

Penggabungan Model SVM dan *Naive Bayes* dengan Pendekatan *Soft Voting* untuk Analisis Sentimen *Tong Tji Tea House*

Combining SVM and Naive Bayes Models using a Soft Voting Approach for Sentiment Analysis of Tong Tji Tea House

¹Fendi Pradana Saputra*, ²Ozzi Suria

^{1,2}Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Mercu Buana
Yogyakarta

^{1,2}Gg. Jemb. Merah No.84C, Soropadan, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah
Istimewa Yogyakarta 55283

*e-mail: saputrafendi460@gmail.com, ozzisoria@mercubuana-yogya.ac.id

(received: 1 July 2025, revised: 12 July 2025, accepted: 13 July 2025)

Abstrak

Di era teknologi digital dan media sosial saat ini mempengaruhi masyarakat untuk lebih aktif memberikan review terhadap layanan restoran yang memuat berbagai pendapat dan persepsi pelanggan. Penelitian ini bertujuan untuk menganalisis sentimen dalam teks ulasan berbahasa Indonesia dengan menerapkan tiga model machine learning yaitu Support Vector Machine (SVM), Naive Bayes (NB), dan kombinasi keduanya melalui metode Ensemble Soft Voting Classifier. Objek penelitian terfokus pada review pengguna restoran Tong Tji Tea House yang diperoleh dari platform Google Maps, dengan sebaran data terdiri dari sentimen positif (2.676 data), netral (670 data), dan negatif (251 data). Ketidakseimbangan distribusi kelas menjadi tantangan dalam membangun model klasifikasi yang optimal. Untuk mengatasi hal tersebut digunakan optimasi parameter menggunakan metode grid search. Model SVM dengan kernel linier dan parameter $C=10$ menghasilkan akurasi sebesar 0,9289 dan skor F1 positif sebesar 0,9289. Model NB mencatatkan akurasi sebesar 0,8340 dengan skor F1 sebesar 0,9102. Sedangkan kombinasi model Ensemble pendekatan soft voting dengan bobot 4:1 (SVM:NB) menunjukkan performa terbaik dengan akurasi 0,9344 dan skor F1 positif 0,9750. Hasil tersebut menunjukkan bahwa metode Ensemble mampu meningkatkan akurasi dan ketahanan model dalam menangani ketidakseimbangan data.

Kata kunci: analisis sentimen, tong Tji tea house, machine learning, kombinasi ensemble soft , ulasan restoran.

Abstract

In today's digital technology and social media era, people are increasingly influenced to actively share reviews of restaurant services, expressing a wide range of customer opinions and perceptions. This study aims to analyze sentiment in Indonesian-language review texts using three machine learning models: Support Vector Machine (SVM), Naive Bayes (NB), and a combination of both through an Ensemble Soft Voting Classifier approach. The research focuses on user reviews of the Tong Tji Tea House, collected from the Google Maps platform, with sentiment data distributed as follows: positive (2,676 entries), neutral (670 entries), and negative (251 entries). The class imbalance poses a significant challenge in developing an optimal classification model. To address this, parameter optimization was carried out using the Grid Search method. The SVM model with a linear kernel and $C=10$ parameter achieved an accuracy of 0.9289 and a positive F1-score of 0.9289. The NB model recorded an accuracy of 0.8340 with an F1-score of 0.9102. Meanwhile, the Ensemble model with a soft voting approach and a 4:1 weight ratio (SVM:NB) demonstrated the best performance, achieving an accuracy of 0.9344 and a positive F1-score of 0.9750. These results indicate that the Ensemble method effectively enhances model accuracy and robustness in handling imbalanced data.

Keywords: *sentiment analysis, tong Tji tea house, machine learning, ensemble soft combination, restaurant reviews*

1 Pendahuluan

Dalam masa transformasi digitalisasi saat ini, volume Materi teks atau ketikan kata yang berasal dari platform digital internet dan web aplikasi lainnya sangat besar, menyimpan opini dan sentimen publik yang berharga untuk pengambilan keputusan di berbagai bidang seperti bisnis dan kebijakan publik [1]. Melalui analisis sentimen, opini-opini tersebut secara otomatis bisa dimasukkan ke dalam kelompok positif, negatif, ataupun netral, sehingga memberikan pemahaman penting mengenai reaksi masyarakat terhadap produk atau layanan tertentu [2]. Penggunaan media sosial sebagai sumber informasi real-time semakin meningkat, sehingga penting untuk mengembangkan metode analisis sentimen yang dapat menangani dinamika opini yang berubah dengan cepat.

Metode Machine learning SVM dan *Naïve Bayes* sangat luas digunakan analisis sentimen. SVM lebih unggul terhadap akurasi klasifikasi terhadap dataset berdimensi tinggi [2], sedangkan *Naïve Bayes* lebih cepat dan sederhana dalam pemodelan [3]. Namun, setiap metode memiliki keterbatasan, sehingga menggabungkan model dengan pendekatan Soft Voting Ensemble mulai digunakan untuk meningkatkan kinerja klasifikasi [4]. Pendekatan ini menggabungkan probabilitas keluaran model guna menghasilkan prediksi yang lebih baik dan akurat, yang sangat relevan dalam konteks data media sosial yang bervariasi [5].

SVM dan *Naïve Bayes* adalah algoritma yang kerap dimanfaatkan untuk proses analisis sentimen karena keunggulannya dalam klasifikasi teks. SVM, dengan kemampuannya untuk menemukan hyperplane optimal yang memisahkan data, sering menunjukkan kinerja yang superior dalam situasi dataset yang kompleks dan berukuran besar [6]. Sementara itu, *Naïve Bayes* bekerja berdasarkan asumsi independensi antar fitur, sehingga memudahkan perhitungan probabilitas dan memungkinkan model ini untuk beroperasi dengan baik meskipun dilakukan pada beberapa data yang lebih terbatas [7].

Penelitian yang menggabungkan kedua metode ini dalam suatu strategi Ensemble Voting menunjukkan hasil yang menjanjikan. Metode Ensemble, yang mencakup penggabungan beberapa algoritma, terbukti meningkatkan akurasi dan ketahanan model. Misalnya, hasil analisis oleh Alqaraleh menunjukkan kinerja yang lebih baik saat menggunakan penggabungan Random Forest (RF) dan AdaBoost (ADB) [8]. Selain itu, menggabungkan teknik Bagging, SVM, dan *Naïve Bayes* dalam strategi voting mayoritas untuk meningkatkan deteksi sentimen [8]. Penelitian lain yang menggunakan Ensemble Stacking juga mengonfirmasi temuan ini, di mana penggabungan SVM dengan teknik lain meningkatkan kinerja model secara keseluruhan [9].

Penggunaan teknik Ensemble sangat relevan dalam kerangka analisis sentimen yang sering kali melibatkan data berdimensi tinggi dan heterogen, seperti yang ditunjukkan dalam penelitian oleh P dan Rao, yang menekankan perlunya pelatihan iteratif untuk mencapai kinerja optimal dalam analisis sentimen [10]. Lebih lanjut, diversitas dalam teknik Ensemble dianggap Mampu menekan tingkat kesalahan di antara algoritma pengelompokan, seperti yang diungkapkan oleh Leidiyana, di mana kombinasi pengklasifikasi tunggal menghasilkan performa yang lebih robust dibandingkan model tunggal [11]. Dengan demikian, kombinasi SVM, *Naïve Bayes*, dan pendekatan Ensemble tidak hanya menangani masalah kerumitan data, tetapi juga memberikan solusi yang lebih handal dan akurat dalam menganalisis sentimen publik.

Belum ada studi yang membandingkan tiga algoritma klasifikasi (*Naïve Bayes*, SVM, dan Ensemble Voting) dalam konteks ulasan restoran Tong Tji Tea House. Oleh sebab itu, peneliti meneliti ini dimaksudkan untuk melakukan komparasi performa ketiga algoritma tersebut dalam menganalisis sentimen ulasan pelanggan restoran Tong Tji Tea House pada platform Google Maps. Adapun rumusan masalah yang diajukan adalah: pertama, bagaimana performa masing-masing algoritma dalam mengklasifikasikan sentimen ulasan; kedua, algoritma mana yang memberikan hasil klasifikasi terbaik; dan ketiga, bagaimana kontribusi masing-masing metode terhadap peningkatan akurasi klasifikasi sentimen.

Secara keseluruhan, pemahaman mendalam mengenai metode SVM, *Naïve Bayes*, dan strategi Voting Ensemble dalam analisis sentimen tidak hanya memberikan wawasan mengenai dinamika publik namun juga memperkuat kemampuan kita dalam merespons dan menafsirkan opini publik

terhadap berbagai isu kontemporer. Keterlibatan dalam penelitian dan pengembangan lebih lanjut di bidang ini diharapkan dapat menghasilkan metode yang lebih efisien yang dapat dikaitkan dengan berbagai penerapan dalam industri, kesehatan, dan kebijakan publik.

2 Tinjauan Literatur

Pada studi yang sebelumnya membahas penerapan algoritma pengoperasian algoritma Naive Bayes dan SVM untuk mengkaji opini pengguna aplikasi TikTok. Pembobotan pada penelitian ini di hitung dari setiap kata berdasarkan kemunculan dalam dokumen menggunakan metode TF-IDF. Pada tahap pra-pemrosesan data di bagi di bagi menjadi 0,2 untuk data ulasan dan 0,8 untuk data training. Selanjutnya Hasil pengujian penelitian ini memuat Confusion matrix serta akurasi 79%, dengan precision yang diperoleh sebesar 86%. Untuk metode Support Vector Machine, akurasi yang dicapai adalah 84% dan nilai precision rata-ratanya 77%. Kala di lihat dari hasil akurasi yang telah di dapatkan dapat di simpulkan bahwa hasil metode Support Vector Mechine lebih baik dari pada Naive Bayes [12].

Penelitian terbaru mengenai analisis sentimen media sosial khususnya Twitter menggunakan model klasifikasi tradisional seperti Regresion Logistik, Naive Bayes, dan SVM dengan akurasi mendekati 89% [13]. Penggunaan metode Ensemble seperti Soft Voting dan Random Forest terbukti mampu meningkatkan akurasi hingga 91% dan 85.88% dalam kasus analisis sentimen harga tiket pesawat dan isu kesehatan COVID-19 [14].

Peneliti terdahulu juga melakukan analisis sentimen terhadap cuitan Twitter terkait seleksi CPNS 2021 dengan menggunakan metode klasifikasi k-NN, C4.5, Naive Bayes, dan Ensemble Vote. Studi ini menggunakan model CRISP-DM dan data dari akun resmi @BKNGoid. Pembersihan teks, Tokenisasi, penghapusan Stopword, dan Stemming adalah bagian dari proses preprocessing. Dapat disimpulkan dari hasil bahwa akurasi algoritma Naive Bayes mencapai 100%, tertinggi di antara model lain, diikuti oleh C4.5 (85,71%), Vote Ensemble (66,67%), dan k-NN (61,90%). Hasil studi mengindikasikan bahwa algoritma Naive Bayes bekerja dengan baik dalam tugas klasifikasi opini publik pada media sosial.[15]

Pra-pemrosesan data, yang mencakup Cleaning, Tokenize, Stopword, dan Stemming, sangat penting untuk meningkatkan kualitas input model.. Pembobotan istilah berbasis pengawasan, seperti Regularized Entropy (RE), juga memiliki dampak positif yang signifikan terhadap kinerja klasifikasi data teks pendek [16].

Pengembangan lebih lanjut disarankan dalam penggunaan model pembelajaran mendalam dan embeddings yang telah dilatih sebelumnya, misalnya BERT, untuk mengatasi kompleksitas bahasa dan variasi idiomatik, serta dalam pengujian lintas domain dan aplikasi real-time [17].

3 Metode Penelitian

3.1 Analisis sentiment

Analisis sentimen adalah cara untuk menemukan dan mengukur opini atau perasaan seseorang terhadap suatu entitas. Pada review produk online, proses ini dilakukan menilai opini kolektif konsumen terhadap produk yang dibahas [18]. Analisis persepsi juga digunakan untuk mengetahui kecenderungan sentimen atau ulasan yang terdapat pada riview Google Maps Tong Tji Tea House apakah bersifat negetif, positif, atau netral. Teknik analisis sentimen banyak dimanfaatkan diberbagai bidang bisnis untuk mengevaluasi kepuasan pelanggan melalui ulasan daring.

Proses analisis text cloud beserta penyajian jumlah data yang akan dikelola telah selesai, kemudian langkah selanjutnya yaitu menganalisis data yang sudah dikumpulkan.[19]

3.2 Web Scrapping

Pada tahap ini dataset ulasan Tong Tji Tea House dari Google maps di dapatkan dengan cara scrapping dengan menggunakan bahasa pemrograman phyton. Ulasan pengunjung Tong Tji Tea House di ambil pada Google maps sebagai dataset yang akan di ganakan dalam proses analisis sentimen. Proses scrapping data akan berfokus pada kategori data ulasan yang terbaru.

3.3 Preprocessing Data

Langkah pra-pemrosesan adalah proses awal dalam menyiapkan data teks dengan tujuan untuk membuat data teks menjadi lebih tersistematis dan terlihat rapi, sehingga analisis dapat dilakukan

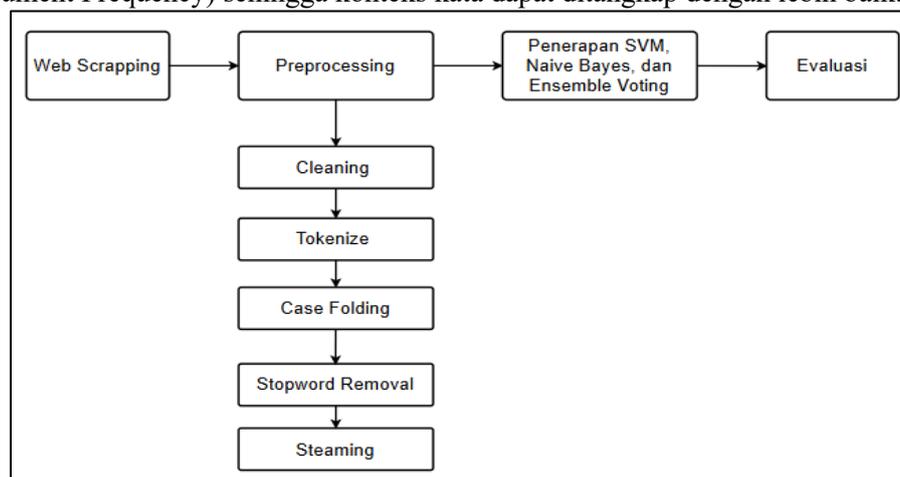
dengan mudah dan lebih akurat[20]. Data yang dikumpulkan merupakan data yang tidak terstruktur sehingga diperlukan Prapemrosesan. Berikut yang dilakukan pada tahap Preprocessing:

- a. **Cleaning**
Peneliti pada proses ini melakukan serangkaian langkah pembersihan teks dengan menghapus seluruh karakter yang bukan merupakan huruf alfabet. Tahap ini dimaksudkan untuk menghilangkan komponen analisis sentimen yang tidak penting..
- b. **Tokenize**
Pada titik ini, proses pemisahan teks menjadi bagian-bagian kecil yang dikenal sebagai token dilakukan..Proses ini juga menghapus tanda baca yang tidak diperlukan misalnya “ penyajian makanan cepat dan sangat bersih” akan di pisah menjadi token token seperti [“ penyajian” “makanan”, “cepat”, “dan”, “sangat”, “bagus”.
- c. **Case Folding**
Pada tahapan ini transformasi huruf dilakukan dengan mengubah seluruh karakter menjadi huruf kecil. Tujuannya agar tidak ada perbedaan pengolahan antara kata yang di tulis dengan huruf besar maupun kecil. Contoh: “Tong Tji Tea House harganya sangat terjangkau” akan diubah menjadi “tong tji tea house harganya sangat terjangkau”.
- d. **Stopword Removal**
Proses ini dikenal sebagai penghilangan stopwords, dimana kata-kata yang dianggap kurang bermakna dalam konteks analisis akan dihilangkan untuk meningkatkan kualitas dan akurasi hasil analisis sebagainya. Contoh: “Tong Tji Tea house adalah salah satu restoran yang bersih di Amplaz” setelah stopwords beberapa kata akan di hapus menjadi “tong tji tea house restoran bersih di amplaz”.
- e. **Steaming**
Tahapan terakhir pada Preprocessing data adalah steaming yang berfungsi untuk mengembalikan kata yang berimbuhan ke bentuk dasarnya. Langkah ini sangat penting karena dalam tahapan pemrosesan teks Dapat menyederhanakan variasi kata dan meningkatkan kinerja algoritma klasifikasi. Contoh : "Makanan di restoran ini sangat lezat dan pelayanannya memuaskan, saya sudah merekomendasikannya ke teman-teman". Akan di ubah menjadi “makan”, “di”, “restoran”, “ini”, “sangat”, “enak”, “dan”, “layanan”, “puas”, “saya”, “sudah”, “rekomendasi”, “kepada”, “teman”.

3.4 Penerapan Algoritma SVM, Naïve Bayes, dan Ensemble Voting

Studi ini menerapkan analisis sentimen pada ulasan pengguna dengan tiga algoritma klasifikasi: Support Vector Machine (SVM), Naive Bayes, dan Ensemble Voting. Tujuan utama penerapan ketiga algoritma ini adalah untuk membandingkan bagaimana masing-masing model bekerja untuk mengkategorikan secara akurat sentimen positif, negatif, dan netral pada teks berbahasa Indonesia..

Proses diawali dengan pra-pemrosesan data yang meliputi pelipatan huruf, penghilangan tanda baca dan angka, tokenisasi, penghilangan stopwords, dan stemming menggunakan pendekatan berbasis Sastrawi. Setelah teks dibersihkan, fitur diserap menggunakan model TF-IDF (Term Frequency - Inverse Document Frequency) sehingga konteks kata dapat ditangkap dengan lebih baik.



Gambar 1 Alur penelitian

Gambar 1 menggambarkan alur penelitian ini dimulai dengan pengumpulan data melalui proses web, scraping untuk mendapatkan ulasan pengguna dari platform digital. Data yang diperoleh kemudian Data tersebut diproses melalui beberapa tahap pra-pemrosesan yang meliputi pembersihan data (Cleaning), pemecahan teks menjadi token (Tokenisasi), pengumpulan huruf menjadi huruf kecil (Case Folding), menghapus kata umum yang tidak penting (Stopword Removal), dan perubahan kata ke bentuk dasarnya Selanjutnya, tiga metode klasifikasi—Ensemble Voting, Support Vector Machine (SVM), dan Naive Bayes yang digunakan untuk mengkategorikan sentimen dalam ulasan tersebut. Proses lanjutan hasil klasifikasi dianalisis menggunakan matrik akurasi, presisi, recall, dan skor F1. Ini dilakukan agar mengetahui efektifitas masing-masing model dalam mengidentifikasi sentimen.

4 Hasil dan Pembahasan

4.1 Preprocessing

Sebelum proses selanjutnya, data analisis harus diperiksa untuk menghilangkan karakter karater yang tidak relevan. Tahap awal preprocessing merupakan langkah penting sebelum melakukan analisis data, karena bertujuan untuk mentransformasikan data tidak terstruktur menjadi bentuk yang lebih sistematis, sehingga dapat mendukung kelancaran proses analisis serta mengurangi risiko kesalahan interpretasi [21].

Tabel 1 Hasil text cleaning

Sesudah	Sebelum
Makanan cukup enak!! Tapi kadang pelayanan lama 😞 😞	Makanan cukup enakTapi kadang pelayanan lama

Pada Tabel 1 Proses membersihkan data mencakup menghilangkan komponen yang tidak penting atau mengganggu seperti noise berupa hashtag, nama pengguna (username), emotikon, serta tanda baca yang tidak mempengaruhi pemeriksaan sentimen[22]. Langkah ini penting di lakukan untuk merapikan struktur teks agar lebih mudah di jelaskan, terutama pada data dari media sosial yang sering kali mengandung Noise.

Tabel 2 Hasil tokenize

Sesudah	Sebelum
Makanan cukup enakTapi kadang pelayanan lama	['Makanan', 'cukup', 'enak', 'Tapi', 'kadang', 'pelayanan', 'lama']

Selanjutnya pada Tabel 2 tahap Memecah kalimat ulasan menjadi potongan kata dikenal sebagai tokenisasi.[23]. Fungsi proses ini adalah untuk mengubah kalimat mentah menjadi bentuk yang lebih tertata dan rapi, Jadi, data tersebut dapat dianalisis secara komputer dengan lebih efektif. Dalam konteks analisis sentimen, tokenisasi sangat penting karena membantu model dalam mengenali dan memproses setiap kata sebagai unit informasi yang berdiri sendiri.

Tabel 3 Hasil case folding

Sesudah	Sebelum
'Makanan', 'cukup', 'enak', 'Tapi', 'kadang', 'pelayanan', 'lama'	'makanan', 'cukup', 'enak', 'tapi', 'kadang', 'pelayanan', 'lama'

Pada tahap selanjutnya di Tabel 3 yaitu proses mengubah semua karakter huruf dalam teks menjadi huruf kecil dikenal sebagai case folding.[24]. Proses ini dilakukan karena teks sering kali tidak konsisten dalm penggunaan huruf kapital, yang Dapat mempengaruhi hasil pencarian dan analisis. Dengan *Case Folding*, teks menjadi lebih seragam dan analisis Dapat di lakukan dengan lebih akurat dan efisien.

Tabel 4 Hasil stopword removal

Sesudah	Sebelum
'makanan', 'cukup', 'enak' 'tapi', 'kadang', 'pelayanan', 'lama'	'makanan', 'cukup', 'enak', 'kadang', 'pelayanan', 'lama'

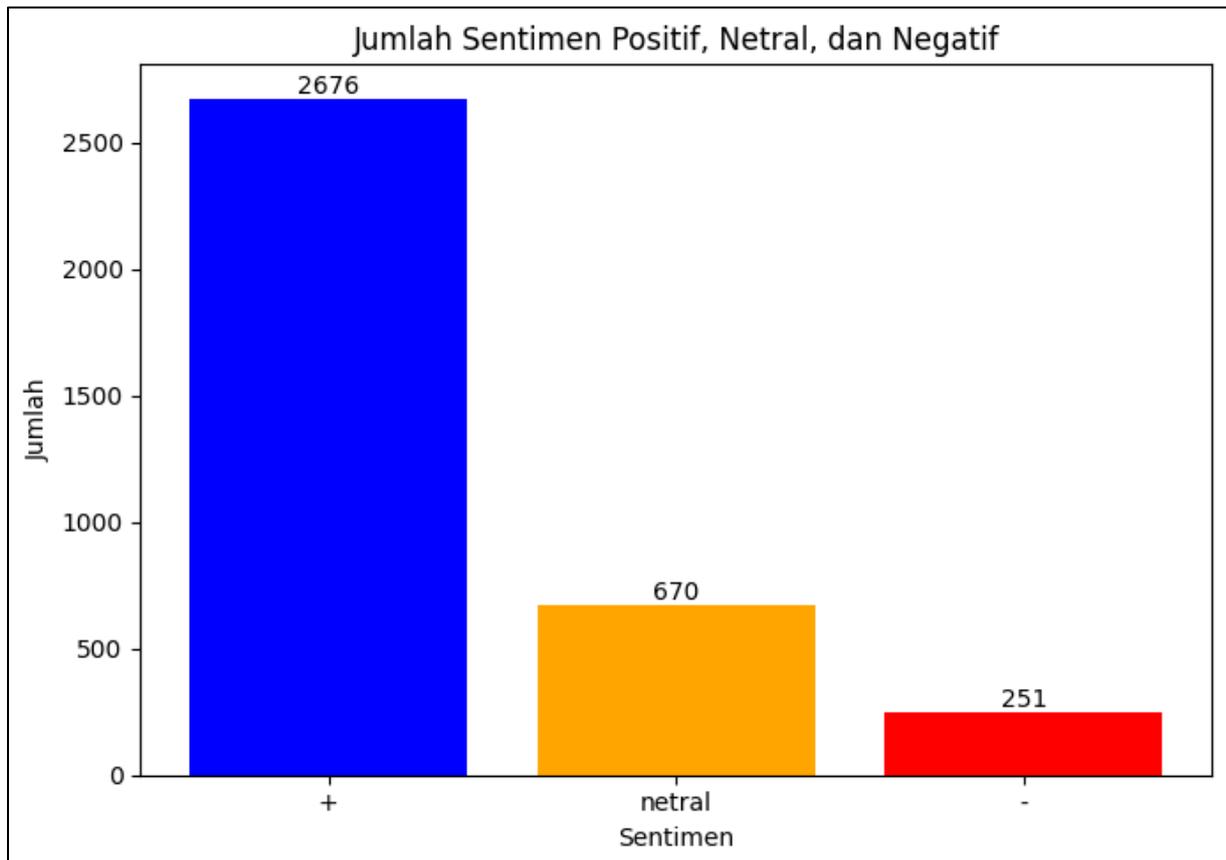
Selanjutnya pada Tabel 4 Penghapusan stopwords adalah proses menghapus kata-kata dan karakter yang terdapat dalam kategori stopwords dan tanda baca. Stopword sendiri adalah kata-kata yang sering digunakan tetapi tidak banyak mempengaruhi emosi yang terkandung dalam sebuah kalimat[25]. Pada umumnya *Stopword Removal* menggunakan perpustakaan digital yang berisi daftar kata umum. Namun, tidak semua kata dalam daftar tersebut relevan dengan semua jenis dokumen, sehingga dapat mempengaruhi hasil pemrosesan.

Tabel 5 Hasil steaming

Sesudah	Sebelum
"makanan cukup enak kadang pelayanan lama"	"makanan cukup enak kadang pelayanan lama"

Tabel 5 adalah tahap terakhir pada Preprocessing yaitu Tahap steaming berfungsi untuk merubah kata-kata menjadi bentuk dasar mereka. Proses tersebut dilakukan karena kata berimbuhan memiliki makna yang sama dengan kata dasar[26]. Algoritma *Stemming* yang digunakan adalah algoritma sederhana yang akan menghapus kombinasi awalan (prefix) dan akhiran (suffix) jika awalan dan akhiran tersebut ada di dalam suatu kata.

4.2 Distribusi Sentiment



Gambar 2 Jumlah distribusi sentimen

Pada Gambar 2 di atas data sentimen mengungkap dominasi sentimen positif yang signifikan dengan 2676 sampel, yang secara signifikan lebih tinggi daripada sentimen netral sekitar 670 sampel dan sentimen negatif hanya 251 sampel. Kondisi ketidakseimbangan ini menyoroti tantangan yang dihadapi dalam mengembangkan model klasifikasi untuk sentimen, khususnya dalam mengidentifikasi minoritas seperti netral dan negatif. Ukuran sampel yang sangat besar ini dapat mengakibatkan model yang bias untuk mayoritas review Tong Tji, membuat kemampuan klasifikasi untuk kelas minoritas kurang ideal. Karena itu, perlu menggunakan teknik adaptasi komputasional seperti Hyperparameter Tuning dan Class Weight Soft Voting untuk distribusi yang tidak seragam sehingga model bisa mendapatkan prediksi yang baik dan konsisten pada semua nilai sentimen.

4.2.1 Hipertuned kedua metode

Metode klasifikasi teks menggunakan model Naive Bayes (NB) dan Support Vector Machine (SVM) digunakan untuk melakukan analisis sentimen ulasan restoran Tong Tji. Melalui proses *Grid Search* otomatis, diperoleh konfigurasi Hyperparameter yang optimal untuk masing-masing model guna meningkatkan performa klasifikasi.

Tabel 6 Classifier hypermeters

Model	Hyperparameter
NB	Alpha=0,1
SVM	'C': 10, 'gamma': 'scale', 'kernel': 'linear' probability=true}

Pada Tabel 6 di atas, model NB disetel dengan parameter Smoothing Alpha sebesar 0,1, yang berfungsi untuk mengatasi masalah sparsity dalam data teks serta menghindari probabilitas nol pada fitur kata yang tidak muncul dalam data pelatihan. Pengaturan ini meningkatkan kemampuan generalisasi tanpa menurunkan sensitivitas model terhadap variasi distribusi kata-kata dalam ulasan. Sementara itu, model SVM dioptimalkan dengan kombinasi hyperparameter regulasi C sebesar 10, kernel linear, dan gamma yang disetel pada skema 'scale', serta pengaktifan estimasi probabilitas (probability=True).

Konfigurasi tersebut memungkinkan SVM untuk secara efektif mengonstruksi hyperplane pemisah di ruang fitur berdimensi tinggi yang dihasilkan oleh transformasi TF-IDF, sekaligus menyediakan estimasi probabilitas kelas yang penting dalam mekanisme penggabungan model berbasis voting lembut (Soft Voting). Pendekatan ini menyeimbangkan antara kompleksitas model dan kemampuan generalisasi, sehingga menghasilkan klasifikasi sentimen yang robust dan akurat dalam konteks analisis review yang kaya dan beragam pada restoran Tong Tji Tea House.

Tabel 7 Akurasi SVM

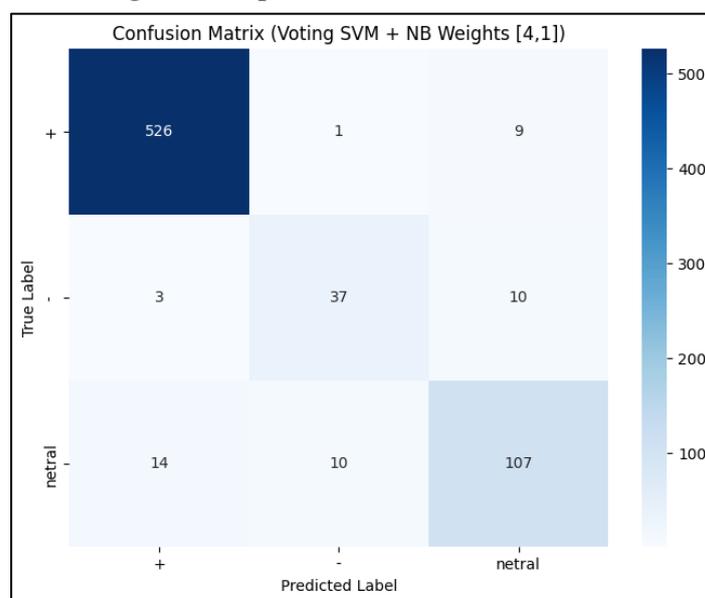
	Precision	Recall	F1-score	support
+	0.97	0.98	0.97	536
-	0.78	0.64	0.70	50
netral	0.80	0.85	0.83	131
accuracy			0.93	717
macro avg	0.85	0.82	0.83	717
Weighted avg	0.93	0.93	0.93	717

Tabel 8 Akurasi naïve bayes

	Precision	Recall	F1-score	support
+	0.85	0.98	0.91	536
-	0.60	0.58	0.59	50
netral	0.83	0.34	0.49	131
accuracy			0.83	717
macro avg	0.76	0.63	0.66	717
Weighted avg	0.83	0.83	0.81	717

Pada Tabel 7 dan 8 ini menjelaskan report dari hasil masing masing metode yang telah di lakukan penelitian baik dari metode SVM maupun metode Naïve Bates. Model Naive Bayes (NB) dengan hyperparameter alpha sebesar 0,1 dan model Support Vector Machine (SVM) dengan konfigurasi parameter C=10, gamma = 'scale', kernel linear, dan probability = True menunjukkan perbedaan signifikan dalam performa klasifikasi. Dari hasil evaluasi, model SVM mampu mencapai nilai akurasi 0.93 dengan weighted F1-score 0.93, yang mengindikasikan kemampuan generalisasi yang lebih baik dibandingkan model NB yang memperoleh akurasi 0.83 dan weighted F1-score 0.81. Analisis metrik Presisi, Recall, dan F1-score per kelas juga memperkuat temuan ini, di mana SVM unggul terutama pada kelas positif dengan precision dan recall mencapai 0.97 dan 0.98, sementara NB menunjukkan performa yang lebih rendah terutama pada kelas negatif dan netral. Pemilihan Hyperparameter untuk kedua model dilakukan melalui Grid Search otomatis guna memperoleh konfigurasi terbaik yang mempertimbangkan keseimbangan antara bias dan varians. Hasil ini menunjukkan bahwa meskipun NB merupakan model yang sederhana dan cepat, SVM dengan kernel linear memberikan performa yang lebih efektif dan stabil dalam konteks analisis sentimen terhadap review restoran Tong Tji Tea House.

4.2.2 Penerapan Soft Voting Terhadap Kedua Metode



Gambar 5 Confusion matrix soft voting SVM + NB

Pada Gambar 5 adalah hasil confusion matrix penggabungan model Ensemble Soft Voting SVM dan Naïve bayes. Matriks konfusi menunjukkan bahwa sebagian besar sampel kelas positif dan netral diklasifikasikan dengan benar, sedangkan kelas negatif memiliki beberapa kesalahan lagi. Bobot

<http://sistemasi.ftik.unisi.ac.id>

dominan pada SVM mencerminkan perannya sebagai prediktor utama, sedangkan NB berperan sebagai pelengkap yang menstabilkan hasil Ensemble. Pendekatan ini menggabungkan kekuatan kedua algoritma sehingga kinerjanya lebih baik dibandingkan model tunggal.

Confusion matrix menyajikan informasi yang lebih mendetail terkait sebaran prediksi yang tepat maupun keliru pada masing-masing kelas. Dari 536 sampel kelas positif, sebanyak 526 berhasil diklasifikasikan dengan benar, sementara sebagian kecil prediksi salah jatuh pada kelas negatif dan netral. Untuk kelas negatif yang terdiri dari 50 sampel, 37 terklasifikasi dengan tepat, sedangkan beberapa prediksi meleset ke kelas positif dan netral. Kelas netral, dengan jumlah 131 sampel, menunjukkan sejumlah kesalahan yang lebih tersebar, di mana 107 sampel diprediksi secara akurat. Akan tetapi, masih terdapat beberapa data yang keliru diklasifikasikan ke dalam kelas positif maupun negatif.

Model Ensemble Soft Voting yang menggabungkan Support Vector Machine (SVM) dan Naive Bayes (NB) dengan bobot 4:1 menunjukkan performa tinggi pada dataset review restoran Tong Tji Tea House. Akurasi dan Weighted F1-score mencapai 0.93, dengan kelas positif memperoleh presisi 0.97 dan recall 0.98, menunjukkan prediksi yang sangat akurat. Kelas negatif dan netral memiliki presisi dan recall masing-masing relatif lebih rendah namun masih memadai.

Rasio 4:1 dipilih berdasarkan optimasi Auto Grid Search yang menguji berbagai kombinasi bobot dan memilih yang memberikan performa terbaik pada metrik akurasi dan F1-score. Dengan demikian, bobot tinggi diberikan pada SVM yang lebih presisi, dan bobot lebih kecil pada NB yang cepat dan sederhana, menghasilkan model Ensemble yang efisien dan unggul.

Tabel 9 Model ensemble soft voting SVM + NB

	Precision	Recall	F1-score	support
+	0.97	0.98	0.97	536
-	0.77	0.74	0.76	50
netral	0.85	0.82	0.83	131
accuracy			0.93	717
macro avg	0.86	0.85	0.85	717
Weighted avg	0.93	0.93	0.93	717

Accuracy on test set: 0,9344

Pada Tabel 9 di atas, model Ensemble Soft Voting yang menggabungkan model Support Vector Machine (SVM) dan Naive Bayes (NB) dengan bobot penentuan 4:1 menunjukkan performa klasifikasi yang sangat baik pada dataset review restoran Tong Tji Tea House. Hasil evaluasi menggunakan classification report memperlihatkan nilai akurasi keseluruhan sebesar 0.93 dan weighted F1-score sebesar 0.93, menandakan bahwa model mampu memberikan prediksi yang akurat dan seimbang di sebagian besar kelas.

Analisis metrik kinerja per kelas menunjukkan bahwa kelas positif (+) memperoleh nilai precision sebesar 0.97 dan recall sebesar 0.98, yang mengindikasikan kemampuan model dalam meminimalkan kesalahan prediksi salah positif dan salah negatif secara signifikan pada kelas ini. F1-score yang tinggi sebesar 0.97 memperkuat keandalan prediksi pada kelas positif yang dominan dalam dataset. Sementara itu, untuk kelas negatif (-), precision dan recall masing-masing sebesar 0.77 dan 0.74, menunjukkan performa yang masih cukup baik namun dengan tingkat kesalahan yang lebih tinggi dibandingkan kelas positif. Kelas netral memperoleh Precision 0.85 dan Recall 0.82, memperlihatkan bahwa model masih mampu menangani kelas dengan distribusi data yang lebih seimbang meskipun terdapat beberapa misclassifikasi.

Dominasi bobot pada model SVM merefleksikan peran utama SVM dalam memberikan prediksi yang lebih presisi, sementara kontribusi NB berfungsi sebagai penguat yang membantu menstabilkan keputusan model secara keseluruhan. Pendekatan Ensemble ini memungkinkan pemanfaatan kekuatan

masing-masing algoritma sehingga tercapai performa klasifikasi sentimen yang lebih baik dibandingkan penggunaan model secara tunggal. Meskipun secara umum performa model menunjukkan hasil yang baik, tingkat kesalahan klasifikasi yang masih ditemukan pada kelas negatif dan netral menunjukkan perlunya penerapan pendekatan lanjutan guna mengatasi permasalahan ketidakseimbangan kelas serta meningkatkan kemampuan diskriminatif model terhadap masing-masing kategori sentimen.

Secara umum, temuan penelitian ini mengindikasikan bahwa penerapan metode Soft Voting dengan pemberian bobot yang telah dioptimalkan mampu memberikan peningkatan kinerja dalam proses klasifikasi sentimen pada dataset yang digunakan, sekaligus memberikan fleksibilitas untuk menggabungkan kekuatan beberapa model dalam satu kerangka Ensemble yang efektif.

4.2.3 Evaluasi Diskusi Beserta State of Art

Tabel 10 Hasil penelitian saat ini

Author	Model	Akurasi	Precision (+)	Recall (+)	F1 score (+)
Workcurrent Author	NB	0.8340	0.8512	0.9776	0.9102
	SVM	0.9289	0.9710	0.9753	0.9732
	SoftVoting SVM+NB	0.9344	0.9687	0.9813	0.9750

Pada Tabel 10 di atas, adalah hasil penelitian saat ini menggunakan Hyperparamter dari model SVM dan NB di kombinasikan dengan model Ensemble Soft Voting menunjukkan hasil akurasi yang masih cukup bagus.

Tabel 11 Hasil penelitian terdahulu

AUTHOR	MODEL	AKURASI	PRECISION (+)	RECALL (+)	F1 SCORE (+)
[27]	NB	0.72	0.70	0.97	0.81
[28]	SVM	0.8444	0.9289	0.9150	0.9219
[29]	HardVoting (SVM+DT+LR)	0.892	0.8808	0.8967	0.8887
	HardVoting(SVM+DT+NB)	0.8987	0.8876	0.9036	0.8955
	HardVoting (DT+LR+NB)	0.8947	0.8871	0.8961	0.8919
	HardVoting(SVM+LR+NB)	0.8947	0.8808	0.9017	0.8911
	NB	0.89	0.8955	0.8858	0.8856
	SVM	0.895	0.8794	0.8975	0.8871
	DT	0.807	0.8154	0.8044	0.8042
	LR	0.8867	0.8774	0.8757	0.8757

Pada Tabel 11 di atas dapat dilihat bahwa hasil penelitian terdahulu tanpa kombinasi model Ensemble khususnya pada metode NB memiliki hasil akurasi yang kurang bagus, yaitu di angka 0.72.

Tabel evaluasi yang disajikan menunjukkan perbandingan kinerja beberapa model klasifikasi yang meliputi Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR), serta beberapa teknik Ensemble seperti Soft Voting dan Hard Voting yang menggabungkan beberapa model. Pengukuran performa model dilakukan menggunakan matrix utama yaitu Akurasi, Presisi (Precision +), (Recall +), dan F1 score (+).

Pada tahap awal evaluasi, model SVM dan Naive Bayes (NB) diuji secara individual serta dalam gabungan menggunakan pendekatan Soft Voting. Pendekatan Soft Voting tersebut menghasilkan performa terbaik dengan F1 score tertinggi sebesar 0.9750, serta menunjukkan keunggulan peningkatan akurasi sebesar 0,535 dalam memadukan kekuatan masing-masing model untuk meningkatkan keseimbangan antara presisi dan recall.

Pada tabel kedua, performa model dari beberapa studi dan dokumen jurnal dibandingkan kembali. Di sini, model Hard Voting yang menggabungkan SVM, Decision Tree, dan Logistic Regression secara konsisten menunjukkan performa superior dengan F1 score sekitar 0.89 hingga 0.90, yang lebih tinggi dibandingkan model individual SVM, NB, DT, dan LR. Hal ini mengindikasikan bahwa metode Ensemble melalui Hard Voting dapat meningkatkan Robustness dan generalisasi model dibandingkan penggunaan satu model tunggal.

Selain itu, diperoleh temuan bahwa model SVM cenderung mempertahankan nilai presisi yang tinggi, sedangkan NB memberikan nilai recall yang relatif tinggi. Dengan demikian, kombinasi model dalam Ensemble menjadi strategi efektif untuk mengatasi trade-off antara presisi dan recall.

5 Kesimpulan

Analisis sentimen ulasan restoran Tong Tji Tea House menunjukkan dominasi sentimen positif yang menyebabkan ketidakseimbangan kelas, menimbulkan bias klasifikasi.

Pemodelan menggunakan Support Vector Machine (SVM) dan Naive Bayes (NB) yang dioptimasi hyperparameternya menunjukkan bahwa SVM dengan kernel linear dan konfigurasi tertentu unggul dalam akurasi dan F1-score. Kombinasi Ensemble Soft Voting antara SVM dan NB menghasilkan meningkatkan 0.535% klasifikasi yang lebih akurat dan Robust dengan F1-weighted mencapai 0,93. Hard Voting Ensemble juga menunjukkan performa yang baik dengan kombinasi SVM, Decision Tree, dan Logistic Regression. Penelitian ini menegaskan pentingnya pemilihan model, optimasi hyperparameter, dan penerapan teknik Ensemble dalam klasifikasi sentimen teks yang memiliki distribusi kelas tidak seimbang. Temuan ini memperkuat bukti peneliti sebelumnya mengatakan bahwa kombinasi beberapa model dalam Ensemble dapat meningkatkan Robustness dan kemampuan generalisasi model, sehingga menghasilkan performa klasifikasi yang lebih stabil dan akurat dalam analisis sentimen. Saran dari peneliti sekarang, disarankan mengembangkan teknik augmentasi data yang lebih adaptif seperti varian SMOTE dan Deep learning, serta mengimplementasikan model berbasis deep learning seperti IndoBERT untuk representasi kontekstual yang lebih baik. Pendekatan Ensemble yang lebih kompleks seperti Stacking dan studi longitudinal juga dianjurkan untuk menguji ketahanan model.

Referensi

- [1] A. Nikseresht, M. H. Raeisi, and H. A. Mohammadi, 'Decision Making for Celebrity Branding: An Opinion Mining Approach based on Polarity and Sentiment Analysis using Twitter Consumer-Generated Content (CGC)', "arxiv, September, 2021. doi: 10.48550/arxiv.2109.12630.
- [2] N. Singh and U. C. Jaiswal, 'Sentiment Analysis using Machine Learning', ADCAIJ, Vol.12, December 2023. doi: 10.14201/adcaij.26785
- [3] F. Rustam, I. Ashraf, A. Mehmood, S. Ullah, and G. S. Choi, 'Tweets Classification on the Base of Sentiments for US Airline Companies', 2019 *Mdpi*. doi: 10.3390/e21111078.
- [4] Y. B. Lasotte, E. J. Garba, Y. M. Malgwi, and M. A. Buhari, 'An Ensemble Machine Learning Approach for Fake News Detection and Classification using a Soft Voting Classifier', 2022. doi: 10.24018/ejece.2022.6.2.409.

- [5] V. Rupapara, F. Rustam, A. Amaar, P. B. Washington, E. Lee, and I. Ashraf, 'Deepfake Tweets Classification using Stacked Bi-LSTM and Words Embedding', 2021. doi: 10.7717/peerj-cs.745.
- [6] A. Supian, B. T. Revaldo, N. Marhadi, L. Efrizoni, and R. Rahmaddeni, 'Perbandingan Kinerja Naïve Bayes dan SVM pada Analisis Sentimen Twitter Ibukota Nusantara', 2024, *Jif*. doi: 10.33884/jif.v12i01.8721.
- [7] A. R. Isnain, N. S. Marga, and D. Alita, 'Sentiment Analysis of Government Policy on Corona Case using Naive Bayes Algorithm', *IJCCS*. 2021, doi: 10.22146/ijccs.60718.
- [8] A. Özçift and M. Bozuyula, 'Evaluation of Ensemble Algorithms and Deep Learning Transformers in Medical Sentiment Prediction', *Dergipark*. 2021, doi: 10.31590/ejosat.1010241.
- [9] Y. Setiawan, J. Jondri, and W. Astuti, 'Twitter Sentiment Analysis on Online Transportation in Indonesia using Ensemble Stacking', *ejurnal.stmik-budidarma*, 2022, doi: 10.30865/mib.v6i3.4359.
- [10] V.R.N. Seileela and N. N. M. Rao, 'Iterative Ensemble Learning Over High Dimensional Data for Sentiment Analysis', *scpe*, 2024, doi: 10.12694/scpe.v25i2.2650.
- [11] H. Leidiyana, 'Ensemble Stacking dalam Analisa Sentimen Reaksi Veteran Militer as terhadap Pengambilalihan Afghanistan Oleh Taliban', *inti*, 2023, doi: 10.33480/inti.v18i1.4175.
- [12] F. A. Indriyani, A. Fauzi, and S. Faisal, 'Analisis sentimen aplikasi Tiktok menggunakan Algoritma Naïve Bayes dan Support Vector Machine', *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, Vol. 10, no. 2, pp. 176–184, Jul. 2023, doi: 10.37373/tekno.v10i2.419.
- [13] N. Agustina dan C. N. ihsan , and N. M. A. Al-Hashedi, 'Ense2 Tampilan Pendekatan Ensemble untuk Analisis Sentimen Covid19 menggunakan Pengklasifikasi Soft Voting', *Applied Computational Intelligence and Soft Computing* , 2023.
- [14] D. Triyana, M. M. Al Haromainy, H. Maulana 'Ensembl3View of Implementasi Metode Ensemble Majority Vote pada Algoritma Naïve Bayes dan Random Forest untuk Analisis Sentimen Twitter Harga Tiket Pesawat Domestik', *Jurnal Mahasiswa teknik Iformatika*, Vol. 8, 2024.
- [15] Susi, 'Analisis Sentimen Cuitan Twitter Pada Proses Seleksi CPNS menggunakan Algoritma Ensembles Vote' *IKRAM: Jurnal Ilmu Komputer Al Muslim*, Vol. 1, No. 1, 2022.
- [16] A. Onan, 'View of Ensemble of Classifiers and Term Weighting Schemes for Sentiment Analysis in Turkish', *Scientific Research Communications*, Vol. 1, No. 1, 2021.
- [17] Y. A. Mustofa, I. Surya, and K. Idris, 'Pendekatan Ensemble pada Analisis Sentimen Ulasan Aplikasi Google Play Store Ensemble Approach to Sentiment Analysis of Google Play Store App Reviews', *Jambura Journal of Electrical and Electronics Engineering*, Vol. 6, 2024.
- [18] B. S. Amalia, Y. Umaidah, R. Mayasari, S. Karawang Jl HSRonggo Waluyo, K. Telukjambe Timur, and K. Karawang, 'Analisis Sentimen Riview Pelanggan Restoran menggunakan Algoritma Supprt Vector Machine dan K-Nearest Neighbor', *Jurnal Sains Teknologi dan Industri*, Vol. 19, No. 1, pp. 28–34, 2021.
- [19] O. Fanny and H. Suroyo, 'Analisis Sentimen Pengguna Media Sosial terhadap Omnibus Law berdasarkan Hashtag di Twitter Analysis of Social Media Users Sentiments Against Omnibus Law Based on Hashtags on Twitter', *SISTEMASI: Jurnal Sistem Informasi* 2022. DOI: <https://doi.org/10.32520/stmsi.v11i1.1685>.
- [20] M. Fariz Firdaus, D. E. Ratnawati, N. Y. Setiawan, and P. Korespondensi, 'Analisis Sentimen berbasis Aspek pada Ulasan Pelanggan Restoran menggunakan Algoritma Support Vector Machine (Studi Kasus: Depot Bamara)', *Jurnal Teknologi informasi dan ilmu komputer*, Vol. 11, No. 6, pp. 1265–1272, 2024, doi: 10.25126/jtiik.2024117564.
- [21] P. Gracezando Yehova and M. Ihsan Jambak, 'Perbandingan Metode Naïve Bayes dan SVM pada Ulasan Google Playstore Mobile Legends Bang Bang' , *Jurnal Device*, Vol. 14, No. 1, pp. 140–149, 2024.
- [22] S. Riyadiiban and S. Riyadi, 'Analisis Sentimen Opini Masyarakat terhadap Stadion Jakarta Internasional Stadium (JIS) pada Twitter ... Analisis Sentimen Opini Masyarakat terhadap <http://sistemasi.ftik.unisi.ac.id>

- Stadion Jakarta Internasional Stadium (JIS) pada Twitter dengan Perbandingan Metode *Naive Bayes* dan *Support Vector Machine*', *Jurnal Sains dan Teknologi*, Vol. 5, No. 3, p. 2024, 2024, doi: 10.55338/saintek.v5i3.2962.
- [23] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, 'Perbandingan Metode *Naive Bayes* dan *Support Vector Machine* pada Analisis Sentimen *Twitter*', *SMATIKA Jurnal*, 2024.
- [24] S. N. Fadhilah and F. S. Utomo, 'Algoritma *Naive Bayes* untuk Analisis Sentiment *Review Blibli.com* di *Google Play Store* *Naive Bayes Algorithm for Sentiment Analysis of Blibli.com Review on Google Play Store*', *SISTEMASI: Jurnal Sistem Informasi* 2021. DOI: <https://doi.org/10.32520/stmsi.v13i2.3887>
- [25] L. B. Ilmawan and M. A. Mude, 'Perbandingan Metode Klasifikasi *Support Vector Machine* dan *Naive Bayes* untuk Analisis Sentimen pada Ulasan Tekstual di *Google Play Store*', *ILKOM Jurnal Ilmiah*, Vol. 12, No. 2, pp. 154–161, Aug. 2020, doi: 10.33096/ilkom.v12i2.597.154-161.
- [26] N. P. G. Naraswati, D. C. Rosmilda, D. Desinta, F. Khairi, R. Damaiyanti, and R. Nooraeni, 'Analisis Sentimen Publik dari *Twitter* tentang Kebijakan Penanganan Covid-19 di Indonesia dengan *Naive Bayes Classification*', *SISTEMASI: Jurnal Sistem Informasi*, 2021. <https://doi.org/10.32520/stmsi.v10i1.1179>
- [27] A. F. Alkindi and N. Nasution, 'Analisis Sentimen Ulasan Penggunaan pada Game *Roblox* dengan Metode *Support Vector Machine* dan *Naive Bayes*', *J-Com (Journal of Computer)*, Vol. 4, No. 2, pp. 164–177, Jul. 2024, doi: 10.33330/j-com.v4i2.3319.
- [28] Salahuddin, I. Sabila, 'Analisis dan Implementasi Sistem Penilaian Ulasan dengan Teknik *Sentiment Analysis* berbasis *Machine Learning* untuk Peningkatan *Feedback* Pemilik Cafe," *Infomedia*, Vol.9 No.2,2024. DOI: <http://dx.doi.org/10.30811/jim.v9i2.5980>
- [29] R. A. M. San Ahmed, 'Hard Voting Approach using *SVM*, *Naive Bays* and *Decision Tree* for *Kurdish Fake News Detection*', *Iraqi Journal for Computer Science and Mathematics*, Vol. 4, No. 3, pp. 25–33, 2023, doi: 10.52866/ijcsm.2023.02.03.003.