

Klasifikasi Emosi pada Teks Berbahasa Inggris Menggunakan Pendekatan *Ensemble Bagging*

Erfian Junianto^{1,2}, Mila Puspitasari^{1,3}, Salman Ilyas Zakaria^{1,3}, Toni Arifin^{1,2}, Ignatius Wiseto Prasetyo Agung^{1,2}

¹ARS Digital Research and Innovation (ADRI), Universitas Adhirajasa Reswara Sanjaya, Bandung, Jawa Barat, 40282, Indonesia

²Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Adhirajasa Reswara Sanjaya, Bandung, Jawa Barat, 40282, Indonesia

³Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Adhirajasa Reswara Sanjaya, Bandung, Jawa Barat, 40282, Indonesia

[Diserahkan: 14 Agustus 2024, Direvisi: 30 Oktober 2024, Diterima: 19 November 2024]
Penulis Korespondensi: Mila Puspitasari (email: mileuups14@gmail.com)

INTISARI — Penelitian ini menyoroti pentingnya klasifikasi emosi dalam teks berbahasa Inggris, khususnya dalam konteks interaksi manusia di media sosial yang sering melibatkan data tidak terstruktur. Emosi memainkan peran krusial dalam komunikasi dan pemahaman yang lebih baik terhadap emosi ini dapat membantu dalam analisis perilaku pengguna. Tujuan utama dari penelitian ini yaitu untuk meningkatkan akurasi, *recall*, *precision*, dan *f1-score* dalam klasifikasi emosi dengan menerapkan pendekatan *ensemble bagging* yang menggabungkan algoritma *naïve Bayes*, *logistic regression*, dan *k-nearest neighbors* (KNN). Metodologi yang diterapkan mencakup pengumpulan data dari berbagai sumber, diikuti dengan pembersihan dan analisis data menggunakan teknik penambangan teks dan pembelajaran mesin. Data yang diperoleh kemudian dianalisis untuk mendeteksi emosi, seperti marah, bahagia, sedih, terkejut, malu, jijik, dan takut. Evaluasi kinerja dilakukan dengan membandingkan hasil metode *ensemble bagging* dengan algoritma individual untuk mengukur efektivitasnya. Hasil penelitian menunjukkan bahwa metode *logistic regression* mencapai akurasi tertinggi sebesar 98,76%, diikuti oleh *naïve Bayes* dan KNN. Metode *ensemble* ini berhasil mengatasi keterbatasan masing-masing algoritma individu dengan meningkatkan stabilitas dan keandalan klasifikasi secara keseluruhan. Temuan ini memberikan wawasan berharga tentang teknik analisis emosi berbasis teks dan menunjukkan potensi metode *ensemble* untuk meningkatkan akurasi klasifikasi. Penelitian ini juga merekomendasikan arah penelitian di masa depan untuk mengeksplorasi teknik *ensemble* tambahan dan mengoptimalkan kompleksitas model guna mencapai peningkatan kinerja analisis emosi pada berbagai *dataset* yang lebih luas.

KATA KUNCI — Penambangan Teks, *Bagging*, Klasifikasi Emosi, Pembelajaran Mesin.

I. PENDAHULUAN

Kehidupan manusia selalu terkait erat dengan emosi. Emosi merupakan respons yang muncul sebagai reaksi terhadap orang lain atau suatu peristiwa dan berperan penting dalam aktivitas pada umumnya. Emosi terbagi menjadi dua jenis utama, yaitu emosi positif dan negatif, yang mencakup berbagai kategori, seperti bahagia (*happy*), marah (*anger*), takut (*fear*), dan sedih (*sad*). Studi tentang emosi penting dalam berbagai bidang, seperti psikologi, ilmu kognitif, dan analisis media sosial, karena emosi membantu memahami perilaku dan interaksi sosial manusia [1].

Seseorang dapat dengan mudah mengungkapkan emosinya secara verbal melalui tulisan di media sosial. Emosi tidak hanya merupakan ekspresi individual, tetapi juga pendorong bagi proses kognitif dan pembentukan strategi. Emosi mencakup berbagai jenis bahasa yang umumnya diucapkan dan dipahami [2]. Selain itu, emosi juga memiliki dampak signifikan terhadap ingatan, hubungan sosial, dan bahkan pengambilan keputusan [3].

Identifikasi emosi dapat dilakukan dengan dua cara, yaitu secara verbal dan nonverbal. Emosi verbal disampaikan melalui ucapan atau tulisan, sementara emosi nonverbal diekspresikan melalui bahasa tubuh, seperti ekspresi wajah, gerakan, dan tindakan tangan atau kaki. Tanpa ekspresi wajah atau variasi suara, mendeteksi emosi dalam percakapan teks akan menjadi lebih sulit [4]. Oleh karena itu, diperlukan metode khusus untuk mengidentifikasi emosi dalam dialog berbasis teks. Klasifikasi emosi merupakan proses pengelompokan atau pemetaan dokumen ke dalam kategori emosi yang telah ditentukan sebelumnya [5].

Untuk mengklasifikasikan emosi dalam teks dengan data yang tidak terstruktur, analisis teks perlu dilakukan menggunakan penambangan teks (*text mining*) [6]-[10]. Tujuan penambangan teks adalah untuk mengidentifikasi dan mengeksplorasi pola menarik dalam sumber data, yang terdiri atas sekumpulan dokumen. Pola-pola ini sering kali tidak ditemukan dalam catatan basis data konvensional, tetapi sangat relevan dalam konteks data teks yang tidak terstruktur. Dengan demikian, penambangan teks memainkan peran krusial dalam mengungkap wawasan berharga dari teks yang luas dan beragam [11].

Penambangan teks sering digunakan pada penelitian dalam analisis sentimen [12]-[15] dan klasifikasi *dataset* emosi berbentuk teks [16], [17]. Analisis sentimen merupakan studi pengolahan data terhadap perasaan, sentimen, dan emosi yang terungkap dalam teks [18]. Untuk menganalisis sentimen, terdapat berbagai algoritma yang dapat digunakan, seperti *naïve Bayes* (NB), *logistic regression* (LR), dan *k-nearest neighbors* (KNN) [19]. Ciri utama algoritma-algoritma ini adalah asumsi naif terhadap ketergantungan dari setiap kondisi atau kejadian yang diamati [20]-[25].

Beberapa penelitian sebelumnya telah banyak menggunakan algoritma NB, LR, dan KNN dalam klasifikasi [26]. Meskipun akurasi yang diperoleh dari ketiga algoritma tersebut cukup tinggi, masih ada ruang untuk peningkatan karena KNN rentan terhadap distribusi data yang tidak merata dan adanya *outlier*, NB rentan terhadap asumsi independensi fitur yang sering tidak realistis, dan LR cenderung mengalami *overfitting* pada *dataset* kecil yang mengandung derau. Semua faktor ini dapat memengaruhi akurasi dan generalisasi model

[27]. Oleh karena itu, penelitian ini berfokus pada optimasi *ensemble bagging* yang menggabungkan ketiga algoritma tersebut untuk meningkatkan kinerja klasifikasi.

Penelitian tentang penambahan teks dalam klasifikasi emosi telah banyak melibatkan kombinasi teknik *ensemble bagging* dengan metode NB, LR, dan KNN. Dari berbagai penelitian ini, para peneliti mencatat tingkat akurasi yang bervariasi, antara 46% hingga 77%. Sebagai contoh, sebuah penelitian melakukan klasifikasi jenis vaksin COVID-19 menggunakan metode KNN, yang berhasil mencapai akurasi sebesar 46,20% [28].

Selanjutnya, penelitian mengenai persepsi pengguna Twitter terkait kasus COVID-19 menggunakan metode LR dan mencapai akurasi sebesar 77% [29]. Studi lain tentang identifikasi emosi di Twitter menggunakan metode NB dan penggabungan fitur, yang menunjukkan akurasi 50,55% [30]. Hasil akurasi tersebut dipengaruhi oleh volume data yang besar dan variasi label emosi yang digunakan, termasuk tujuh kelas emosi, seperti marah, bahagia, sedih, takut, jijik (*disgust*), malu (*shame*), dan terkejut (*surprise*), yang digunakan dalam penelitian ini.

Penelitian lain mengenai klasifikasi calon anggota menggunakan metode *bagging* berbasis NB berhasil mencapai nilai akurasi lebih tinggi, yakni 67,92%, dibandingkan dengan penggunaan algoritma NB tunggal yang hanya mencapai akurasi 62,26% [31]. Metode *bagging* menggunakan NB juga mencatat nilai *recall* atau sensitivitas sebesar 0,76, presisi sebesar 0,74, dan *f1-score* sebesar 0,75. Di sisi lain, algoritma NB tunggal mencatat *recall* sebesar 0,65, presisi sebesar 0,85, dan *f1-score* sebesar 0,74. Temuan ini menunjukkan bahwa penggunaan metode *bagging* dapat meningkatkan kinerja klasifikasi secara signifikan dalam konteks studi tersebut.

Penelitian selanjutnya berfokus pada penambahan teks di media sosial untuk mendeteksi emosi pengguna [11]. Pengujian dilakukan menggunakan metode *support vector machine* (SVM) dan KNN, yang menghasilkan rata-rata nilai presisi sebesar 0,4564 dan *recall* sebesar 0,502, dengan akurasi mencapai 0,8104 untuk SVM. Sementara itu, untuk KNN diperoleh rata-rata nilai presisi 0,3421, *recall* sebesar 0,4595, dan akurasi mencapai 0,797. Studi ini memberikan kontribusi penting dalam pengembangan teknik analisis emosi berbasis teks di platform media sosial.

Penelitian lainnya berfokus pada perbandingan metode LR dan *random forest* dalam studi kasus klasifikasi emosi cuitan (*tweet*) [32]. Hasil dari perbandingan kedua metode menunjukkan bahwa LR mencapai akurasi sebesar 78,22%, sedangkan *random forest* mencapai akurasi sebesar 72,41%. Studi ini memberikan wawasan tentang efektivitas dan kinerja relatif kedua pendekatan dalam konteks pengolahan dan analisis teks.

Teknik *ensemble bagging* (*bootstrap aggregating*) merupakan strategi optimasi yang mengintegrasikan multipel model pembelajaran mesin (*machine learning*) menjadi satu entitas. Dalam pendekatan ini, setiap model dilatih secara terpisah dan hasilnya digabungkan setelah proses pembelajaran individu selesai. Penelitian juga mengevaluasi peningkatan kinerja antara *ensemble* dan *non-ensemble* dengan mempertimbangkan skor akurasi, presisi, *recall*, dan *f1-score* sebagai metrik utama [27], [31], [33].

Penelitian ini bertujuan meningkatkan akurasi, *recall*, presisi, dan *f1-score* klasifikasi emosi menggunakan pendekatan *ensemble bagging* yang menggabungkan algoritma NB, LR, dan KNN. Penelitian ini mengevaluasi efektivitas

kombinasi berbagai pendekatan tersebut dalam mengatasi keterbatasan penelitian sebelumnya, terutama dalam hal ketepatan dan keandalan prediksi. Dengan menggunakan metode ini, diharapkan hasil klasifikasi akan lebih akurat dan dapat diandalkan serta mampu memberikan solusi yang lebih efektif dalam analisis data yang kompleks.

II. KLASIFIKASI EMOSI

Pada bagian ini dibahas klasifikasi emosi, *ensemble bagging*, penambahan teks, pembelajaran mesin, NB, KNN, serta LR.

A. KLASIFIKASI

Klasifikasi adalah proses penempatan suatu objek ke dalam label atau kategori yang telah ditentukan, sering kali menggunakan model seperti peraturan klasifikasi (*IF-THEN*), pohon keputusan, formula matematika, dan jaringan saraf. Algoritma klasifikasi dimulai dengan membangun aturan klasifikasi berdasarkan data yang ada, diikuti dengan tahap pembelajaran yang melibatkan proses pengolahan data untuk meningkatkan akurasi model. Setelah itu, dilakukan tahap pengujian menggunakan data uji untuk mengevaluasi dan memvalidasi kinerja algoritma dalam memprediksi atau mengklasifikasikan data yang baru [34].

Pembelajaran dan pengelompokan adalah dua tahap dalam proses klasifikasi. Pada tahap pembelajaran, algoritma klasifikasi menganalisis data untuk membuat model. Model ini kemudian digunakan untuk mengelompokkan data pada tahap berikutnya. Proses menemukan model yang dapat menggambarkan dan membedakan kelas data disebut klasifikasi. Tujuan utamanya adalah agar model tersebut bisa memprediksi kelas objek yang belum diketahui [35].

B. EMOSI

Emosi dapat didefinisikan sebagai kondisi yang biasanya dipicu oleh suatu peristiwa yang signifikan bagi seseorang. Emosi mencakup aspek mental, baik dari sisi keinginan maupun pemikiran. Desain keadaan, proses, dan model komputasi harus disederhanakan untuk menangani kerumitan emosi manusia dalam kehidupan nyata [36].

Tidak ada emosi positif atau negatif karena, pada prinsipnya, emosi adalah reaksi manusia terhadap berbagai situasi. Pada dasarnya, manusia yang memiliki kecerdasan emosi mampu mengidentifikasi, mengenali, dan menilai kualitas diri sebagai individu. Kecerdasan emosi adalah kemampuan mental untuk menilai, merancang, dan menyelesaikan masalah, baik yang sederhana maupun rumit, serta membuat keputusan yang bijaksana [37].

C. ENSEMBLE BAGGING

Metaalgoritma yang dimaksudkan untuk meningkatkan kinerja algoritma pembelajaran mesin dikenal sebagai *bagging* atau *bootstrap aggregating*. Tahap pertama melibatkan *resampling* data latih untuk membuat *sub-dataset* dari *dataset* utama, kemudian tahap kedua melibatkan menggabungkan berbagai nilai prediksi menjadi satu nilai prediksi akhir. Dengan metode ini, variasi dan *overfitting* dapat dikurangi, sehingga menghasilkan model yang lebih stabil dan akurat [27].

Bagging adalah teknik yang diperkenalkan oleh Leo Breiman, yang bertujuan untuk mengurangi variasi dari variabel independen dengan menggunakan pendekatan kelompok. Tujuannya adalah meningkatkan ketepatan prediksi klasifikasi dari model tunggal berbasis pohon. Metode ini disebut *aggregating bootstrap*. Untuk membuat *dataset* baru, metode *bagging* dasar menggunakan sampel acak dengan

penggantian dari *dataset* asli. *Dataset* baru memiliki ukuran yang sama dengan *dataset* instruksi, sedangkan sampel acak diambil dengan pengembalian dari *dataset* instruksi, juga dikenal sebagai *sample bootstrap*. Berikutnya, data baru digunakan untuk menghasilkan pohon klasifikasi dalam beberapa iterasi. Hasil dari setiap iterasi digabungkan untuk menghasilkan prediksi akhir [38].

D. PENAMBANGAN TEKS

Penambahan teks adalah proses yang memanfaatkan dokumen dan berbagai alat analisis untuk mendapatkan pemahaman yang mendalam. Penambahan teks dapat menggunakan data untuk mengekstrak informasi berguna, menemukan, dan mengklasifikasikan data teks, seperti dalam penelitian ini, untuk mengeksplorasi pola yang menarik [17]. Pada dasarnya, mekanisme kerja algoritma penambahan teks mirip dengan mekanisme algoritma penambahan data (*data mining*). Penambahan teks dan penambahan data berbeda karena penambahan teks menggunakan data yang tidak terstruktur, sedangkan penambahan data menggunakan data yang terstruktur. Kegunaan penambahan teks dapat membantu mengatasi masalah klasifikasi, pengklasteran (*clustering*), dan prediksi data tekstual [39].

Penerapan algoritma penambahan teks dilakukan setelah data teks dimodifikasi ke dalam format yang lebih optimal. Metode yang sering digunakan meliputi klasifikasi untuk mengelompokkan teks ke dalam kategori yang sudah ditentukan serta ekstraksi informasi untuk menemukan dan mengambil data penting, seperti nama, hubungan, dan peristiwa. Selain itu, analisis sentimen digunakan untuk mengidentifikasi pandangan atau perasaan dalam teks: positif, negatif, atau netral. Evaluasi dan validasi dilakukan dengan teknik seperti presisi, *recall*, dan *f1-score*, yang digunakan untuk menguji kinerja model dengan data uji guna memastikan hasil yang akurat dan dapat dipercaya [40].

E. PEMBELAJARAN MESIN

Penelitian mengenai algoritma yang mempelajari cara melakukan tugas-tugas spesifik yang biasanya dilakukan secara otomatis oleh manusia dikenal sebagai pembelajaran mesin. Pembelajaran adalah kemampuan menyelesaikan tugas yang sudah pernah dilakukan atau menarik kesimpulan dari pola yang sudah diamati, termasuk kemampuan memproses informasi baru. Fokus utama dari pembelajaran mesin adalah pengembangan algoritma yang memungkinkan sistem belajar secara otonom tanpa perlu campur tangan manusia [41].

Pembelajaran mesin dibagi menjadi dua konsep pembelajaran yang utama. Konsep pertama adalah pembelajaran terawasi (*supervised learning*), yang menggunakan teknik untuk membangun fungsi berdasarkan data latih yang ada. Beberapa contoh algoritma dari konsep pembelajaran terawasi adalah NB, KNN, dan LR. Kedua, pembelajaran tak terawasi (*unsupervised learning*) adalah teknik pembelajaran mesin yang berusaha mengidentifikasi pola dari data latih tanpa adanya klasifikasi pada data masukan tersebut. Berbeda dengan pembelajaran terawasi, teknik ini tidak menggunakan data yang sudah diklasifikasikan sebagai panduan. Salah satu contoh algoritma dalam pembelajaran tak terawasi adalah pengklasteran [42].

F. NAÏVE BAYES

Klasifikasi Bayes adalah metode statistik yang mampu memproyeksikan kelas sebuah anggota dengan mempertimbangkan probabilitasnya. *Naïve Bayes* adalah

metode klasifikasi yang mengasumsikan bahwa setiap atribut dalam suatu kelas tidak saling memengaruhi [34].

Salah satu kelebihan metode klasifikasi *naïve Bayes* adalah bahwa untuk melakukan proses klasifikasi, diperlukan data pelatihan terlebih dahulu. Hal ini disebabkan oleh kebutuhan akan data latih yang relatif sedikit untuk menghitung parameter rata-rata dan varians yang diperlukan dalam proses klasifikasi. Selama proses pelatihan, dokumen dikategorikan (data latih) dan kemudian diproses untuk menghasilkan pengetahuan berupa nilai probabilitas pada setiap kata. Proses ini menghasilkan sebuah kata untuk setiap dokumen yang mengidentifikasi ke dalam kategori tertentu. Metode *naïve Bayes* cocok digunakan dalam konteks ini karena kemampuannya untuk menangani klasifikasi dengan data yang relatif sedikit dan memberikan hasil yang dapat diandalkan berdasarkan probabilitas masing-masing kategori [43].

G. K-NEAREST NEIGHBORS (KNN)

Metode KNN melakukan klasifikasi dengan cara membandingkan objek baru dengan data pelatihan yang paling mirip, berdasarkan nilai k [44]. Dalam pengertian lain, metode ini menggunakan algoritma terawasi, dengan hasil klasifikasi untuk objek baru ditentukan oleh mayoritas tetangga terdekatnya dalam himpunan pelatihan [45].

Kelebihan KNN yaitu tahan terhadap data pelatihan yang memiliki derau dan efektif di lingkungan dengan data pelatihan yang besar. Selain itu, KNN juga relatif sederhana untuk diimplementasikan dan dapat digunakan untuk masalah klasifikasi maupun regresi. Sementara itu, kelemahan KNN termasuk kewajiban untuk menentukan nilai dari parameter k , atau jumlah tetangga terdekat, yang dapat memengaruhi kinerja model. Selain itu, pendekatan berbasis jarak sering kali tidak jelas, yaitu mengenai jenis jarak yang paling tepat digunakan serta atribut yang harus dipilih untuk memberikan hasil terbaik. Selain itu, biaya komputasi KNN cukup tinggi karena memerlukan perhitungan jarak antara setiap *instance* baru dengan seluruh sampel pelatihan [46].

H. LOGISTIC REGRESSION

Salah satu algoritma pembelajaran mesin untuk klasifikasi adalah LR, yang digunakan untuk memprediksi kemungkinan variabel dependen kategoris. Variabel biner yang mengandung data berkode 1 atau 0 disebut variabel dependen dalam analisis LR. Metode ini adalah teknik regresi linier umum yang digunakan untuk mempelajari pemetaan dari sejumlah variabel numerik ke variabel biner atau probabilistik. LR sering digunakan karena kemampuannya dalam memberikan interpretasi yang jelas tentang hubungan antara variabel independen dan variabel dependen serta dalam mengatasi masalah klasifikasi dengan baik [47].

LR telah banyak digunakan dalam berbagai masalah analisis sentimen, seperti memprediksi hasil positif, negatif, atau netral [48]. LR sesuai untuk memprediksi ketika variabel data keluaran hanya memiliki dua nilai, yaitu biner. Namun, LR multinomial lebih baik digunakan untuk memprediksi data dengan lebih dari dua kemungkinan hasil [29].

Analisis LR ordinal digunakan untuk melihat hubungan variabel respons dan variabel prediktor. Model logistik digunakan untuk LR ordinal, yang memiliki skala ordinal dan variabel responsnya bersifat *polychotomous* [49].

LR sederhana ditunjukkan pada (1) [50], sedangkan bentuk umum dari LR biner berganda dituliskan pada (2). Sementara itu, untuk mencari probabilitas dari LR biner berganda, digunakan (3).

$$\ln \left[\frac{p(y=1)}{1-p(y=1)} \right] = \beta_0 + \beta x \quad (1)$$

$$\ln \left[\frac{p(y=1)}{1-p(y=1)} \right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (2)$$

$$p(y = 1) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}} \quad (3)$$

III. METODOLOGI

Gambar 1 merupakan alur atau tahapan model klasifikasi emosi menggunakan pendekatan *ensemble bagging*. Gambar 1 memvisualisasikan langkah-langkah proses dari masukan data teks hingga hasil klasifikasi emosi.

A. PENGUMPULAN DATA

Pengumpulan data adalah tahap pertama dari penelitian ini. Data yang dikumpulkan diperoleh dari penelitian sebelumnya [51]. Data tersebut telah dikategorikan ke dalam tujuh kelas emosi, yaitu bahagia, sedih, terkejut, jijik, marah, takut, dan malu. Secara keseluruhan, data yang digunakan terdiri atas 2.414 teks yang menggambarkan ketujuh emosi tersebut.

Tabel I menampilkan distribusi jumlah *dataset* berdasarkan emosi yang diklasifikasikan [32]. Setiap baris di dalam Tabel I merepresentasikan satu jenis emosi yang diidentifikasi dalam teks, yaitu bahagia, sedih, terkejut, jijik, marah, takut, dan malu. Angka-angka yang tercantum menunjukkan banyaknya sampel data yang dikumpulkan untuk masing-masing kategori emosi. Hal ini penting untuk menunjukkan keseimbangan atau ketidakseimbangan data, yang dapat memengaruhi hasil akhir proses klasifikasi.

Tabel II menampilkan sampel atau distribusi data emosi yang digunakan dalam penelitian. Data emosi dikategorikan ke dalam tujuh kelas, yaitu bahagia, sedih, marah, malu, takut, jijik, dan terkejut. Tabel II membantu untuk memahami proporsi masing-masing emosi dalam *dataset* yang digunakan, yang penting untuk analisis dan interpretasi hasil klasifikasi.

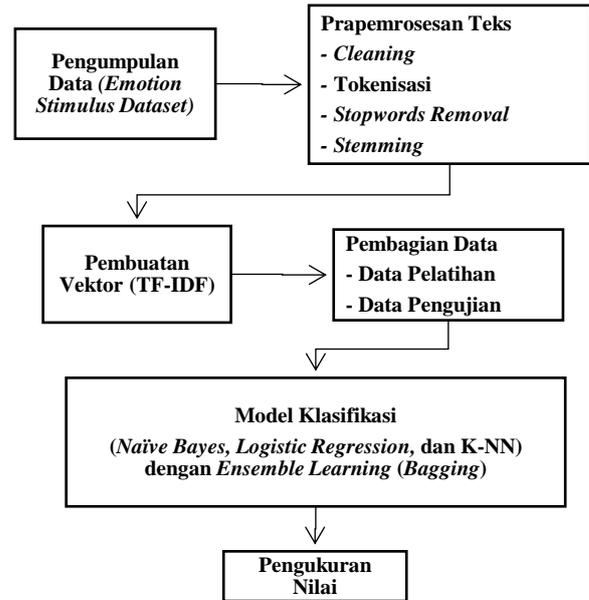
B. PRAPEMROSESAN TEKS

Prapemrosesan teks adalah langkah pertama setelah proses analisis data. Pada tahap ini, data mentah diproses dan disiapkan untuk digunakan dalam proses analisis selanjutnya. Proses pengolahan data ini meliputi berbagai teknik, seperti *cleaning*, tokenisasi, *stopword removal*, dan *stemming*. Langkah-langkah ini bertujuan untuk memastikan bahwa data dalam kondisi optimal sebelum digunakan dalam model pembelajaran mesin [52].

1) CLEANING

Cleaning adalah bagian penting dari prapemrosesan data teks, yang bertujuan untuk menghilangkan derau dan membuat teks lebih konsisten untuk analisis lebih mendalam. Derau adalah data atau informasi yang tidak relevan atau yang dapat mengganggu proses analisis. Pada tahap ini, atribut yang tidak relevan dengan data dibersihkan, seperti karakter atau simbol (!@#\$%^&*():{ } . , ? ~ / []), *hashtag*, URL, *mention*, dan emotikon [45].

Gambar 2 mendefinisikan fungsi `clean_text` untuk membersihkan teks dalam sebuah *DataFrame* dengan cara mengubah semua huruf menjadi huruf kecil dan menghapus tanda baca serta angka menggunakan ekspresi biasa. Fungsi ini kemudian diterapkan ke kolom "teks" dalam *DataFrame* "df", menghasilkan kolom baru "cleaned" berisi teks yang telah dibersihkan. Hasil *cleaning* ditampilkan dengan menggunakan `head` untuk mencetak lima baris pertama dari kolom yang telah dibersihkan. Misalnya, terdapat kalimat sebagai berikut:



Gambar 1. Metode penelitian.

TABEL I
JUMLAH DATASET

Kelas	Jumlah
Bahagia	479
Sedih	575
Terkejut	213
Jijik	95
Marah	483
Takut	423
Malu	146

The ANC, while expressing delight at their decision, nevertheless continued to call for the dissolution of the tricameral parliament, which it regarded as essentially racist diubah menjadi:

The ANC while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist.

2) TOKENISASI

Proses tokenisasi adalah pembagian teks menjadi kalimat, kata, atau token yang memiliki fitur seperti jenis kapitalisasi, kehadiran angka, tanda baca, karakter khusus, dan sejenisnya. Pada titik ini, *string* masukan dipotong atau dipisahkan sesuai dengan kata yang menyusunnya [29].

Tokenisasi juga merupakan proses pemisahan kalimat dalam dokumen menjadi kata-kata individual dengan menggunakan spasi sebagai satu-satunya pemisah. Namun, dalam bahasa Indonesia dan banyak bahasa lainnya, ada kata-kata yang tidak boleh dipisahkan karena pemisahan tersebut dapat mengubah atau bahkan menghilangkan maknanya [53], [54].

Gambar 3 menunjukkan *source code* untuk tokenisasi teks yang sudah dibersihkan dalam sebuah *DataFrame* menggunakan modul `nlTK`. Jika belum tersedia, kode akan mengunduh paket `punkt`. Selanjutnya, fungsi `tokenize_text` digunakan, yang memanfaatkan `word_tokenize` untuk memecah teks menjadi token. Fungsi ini diterapkan ke kolom "cleaned" dalam *DataFrame* "df" untuk membuat kolom baru bernama "tokens". Sebagai hasilnya, teks yang telah dipecah

```
# Lowercase dan hapus tanda baca serta angka
def clean_text(text):
    text = text.lower()
    return re.sub(r'^a-zA-Z\s+', '', text)

df['cleaned'] = df['Text'].apply(clean_text)
print("Cleaned Text:\n", df['cleaned'].head(), "\n")
```

Cleaned Text:

```
0    i suppose i am happy being so tiny it means ...
1    lennox has always truly wanted to fight for th...
2    he was a professional musician now still sens...
3    holmes is happy having the freedom of the hous...
4    i had problems with tutors trying to encourage...
Name: cleaned, dtype: object
```

Gambar 2. Source code proses cleaning dan keluarannya.

TABEL II
SAMPEL DATA TEKS

Kelas	Teks
Bahagia	<i>Lennox has always truly wanted to fight for the world title and was happy taking the tough route</i>
Sedih	<i>She was heartbroken to hear the voice of Camilla</i>
Terkejut	<i>I was astonished at the smallness of the area which these enclosed</i>
Jijik	<i>The cave-man got fed-up with walking</i>
Marah	<i>I felt a stifled anger against her</i>
Takut	<i>I am horrified that somebody would do such a thing</i>
Malu	<i>But I 'm not ashamed of my body</i>

menjadi kata-kata individu akan ditampilkan dalam lima baris pertama kolom “tokens”. Misalnya, kalimat berikut:

The ANC while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist diubah menjadi:

the anc while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist.

3) STOPWORD REMOVAL

Stopword removal adalah proses menghilangkan kata-kata dari dokumen yang dianggap tidak bermakna, tidak mengandung kata emosi, atau tidak memperkuat kata emosi. Proses ini dilakukan dengan cara membandingkan kata-kata dengan kamus *stopword removal* yang sudah ada. Metode ini juga melibatkan penghapusan konjungsi atau kata berimbuhan, seperti “at”, “while”, atau “to”, sehingga hanya kata-kata penting yang tersisa. Selain itu, ada langkah untuk mengoreksi kata-kata dengan kesalahan ejaan dan kata-kata pendek tertentu agar sesuai dengan format data yang digunakan [17], [52]

Stopword removal merupakan salah satu langkah prapemrosesan yang sering digunakan oleh berbagai aplikasi. Tujuannya adalah untuk menghilangkan kata-kata yang sering muncul di setiap dokumen dalam korpus. Awalnya, kata ganti dan awalan termasuk dalam kategori *stopword*. Dalam beberapa fungsi *natural language processing* (NLP) seperti pencarian informasi dan klasifikasi, kata-kata ini tidak signifikan, sehingga antara dokumen dengan lainnya sulit dibedakan [55].

Gambar 4 menunjukkan cara menghilangkan *stopword* dari token dalam sebuah *DataFrame* menggunakan “*nlTK*”. Daftar *stopword* dalam bahasa Inggris dimulai setelah mengimpor “*stopwords*” dari “*nlTK.corpus*” dan mengunduh paket “*stopwords*” jika belum tersedia. Selanjutnya, fungsi “*remove_stopwords*” digunakan untuk menghapus *stopword*

```
import nltk
from nltk.tokenize import word_tokenize

nltk.download('punkt')

# Tokenize teks yang sudah dibersihkan
def tokenize_text(text):
    return word_tokenize(text)

df['tokens'] = df['cleaned'].apply(tokenize_text)
print("Tokens:\n", df['tokens'].head(), "\n")
```

[nlTK_data] Downloading package punkt to /root/nltk_data...
[nlTK_data] Unzipping tokenizers/punkt.zip.
Tokens:
0 [i, suppose, i, am, happy, being, so, tiny, it...
1 [lennox, has, always, truly, wanted, to, fight...
2 [he, was, a, professional, musician, now, stil...
3 [holmes, is, happy, having, the, freedom, of, ...
4 [i, had, problems, with, tutors, trying, to, e...
Name: tokens, dtype: object

Gambar 3. Source code proses tokenizer dan keluarannya.

```
nltk.download('stopwords')

# Inisialisasi stopwords
stop_words = set(stopwords.words('english'))

# Hapus stopwords dari token
def remove_stopwords(tokens):
    return [word for word in tokens if word not in stop_words]

df['tokens_no_stopwords'] = df['tokens'].apply(remove_stopwords)
print("Tokens without Stopwords:\n", df['tokens_no_stopwords'].head(), "\n")
```

Tokens without Stopwords:
0 [suppose, happy, tiny, means, able, surprise, ...
1 [lennox, always, truly, wanted, fight, world, ...
2 [professional, musician, still, sensitive, hap...
3 [holmes, happy, freedom, house]
4 [problems, tutors, trying, encourage, diversit...
Name: tokens_no_stopwords, dtype: object

[nlTK_data] Downloading package stopwords to /root/nltk_data...
[nlTK_data] Unzipping corpora/stopwords.zip.

Gambar 4. Source code proses stopword removal dan keluarannya.

dari daftar token dan diterapkan ke kolom “tokens” dalam *DataFrame* “df”, yang menghasilkan kolom baru “tokens_no_stopwords”. Hasilnya adalah lima baris pertama dari kolom “tokens_no_stopwords” ditampilkan. Kolom ini memuat token yang telah dibersihkan dari *stopword*. Misalnya, kalimat berikut:

the anc while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist diubah menjadi:

anc expressing delight decision nevertheless continued call dissolution tricameral parliament regarded essentially racist.

4) STEMMING

Dalam NLP, *stemming* adalah proses menghilangkan afiks (awalan, akhiran, atau imbuhan lainnya) dari sebuah kata untuk mengembalikan ke bentuk dasarnya, yang dikenal sebagai “stem” atau “akar kata”. Tujuannya adalah untuk memfasilitasi analisis teks dengan mengurangi variasi kata menjadi bentuk dasar yang seragam. Dengan demikian, kata-kata dapat dikelompokkan berdasarkan makna yang serupa atau struktur data teks dapat disederhanakan. Misalnya, kata-kata seperti “regarding” dan “regarded” akan diubah menjadi bentuk akar, yaitu “regard”. Algoritma *stemming* yang paling umum digunakan dalam bahasa Inggris adalah *porter stemmer* [1], [56]

Gambar 5 menunjukkan proses *stemming* menggunakan algoritma *porter stemmer* dari *nlTK* untuk mengubah bentuk dasar token teks. Langkah berikutnya adalah membuat kolom

```

stemmer = PorterStemmer()

def stem_tokens(tokens):
    return [stemmer.stem(word) for word in tokens]

df['stemmed_tokens'] = df['tokens_no_stopwords'].apply(stem_tokens)
df['processed_text'] = df['stemmed_tokens'].apply(lambda tokens: ' '.join(tokens))
print("Stemmed Tokens:\n", df['stemmed_tokens'].head(), "\n")
print("Processed Text:\n", df['processed_text'].head(), "\n")

Stemmed Tokens:
0      [suppos, happi, tini, mean, abl, surpris, peop...
1      [lennox, alway, truli, want, fight, world, tit...
2      [profession, musician, still, sensit, happi, s...
3              [holm, happi, freedom, hous]
4      [problem, tutor, tri, encourag, divers, work, ...
Name: stemmed_tokens, dtype: object

Processed Text:
0      suppos happi tini mean abl surpris peopl gener...
1      lennox alway truli want fight world titl happi...
2      profession musician still sensit happi someth ...
3              holm happi freedom hous
4      problem tutor tri encourag divers work experi ...
Name: processed_text, dtype: object
    
```

Gambar 5. Source code proses stemming dan keluarannya.

baru dalam *DataFrame* yang berisi teks dan token yang telah dikenai proses *stemming* serta mencetak hasilnya untuk analisis lebih lanjut. Misalnya, kalimat berikut:

anc expressing delight decision nevertheless continued call dissolution tricameral parliament regarded essentially racist

Diubah menjadi:

anc express delight decis nevertheless continu call dissolut tricam parliament regard essenti racist

Tabel III menyajikan gambaran secara keseluruhan tahap-tahap prapemrosesan yang diterapkan pada teks sebelum dilakukan proses klasifikasi emosi. Tahapan prapemrosesan ini meliputi *cleaning*, tokenisasi, *stopwords removal*, dan *stemming*. Sebagai contoh, teks asli *“The ANC while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist”* melalui beberapa tahap hingga menjadi bentuk dasar token yang lebih sederhana seperti *“anc express delight decis nevertheless continu call dissolut tricam parliament regard essenti racist”*. Setiap tahap memiliki peran penting dalam mempersiapkan data agar lebih mudah diolah oleh algoritma pembelajaran mesin, dengan tujuan akhir meningkatkan akurasi dalam proses klasifikasi emosi.

C. PEMBUATAN VEKTOR

Pembuatan vektor (*vector creation*) adalah proses mengubah data atau entitas menjadi vektor yang dapat dianalisis atau diproses lebih lanjut dengan menerapkan teknik pembobotan *term frequency-inverse document frequency* (TF-IDF) [56]. Metode TF-IDF banyak digunakan dalam pencarian teks (*text retrieval*) dan prapemrosesan teks. Teknik ini diterapkan pada kata-kata yang telah diekstraksi untuk memperoleh nilai bobotnya. Metode ini membantu dalam menilai pentingnya suatu kata dalam sebuah dokumen relatif terhadap kumpulan dokumen lainnya, sehingga mempermudah proses analisis dan pencarian informasi yang relevan.

TF-IDF juga merupakan metode statistik yang digunakan untuk mengevaluasi tingkat pentingnya sebuah kata dalam sebuah dokumen. *Term frequency* (TF) mengukur besarnya frekuensi kata muncul dalam dokumen tertentu, yang mencerminkan pentingnya kata tersebut dalam konteks dokumen. Sementara itu, *document frequency* (DF) adalah ukuran yang menunjukkan frekuensi kata tersebut muncul dalam dokumen-dokumen koleksi, yang mengindikasikan tingkat keumuman kata tersebut dalam keseluruhan korpus dokumen tersebut. TF-IDF menggabungkan kedua ukuran ini

TABEL III
PRAPEMROSEAN TEKS

Teks	Prapemrosesan
<i>The ANC, while expressing delight at their decision, nevertheless continued to call for the dissolution of the tricameral parliament, which it regarded as essentially racist</i>	Teks
<i>The ANC while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist</i>	Cleaning
<i>the anc while expressing delight at their decision nevertheless continued to call for the dissolution of the tricameral parliament which it regarded as essentially racist</i>	Tokenisasi
<i>anc expressing delight decision nevertheless continued call dissolution tricameral parliament regarded essentially racist</i>	Stopword removal
<i>anc express delight decis nevertheless continu call dissolut tricam parliament regard essenti racist</i>	Stemming

untuk memberikan penilaian yang lebih akurat tentang kepentingan kata dalam konteks dokumen spesifik, sekaligus mengurangi pengaruh kata-kata umum yang sering muncul di banyak dokumen [57].

Tabel IV menunjukkan hasil distribusi nilai TF-IDF untuk setiap kata pada *dataset* emosi. TF-IDF merupakan metrik yang digunakan untuk mengevaluasi pentingnya suatu kata dalam suatu dokumen atau korpus teks. Nilai TF-IDF tinggi menunjukkan bahwa kata tersebut merupakan kata yang penting dan memiliki daya pembeda yang kuat untuk mengidentifikasi emosi tertentu di dalam teks.

D. PEMBAGIAN DATASET

Langkah berikutnya melibatkan pembagian *dataset* menjadi dua bagian, yakni data pelatihan dan data pengujian.

1) DATA PELATIHAN

Data pelatihan adalah bagian dari kumpulan data yang digunakan untuk memprediksi atau menjalankan berbagai fungsi algoritma pembelajaran mesin. Selama proses pelatihan, sampel data pelatihan digunakan sebagai masukan. Algoritma kemudian mempelajari fitur-fitur tersebut untuk membangun model pembelajaran. Mesin yang dilatih dapat menemukan hubungannya sendiri melalui proses pelatihan dengan memberikan instruksi melalui algoritma. Data pelatihan berperan penting dalam mengajarkan mesin untuk mengenali pola-pola dalam data, sehingga model yang dihasilkan mampu membuat prediksi yang akurat atau melakukan tugas tertentu dengan efisien. Melalui iterasi dan penyesuaian parameter, kinerja model dapat ditingkatkan, memastikan bahwa model mampu menangani data baru dengan baik di masa mendatang [58].

2) DATA PENGUJIAN

Data pengujian merupakan bagian vital dalam evaluasi model pembelajaran mesin, yang digunakan untuk menguji prediksi yang telah dibuat setelah proses pelatihan selesai. Dalam proses ini, model digunakan untuk membuat prediksi terhadap data pengujian atau data produksi yang sebelumnya tidak pernah dilihat. Data pengujian ini mencakup sampel-sampel yang merepresentasikan situasi yang dihadapi dalam aplikasi nyata. Evaluasi terhadap model menggunakan data pengujian membantu mengukur akurasi dan kinerja model serta mengidentifikasi kelemahan dan kekuatan yang perlu

TABEL IV
HASIL DISTRIBUSI TF-IDF

No.	Teks	Hasil TF-IDF
1.	<i>lin foh quite at being to so</i>	0,068
2.	<i>many of the as the wore on were to find the door not heard of the noon and so they had all day</i>	0,115
3.	<i>lin yuan was with proud of his and that he had tsu ma 's</i>	0,229
4.	<i>even so was when we were from last</i>	0,351
5.	<i>you 're not just for lucy</i>	0,153
6.	<i>that to when they heard of his 's grim fate</i>	0,272

diperbaiki sebelum model diterapkan secara luas. Dengan demikian, penggunaan data pengujian tidak hanya penting untuk menghindari *overfitting*, tetapi juga untuk memastikan bahwa model dapat bekerja dengan baik dalam mengatasi masalah yang belum pernah ditemui sebelumnya [58]. Dalam penelitian ini, digunakan data pelatihan sebesar 90% dan data pengujian sebesar 10% [32].

Gambar 6 adalah *source code* pembagian *dataset* menjadi data pelatihan dan data pengujian menggunakan fungsi `train_test_split` dari pustaka `sklearn.model_selection`. Pertama, kolom `emotion` dari `DataFrame` `df` disimpan dalam variabel `y`, yang berisi label atau target untuk klasifikasi. Kemudian, *dataset* yang telah diubah menjadi fitur TF-IDF (`X_tfidf`) bersama dengan label `y` dibagi menjadi data pelatihan (90%) dan data pengujian (10%), dengan menggunakan `random_state=42` untuk mempertahankan konsistensi hasil pembagian setiap kali kode dieksekusi.

E. PEMODELAN KLASIFIKASI

Setelah tahapan pembagian *dataset* selesai, langkah berikutnya adalah menerapkan pemodelan klasifikasi kombinasi algoritma pembelajaran mesin, yakni NB, LR, dan KNN, dengan metode *ensemble bagging*. Pada penelitian ini, penggabungan hasil dari ketiga model dilakukan pada tahap pengujian. Setiap algoritma terlebih dahulu memproses data uji secara individual. Hasil prediksi dari masing-masing algoritma kemudian digabungkan melalui teknik *ensemble bagging* untuk menghasilkan prediksi akhir yang lebih stabil dan akurat. Secara matematis, rumus dari *bagging* direpresentasikan dengan (4).

$$\widehat{f}_{bag} = \widehat{f}_1(X) + \widehat{f}_2(X) + \dots + \widehat{f}_b(X). \quad (4)$$

Bagian di sebelah kiri \widehat{f}_{bag} adalah prediksi hasil dari *ensemble bagging* (*bagged prediction*) dan bagian di sebelah kanan $\widehat{f}_b(X)$ adalah prediksi dari masing-masing pembelajar individu (*individual learner*).

F. PENGUKURAN NILAI

Pada tahap terakhir, dilakukan pengukuran untuk mengevaluasi tingkat akurasi proses klasifikasi. Hasilnya disajikan dalam bentuk *confusion matrix* yang menggambarkan nilai-nilai akurasi.

Confusion matrix adalah matriks 2×2 yang merangkum semua hasil klasifikasi yang benar dan salah. Dari kombinasi nilai *true positive* (TP), *false positive* (FP), *false negative* (FN), dan *true negative* (TN), dihasilkan empat variabel pengukuran untuk mengevaluasi kinerja model klasifikasi.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}. \quad (5)$$

```
from sklearn.model_selection import train_test_split

y = df['Emotion']

# Bagi dataset
X_train, X_test, y_train, y_test = train_test_split(X_tfidf, y, test_size=0.1, random_state=42)
```

Gambar 6. Source code pembagian dataset.

Persamaan (5) merupakan rumus akurasi, yang mengukur rasio prediksi benar (positif dan negatif) terhadap semua prediksi.

$$\text{Precision} = \frac{TP}{TP+FP}. \quad (6)$$

Persamaan (6) digunakan untuk menghitung presisi, yang mengukur tingkat ketepatan model dalam memprediksi positif dengan membandingkan TP terhadap semua prediksi positif (TP + FP).

$$\text{Recall} = \frac{TP}{TP+FN}. \quad (7)$$

Persamaan (7) merupakan *recall*, yang mengukur kemampuan model mendeteksi semua data positif sebenarnya, yaitu rasio TP terhadap jumlah keseluruhan positif (TP + FN).

$$F1 \text{ score} = \frac{2(\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}}. \quad (8)$$

Persamaan (8) merupakan *f1-score*, yaitu rata-rata harmonik antara presisi dan *recall*, yang memberikan keseimbangan antara keduanya.

IV. HASIL DAN PEMBAHASAN

Dalam penelitian ini, proses analisis dilakukan menggunakan *tools* Google Colab. Untuk meningkatkan kinerja klasifikasi, penelitian ini menggunakan metode *ensemble bagging* yang digabungkan dengan ketiga algoritma klasifikasi, yaitu NB, KNN, dan LR. Metode ini diharapkan dapat menghasilkan prediksi yang lebih tepat dan tepercaya dibandingkan jika algoritma digunakan secara individu.

Tabel V menunjukkan hasil evaluasi kinerja tiga algoritma klasifikasi yang digunakan dalam penelitian ini, yaitu NB, LR, dan KNN, tanpa menggunakan pendekatan *ensemble bagging*. Hasil yang disajikan mencakup nilai akurasi, *f1-score*, presisi, dan *recall* untuk setiap algoritma. Dari tabel, dapat dilihat bahwa LR memberikan kinerja terbaik dengan akurasi tertinggi sebesar 85,95%, diikuti oleh NB dengan akurasi 66,94% dan KNN dengan akurasi terendah, yaitu 52,48%. Selain itu, LR juga memiliki nilai presisi dan *recall* yang tinggi, masing-masing sebesar 88,08% dan 85,95%. Nilai-nilai *f1-score*, presisi, dan *recall* ini memberikan gambaran lebih lengkap tentang kemampuan setiap model dalam mengklasifikasikan data secara akurat dan konsisten.

Tabel VI menampilkan hasil evaluasi kinerja dari tiga algoritma klasifikasi yang digunakan dalam penelitian ini, yaitu NB, LR, dan KNN, dengan pendekatan *ensemble bagging*. Metode *ensemble bagging* ini bertujuan untuk meningkatkan akurasi prediksi dengan menggabungkan hasil dari beberapa model yang dilakukan pada tahap pengujian. Hasil yang ditunjukkan dalam Tabel VI meliputi nilai akurasi, *f1-score*, presisi, dan *recall* untuk masing-masing algoritma. Dari Tabel VI, dapat dilihat bahwa LR dengan pendekatan *ensemble bagging* mencapai akurasi tertinggi sebesar 98,76%, diikuti oleh KNN dan NB dengan akurasi yang lebih rendah. Nilai presisi, *recall*, dan *f1-score* juga ditunjukkan untuk memberikan gambaran menyeluruh mengenai kinerja setiap algoritma dalam mengklasifikasikan emosi pada teks. Nilai-

TABEL V
LAPORAN KLASIFIKASI TANPA ENSEMBLE BAGGING

Hasil	Akurasi (%)	F1-Score	Presisi	Recall
Naïve Bayes	66,94	64,85	71,65	66,94
Logistic Regression	85,95	83,52	88,08	85,95
KNN	52,48	51,73	52,02	52,48

TABEL VI
LAPORAN KLASIFIKASI DENGAN ENSEMBLE BAGGING

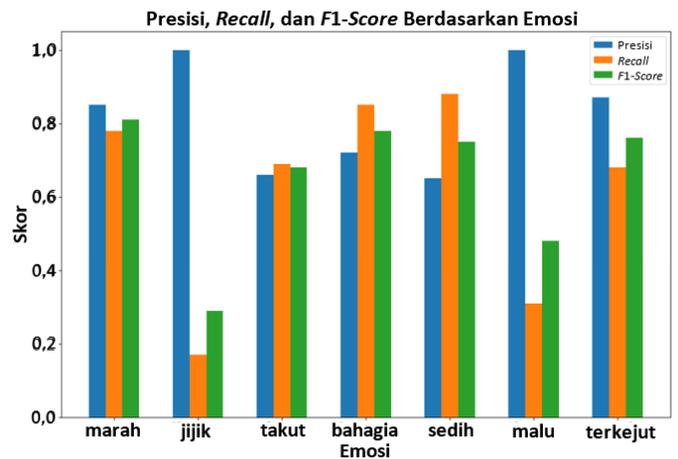
Hasil	Akurasi (%)	F1-Score	Presisi	Recall
Ensemble Bagging Naïve Bayes	73,55	72,03	77,11	73,55
Ensemble Bagging Logistic Regression	98,76	98,76	98,79	98,76
Ensemble Bagging KNN	82,64	81,72	83,98	82,64

nilai ini penting untuk memahami keandalan model dalam melakukan prediksi yang tepat dan konsisten.

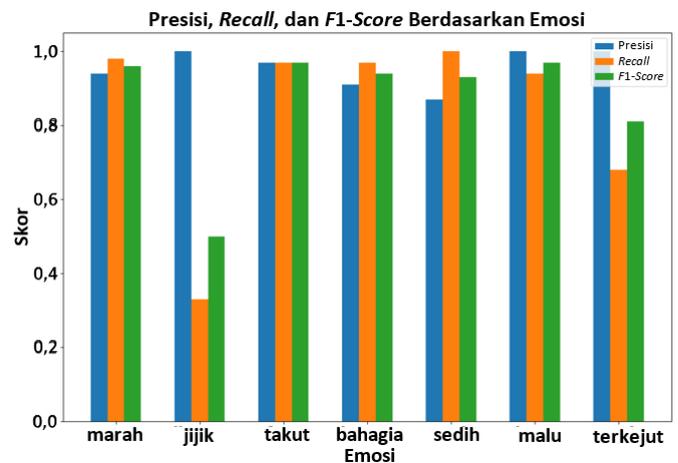
Gambar 7 menunjukkan kinerja algoritma NB dalam klasifikasi emosi pada teks berbahasa Inggris, yang diukur menggunakan tiga metrik utama: presisi, recall, dan f1-score. Pada grafik tersebut, dapat dilihat bahwa emosi “jijik” dan “malu” memiliki presisi tertinggi, sebesar 100%, yang menunjukkan bahwa model ini sangat efektif dalam mengidentifikasi kedua emosi tersebut dengan benar tanpa banyak kesalahan positif. Sementara itu, emosi “sedih” memiliki nilai recall tertinggi, sebesar 88%, yang menunjukkan bahwa model ini sangat sensitif dalam mendeteksi semua instance emosi “sedih” yang sebenarnya dalam dataset. Emosi “marah” memiliki f1-score tertinggi, sebesar 81%, yang menandakan bahwa model ini memiliki keseimbangan yang baik dalam mendeteksi emosi dengan akurat dan lengkap. Analisis ini mengindikasikan bahwa meskipun NB efektif dalam beberapa kategori emosi, terdapat variasi dalam kinerja berdasarkan jenis emosi yang diklasifikasikan.

Gambar 8 memperlihatkan hasil analisis kinerja algoritma LR dalam klasifikasi emosi pada teks berbahasa Inggris, menggunakan metrik presisi, recall, dan f1-score. Grafik ini menunjukkan bahwa emosi “jijik”, “malu”, dan “terkejut” memiliki presisi tertinggi, yaitu sebesar 100%, yang menunjukkan bahwa model LR sangat akurat dalam mengidentifikasi emosi-emosi ini tanpa kesalahan positif. Untuk recall, emosi “sedih” memiliki nilai tertinggi, sebesar 100%, yang menunjukkan bahwa model ini mampu mendeteksi semua instance emosi “sedih” dalam dataset dengan sangat baik. Sementara itu, untuk f1-score, emosi “takut” dan “malu” memiliki nilai tertinggi, yaitu sebesar 97%, yang menunjukkan bahwa model ini memiliki keseimbangan optimal dalam presisi dan recall untuk kedua emosi tersebut. Secara keseluruhan, hasil ini menunjukkan bahwa LR adalah algoritma yang sangat efektif dalam klasifikasi emosi, terutama untuk emosi-emosi dengan nilai presisi dan recall yang tinggi.

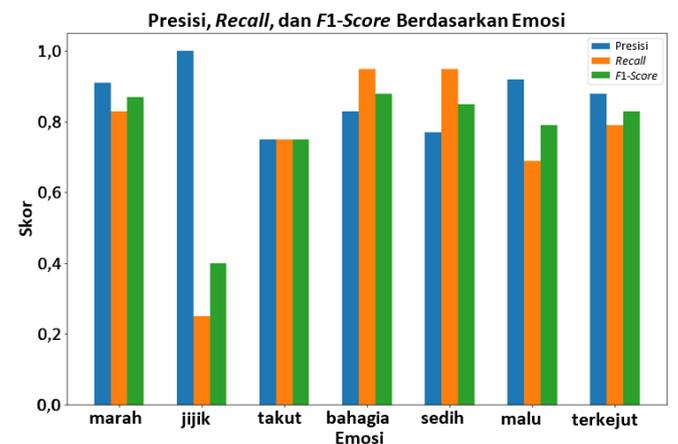
Gambar 9 memperlihatkan kinerja algoritma KNN dalam klasifikasi emosi pada teks berbahasa Inggris berdasarkan metrik presisi, recall, dan f1-score. Grafik ini menunjukkan bahwa emosi “jijik” memiliki presisi tertinggi, sebesar 100%, yang menunjukkan bahwa model KNN sangat baik dalam mengidentifikasi emosi ini dengan akurasi tinggi. Emosi “bahagia” dan “sedih” memiliki nilai recall tertinggi, sebesar 95%, yang menandakan bahwa model ini sangat sensitif dan



Gambar 7. Grafik jumlah presisi, recall, dan f1-score berdasarkan emosi (NB).



Gambar 8. Grafik jumlah presisi, recall, dan f1-score berdasarkan emosi (LR).



Gambar 9. Grafik jumlah presisi, recall, dan f1-score berdasarkan emosi (KNN).

mampu mendeteksi hampir semua instance dari kedua emosi tersebut dalam dataset. Untuk f1-score, emosi “bahagia” memiliki nilai tertinggi, yaitu sebesar 88%, yang menunjukkan bahwa model ini memiliki keseimbangan yang baik antara presisi dan recall dalam mendeteksi emosi “bahagia”. Hasil ini mengindikasikan bahwa meskipun KNN memiliki kinerja yang baik dalam beberapa kategori emosi, ada variasi dalam efektivitasnya berdasarkan jenis emosi yang diklasifikasikan.

V. KESIMPULAN

Penelitian ini membuktikan bahwa teknik ensemble bagging yang menggabungkan algoritma NB, LR, dan KNN dapat meningkatkan kinerja dalam klasifikasi emosi pada teks berbahasa Inggris. Berdasarkan hasil evaluasi, metode

ensemble bagging dengan LR memberikan akurasi tertinggi, sebesar 98,76%, lebih tinggi dibandingkan penggunaan NB dan KNN secara terpisah, yang masing-masing hanya mencapai akurasi 66,94% dan 52,48%. Selain itu, metode *ensemble* ini juga memberikan peningkatan signifikan pada metrik presisi, *recall*, dan *f1-score* dibandingkan dengan hasil dari algoritma individu. Hasil ini mendukung temuan dari penelitian sebelumnya, yang menunjukkan bahwa penggunaan *bagging* pada NB meningkatkan akurasi dari 66,94% menjadi 73,55%. Dengan demikian, pendekatan *ensemble bagging* terbukti lebih efektif dalam meningkatkan keandalan dan akurasi klasifikasi emosi, terutama pada data teks yang tidak terstruktur.

KONTRIBUSI PENULIS

Konseptualisasi, Erfian Junianto; metodologi, Mila Puspitasari; perangkat lunak, Salman Ilyas Zakaria; validasi, Erfian Junianto, Mila Puspitasari, dan Toni Arifin; analisis formal, Mila Puspitasari; investigasi, Mila Puspitasari; sumber daya, Ignatius Wiseto Prasetyo Agung; kurasi data, Salman Ilyas Zakaria; penulisan—penyusunan draf asli, Mila Puspitasari; penulisan—peninjauan dan penyuntingan, Erfian Junianto, Toni Arifin, Ignatius Wiseto Prasetyo Agung; visualisasi, Salman Ilyas Zakaria; pengawasan, Erfian Junianto; administrasi proyek, Erfian Junianto; akuisisi pendanaan, Ignatius Wiseto Prasetyo Agung.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Universitas Adhirajasa Reswara Sanjaya dan Fakultas Teknologi Informasi atas dukungan dan sumber daya yang sangat berharga yang sangat membantu dalam penelitian ini. Ucapan terima kasih juga disampaikan kepada tim *Research and Innovation Digital ARS* yang telah memberikan diskusi yang mendalam dan bantuan teknis selama penelitian ini. Selain itu, penulis berterima kasih atas dorongan terus-menerus dari keluarga dan rekan kerja, yang sangat berperan dalam menyelesaikan penelitian ini.

REFERENSI

- [1] A.T. Bagus, "Klasifikasi emosi pada teks menggunakan metode deep learning," Skripsi, Universitas Islam Indonesia, Sleman, Indonesia, 2022.
- [2] H. Krishnan, M.S. Elayidom, dan T. Santhanakrishnan, "Emotion detection of tweets using naïve Bayes classifier," *Int. J. Eng. Technol. Sci. Res. (IJETSJR)*, vol. 4, no. 11, hal. 457–462, Nov. 2017.
- [3] A.N. Rohman, E. Utami, dan S. Raharjo, "Deteksi kondisi emosi pada media sosial menggunakan pendekatan leksikon dan natural language processing," *J. Eksplor Inform.*, vol. 9, no. 1, hal. 70–76, Sep. 2019, doi: 10.30864/eksplor.v9i1.277.
- [4] A. Chatterjee, N. Narahari, M. Joshi, dan P. Agrawal, "SemEval-2019 task 3: EmoContext contextual emotion detection in text," dalam *Proc. 13th Int. Workshop Semant. Eval. (SemEval-2019)*, 2019, hal. 39–48, doi: 10.18653/v1/S19-2005.
- [5] P.W.A. Wibawa dan C. Pramatha, "Systematic literature review: Machine learning methods in emotion classification in textual data," *J. Sisfokom (Sist. Inf. Komput.)*, vol. 12, no. 3, hal. 425–433, Nov. 2023, doi: 10.32736/sisfokom.v12i3.1787.
- [6] D. Ariyanti dan K. Iswardani, "Teks mining untuk klasifikasi keluhan masyarakat menggunakan algoritma naïve Bayes," *IKRA-ITH Inform. J. Komput Inform.*, vol. 4, no. 3, hal. 125–132, Nov. 2020.
- [7] N. Andriani dan A. Wibowo, "Implementasi text mining klasifikasi topik tugas akhir mahasiswa teknik informatika menggunakan pembobotan TF-IDF dan metode cosine similarity berbasis web," dalam *Proc. Semin. Nas. Mhs. Ilmu Komput. Apl. (SENAMIKA)*, 2021, hal. 130–137.
- [8] Bimananda W. dkk., "Analisis text mining dari cuitan Twitter mengenai infrastruktur di Indonesia dengan metode klasifikasi naïve Bayes," *Eig. Math. J.*, vol. 2, no. 2, hal. 92–101, Des. 2019, doi: 10.29303/emj.v1i2.36.
- [9] T.W.D. Sari, "Penerapan text mining dengan menggunakan algoritma TF-IF untuk klasifikasi genre novel," *Pelita Inform., Inf. Inform.*, vol. 10, no. 1, hal. 29–37, Jul. 2021.
- [10] M. Afdal dan L.R. Elita, "Penerapan text mining pada aplikasi Tokopedia menggunakan algoritma k-nearest neighbor," *J. Ilm. Rekayasa Manaj. Sist. Inf.*, vol. 8, no. 1, hal. 78–87, Feb. 2022, doi: 10.24014/rmsi.v8i1.16595.
- [11] D. Ardiada, M. Sudarma, dan D. Giriantari, "Text mining pada sosial media untuk mendeteksi emosi pengguna menggunakan metode support vector machine dan k-nearest neighbour," *Maj. Ilm. Teknol. Elekt.*, vol. 18, no. 1, hal. 55–60, Jan.-Apr. 2019, doi: 10.24843/mite.2019.v18i01.p08.
- [12] T. Ridwansyah, "Implementasi text mining terhadap analisis sentimen masyarakat dunia di Twitter terhadap Kota Medan menggunakan k-fold cross validation dan naïve Bayes classifier," *Kajian Ilm. Inform. Komput.*, vol. 2, no. 5, hal. 178–185, Apr. 2022, doi: 10.30865/klik.v2i5.362.
- [13] S.S. Berutu, "Text mining dan klasifikasi sentimen berbasis naïve Bayes pada opini masyarakat terhadap makanan tradisional," *J. Sist. Komput. Inform. (JSON)*, vol. 4, no. 2, hal. 254–262, Des. 2022, doi: 10.30865/json.v4i2.5138.
- [14] D.T. Alamanda dkk., "Sentiment analysis using text mining of Indonesia tourism reviews via social media," *Int. J. Humanit. Arts Soc. Sci.*, vol. 5, no. 2, hal. 43–53, Apr. 2019, doi: 10.20469/ijhss.5.10005-2.
- [15] F.F. Mailo dan L. Lazuardi, "Analisis sentimen data Twitter menggunakan metode text mining tentang masalah obesitas di Indonesia," *J. Inf. Syst. Public Health*, vol. 6, no. 1, hal. 44–51, Apr. 2021, doi: 10.22146/jisph.44455.
- [16] E. Indrayuni, "Klasifikasi text mining review produk kosmetik untuk teks bahasa Indonesia menggunakan algoritma naïve Bayes," *J. Khatulistiwa Inform.*, vol. 7, no. 1, hal. 29–36, Jun. 2019, doi: 10.31294/jki.v7i1.5740.
- [17] W. Hermanto, B. Irawan, dan C. Setianingsih, "Klasifikasi emosi pada lirik lagu menggunakan algoritma support vector machine dan optimasi particle swarm optimization," *e-Proc. Eng.*, vol. 8, no. 5, hal. 6307–6327, Okt. 2021.
- [18] N. Anggraini, E.S.N. Harahap, dan T.B. Kurniawan, "Text mining-text analysis related to COVID-19 vaccination issues," *J. IPTEK-KOM (J. Ilmu Pengetah. Teknol. Komun.)*, vol. 23, no. 2, hal. 141–153, Des. 2021, doi: 10.33169/iptekom.23.2.2021.141-153.
- [19] H.P. Doloksaribu dan Y.T. Samuel, "Komparasi algoritma data mining untuk analisis sentimen aplikasi Pedulilindungi," *J. Teknol. Inf., J. Keilmuan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, hal. 1–11, Jan. 2022, doi: 10.47111/jti.v16i1.3747.
- [20] S.D. Pramukti, A. Nugroho, dan A.S. Sunge, "Analisis sentimen masyarakat dengan metode naïve Bayes dan particle swarm optimization," *Techno.Com*, vol. 21, no. 1, hal. 62–75, Feb. 2022, doi: 10.33633/te.v21i1.5332.
- [21] R. Fajar, "Implementasi algoritma naïve Bayes terhadap analisis sentimen opini film pada Twitter," *J. Inovtek Polbeng Seri Inform.*, vol. 3, no. 1, hal. 50–59, Jun. 2018, doi: 10.35314/isi.v3i1.335.
- [22] S. Budi, "Text mining untuk analisis sentimen review film menggunakan algoritma k-means," *Techno.Com*, vol. 16, no. 1, hal. 1–8, Feb. 2017, doi: 10.33633/te.v16i1.1263.
- [23] A.K. Fauziyyah, "Analisis sentimen pandemi COVID19 pada streaming Twitter dengan text mining python," *J. Ilm. Sinus*, vol. 18, no. 2, hal. 31–42, Jul. 2020, doi: 10.30646/sinus.v18i2.491.
- [24] R. Wahyudi dan G. Kusumawardhana, "Analisis sentimen pada review aplikasi Grab di Google Play Store menggunakan support vector machine," *J. Inform.*, vol. 8, no. 2, hal. 200–207, Sep. 2021, doi: 10.31294/ji.v8i2.9681.
- [25] R. Siringoringo dan Jamaluddin, "Text mining dan klusterisasi sentimen pada ulasan produk toko online," *J. Penelit. Tek. Inform.*, vol. 2, no. 1, hal. 314–319, Apr. 2019, doi: 10.34012/jutikomp.v2i1.456.
- [26] H.P. Koapaha dan N. Ananto, "Bagging based ensemble analysis in handling unbalanced data on classification modelling," *Klabat Account. Rev.*, vol. 2, no. 2, hal. 165–178, Sep. 2021, doi: 10.60090/kar.v2i2.589.165-178.
- [27] R. Siringoringo dan I.K. Jaya, "Ensemble learning dengan metode SMOTEBagging pada klasifikasi data tidak seimbang," *J. Inf. Syst. Dev. (ISD)*, vol. 3, no. 2, hal. 75–81, Jul. 2018.
- [28] P. Arsi, I. Prayoga, dan M.H. Asyari, "Klasifikasi sentimen publik terhadap jenis vaksin COVID-19 yang tersertifikasi WHO berbasis NLP dan KNN," *J. Media Inform. Budidarma*, vol. 7, no. 1, hal. 260–266, Jan. 2023, doi: 10.30865/mib.v7i1.5418.
- [29] A.K. Santoso dkk., "Klasifikasi persepsi pengguna Twitter terhadap kasus COVID-19 menggunakan metode logistic regression," *J. Inform. Kaputama (JIK)*, vol. 5, no. 2, hal. 234–241, Jul. 2021, doi: 10.59697/jik.v5i2.247.

- [30] F. Fanesya, R.C. Wihandika, dan Indriati, "Deteksi emosi pada Twitter menggunakan metode naïve Bayes dan kombinasi fitur," *J. Pengembangan Teknol. Inf. Ilmu Komput.*, vol. 3, no. 7, hal. 6678–6686, Jul. 2019.
- [31] S. Shofiyah, "Studi komparasi metode naïve Bayes dan metode ensemble bagging," Skripsi, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta, Indonesia, 2020.
- [32] W.O. Simanjuntak, A.B.P. Negara, dan R. Septriana, "Perbandingan algoritma logistic regression dan random forest (Studi kasus: Klasifikasi emosi tweet)," *J. Apl. Ris. Inform.*, vol. 1, no. 2, hal. 160–164, Agu. 2023, doi: 10.26418/juara.v2i1.69682.
- [33] A. Helmut, Adiwijaya, dan D.T. Murdiansyah, "Klasifikasi email multi kelas menggunakan ensemble bagging," *e-Proc. Eng.*, vol. 7, no. 1, hal. 2498–2505, Apr. 2020.
- [34] Mursyidah, H.T. Hidayat, dan D.M. Sari, "Klasifikasi teks emosi bahasa Aceh menggunakan metode term frekuensi / invers dokument frekuensi," *J. Infomedia, Tek. Inform. Multimed. Jar.*, vol. 2, no. 1, hal. 14–19, Mar. 2017, doi: 10.30811/v2i1.462.
- [35] H. Susana, N. Suarna, Fathurrohman, dan Kaslani, "Penerapan model klasifikasi metode naïve Bayes terhadap penggunaan akses internet," *J. Ris. Sist. Inf. Teknol. Inf. (JURISISTEKNI)*, vol. 4, no. 1, hal. 1–8, Jan. 2022, doi: 10.52005/jursistekni.v4i1.96.
- [36] I. Destuardi dan S. Sumpeno, "Klasifikasi emosi untuk teks bahasa Indonesia menggunakan metode naïve Bayes," dipresentasikan pada Semin. Nas. Pascasarj. IX – ITS, Surabaya, Jawa Timur, Indonesia, 12 Agu. 2015.
- [37] A. Ansori, "Kepribadian dan emosi," *J. Literasi Pendidik. Nusant.*, vol. 1, no. 1, hal. 41–54, Jan.-Jun. 2020.
- [38] Y.A. Jatmiko, S. Padmadisastra, dan A. Chadidjah, "Analisis perbandingan kinerja cart konvensional, bagging dan random forest pada klasifikasi objek: Hasil dari dua simulasi," *Media Stat.*, vol. 12, no. 1, hal. 1–12, Jul. 2019, doi: 10.14710/medstat.12.1.1-12.
- [39] A. Priyanto dan M.R. Ma'arif, "Implementasi web scraping dan text mining untuk akuisisi dan kategorisasi informasi laman web tentang hidroponik," *Indones. J. Inf. Syst.*, vol. 1, no. 1, hal. 25–33, Agu. 2018, doi: 10.24002/ijis.v1i1.1664.
- [40] M.P.R. Putra dan K.R.N. Waedani, "Penerapan text mining dalam menganalisis kepribadian pengguna media sosial," *JUTIM (J. Tek. Inform. Musirawas)*, vol. 5, no. 1, hal. 63–71, Jun. 2020, doi: 10.32767/jutim.v5i1.
- [41] S. Shalev-Shwartz dan S. Ben-David, *Understanding Machine Learning : From Theory To Algorithms*. New York, NY, AS: Cambridge University Press, 2014.
- [42] A. Fathurohman, "Machine learning untuk pendidikan: Mengapa dan bagaimana," *J. Inform. Teknol. Komput.*, vol. 1, no. 3, hal. 57–62, Nov. 2021, doi: 10.55606/jitek.v1i3.306.
- [43] R.N. Devita, H.W. Herwanto, dan A.P. Wibawa, "Perbandingan kinerja metode naïve Bayes dan k-nearest neighbor untuk klasifikasi artikel berbahasa Indonesia," *J. Teknol. Inf. Ilmu Komput.*, vol. 5, no. 4, hal. 427–434, Sep. 2018, doi: 10.25126/jtiik.201854773.
- [44] P.D. Nugraha, S.A. Faraby, dan Adiwijaya, "Klasifikasi dokumen menggunakan metode k-nearest neighbor (KNN) dengan information gain document," *e-Proc. Eng.*, vol. 5, no. 1, hal. 1541–1550, Mar. 2018.
- [45] A.S. Rezeki, "Klasifikasi emosi pada Twitter dengan metode k-nearest neighbor (KNN)," Skripsi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia, 2021.
- [46] Y. Kustiyahningsih dan N. Syafa'ah, "Sistem pendukung keputusan untuk menentukan jurusan pada siswa SMA menggunakan metode KNN dan smart," *J. Sist. Inf. Indones.*, vol. 1, no. 1, hal. 19–28, Apr. 2015.
- [47] I.F. Ramadhy dan Y. Sibaroni, "Analisis trending topik Twitter dengan fitur ekspansi fasttext menggunakan metode logistic regression," *JURIKOM (J. Ris. Komput.)*, vol. 9, no. 1, hal. 1–7, Feb. 2022, doi: 10.30865/jurikom.v9i1.3791.
- [48] A.K. Santoso, "Analisis sentimen Twitter bahasa Indonesia menggunakan pendekatan machine learning," *J. Inform. Kaputama (JIK)*, vol. 6, no. 2, hal. 129–136, Jul. 2022.
- [49] R. Susetyoko, W. Yuwono, E. Purwanti, dan N. Ramadijanti, "Perbandingan metode random forest, regresi logistik, naïve Bayes, dan multilayer perceptron pada klasifikasi uang kuliah tunggal (UKT)," *J. Infomedia, Tek. Inform. Multimed. Jar.*, vol. 7, no. 1, hal. 8–16, Jun. 2022, doi: 10.30811/jim.v7i1.2916.
- [50] N.K. Hasibuan, S. Dur, dan I. Husein, "Faktor penyebab penyakit diabetes melitus dengan metode regresi logistik," *G-Tech, J. Teknol. Terap.*, vol. 6, no. 2, hal. 257–264, Okt. 2022, doi: 10.33379/gtech.v6i2.1696.
- [51] D. Ghazi, D. Inkpen, dan S. Szpakowicz, "Detecting emotion stimuli in emotion-bearing sentences," dalam *Comput. Linguist. Intell. Text Process.*, 2015, hal. 152–165. doi: 10.1007/978-3-319-18117-2_12.
- [52] S. Khairunnisa, K. Adiwijaya, dan S.A. Faraby, "Pengaruh text preprocessing terhadap analisis sentimen komentar masyarakat pada media sosial Twitter (Studi kasus pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, hal. 406–414, Apr. 2021, doi: 10.30865/mib.v5i2.2835.
- [53] R.I. Pristiyanti, M.A. Fauzi, dan L. Muflikhah, "Sentiment analysis peringkasan review film menggunakan metode information gain dan k-nearest neighbor," *J. Media Inform. Budidarma*, vol. 2, no. 3, hal. 1179–1186, Mar. 2018.
- [54] J. Petrus, "Kerangka kerja tokenisasi berdasarkan struktur kalimat bahasa Indonesia," Skripsi, Universitas Sriwijaya, Palembang, Indonesia, 2023.
- [55] I. Gotama, S. Hariyanto, dan H. Wijaya, "Klasifikasi berita hoaks topik COVID-19 dengan klasifikasi Rocchio dan cosine similarity," *Algor*, vol. 2, no. 1, hal. 84–92, Nov. 2020.
- [56] E. Junianto dan R. Rachman, "Implementation of text mining model to emotions detection on social media comments using particle swarm optimization and naïve Bayes classifier," dalam *2019 7th Int. Conf. Cyber IT Serv. Manag. (CITSM)*, 2019, hal. 1–6, doi: 10.1109/CITSM47753.2019.8965382.
- [57] J.A. Septian, T.M. Fachrudin, dan A. Nugroho, "Analisis sentimen pengguna Twitter terhadap polemik persepakbolaan Indonesia menggunakan pembobotan TF-IDF dan k-nearest neighbor," *J. Intell. Syst. Comput.*, vol. 1, no. 1, hal. 43–49, Agu. 2019, doi: 10.52985/insyst.v1i1.36.
- [58] A.C. Darmawan, "Pengembangan aplikasi berbasis web dengan python flask untuk klasifikasi data menggunakan metode decision tree C4.5," Skripsi, Universitas Islam Indonesia, Sleman, Indonesia, 2023.