

dapat meningkatkan akurasi pelabelan sentimen sesuai dengan konteks penelitian.

D. Split Data

Dataset dibagi menjadi data pelatihan dan data pengujian pada tahap pembagian data untuk kebutuhan pengembangan dan evaluasi model. Pembagian dilakukan dengan rasio 20% sebagai data uji dan 80 persen sebagai data latih, sehingga dari 533 data diperoleh 426 data latih dan 107 data pengujian. Tahap ini bertujuan untuk memastikan bahwa model *SVM* dapat dievaluasi secara objektif menggunakan data yang belum pernah dipelajari sebelumnya, sehingga kinerja klasifikasi dapat diukur secara akurat.

Tabel 3 berikut merupakan hasil split data yang diambil dari 8 data sebagai representasi dari keseluruhan dataset. Contoh ini digunakan untuk menggambarkan proses pemisahan data dibagi menjadi dua kelompok: data latih dan data uji.

Tabel 3. Split Data

Data	Stemming	Sentimen
Data Latih	[kurang, karena, jaringan, baik]	Negatif
	[cukup, baik, tingkat]	Positif
	[tidak, terlalu]	Negatif
	[kurang, dukung, tingkat, jaringan, kampus, stabil]	Negatif
	[sekadar, upload, tugas, bilang, bagus]	Positif
	[belum]	Negatif
Data Uji	[jaringan, internet, kampus, kurang, dukung, aktivitas, akademik, mahasiswa, hari, hari]	?
	[kurang, dukung, pakai, paket, internet, sendiri]	?

E. Support Vector Machine (SVM)

Pada tahap berikutnya, algoritma *SVM* digunakan untuk melakukan klasifikasi sentimen. Dataset yang dihasilkan dari proses *preprocessing* dibagi menjadi data pelatihan dan data pengujian. Data pelatihan digunakan untuk membangun model. dalam mengenali pola sentimen, sedangkan data pengujian dimanfaatkan untuk menilai kinerja model pada data yang belum pernah dipelajari. Tujuan dari proses ini adalah untuk memastikan kemampuan generalisasi model terhadap data baru. Sebelum masuk ke proses klasifikasi *SVM*, nilai pembobotan kata menggunakan metode *TF-IDF* terlebih dahulu dinormalisasikan agar seluruh fitur berada pada

rentang nilai yang seimbang sehingga *SVM* dapat bekerja secara optimal dalam membentuk *hyperplane* pemisah antar kelas sentimen.

$$TF_{norm} = \frac{0.275}{\sqrt{0.275^2 + 0.448^2 + 0.275^2 + 0.275^2}} = \frac{0.275}{0.6539} = 0.421$$

Selanjutnya dilakukan perhitungan dengan menggunakan fungsi kernel linear sebagai fungsi pemisah antar kelas sentimen. Kernel linear digunakan karena sesuai untuk data teks dengan representasi numerik tinggi seperti *TF-IDF*, serta mampu menghasilkan *hyperplane* terbaik untuk memisahkan data berdasarkan margin maksimum antar kelas [19]. Pada tahap ini, nilai hasil perkalian vektor fitur antara data pelatihan dan data pengujian dimanfaatkan untuk menentukan kelas sentimen berdasarkan posisi relatifnya terhadap *hyperplane*.

$$K(x_i, x_j) = x_i \cdot x_j$$

$$K(x_1, x_1) = (t1d1 * t1d1) + (t1d2 * t1d2) + (t1d3 * t1d3) + (t1d4 * t1d4) + (t1d5 * t1d5) + (t1d6 * t1d6)$$

$$K(x_1, x_1) = (0,421 * 0,421) + (0 * 0) + (0 * 0) + (0,302 * 0,302) + (0 * 0) + (0 * 0)$$

$$K(x_1, x_1) = 0,1772 + 0 + 0 + 0,0912 + 0 + 0$$

$$D_{1,1} = 0,2684$$

$$K(x_1, x_2) = (t1d1 * t2d1) + (t1d2 * t2d2) + (t1d3 * t2d3) + (t1d4 * t2d4) + (t1d5 * t2d5) + (t1d6 * t2d6)$$

$$K(x_1, x_2) = (0,421 * 0,685) + (0 * 0) + (0 * 0) + (0,302 * 0) + (0 * 0) + (0 * 0)$$

$$K(x_1, x_2) = 0,2884 + 0 + 0 + 0 + 0 + 0$$

$$D_{1,2} = 0,2884$$

Proses perhitungan matriks Hessian diawali dengan penetapan nilai awal α sebesar 0. Proses ini dilakukan untuk memulai optimasi pada fungsi dual *SVM* yang bertujuan mencari nilai parameter α i optimum sebagai dasar pembentukan *hyperplane* pemisah antar kelas sentimen. *Matriks Hessian* berfungsi untuk memetakan hubungan antar data berdasarkan hasil perkalian label kelas dan nilai kernel linear, sehingga dapat menentukan struktur margin maksimum yang menjadi dasar pengambilan keputusan klasifikasi sentiment [20].

$$D_{1,1} = y_i \cdot y_j (K(x_i x_j) + \lambda^2)$$

$$D_{1,1} = -1 * -1 (0.2684 + 0.5^2)$$

$$D_{1,1} = 0.5184$$

Tabel 4. Nilai x^- dan x^+

x^+	x^-
1.2111	-1.2364

Data Uji 1:

[jaringan, internet, kampus, kurang, dukung, aktivitas, akademik, mahasiswa, hari, hari]

$$f(x) = w \cdot x + b$$

$$= \sum \alpha_i \cdot y_i \cdot K(x_i, x_j) + b$$

$$= ((-0.0137) + (-0.0446) + (-0.0223) + (-0.0273) + (-0.04458) + (-0.0223) + (-0.0223) + (-0.0223) + (-0.0446)) + (0.01265)$$

$$= -0.15246 + (-0.01265)$$

$$= -0.16511$$

$$\text{Fungsi Klasifikasi} = \text{sign}(-0.16511)$$

$$= -1$$

Data Uji 2

[kurang, dukung, pakai, paket, internet, sendiri]

$$f(x) = w \cdot x + b$$

$$= \sum \alpha_i \cdot y_i \cdot K(x_i, x_j) + b$$

$$= ((-0.0273) + (-0.0446) + (-0.0223) + (-0.0223) + (-0.04458) + (-0.0223) + (0.01265))$$

$$= -0.1834 + (-0.01265)$$

$$= -0.196025$$

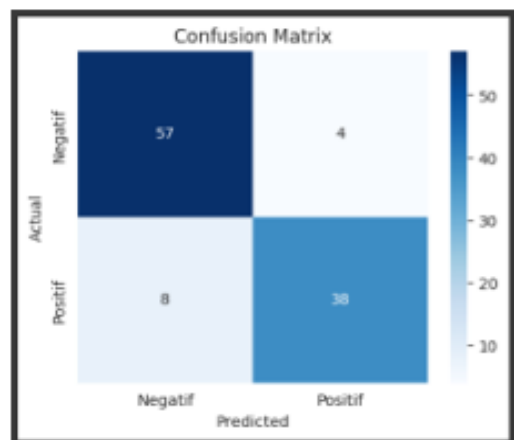
$$\text{Fungsi Klasifikasi} = \text{sign}(-0.196025)$$

$$= -1$$

Setelah uji data 1 dan 2 selesai, diperoleh hasil bahwa nilai dari fungsi klasifikasi pada kedua data tersebut adalah -1. Berdasarkan hasil tersebut, kedua data uji dikategorikan ke dalam kelas -1, yang merepresentasikan kelas sentimen negatif. Hal ini menunjukkan bahwa sistem mampu mengenali pola teks yang menunjukkan kecenderungan sentimen negatif sesuai dengan label yang telah ditentukan. Oleh karena itu, algoritma SVM untuk melakukan proses klasifikasi berhasil berjalan sesuai harapan dalam membedakan sentimen negatif pada data uji.

F. Evaluasi

Evaluasi model dilakukan dengan membandingkan nilai sebenarnya dibandingkan dengan hasil prediksi. Dalam penelitian ini, Confusion Matrix digunakan untuk mengukur kinerja model klasifikasi *Support Vector Machine* dalam memprediksi data. Gambar 3 dibawah merupakan hasil dari klasifikasi menggunakan algoritma *SVM*.



Gambar 3. Confusion Matrix

Hasil *confusion matrix* menunjukkan bahwa dari total 107 data yang diuji, terdapat 57 data sentimen

negatif berhasil diklasifikasikan dengan tepat sebagai negatif (True Negative), dan 38 data sentimen positif di prediksi dengan benar sebagai positif (True Positive). Namun, masih terdapat kesalahan prediksi, yaitu 4 data negatif diklasifikasikan sebagai positif (False Positive) serta 8 data positif salah dikategorikan sebagai negatif (False Negative). Jumlah kesalahan klasifikasi yang relatif rendah mengindikasikan bahwa model memiliki tingkat generalisasi cukup baik. Secara umum, kombinasi jumlah prediksi benar yang lebih tinggi dibanding prediksi salah menunjukkan bahwa algoritma SVM dengan TF-IDF mampu memberikan hasil klasifikasi sentimen yang efektif untuk menganalisis opini mahasiswa terkait kualitas jaringan internet kampus.

```

Akurasi: 0.8878504672897196
Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.88	0.93	0.90	61
Positif	0.90	0.83	0.86	46
accuracy			0.89	107
macro avg	0.89	0.88	0.88	107
weighted avg	0.89	0.89	0.89	107

Gambar 4. Akurasi SVM

Berdasarkan hasil evaluasi menggunakan *classification report*, diperoleh tingkat akurasi sebesar 0.8878 atau 88%, yang mengindikasikan bahwa model SVM dapat mengklasifikasikan komentar sentimen mahasiswa terhadap kualitas jaringan internet UINSU Tuntungan Medan dengan cukup baik. Untuk kelas Negatif, model menghasilkan nilai tingkat ketepatan 0.88, tingkat pengenalan 0.93, dan skor F1 0.90, yang berarti sebagian besar komentar negatif berhasil diprediksi dengan benar oleh model. Sementara itu, pada kelas Positif, *precision* yang diperoleh 0.90, pengenalan 0.83, dan nilai F1 0.86, yang mengindikasikan bahwa model juga memiliki performa bagus. Namun, masih ada beberapa data positif yang salah dianggap negatif. Nilai *macro avg* dan *weighted avg* yang relatif seimbang memperkuat bahwa model cukup stabil dan mampu menangani distribusi data yang tidak seimbang.

```

y_pred
array(['Positif', 'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Positif',
      'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Positif',
      'Positif', 'Positif', 'Negatif', 'Positif', 'Negatif', 'Positif',
      'Positif', 'Positif', 'Positif', 'Negatif', 'Positif', 'Positif',
      'Negatif', 'Positif', 'Positif', 'Positif', 'Positif', 'Positif',
      'Negatif', 'Negatif', 'Positif', 'Negatif', 'Positif', 'Negatif',
      'Negatif', 'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      'Negatif', 'Positif', 'Negatif', 'Positif', 'Positif', 'Positif',
      'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      'Negatif', 'Negatif', 'Positif', 'Negatif', 'Negatif', 'Negatif',
      'Positif', 'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      'Negatif', 'Positif', 'Negatif', 'Negatif', 'Positif', 'Positif',
      'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      'Positif', 'Positif', 'Negatif', 'Negatif', 'Positif', 'Negatif',
      'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif', 'Negatif',
      dtype=object)

```

Gambar 5. Hasil Prediksi SVM

Gambar 5 menampilkan hasil prediksi yang diperoleh dari penerapan algoritma (SVM).

Berdasarkan proses klasifikasi yang dilakukan, model SVM mampu memisahkan data ke dalam masing-masing kelas berdasarkan pola dan karakteristik fitur yang telah diekstraksi sebelumnya. Hasil ini menunjukkan bagaimana SVM bekerja dalam menentukan kategori suatu data uji sesuai dengan nilai fungsi keputusan yang diperoleh.

G. Visualisasi Wordcloud

Visualisasi wordcloud digunakan untuk menggambarkan distribusi kata pada teks dengan sentimen positif dan negatif. Ukuran kata pada wordcloud merepresentasikan tingkat frekuensi kemunculan, di mana kata-kata yang sering muncul ditampilkan dalam ukuran yang lebih besar. Gambar 6 dan 7 menunjukkan visualisasi tersebut.



Gambar 6. World Sentimen Positif



Gambar 7. Worldcloud Sentimen Negatif

Hasil visualisasi wordcloud menunjukkan perbedaan fokus utama antara sentimen positif dan negatif. Pada sentimen positif, kata-kata seperti “tugas”, “akses”, “kuliah”, dan “referensi” muncul sebagai kata yang paling dominan, mencerminkan bahwa mahasiswa menilai jaringan internet cukup membantu dalam menunjang aktivitas akademik seperti pengumpulan tugas dan akses materi

pembelajaran. Sementara itu, pada sentimen negatif terlihat dominasi kata seperti “tunggu”, “lambat”, “internet”, dan “kualitas jaringan”, yang menggambarkan keluhan mahasiswa terkait koneksi yang lambat, tidak stabil, serta kurang mendukung kelancaran kegiatan perkuliahan daring maupun pencarian referensi akademik.

H. Analisis Keterkaitan Hasil *Wordcloud* dan *Confusion Matrix* dengan Kondisi Riil Kampus

Hasil visualisasi *wordcloud* dan evaluasi dari model menggunakan *confusion matrix* menyajikan gambaran yang saling melengkapi tentang persepsi mahasiswa terhadap kualitas jaringan internet di UINSU Kampus IV Tuntungan. Melalui *wordcloud* sentimen negatif, kata-kata “lambat”, “tunggu”, “internet”, dan “kualitas jaringan” timbul dengan ukuran dominan. Dominasi kata-kata tersebut mencerminkan kondisi riil yang sering dialami mahasiswa, yaitu keterbatasan kecepatan dan stabilitas jaringan internet kampus, terutama saat digunakan secara bersamaan untuk kegiatan perkuliahan daring, pengunggahan tugas, dan akses referensi akademik.

Sebaliknya, pada *wordcloud* sentimen positif terlihat dominasi kata seperti “tugas”, “akses”, “kuliah”, dan “referensi”. Hal ini menunjukkan bahwa meskipun terdapat banyak keluhan, sebagian mahasiswa masih merasakan manfaat jaringan internet kampus dalam mendukung aktivitas akademik tertentu, terutama pada kondisi penggunaan yang tidak padat atau untuk kebutuhan akses dasar.

Keterkaitan antara hasil *wordcloud* dan *confusion matrix* dapat dilihat dari dominasi sentimen negatif yang berhasil diklasifikasikan dengan baik oleh model SVM. *Confusion matrix* menunjukkan bahwa sebagian besar data sentimen negatif dapat diprediksi secara tepat (True Negative yang tinggi), yang mengindikasikan bahwa pola kata-kata keluhan yang muncul pada kondisi nyata kampus berhasil dikenali secara konsisten oleh sistem. Hal ini sejalan dengan temuan lapangan bahwa masalah utama jaringan internet kampus bersifat berulang dan memiliki karakteristik teks yang serupa, seperti keluhan mengenai kecepatan dan kestabilan koneksi.

Sementara itu, masih ditemukannya kesalahan klasifikasi pada sebagian data sentimen positif dan negatif (False Positive dan False Negative) dapat dikaitkan dengan kondisi riil di mana pengalaman mahasiswa terhadap jaringan internet kampus bersifat situasional. Pada waktu dan lokasi tertentu jaringan dapat berfungsi dengan baik, namun pada kondisi lain mengalami gangguan. Variasi pengalaman ini menyebabkan munculnya komentar yang ambigu atau memiliki makna ganda, sehingga

menyulitkan model dalam menentukan kelas sentimen secara tepat.

Secara keseluruhan, keterkaitan antara hasil *wordcloud*, *confusion matrix*, dan kondisi nyata di lapangan menunjukkan bahwa metode analisis sentimen yang menggunakan algoritma SVM mampu merepresentasikan persepsi mahasiswa secara objektif. Temuan ini memperkuat bahwa hasil analisis tidak hanya valid secara teknis, tetapi juga relevan dengan kondisi riil kualitas jaringan internet di UINSU Kampus IV Tuntungan, sehingga dapat dijadikan dasar pertimbangan bagi pihak kampus dalam melakukan evaluasi dan peningkatan layanan jaringan internet.

I. Diskusi

Meskipun pendekatan *lexicon-based* yang dipakai dalam penelitian ini mampu memberikan pelabelan sentimen secara otomatis dan efisien, metode ini menunjukkan beberapa keterbatasan yang relevan untuk diperhatikan. Salah satu keterbatasan signifikan adalah ketergantungan pada daftar kata dalam kamus sentimen. Jika suatu kata atau ekspresi tidak tercantum dalam kamus, maka sistem tidak dapat menangkap makna sentimen yang terkandung dalam komentar tersebut secara optimal.

Selain itu, pendekatan *lexicon-based* belum sepenuhnya mampu memahami konteks kalimat secara menyeluruh, seperti penggunaan ironi, sarkasme, atau makna ganda dalam bahasa alami. Sebagai contoh, kata yang secara umum bersifat positif dapat memiliki makna negatif ketika digunakan dalam konteks tertentu, dan sebaliknya. Kondisi ini berpotensi menyebabkan bias pada proses pelabelan sentimen, khususnya pada data teks yang bersifat subjektif dan kontekstual.

Penggunaan kamus sentimen buatan juga memiliki potensi bias subjektivitas, karena proses penentuan polaritas kata dilakukan berdasarkan interpretasi peneliti. Meskipun kamus telah disusun secara sistematis dan disesuaikan dengan konteks penelitian, tetap terdapat kemungkinan bahwa beberapa kata memiliki makna sentimen yang berbeda bagi responden tertentu.

Namun demikian, keterbatasan tersebut dapat diminimalkan dengan penggunaan dataset yang relatif besar, yaitu 533 data komentar mahasiswa, serta penerapan algoritma SVM yang memiliki kinerja generalisasi yang handal dalam mempelajari pola data berlabel. Temuan dari evaluasi model yang menunjukkan akurasi, precision, recall, dan nilai F1 yang cukup tinggi mengindikasikan bahwa *bias* pelabelan yang mungkin terjadi tidak secara signifikan menurunkan kinerja model secara keseluruhan.

Sebagai pengembangan di masa mendatang, penelitian selanjutnya disarankan untuk mengombinasikan pendekatan *lexicon-based* dengan metode pelabelan berbasis pembelajaran mesin atau anotasi manual (human annotation), serta membandingkan hasilnya dengan pendekatan berbasis deep learning untuk mengurangi potensi bias label dan meningkatkan akurasi analisis sentimen.

IV. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen mahasiswa UINSU Kampus IV Tuntungan terhadap kualitas jaringan internet Dengan menggunakan SVM, dapat disimpulkan bahwa persepsi mahasiswa cenderung menunjukkan sentimen negatif, terutama terkait jaringan yang dianggap lambat dan kurang stabil dalam menunjang kegiatan akademik, meskipun terdapat sebagian tanggapan positif. Model klasifikasi sentimen berbasis SVM dengan pembobotan TF-IDF berhasil bekerja secara efektif dengan tingkat akurasi 88,78%, dan nilai presisi, *recall*, dan nilai F1 yang seimbang, sehingga dapat diandalkan untuk mengidentifikasi opini positif dan negatif. Untuk penelitian berikutnya, disarankan menambah jumlah dan variasi sumber data, serta membandingkan performa model dengan algoritma lain, seperti *Naïve Bayes*, *Logistic Regression*, *Random Forest*, atau *Deep Learning* guna memperoleh hasil lebih komprehensif dan akurat.

REFERENSI

- [1] N. Hidayati, "PEMANFAATAN WEBSITE SEKOLAH SEBAGAI STRATEGI DIGITAL MARKETING," 2020.
- [2] W. Agustiono, M. Cahyani Fajrin, F. Hastarita Rachman, J. P. Raya Telang BOX, and K. Bangkalan, "SISTEMASI: Jurnal Sistem Informasi Rencana Strategi Teknologi Informasi pada Perguruan Tinggi di Indonesia: Sebuah Tinjauan Pustaka," 2021. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [3] B. Harahap and M. Y. Saragih, "Urgensitas jurnalistik media sosial whatsapp pada mahasiswa fakultas ilmu sosial universitas islam negeri sumatera utara," JRTI (Jurnal Riset Tindakan Indonesia), vol. 10, no. 1, pp. 1–10, Feb. 2025, doi: 10.29210/30035459000.
- [4] A. P. Sinaga, I. Syahputra, Melati, and Nurbaiti, "Optimalisasi Jaringan Wifi (Wireless Fidelity) sebagai Fasilitas Pendukung Akademik Mahasiswa (Studi Kasus di UINSU)," *Cognoscere: Jurnal Komunikasi dan Media Pendidikan*, vol. 2, no. 4, Dec. 2024, doi: 10.61292/cognoscere.244.
- [5] I. S. Aisah, I. Bambang, and T. Suprapti, "ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK ANALISIS SENTIMEN ULASAN APLIKASI AL QUR'AN DIGITAL," *JATI (Jurnal Mahasiswa Teknik Informatika)*, 2024.
- [6] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 375–384, Feb. 2024, doi: 10.57152/malcom.v4i2.1206.
- [7] M. F. Fachrudin, C. V. Angkoso, and D. A. Fatah, "Analisis Sentimen Pada Sosial Media Twitter Terhadap Kualitas Jaringan Internet Telkomsel Menggunakan Ensemble K-Nearest Neighbour - Support Vector Machine," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 6, pp. 1253–1264, Dec. 2024, doi: 10.25126/jtiik.2024118713.
- [8] Hermanto, Mustopa, and Kuntoro, "Algoritma Klasifikasi Naive Bayes Dan Support Vector.," *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, vol. 5, no. 2, pp. 211–220, 2020.
- [9] A. N. Indraini and I. Ernawati, "Analisis Sentimen Terhadap Pembelajaran Daring Di Indonesia Menggunakan Support Vector Machine (SVM)," *Jurnal Ilmiah FIFO*, vol. 14, no. 1, p. 68, Jul. 2022, doi: 10.22441/fifo.2022.v14i1.007.
- [10] D. Sari and R. Kurniawan, "Analisis Sentimen Terhadap Kinerja Program Walikota Medan pada Media Sosial X Menggunakan Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1539–1548, Oct. 2024, doi: 10.57152/malcom.v4i4.1685.
- [11] A. S. Lubis and R. A. Putri, "Analisis Sentimen Mahasiswa Terhadap Penggunaan E-Learning dengan Algoritma Support Vector Machine (SVM)," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 5, no. 2, pp. 778–788, Jul. 2025, doi: 10.51454/decode.v5i2.1247.
- [12] S. Desmon, A. Nainggolan, and I. Yuadi, "Classification of Wrist Accessories: Advanced Watches with Logistic Regression, SVM, and Deep Features from Inception V3 and VGG-19," vol. 11, pp. 25–34, 2025, [Online]. Available: <http://http://jurnal.unmer.ac.id/index.php/jtmi>
- [13] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN," *Jurnal KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [14] S. Ratnaswari, N. C. Wibowo, and D. S. Y. Kartika, "ANALISIS SENTIMEN MENGGUNAKAN METODE LEXICON-BASED DAN SUPPORT VECTOR MACHINE PADA PRESIDEN DAN WAKIL PRESIDEN INDONESIA PERIODE 2024–2029," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5604.
- [15] Y. A. Prasetyo, E. Utami, and A. Yaqin, "Pengaruh Komposisi Split Data Terhadap Performa Akurasi Analisis Sentimen Algoritma Naive Bayes dan

- SVM,” Journal homepage: Journal of Electrical Engineering and Computer (JEECOM), vol. 6, no. 2, 2024, doi: 10.33650/jeeecom.v4i2.
- [16] D. Septiani and I. Isabela, “ANALISIS TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DALAM TEMU KEMBALI INFORMASI PADA DOKUMEN TEKS,” SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia, 2022.
- [17] Sriani, A. H. Lubis, and L. P. A. Lubis, “Sentiment analysis on twitter about the death penalty using the support vector machine method,” TEKNOSAINS : Jurnal Sains, Teknologi dan Informatika, vol. 11, no. 2, pp. 312–321, Jul. 2024, doi: 10.37373/tekno.v11i2.1096.
- [18] I. Widhi Saputro and B. Wulan Sari, “Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa Naïve Bayes Algorithm Performance Test for Student Study Prediction,” Citec Journal, vol. 6, no. 1, 2019.
- [19] N. Arifin, U. Enri, and N. Sulistiyowati, “STRING (Satuan Tulisan Riset dan Inovasi Teknologi) PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) DENGAN TF-IDF N-GRAM UNTUK TEXT CLASSIFICATION,” 2021.
- [20] E. R. Wulan, “Aplikasi Matriks Hessian Pada Model EPQ (Economic Production Quantity) dengan Kendala Rework,” vol. 1, no. 1, 2015.