

ANALISIS KINERJA SUPPORT VECTOR MACHINE DAN NAÏVE BAYES CLASSIFIER DALAM KLASIFIKASI SENTIMEN DAN EMOSI PADA UMPAN BALIK MAHASISWA TERHADAP KINERJA DOSEN

Teguh Syafrudin¹, Arief Hermawan², Donny Avianto³

^{1,2}Magister Teknologi Informasi, Universitas Teknologi Yogyakarta, Sleman, Indonesia

³Program Studi Informatika, Universitas Teknologi Yogyakarta, Sleman, Indonesia

e-mail: teguhsyafrudin21@gmail.com¹, ariefdb@uty.ac.id², donny@uty.ac.id³

Abstrak

Evaluasi kinerja dosen melalui umpan balik mahasiswa merupakan komponen penting dalam meningkatkan mutu pembelajaran di perguruan tinggi. Penelitian ini bertujuan untuk mengembangkan model klasifikasi sentimen dan emosi menggunakan pendekatan hybrid machine learning dengan mengombinasikan algoritma Naïve Bayes dan Support Vector Machine (SVM). Dataset berasal dari umpan balik mahasiswa yang telah dianotasi secara manual ke dalam tiga kategori sentimen positif, netral, negatif dan delapan kategori emosi. Proses preprocessing dilakukan melalui tokenisasi, stemming, dan transformasi data ke dalam bentuk TF-IDF. Hasil klasifikasi menunjukkan bahwa SVM memiliki performa terbaik untuk klasifikasi sentimen dengan akurasi mencapai 90%, mengungguli Naïve Bayes yang hanya memperoleh akurasi 80%. Sebaliknya, performa klasifikasi emosi jauh lebih rendah, dengan akurasi maksimum 35% pada model SVM dan 20% pada Naïve Bayes. Beberapa emosi seperti “marah”, “termotivasi”, dan “senang” tidak dapat dikenali oleh model karena ketidakseimbangan distribusi data dan konteks emosi yang sulit ditangkap dari teks pendek. Temuan ini menunjukkan bahwa pendekatan hybrid efektif untuk klasifikasi sentimen dalam kondisi data terbatas, namun klasifikasi emosi memerlukan pendekatan lanjutan seperti reduksi label atau penggunaan model berbasis konteks untuk mencapai hasil yang lebih baik.

Kata kunci: Kinerja Dosen, Naïve Bayes Classifier (NBC), Support Vector Machine (SVM), Umpan Balik Mahasiswa.

Abstract

Evaluating lecturer performance through student feedback is essential in improving the quality of learning in higher education. This study aims to develop a sentiment and emotion classification model using a hybrid machine learning approach by combining the Naïve Bayes and Support Vector Machine (SVM) algorithms. The dataset is derived from student feedback manually annotated into three sentiment categories (positive, neutral, negative) and eight emotion categories. Preprocessing involves tokenization, stemming, and data transformation into TF-IDF format. The classification results show that SVM has the best performance for sentiment classification with an accuracy of 90%, outperforming Naïve Bayes, which only achieved an accuracy of 80%. Conversely, emotion classification performance is much lower, with a maximum accuracy of 35% for the SVM model and 20% for Naïve Bayes. Some emotions such as ‘angry,’ ‘motivated,’ and ‘happy’ could not be recognized by the model due to the imbalance in data distribution and the difficulty of capturing emotional context from short texts. These findings indicate the hybrid approach is practical for sentiment classification under limited data conditions. Still, emotion classification requires further approaches, such as label reduction or context-based models to achieve better results.

Keywords: Lecturer Performance, Naïve Bayes Classifier (NBC), Support Vector Machine (SVM), Student Feedback.

1. PENDAHULUAN

Pendidikan tinggi memegang peran penting dalam membentuk sumber daya manusia yang unggul dan adaptif terhadap perkembangan zaman. Salah satu faktor penentu kualitas pendidikan di perguruan tinggi dapat ditentukan melalui kinerja dosen dalam proses pembelajaran. Evaluasi kinerja dosen menjadi bagian penting dalam proses belajar mengajar karena mencerminkan efektivitas penyampaian materi, interaksi dosen dengan mahasiswa, serta kualitas metodologi pengajaran yang digunakan. Evaluasi yang dilakukan secara berkala dapat memberikan umpan balik konstruktif yang digunakan untuk meningkatkan kualitas pembelajaran, baik dari sisi pedagogik maupun konten materi. Evaluasi kinerja dosen umumnya dilakukan melalui kuesioner atau umpan balik mahasiswa pada akhir semester sebagai bagian dari sistem penjaminan mutu internal institusi. Namun, pendekatan konvensional yang hanya mengandalkan skor numerik sering kali tidak cukup untuk menangkap kedalaman persepsi mahasiswa[1]. Skor numerik hanya memberikan representasi kuantitatif dari penilaian, namun tidak mencerminkan alasan di balik penilaian tersebut, seperti apakah mahasiswa merasa terbantu secara emosional, mengalami kebingungan terhadap materi, atau memiliki saran konstruktif terhadap gaya mengajar dosen. Persepsi yang mendalam sering kali tercermin dalam komentar terbuka, di mana mahasiswa menyampaikan pengalaman belajar mereka secara lebih naratif dan kontekstual. Memahami persepsi secara mendalam sangat penting karena dapat mengungkap aspek-aspek tersembunyi yang tidak terjangkau oleh angka, seperti kejelasan komunikasi, empati dosen, motivasi belajar, serta dinamika kelas yang tidak kondusif. Informasi ini sangat berharga bagi institusi dalam merancang intervensi yang lebih tepat sasaran, baik dalam bentuk pelatihan dosen, penyesuaian metode pembelajaran, maupun pengembangan kebijakan akademik yang lebih adaptif terhadap kebutuhan mahasiswa. Oleh karena itu, analisis terhadap komentar atau saran terbuka dalam kuesioner evaluasi menjadi penting untuk memahami kualitas pengajaran dari sudut pandang mahasiswa secara lebih komprehensif.

Seiring dengan meningkatnya volume umpan balik teks terbuka dari mahasiswa disetiap semesternya, analisis manual menjadi tidak efisien dan rentan terhadap bias subjektif. Untuk mengatasi tantangan tersebut, teknologi pemrosesan *Natural Language Processing (NLP)* hadir sebagai solusi yang efektif [2]. NLP memungkinkan komputer untuk memahami, menginterpretasikan, dan menganalisis bahasa manusia secara otomatis. Dalam konteks pendidikan, NLP dapat digunakan untuk mengekstrak informasi bernilai dari komentar mahasiswa, baik dalam bentuk sentimen maupun emosi yang tersirat dengan menggunakan beberapa algoritma [3]. Analisis sentimen secara khusus berperan penting dalam mengidentifikasi kecenderungan opini mahasiswa, apakah bersifat positif, netral, atau negatif, terhadap aspek-aspek pengajaran yang diberikan[4]. Informasi ini dapat menjadi dasar pengambilan keputusan strategis bagi institusi dalam memperbaiki metode pembelajaran, gaya komunikasi dosen, serta peningkatan kualitas layanan akademik secara menyeluruh[5]. Sementara klasifikasi emosi dalam konteks pendidikan menjadi semakin penting karena emosi memiliki peran signifikan dalam proses pembelajaran dan interaksi antara mahasiswa dan dosen. Emosi positif seperti senang, puas, dan termotivasi dapat meningkatkan motivasi belajar dan keterlibatan mahasiswa, sementara emosi negatif seperti kecewa, marah, atau frustrasi dapat menjadi indikator adanya hambatan dalam proses belajar mengajar. Oleh karena itu, memahami ekspresi emosional dalam umpan balik mahasiswa tidak hanya dapat membantu institusi dalam menilai kinerja dosen, tetapi juga memberikan wawasan yang lebih mendalam untuk merancang strategi peningkatan kualitas pembelajaran yang lebih adaptif dan responsif terhadap kebutuhan emosional mahasiswa. Penelitian terdahulu telah banyak menerapkan algoritma *Machine Learning* untuk penelitian ini, seperti *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* karena sifatnya yang ringan dan efektif untuk klasifikasi teks dalam jumlah besar [6][7]. Misalnya, penelitian oleh Shaika et al. (2023) mengkaji klasifikasi opini mahasiswa terhadap mata kuliah menggunakan pendekatan *supervised learning* dan menunjukkan bahwa SVM mampu mengungguli NBC dalam presisi klasifikasi [4]. Beberapa studi telah menggabungkan metode berbasis leksikon dan *supervised learning* untuk memperbaiki akurasi klasifikasi teks pendidikan [8][9][10], namun masih sedikit yang mengkaji perbandingan performa model pada data umpan balik kinerja dosen dalam bahasa Indonesia. Selain itu, klasifikasi emosi pada teks pendek, seperti komentar mahasiswa, memiliki tantangan tersendiri karena keterbatasan konteks dan ketidakseimbangan data antar label emosi [11][12].

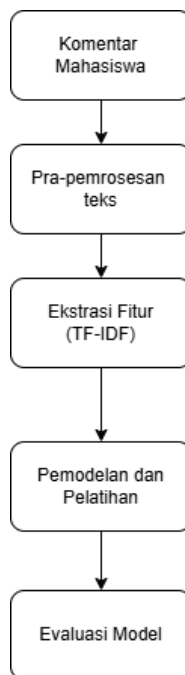
Penelitian ini memiliki keunikan dibandingkan studi sebelumnya karena menggabungkan pendekatan *hybrid machine learning* pada dataset komentar mahasiswa dalam bahasa Indonesia, yang telah dianotasi secara manual untuk dua aspek sentimen yaitu positif, negatif, netral dan emosi dengan 8 label emosi berupa senang, puas, kecewa, frustrasi, biasa saja, bingung, marah, dan termotivasi. Penelitian ini membandingkan dua algoritma klasik, SVM dan *Naïve Bayes*, secara kuantitatif dalam aspek klasifikasi sentimen dan emosi menggunakan pendekatan *lexicon-based* dan fitur *TF-IDF*. Selain itu, tidak hanya menggabungkan dua algoritma klasik, penelitian ini juga mengeksplorasi efektivitas pendekatan hybrid

dalam kondisi terbatas, seperti dataset kecil dan dominasi teks pendek, yang menjadi tantangan umum dalam pendidikan tinggi di Indonesia

Berdasarkan uraian tersebut, tujuan dari penelitian ini adalah untuk membandingkan dan mengukur performa dari algoritma *Support Vector Machine (SVM)* dan *Naïve Bayes Classifier (NBC)* dalam melakukan klasifikasi sentimen dan emosi terhadap umpan balik mahasiswa tentang kinerja dosen. Penelitian ini diharapkan dapat memberikan rekomendasi algoritma yang paling sesuai untuk analisis otomatis terhadap umpan balik berbasis teks pendek dalam konteks pendidikan tinggi di Indonesia.

2. METODE PENELITIAN

Penelitian ini dilakukan dalam beberapa tahapan dimulai dari pengumpulan data berupa komentar mahasiswa, pra-pemrosesan teks, ekstraksi fitur, pemodelan, pelatihan, dan evaluasi model. Pendekatan ini dirancang secara sistematis agar setiap tahapan berkontribusi terhadap pembentukan model klasifikasi yang optimal. Pengumpulan data dilakukan secara terstruktur dari sumber yang relevan, sementara tahap pra-pemrosesan bertujuan untuk membersihkan dan menyiapkan data mentah menjadi format yang dapat digunakan dalam proses pelatihan algoritma. Ekstraksi fitur dilakukan dengan memilih representasi teks yang sesuai untuk memaksimalkan akurasi klasifikasi. Model yang dibangun kemudian dievaluasi menggunakan metrik performa standar agar dapat dibandingkan secara objektif. Setiap tahapan saling terintegrasi dan berperan penting dalam menghasilkan model klasifikasi sentimen dan emosi yang dapat diandalkan



Gambar 1. Tahapan Penelitian

2.1. Sumber dan Jumlah Data

Data diperoleh dari kolom saran terbuka pada formulir evaluasi kinerja dosen berbasis SIAKAD dari dua semester ganjil tahun akademik 2023–2024. Total terdapat 500 data komentar teks pendek yang telah diberi anotasi secara manual oleh dua annotator ahli. Proses anotasi dilakukan untuk dua klasifikasi sentimen terdiri dari positif, netral, negatif. Sementara untuk klasifikasi **emosi** terdiri dari senang, puas, kecewa, frustrasi, biasa saja, bingung, marah, dan termotivasi.

Tabel 1. Dataset sebelum Pra-Pemrosesan

Feedback Mahasiswa	Sentimen	Emosi
Saya merasa semangat belajar karena metode...	Positif	Puas
Materi tidak disampaikan secara sistematis	Negatif	Kecewa
Saya merasa bingung dengan instruksi yang dibe...	Negatif	Frustrasi
....		
Dosen selalu memberi umpan balik yang membangun	Positif	Termotivasi
Tugas terlalu banyak dan tidak proporsional de...	Negatif	Marah

2.2. Pra-Pemrosesan Teks

Tahap berikutnya adalah pra-pemrosesan teks, proses ini digunakan untuk menyeleksi data *Feedback* Mahasiswa dan mengubahnya menjadi data yang lebih terstruktur. Tahapan pertama adalah *Cleaning*, yaitu proses pembersihan teks dari karakter-karakter yang tidak relevan seperti angka, tanda baca, simbol, dan tautan URL yang tidak memberikan kontribusi terhadap makna utama kalimat. Selanjutnya, dilakukan *Case Folding*, yaitu proses mengubah seluruh huruf kapital menjadi huruf kecil agar kata-kata dengan bentuk yang sama namun berbeda kapitalisasi seperti “Bagus” dan “bagus” dapat dikenali sebagai satu entitas yang sama. Setelah itu, teks diproses melalui *Tokenizing*, yaitu pemisahan teks menjadi unit-unit kata atau token. Tahap ini penting untuk memudahkan analisis kata secara individual. Langkah berikutnya adalah *Stemming*, yaitu proses mengembalikan kata ke bentuk dasarnya. Misalnya, kata “mengajar”, “pengajaran”, dan “diajarkan” akan dikembalikan menjadi “ajar”. *Stemming* penting untuk mengurangi kompleksitas bahasa dan menyatukan variasi kata yang merujuk pada konsep yang sama. Terakhir, dilakukan *Stopword Removal*, yaitu penghapusan kata-kata umum yang sering muncul tetapi tidak memiliki makna signifikan dalam klasifikasi, seperti “dan”, “yang”, “ini”, “itu”, dan sebagainya. *Stopword* yang tidak dibuang dapat menyebabkan model bias terhadap kata-kata yang sering muncul, padahal tidak bersifat informatif. Teknik ini merujuk pada praktik terbaik *preprocessing* dalam analisis teks bahasa Indonesia[6][13][14][15].

Tabel 2. Pra-Pemrosesan Teks *Feedback* Mahasiswa

Feedback Mahasiswa	Feedback Preprocessed
Saya merasa semangat belajar karena metode pen...	Semangat ajar metode ajar interaktif
Materi tidak disampaikan secara sistematis	Materi sistematis
Saya merasa bingung dengan instruksi yang dibe...	Bingung instruksi
....	
Dosen selalu memberi umpan balik yang membangun	Dosen umpan bangun
Tugas terlalu banyak dan tidak proporsional de...	Tugas proporsional

2.3 Ekstraksi Fitur dan Pelabelan

Ekstraksi fitur dilakukan dengan metode pembobotan menggunakan *Term Frequency–Inverse Document Frequency (TF-IDF)* untuk mewakili setiap teks sebagai vektor numerik. Untuk pelabelan awal dalam klasifikasi sentimen, digunakan pendekatan *lexicon-based* menggunakan *InSet Lexicon*, yaitu kamus sentimen berbahasa Indonesia yang dikembangkan oleh Koto dan Adriani (2020) [10]. Aturan pelabelan adalah skor < 0 = negatif, skor = 0 = netral, skor > 0 = positif. Untuk label emosi, data dikategorikan secara manual berdasarkan makna dominan dalam setiap komentar, dengan mengacu pada studi-studi klasifikasi emosi serupa[12]. *TF-IDF* menjadi metode representasi fitur yang umum digunakan dalam klasifikasi teks karena mampu menyeimbangkan antara frekuensi kemunculan kata (*TF*) dan seberapa penting kata tersebut secara global (*IDF*) dalam seluruh dokumen. Kata-kata yang terlalu sering muncul di seluruh dokumen,

seperti kata hubung atau kata bantu, akan mendapatkan bobot rendah, sementara kata-kata unik yang muncul secara spesifik dalam dokumen tertentu mendapatkan bobot tinggi [16]. Hasil dari proses ini adalah sebuah matriks dokumen-kata dengan nilai bobot numerik yang menggambarkan pentingnya setiap kata dalam dokumen. Representasi ini kemudian menjadi masukan bagi algoritma klasifikasi untuk belajar mengenali pola sentimen dan emosi berdasarkan distribusi kata [17].

Dalam proses pelabelan emosi, dua annotator independen dilibatkan untuk meningkatkan reliabilitas klasifikasi manual. Setiap komentar mahasiswa dianalisis dan dipetakan ke dalam salah satu dari delapan label emosi: senang, puas, kecewa, frustrasi, biasa saja, bingung, marah, dan termotivasi. Proses ini mempertimbangkan konteks kalimat, nada ekspresi, serta kata kunci yang menunjukkan muatan emosional tertentu. Ketika terdapat perbedaan hasil antara annotator, dilakukan diskusi untuk mencapai kesepakatan akhir. Proses ini bersifat krusial karena label emosi yang akurat akan sangat memengaruhi kinerja model pembelajaran mesin dalam tahap pelatihan dan prediksi. Selain itu, untuk menjaga konsistensi dan akurasi, dilakukan validasi silang terhadap label sentimen dan emosi yang diberikan. Validasi ini bertujuan untuk mendeteksi kemungkinan bias dalam anotasi serta memastikan bahwa setiap label benar-benar merepresentasikan isi dari komentar. Proses pelabelan yang hati-hati ini menjadi bagian penting dalam pembangunan model klasifikasi yang handal, terutama dalam menghadapi karakteristik teks pendek dan ekspresi bahasa alami yang bersifat kompleks serta kontekstual seperti dalam umpan balik mahasiswa.

Tabel 3. Ekstraksi Fitur TF-IDF

adil	ajar	antusias	bangun	...	Dosen
0.0	0.571597	0.0	0.000000		0.000000
0.0	0.000000	0.0	0.000000		0.000000
0.0	0.000000	0.0	0.000000		0.000000
0.0	0.000000	0.0	0.659267		0.361572
0.0	0.000000	0.0	0.000000		0.000000

2.4. Pemodelan dan Pelatihan

2.4.1 Support Vector Machine

Support Vector Machine (SVM) merupakan sebuah metode *supervised learning* yang bisa digunakan untuk proses pengklasifikasian [13]. Metode SVM sering digunakan untuk mengklasifikasikan teks karena menghasilkan akurasi yang tinggi atau lebih baik [14] [18]. Metode SVM merupakan algoritma yang digunakan untuk dapat mengenali pola dan juga mengklasifikasikan data [19] [20]. Metode SVM menjalankan tugasnya dengan cara mencari *hyperlane* yang akan memisahkan kelas data dengan margin maksimal, dengan begitu memungkinkan prediksi yang lebih akurat [21]. Salah satu keunggulan utama SVM adalah kemampuannya dalam menangani data berdimensi tinggi, seperti teks, yang direpresentasikan dalam bentuk vektor TF-IDF. Dalam konteks klasifikasi teks pendek, SVM mampu membentuk hyperplane optimal yang memisahkan kelas-kelas secara jelas meskipun data memiliki banyak fitur namun relatif sedikit dalam jumlah [1].

2.4.2 Naive Bayes Classifier (NBC)

Naive Bayes Classifier (NBC) merupakan sebuah metode klasifikasi berdasarkan pada *teorema bayes*. Metode klasifikasi ini memprediksi suatu data berdasarkan distribusi probabilitas dan fitur-fitur yang tersedia. Algoritma ini bekerja dengan menghitung probabilitas suatu data masuk ke dalam kelas tertentu berdasarkan fitur-fitur yang dimilikinya. Istilah "naïve" merujuk pada asumsi dasar algoritma ini, yaitu menganggap bahwa semua fitur atau atribut bersifat independen satu sama lain dalam memengaruhi kelas target [12]. Meskipun asumsi ini tidak selalu terpenuhi dalam dunia nyata, khususnya dalam data teks, NBC tetap menunjukkan performa yang cukup baik dan efisien terutama dalam tugas klasifikasi berbasis teks pendek seperti komentar mahasiswa. Proses NBC mengacu dari kehadiran suatu fitur di kelas tidak terkait dengan kehadiran fitur lain di kelas yang sama. Metode ini menggunakan probabilitas dan statistik untuk memperkirakan peluang berdasarkan pengalaman masa sebelumnya [22] [23].

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Analisis Kinerja Support vector machine dan Naïve Bayes Classifier dalam Klasifikasi Sentimen dan Emosi pada Umpan Balik Mahasiswa terhadap Kinerja Dosen-(Teguh Syafrudin, Arief Hermawan, Donny Avianto)

1)

$P(H|X)$ adalah probabilitas hipotesis H berdasarkan kondisi x (posteriori probabilitas), sedangkan $P(X|H)$ adalah probabilitas X berdasarkan kondisi, $P(H)$ untuk probabilitas hipotesis H (prior probabilitas) dan $P(X)$ untuk probabilitas dari X [9][23][24][25].

2.5. Evaluasi Model

Tahapan evaluasi menjadi tahapan untuk mengetahui mana metode yang memberikan evaluasi klasifikasi pada kinerja metode. Proses ini menggunakan confusion matrix untuk mengevaluasi performa model melalui persamaan berikut.

$$accuracy = \frac{TP+TN}{Total} \quad (1)$$

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

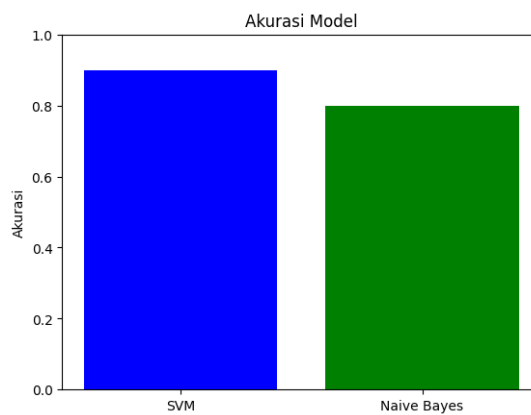
Dengan begitu akan mendapatkan nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Dengan dilakukannya evaluasi klasifikasi peneliti bisa mengetahui model mana yang mempunyai akurasi tinggi apakah *Support Vector Machine* atau *Naïve Bayes Classifier* dalam melakukan klasifikasi [10]. Dalam konteks klasifikasi sentimen, hal ini membantu mengetahui apakah model terlalu bias terhadap sentimen positif atau negatif. Sedangkan untuk klasifikasi emosi, confusion matrix dapat memperlihatkan apakah emosi tertentu, seperti "bingung" atau "puas", sering tertukar dengan emosi lain karena kesamaan makna atau ekspresi. Pendekatan evaluasi ini memberikan gambaran lebih holistik mengenai kinerja model, tidak hanya dari sisi angka agregat tetapi juga pola kesalahan yang terjadi selama prediksi[12].

Evaluasi model perlu dilakukan karena keberhasilan suatu model *Machine Learning* tidak hanya dilihat dari kemampuannya untuk melakukan prediksi, tetapi juga dari seberapa akurat dan andal prediksi tersebut dalam berbagai kondisi data. Tanpa evaluasi yang tepat, model mungkin terlihat bekerja dengan baik pada data pelatihan, namun gagal menangani data baru yang belum pernah dilihat sebelumnya [24]. Oleh karena itu, penggunaan metrik evaluasi seperti akurasi, precision, recall, dan F1-score sangat penting untuk memberikan gambaran menyeluruh mengenai kekuatan dan kelemahan masing-masing algoritma dalam klasifikasi.

Selain itu, metrik seperti precision dan recall sangat penting dalam konteks klasifikasi multi-kelas, terutama untuk tugas klasifikasi emosi yang memiliki label yang lebih banyak dan distribusi tidak seimbang. Precision mengukur ketepatan prediksi positif yang dibuat model, sedangkan recall mengukur kemampuan model dalam menangkap seluruh data aktual dari suatu kelas. F1-score sebagai harmonisasi antara precision dan recall digunakan untuk mengevaluasi keseimbangan performa, terutama ketika terjadi trade-off antara kedua metrik tersebut. Penggunaan macro average dan weighted average juga membantu memberikan perspektif yang lebih adil terhadap distribusi antar kelas. Lebih lanjut, confusion matrix tidak hanya berfungsi sebagai alat evaluasi kuantitatif, tetapi juga dapat digunakan untuk melakukan analisis kualitatif terhadap kesalahan prediksi yang terjadi. Misalnya, jika suatu model sering memprediksi label "bingung" padahal seharusnya "kecewa", maka perlu dianalisis apakah hal itu disebabkan oleh kemiripan kata dalam komentar atau kekurangan dalam representasi fitur. Informasi seperti ini dapat menjadi masukan berharga untuk perbaikan model ke depan, baik melalui optimasi parameter, peningkatan kualitas fitur, atau penambahan data pelatihan yang lebih seimbang [26].

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis sentimen dengan algoritma *Support Vector Machine (SVM)* dan *Naive Bayes*, model SVM mendapatkan hasil performa yang lebih unggul dibandingkan *Naive Bayes* dalam mengklasifikasikan sentimen *feedback* mahasiswa yang digambarkan pada Gambar 2.



Gambar 2. Perbandingan Akurasi Model.

Adapun hasil akurasi perbandingan algoritma SVM dan Naïve Bayes pada analisis sentimen digambarkan sebagai berikut:

Tabel 4. Hasil perbandingan akurasi SVM dan Naïve Bayes

Metrik	SVM	Naive Bayes
Akurasi	0.90	0.80
F1-Score (Macro Avg)	0.91	0.79
F1-Score (Weighted Avg)	0.90	0.80
Precision (Macro Avg)	0.93	0.83
Recall (Macro Avg)	0.90	0.82

Pada Tabel 4. SVM mencapai akurasi sebesar 90%, yang berarti model tersebut mampu memprediksi sentimen dengan benar pada 90% kasus. Selain itu, nilai *F1-score macro average* dan *weighted average* SVM yang masing-masing sebesar 0.91 dan 0.90 mengindikasikan keseimbangan yang baik antara precision dan recall untuk semua kelas sentimen. Hal ini menunjukkan bahwa SVM mampu meminimalkan kesalahan *false positive* dan *false negative* secara keseluruhan. Di sisi lain, *Naive Bayes* memiliki akurasi yang lebih rendah, yaitu 80%. Meskipun precision untuk kelas "Positif" mencapai 1.00, namun recall-nya hanya 0.57, yang berarti model ini cenderung melewatkan beberapa *feedback* positif yang sebenarnya. Nilai *F1-score macro average* dan *weighted average Naive Bayes* yang lebih rendah (0.79 dan 0.80) juga menunjukkan performa yang kurang optimal dibandingkan SVM. Untuk dataset ini, model SVM lebih direkomendasikan karena memiliki tingkat akurasi dan *F1-score* yang lebih tinggi, yang mengindikasikan kemampuan klasifikasi sentimen yang lebih baik dan lebih seimbang untuk semua kelas sentimen. Meskipun demikian, perlu diingat bahwa performa model dapat bervariasi tergantung pada karakteristik dataset dan parameter yang digunakan.

3.1 Analisis Emosi

Berdasarkan hasil analisis emosi menggunakan algoritma *Support Vector Machine (SVM)* dan *Naive Bayes* dengan fitur *TF-IDF* yang digambarkan pada Tabel 5 dan Tabel 6.

Tabel 5. Hasil Analisis Emosi dengan SVM

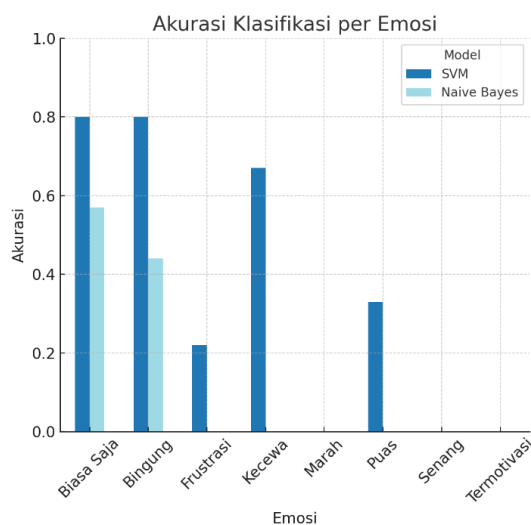
Emosi	Precision	Recall	F1-Score	Support
Biasa Saja	0.67	1.00	0.80	2
Bingung	1.00	0.67	0.80	3
Frustrasi	0.17	0.33	0.22	3
Kecewa	1.00	0.50	0.67	2
Marah	0.00	0.00	0.00	3
Puas	0.50	0.25	0.33	4
Senang	0.00	0.00	0.00	0
Termotivasi	0.00	0.00	0.00	3
Accuracy			0.35	20
Macro Avg	0.42	0.34	0.35	20
Weighted Avg	0.44	0.35	0.37	20

Tabel 6. Hasil Analisis Emosi dengan Naïve Bayes

Emosi	Precision	Recall	F1-Score	Support
Biasa Saja	0.40	1.00	0.57	2
Bingung	0.33	0.67	0.44	3
Frustrasi	0.00	0.00	0.00	3
Kecewa	0.00	0.00	0.00	2
Marah	0.00	0.00	0.00	3
Puas	0.00	0.00	0.00	4
Senang	0.00	0.00	0.00	0
Termotivasi	0.00	0.00	0.00	3
Accuracy			0.20	20
Macro Avg	0.09	0.21	0.13	20
Weighted Avg	0.09	0.20	0.12	20

Model SVM menampilkan hasil performa yang relatif lebih baik dibandingkan *Naive Bayes*, meskipun akurasi keduanya masih tergolong rendah. SVM masih menunjukkan kemampuan dalam mendeteksi beberapa emosi seperti "Biasa Saja" dan "Bingung" dengan recall 1.00 dan 0.67. Akan tetapi, untuk emosi lainnya, performa SVM masih perlu ditingkatkan, terutama untuk emosi "Marah", "Senang", dan "Termotivasi" di mana model tidak dapat memprediksi dengan benar (precision dan recall 0.00). *Naive Bayes* dengan fitur TF-IDF menunjukkan performa yang sangat rendah dengan akurasi hanya 20%. Model

ini hanya mampu memprediksi emosi "Biasa Saja" dengan recall 1.00, sementara untuk emosi lainnya model kesulitan dalam melakukan klasifikasi seperti yang tergambar pada Gambar 3.



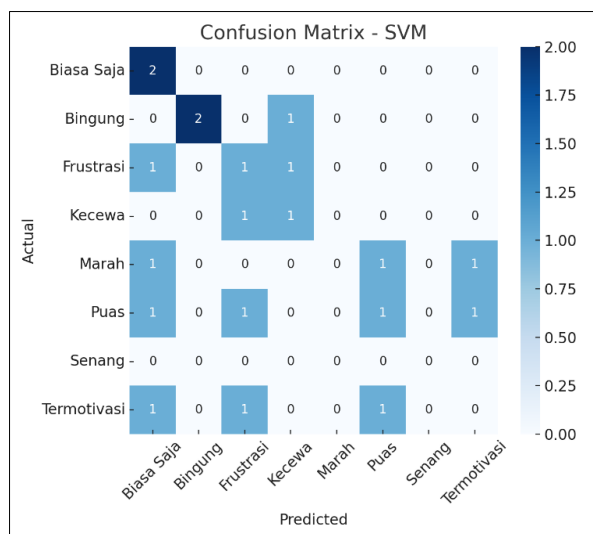
Gambar 3. Akurasi Klasifikasi per Emosi

Adapun hasil perbandingan analisis sentimen digambarkan pada Tabel 7.

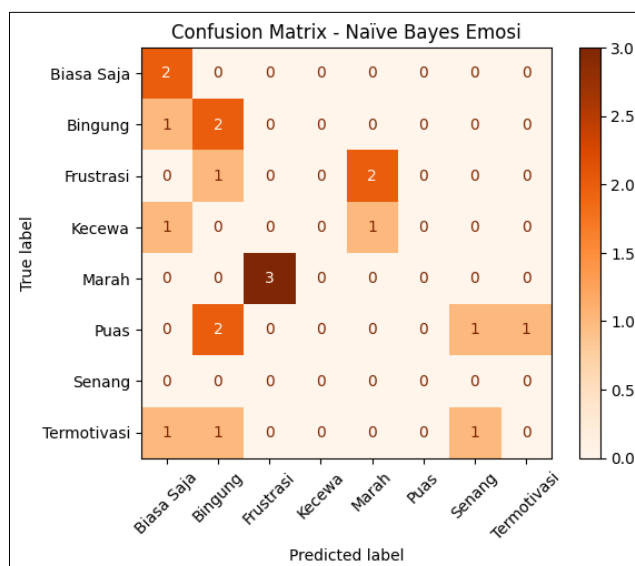
Tabel 7. Perbandingan Hasil Model SVM dan Naïve Bayes

Metrik	SVM	Naive Bayes (TF-IDF Only)
Akurasi	0.35	0.20
F1-Score (Macro Avg)	0.35	0.13
F1-Score (Weighted Avg)	0.37	0.12
Precision (Macro Avg)	0.42	0.09
Recall (Macro Avg)	0.34	0.21

Dari hasil yang analisa, baik SVM yang mendapatkan akurasi sebesar 35% maupun *Naive Bayes* yang hanya 20% masih perlu dioptimalkan untuk meningkatkan performa klasifikasi emosi. Rendahnya akurasi dan F1-score, terutama pada *Naive Bayes*, mengindikasikan bahwa model kesulitan dalam membedakan berbagai emosi dalam *feedback* mahasiswa SVM menunjukkan performa yang sedikit lebih baik dibandingkan *Naive Bayes* dalam klasifikasi delapan label emosi, meskipun secara umum akurasi dan F1-Score masih tergolong rendah. Hal ini disebabkan oleh keterbatasan data, distribusi label yang tidak seimbang, dan ekspresi emosi yang ambigu dalam teks pendek seperti yang digambarkan *Confusion Matrix* pada Gambar 4 dan 5



Gambar 4. Confusion Matrix SVM



Gambar 5. Confusion Matrix Naïve Bayes

3.1 Diskusi

Hasil penelitian ini memberikan pengetahuan bahwa model *Support Vector Machine (SVM)* mencapai tingkat keakurasian yang lebih tinggi dibandingkan dengan *Naïve Bayes* dalam klasifikasi sentimen dan emosi pada *feedback* mahasiswa. Temuan ini sejalan dengan beberapa penelitian terdahulu yang juga menunjukkan keunggulan SVM dalam analisis sentimen. Penelitian lain juga melaporkan hasil serupa pada analisis sentimen *tweet* dengan SVM mencapai akurasi yang lebih tinggi daripada *Naïve Bayes* dan algoritma *machine learning* lainnya [7]. Keunggulan SVM ini disebabkan oleh kemampuannya dalam mengatasi data berdimensi tinggi dan memisahkan kelas dengan margin yang optimal. Lebih lanjut, penggunaan fitur *hybrid* yang menggabungkan *TF-IDF* dan *word embeddings* juga berkontribusi pada peningkatan performa SVM. Dalam penelitian ini juga menunjukkan bahwa kombinasi *TF-IDF* dan *Word2Vec* menghasilkan representasi teks yang lebih baik dan meningkatkan akurasi klasifikasi sentimen.

Penelitian lain juga menyoroti manfaat penggunaan *word embeddings* dalam menangkap semantik dan konteks kata, yang penting untuk analisis sentimen [8]. Meskipun SVM menunjukkan performa yang lebih baik, akurasi klasifikasi emosi masih tergolong rendah. Hal ini sejalan dengan tantangan yang dihadapi dalam analisis emosi teks, di mana emosi merupakan konsep yang kompleks dan dapat diekspresikan dengan berbagai cara bahwa akurasi klasifikasi emosi teks umumnya lebih rendah dibandingkan dengan klasifikasi sentimen [1]. Penelitian lain menunjukkan bahwa model berbasis *attention* seperti *transformers* mencapai hasil yang menjanjikan dalam klasifikasi emosi [14].

Temuan ini juga mengindikasikan bahwa pendekatan hybrid machine learning yang menggabungkan teknik klasik dengan representasi fitur yang tepat dapat menjadi solusi praktis untuk kondisi dengan keterbatasan data dan sumber daya komputasi. Meskipun hasil klasifikasi emosi belum optimal, pendekatan awal ini penting sebagai fondasi eksplorasi model yang lebih kompleks. Penelitian oleh Zidny et al. (2024) juga menunjukkan bahwa kombinasi antara SVM dan representasi fitur seperti TF-IDF mampu memberikan performa klasifikasi yang cukup baik pada data komentar mahasiswa meskipun tanpa menggunakan model berbasis deep learning [1]. Selain itu, penelitian oleh Firdaus dan Suryono (2025) menegaskan bahwa SVM tetap relevan untuk digunakan dalam tugas klasifikasi sentimen jika dataset bersifat terstruktur dan konteks domainnya terbatas, seperti dalam penilaian kinerja dosen[18].

Dengan mempertimbangkan keterbatasan yang ada, hasil dari penelitian ini dapat menjadi dasar untuk pengembangan lebih lanjut yang mengarah pada integrasi model-model berbasis deep learning seperti BERT atau RoBERTa, khususnya untuk klasifikasi emosi yang membutuhkan pemahaman semantik lebih dalam terhadap konteks bahasa alami. Oleh karena itu, penelitian selanjutnya diharapkan dapat menerapkan pendekatan-pendekatan tersebut guna meningkatkan performa model dalam klasifikasi emosi pada *feedback* mahasiswa.

4. KESIMPULAN

Hasil klasifikasi menunjukkan bahwa SVM memiliki performa terbaik untuk klasifikasi sentimen dengan akurasi mencapai 90%, mengungguli Naïve Bayes yang hanya memperoleh akurasi 80%. Sebaliknya, performa klasifikasi emosi jauh lebih rendah, dengan akurasi maksimum 35% pada model SVM dan 20% pada Naïve Bayes. Beberapa emosi seperti “marah”, “termotivasi”, dan “senang” tidak dapat dikenali oleh model karena ketidakseimbangan distribusi data dan konteks emosi yang sulit ditangkap dari teks pendek. Temuan ini menunjukkan bahwa pendekatan berbasis SVM efektif untuk klasifikasi sentimen dalam teks pendek, namun kurang tepat untuk klasifikasi emosi yang lebih kompleks. Oleh karena itu, disarankan agar penelitian selanjutnya mempertimbangkan pendekatan berbasis *deep learning* dan perluasan dataset yang lebih seimbang untuk meningkatkan akurasi klasifikasi emosi. Dengan upaya tersebut, analisis umpan balik mahasiswa diharapkan mampu memberikan wawasan yang lebih kaya dan bermanfaat bagi peningkatan kualitas pembelajaran.

DAFTAR PUSTAKA

- [1] A. H. Zidny, S. H. Wijoyo, and S. A. Wicaksono, “Analisis Kinerja Support Vector Machine dan Naïve Bayes Classifier Pada Data Komentar Kualitas Mengajar Dosen (Studi Kasus : Pendidikan Teknologi Informasi Universitas Brawijaya),” vol. 1, no. 22, pp. 1–6, 2024.
- [2] F. Rumaisa, Y. Puspitarani, A. Rosita, A. Zakiah, and S. Violina, “Penerapan Natural Language Processing (NLP) di bidang pendidikan,” *J. Inov. Masy.*, vol. 1, no. 3, pp. 232–235, 2021.
- [3] A. Koufakou, “Deep learning for opinion mining and topic classification of course reviews,” *Educ. Inf. Technol.*, vol. 29, no. 3, pp. 2973–2997, 2023.
- [4] L. G. Thanveer Shaika, Xiaohui Taoa,, Christopher Dannb, Haoran Xiec, Yan Lia, “Sentiment analysis and opinion mining on educational data: A survey,” *Nat. Lang. Process. J.*, vol. 2, no. Yan Li, p. 100003, 2023.
- [5] S. Wu *et al.*, “A Comprehensive Exploration of Personalized Learning in Smart Education: From Student Modeling to Personalized Recommendations,” *J. ACM*, vol. 37, no. 4, 2024.
- [6] S. Lende *et al.*, “Analisis Sentimen Siswa Terhadap Pelajaran Informatika di SMPK St. Yohanes Kalembo Lona Dengan Metode Naive Bayes Classifier,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 6, pp. 217–225, 2023.
- [7] D. Asti and R. A. Putri, “Analysis of Vina Film Sentiment on Social Media X Using The Naïve Bayes Method,” *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 6, no. 3, pp. 1278–1289, 2024.
- [8] U. Khairani, V. Mutiawani, and H. Ahmadian, “Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 887–894, 2024.
- [9] T. V. Meiyanti, M. Hatta, and A. Sevtiana, “Analisis Sentimen Mahasiswa Dengan Dosen Menggunakan Metode Naive Bayes Classifier Pada Kuesioner Dosen,” *J. Manaj. Sist. Inf.*, vol. 1, no.

- 2, pp. 55–59, 2023.
- [10] S. S. Tandiapa and G. C. Rorimpandey, “Analisis Sentimen Ulasan Pengguna Pada Aplikasi Threads Dengan Metode Lexicon Based dan Naive Bayes Classifier,” *J. Cahaya Mandalika*, vol. 3, no. 1, pp. 339–353, 2024.
- [11] S. KHAIRUNNISA, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *Telkom Open Libr.*, 2021.
- [12] F. S. Pamungkas and I. Kharisudin, “Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter,” *Pros. Semin. Nas. Mat.*, vol. 4, pp. 1–7, 2021.
- [13] A. J. S. Muhamad Hadi Arfian, Hendrian Sofu Maruhawa, Aulia Aisyah Putri, Dinah Ratulugina, Ellsza Ridzky Khoirunnisa, Verell Hermawan, “Analisis Sentimen Pada Media Sosial Menggunakan Metode Support Vector Machine,” *J. Ilmu Tek. dan Kompu*, vol. 8, no. 3, p. 1724, 2024.
- [14] L. G. Thanveer Shaika., Xiaohui Taoa., Christopher Dannb, Haoran Xiec, Yan Lia, “A Comparative Sentiment Analysis of Computer Engineering Student Feedback Using Decision Trees and SVM,” vol. 10, no. 1, pp. 71–82, 2025.
- [15] N. Dhianur Alam Putra, Didi Juardi, Arif Sholehudin, and Slamet Abadi, “Application Penerapan support vector machine pada analisis sentimen kebijakan pembelajaran tatap muka saat pandemi 2021,” *INFOTECH J. Inform. Teknol.*, vol. 4, no. 2, pp. 190–199, 2023.
- [16] “Advances in Multimedia - 2022 - Xiang - Application of an Improved TF-IDF Method in Literary Text Classification.pdf.” .
- [17] D. E. Cahyani and I. Patasik, “Performance comparison of tf-idf and word2vec models for emotion text classification,” *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2780–2788, 2021.
- [18] N. D. Firdaus and R. R. Suryono, “Perbandingan Algoritma Naive Bayes dan SVM dalam Analisis Sentimen Pengguna AI di Platform X,” vol. 6, no. 4, pp. 2624–2634, 2025.
- [19] I. Zulfahmi, H. Syahputra, S. I. Naibaho, M. A. Maulana, and E. P. Sinaga, “Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 10, no. 1, p. 52, 2023.
- [20] N. C. Ramadani, I. Tahyudin, and A. Shouni Barkah, “Perbandingan Algoritma Support Vector Machine, Decision Tree, dan Logistic Regresion Pada Analisis Sentimen Ulasan Aplikasi Netflix,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 10, no. 2, pp. 110–117, 2024.
- [21] H. Setiawan, E. Utami, and S. Sudarmawan, “Analisis Sentimen Twitter Kuliah Online Pasca Covid-19 Menggunakan Algoritma Support Vector Machine dan Naive Bayes,” *J. Komtika (Komputasi dan Inform.*, vol. 5, no. 1, pp. 43–51, 2021.
- [22] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa’ah, “Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021.
- [23] D. Wijaya, R. A. Saputra, and F. Irwiensyah, “Analisis Sentimen Ulasan Aplikasi Samsat Digital Nasional Pada Google Playstore Menggunakan Algoritma Naïve Bayes,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 4, no. 4, pp. 2369–2380, 2024.
- [24] T. Angelya, A. Rahman, and I. Pradesan, “Implementasi Sistem Klasifikasi Tim Kerja Menggunakan Algoritma Naïve Bayes (Studi Kasus: PT.MNL),” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 86–94, 2023.
- [25] K. A. Nugraha, “Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan Naïve Bayes,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 3, pp. 715–721, 2021.
- [26] M. F. Saleh and R. Imanda, “Public Sentiment Analysis of the Free Meal Program : A Comparison of Naive Bayes and Support Vector Machine Methods on the Twitter (X) Social Media Platform,” vol. 9, no. 1, pp. 131–139, 2025.