Smart Comp: Jurnalnya Orang Pintar Komputer
Volume 15, Nomor 3, Juli 2026, hlm. 650-660
Terakreditasi Sinta, Peringkat 4, SK No. 295/C/C3/KPT/2026
DOI : 10.30591/smartcomp.v15i3.9776

P-ISSN : 2089-676X
E-ISSN : 2549-0796

Perbandingan Model Pohon Keputusan dan Random Forest dalam Mendeteksi Sentimen Isu Pajak Pertambahan Nilai

Janji Armanda Fabian[*1], Galih Hendro Martono[2], Ismarmiaty[3]
[1,2,3)]Ilmu Komputer, Universitas Bumigora Mataram
Email: [*1]jafabian599@gmail.com, [2]galih.hendro@universitasbumigora.ac.id,
[3]ismarmiaty@universitasbumigora.ac.id

(Naskah masuk: *22 Oktober 2025*, diterima untuk diterbitkan: *15 Juni 2026*)

Abstrak: *Kenaikan PPN menjadi 12% memunculkan berbagai tanggapan dari masyarakat yang diekspresikan secara luas melalui media sosial, termasuk platform media sosial paling terkenal yaitu Youtube. Fenomena ini menunjukkan bahwa opini publik di dunia maya dapat menjadi bahan kajian penting dalam memahami respon masyarakat terhadap kebijakan fiskal pemerintah. Penelitian ini bertujuan untuk mengklasifikasikan emosi masyarakat terhadap isu kenaikan PPN 12% dengan memanfaatkan komentar YouTube sebagai sumber data. Tahapan penelitian ini dimulai dari pengumpulan data dengan metode web scraping, diikuti tahap pra-pemrosesan data, penerjemahan teks ke dalam bahasa Inggris, serta pelabelan emosi menggunakan NRC Lexicon. Selanjutnya, fitur teks diekstraksi menggunakann TF-IDF untuk kemudian dianalisis dengan Random Forest dan Decision Tree. Didapatkan hasil yang menunjukkan bahwa emosi dominan yang muncul pada komentar masyarakat adalah netral, positif, dan negatif. Model Decision Tree menghasilkan akurasi 57,35% dan macro F1-Score 0,53, sementara Random Forest memperoleh akurasi lebih tinggi sebesar 65,43% dengan macro F1-Score 0,57. Berdasarkan hasil, dapat dilihat Random Forest menunjukkan performa yang lebih tinggi daripada Decision Tree untuk klasifikasi teks.*

Kata Kunci – *analisis sentimen; klasifikasi emosi; random forest; decision tree*

Comparison of Decision Tree and Random Forest Models in Detecting Value Added Tax Issue Sentiment

Abstract: *The increase in VAT to 12% has given rise to various responses from the public, which have been widely expressed through social media, including the most famous social media platform, namely YouTube. This phenomenon shows that public opinion in the virtual world can be an important subject of study in understanding the public's response to government fiscal policy. The purpose of this study is to analyze and calssify public opinions on the 12% VAT issue by utilizing YouTube comments as a data source. The study began with data collection using web scraping techniques, followed by data pre-processing, text translation into English, and emotion labeling using the NRC Lexicon. Next, text features were extracted using the TF-IDF and then analyzed with Decision Tree and Random Forest. The study results reveal that the most significant emotions that appear in public comments are neutral, positive, and negative. The Decision Tree model produced an accuracy of 57.35% with a macro F1-Score of 0.53, while Random Forest obtained a higher accuracy of 65.43% with a macro F1-Score of 0.57. Based on the results, it can be seen that Random Forest produced better performance than Decision Tree in text classification.*

Keywords – *sentiment analysis; emotion classification; random forest; decision tree*

1. PENDAHULUAN

Pajak Pertambahan Nilai (PPN) merupakan salah satu sumber pemasukan negara yang penting di Indonesia. Pemerintah Indonesia menerapkan kenaikan tarif PPN dari 11% menjadi 12% yang akan diterapkan tahun 2025. Ini menimbulkan pro dan kontra dari masyarakat karna dampaknya terhadap daya beli. Di era digital, media sosial menjadi wadah utama untuk mengungkapkan sikap terhadap kebijakan publik [1]. *Platform* media sosial tersebut dapat dimanfaatkan sebagai sumber opini masyarakat, konten yang dikumpulkan dari media sosial dapat dijadikan acuan dalam mengukur apakah tanggapan masyarakat terhadap suatu isu bersikap

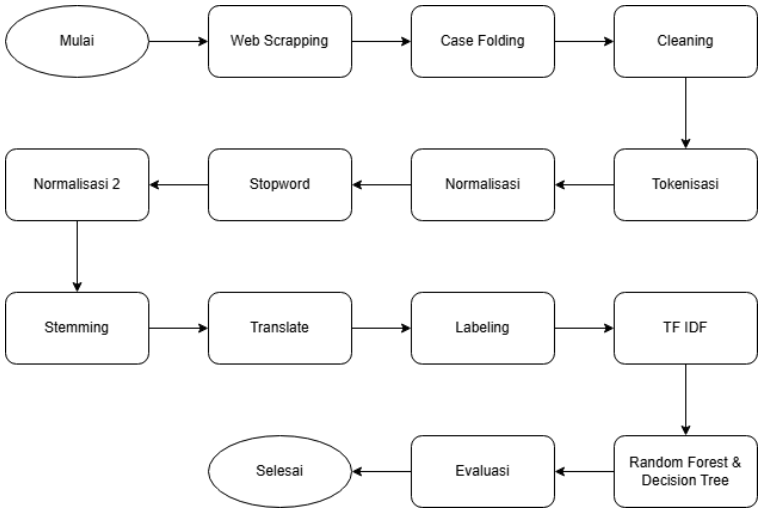
positif, netral, atau negatif [2]. *YouTube* merupakan sosial media yang paling sering dimanfaatkan secara global dengan basis pengguna bulanan aktif sekitar 2,3, sehingga membuatnya sangat populer dan menjadikannya sumber data yang representatif untuk mengamati opini publik [3].

Sentimen analisis merupakan metode otomatis dalam memproses data teks agar mengetahui kandungan sentimen yang ada pada suatu opini atau kalimat [4], untuk mengetahui emosi yang tertuang dalam teks, umum nya analisis sentimen dilakukan dengan cara klasifikasi [5]. Klasifikasi merupakan teknik mengkategorikan dataset untuk melakukan analisis dan prediksi yang lebih akurat dalam *data mining* [6]. Untuk melakukan klasifikasi emosi pada data media sosial yang besar dan beragam, metode pembelajaran mesin dengan pohon keputusan seperti *Random Forest* dan *Decision Tree* banyak dipakai karena fleksibilitas dan akurasinya dalam menangani data yang kompleks [7].

Studi ini memiliki tujuan agar dapat mengembangkan model klasifikasi menggunakan *Random Forest* dan *Decision Tree* untuk analisis sentimen publik pada isu kenaikan PPN 12%. Dengan memanfaatkan komentar *YouTube* sebagai sumber data. Studi ini berfokus untuk mengidentifikasi komentar yang dikategorikan kedalam 12 jenis emosi, sekaligus membandingkan performa dari kedua model. Diharapkan penelitian ini dapat menggambarkan persepsi dan emosi masyarakat secara luas dan akurat terhadap kebijakan kenaikan PPN 12%.

2. METODE PENELITIAN

Dalam penelitian ini, tahapan nya dimulai dari pengumpulan data sebelum akhirnya mendapatkan kesimpulan. Penelitian ini memiliki alur yang digambarkan pada Gambar 1.



Gambar 1. Alur Penelitian

Gambar 1 menggambarkan alur penelitian yang dimulai dari pengumpulan data menggunakan teknik *web scrapping*, dilanjutkan tahap *preprocessing* yang terdiri dari *case folding, cleaning,* tokenisasi, normalisasi, *stopword, stemming,* dan *trasnlate*, selanjutnya labeling menggunakan *NRC Lexicon*, ekstraksi fitur menggunakan *TF-IDF*, pemodelan *RF* dan *DF*, dan terakhir evaluasi.

2.1. Pengumpulan Data

Dalam tahap ini, data dikumpulkan menggunakan metode *web scraping*. *Web scraping* merupakan suatu cara mengumpulkan informasi dan data dari internet secara otomatis, umumnya ini dilakukan pada *website* yang menggunakan yang berbasis *HTML* atau *XHTML* yang datanya dianalisis untuk kebutuhan dan tujuan tertentu [8]. *Web scraping* umum digunakan untuk pengumpulan konten, yang melibatkan perolehan konten situs *web* untuk berbagai tujuan, termasuk penambangan data, *mashup web*, dan integrasi data [9].

2.2. *Pre-processing*

Preprocessing merupakan teknik yang diterapkan pada dataset mentah untuk menghasilkan dataset yang terformat dan bebas dari kesalahan [10]. *Preprocessing* dalam penelitian ini meliputi 9 tahap, mulai dari *case folding, cleaning,* tokenisasi, normalisasi, *stopword, stemming,* dan *translate*. Namun tahap *preprocessing* selalu mengandung risiko menghilangnya informasi yang relevan atau variasi yang berkaitan dengan sifat yang diinginkan [11]. Berikut tahapan *preprocesssing* yang dilakukan:

a. *Case Folding*
   Pada tahap *case folding* dilakukan pengubahan semua huruf menjadi huruf kecil dan pada proses ini juga menghapus karakter, simbol, *URL*, dan angka yang tidak diperlukan pada dataset [12].
b. *Cleaning*
   Cleaning atau pembersihan teks adalah langkah praproses yang menghilangkan kata atau komponen lain yang tidak mengandung informasi relevan, sehingga dapat mengurangi efektivitas sentiment [13].
c. Tokenisasi
   Tokenisasi merupakan proses pemecahan teks menjadi potongan kecil atau token untuk di proses oleh model [14]. Tokenisasi didefinisikan sebagai proses membagi seluruh paragraf menjadi potongan kata yang lebih kecil dari sebuah paragraf [15].
d. Normalisasi
   Normalisai teks merupakan proses mengubah teks kedalam bentuk yang lebih konsisten sebelum analisis lebih lanjut, ini meliputi perbaikan ejaan, ekspansi singkatan, penyesuaian tanda baca, serta variasi linguistik lainnya [16]. Tujuannya untuk mengurangi variasi leksikal dan ortografis sehingga teks menjadi lebih mudah diproses, dianalisis, dan dipahami. Normalisasi dalam penelitian ini dilakukan 2 kali agar hasil lebih optimal.
e. *Stopword*
   Sesuai namanya, *stopword* yaitu tahap menghapus kata yang umim yang sering sekali muncul tapi tidak bermakna [17]. Dengan menghilangkan kata kata tersebut, kualitas data dalam korpus teks meningkat, sehingga istilah-istilah penting menjadi lebih menonjol dalam analisis.
f. *Stemming*
   Tahap ini merupakan proses mengurangi variasi kata dengan mengembalikan kata menjadi bentuk dasar nya, menghapus imbuhan seperti awalan atau akhiran untuk menyatukan kata kata dengan akar yang sama [18]. Dengan demikian, ukuran kosakata berkurang dan kata-kata serupa dapat diperlakukan sama, meningkatkan efektivitas algoritma analisis teks.
g. *Translate*
   Tahap ini melakukan translasi teks otomatis *(machine translation)* yang merupakan proses menerjemahkan teks dari satu bahasa ke bahasa lain secara otomatis [19].

2.3. *Labeling*

Setelah melakukan tahap *preprocessing,* berikutnya adalah labeling atau pelabelan pada dataset menggunakan *NRC Lexicon*. Kamus berbasis leksikon merupakan serangkaian kata yang terkait dengan kategori emosi yang telah ditentukan sebelumnya, yang dapat langsung digunakan untuk memberi label teks sesuai dengan makna emosional dari istilah istilahnya [20]. Dengan kata lain, teks dilabeli berdasarkan makna emosional dari kata-kata yang muncul dalam leksikon.

2.4. *TF-IDF*

Term Frequency-Inverse Document Frequency atau TF-IDF merupakan proses vektorisasi teks yang mengubah dokumen teks menjadi fitur numerik. Metode ini memberi bobot untuk setiap *term* dari seberapa sering *term* itu muncul pada sebuah dokumen dan seberapa tidak sering *term* tersebut muncul pada korpus [18]. Dengan demikian, *term* yang sering muncul dalam satu dokumen namun jarang di dokumen lain akan memiliki bobot lebih tinggi, sehingga *TF-IDF* efektif dalam menyoroti istilah yang informatif,

2.5. Pemodelan

Penelitian ini membandingkan dua algoritma dalam membuat model yaitu *Random Forest* dan *Decision Tree*. *Decision Tree* merupakan metode *machine learning* yang membuat model untuk prediksi dengan menyerupai seperti bentuk pohon. Setiap *node* internal merepresentasikan uji atribut, setiap cabang menunjukkan hasil dari pengujian itu, dan setiap daun menyatakan label kelas yang diprediksi. Algoritma ini menyusun *decision rules* secara hierarkis berdasarkan fitur-fitur data untuk menyimpulkan keputusan akhir. Struktur pohon ini intuitif dan mudah diinterpretasikan, menjadikannya menjadi metode yang banyak digunakan dalam klasifikasi teks maupun tabular [21].

Sedangkan *Random Forest* merupakan ansambel yang terdiri atas pohon keputusan yang banyak. Setiap pohon dibangun dari *subset* acak data *(bootstrap sampling)* dan hanya menggunakan subset acak fitur pada tiap pembelahan. Hasil prediksi akhir diperoleh berdasarkan *majority voting* di antara semua pohon dalam *forest*. Pendekatan ini berfungsi mengurangi risiko *overfitting* dan meningkatkan generalisasi model dibanding *Decision Tree* tunggal [22]. Berikut rumus dari kedua algoritma tersebut:

a. *Random Forest*

$$Forest = \{h(x, \theta_k), k = 1, \dots\} \quad (1)$$

Dengan,

h : Hipotesa atau klasifikasi
x : Input *vector*
θ_k : *Independent and identically random vectors*

Persamaan (1) menunjukkan bahwa algoritma *Random Forest* membangun sekumpulan pohon keputusan atau forest. Setiap pohon dibangun berdasarkan vektor acak yang independen dan identik (θ_k). Proses ini menghasilkan variasi dalam setiap pohon, sehingga prediksi akhir ditentukan melalui mekanisme voting atau agregasi dari seluruh pohon. Dengan demikian, *Random Forest* mampu menurunkan risiko *overfitting* dan menambah akurasi dibandingkan dengan satu pohon keputusan yang tunggal.

b. *Decision Tree*

$$Entropy(\mathrm{S}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} - \mathrm{pi} * \log_2(\mathrm{pi}) \quad (2)$$

$$Gain(\mathrm{S}, \mathrm{A}) = Entropy(\mathrm{S}) - \sum_{\mathrm{i}=1}^{\mathrm{n}} - \frac{|\mathrm{Si}|}{|\mathrm{S}|} * \mathrm{Entropy}(\mathrm{S}) \quad (3)$$

Dengan,

S : Himpunan kasus
n : Jumlah partisi S
pi : Jumlah kasus pada partisi ke- i
A|Si| : Jumlah kasus pada partisi ke- i
|S| : Jumlah kasus dalam S

Persamaan (2) menunjukkan konsep yang dipakai dalam mengukur keacakan pada suatu kumpulan data yaitu entropy. Semakin tinggi skor *entropy*, maka semakin acak distribusi datanya. Sedangkan persamaan (3) adalah *information gain* yang dihitung dengan mengurangi nilai *entropy* total dengan rata-rata *entropy* dari partisi-partisi yang terbentuk berdasarkan atribut A. Nilai *information gain* digunakan untuk memilih atribut terbaik dalam membangun cabang pohon. Dengan kata lain, *Decision Tree* bekerja dengan memilih atribut yang paling informatif pada setiap langkah pembentukan pohon, sehingga menghasilkan model yang dapat melakukan klasifikasi dengan struktur hierarkis yang mudah dipahami.

2.6. Evaluasi

Untuk mengevaluasi kinerja model yang digunakan paa studi ini, diterapkan sejumlah metrik evaluasi, yaitu *accuracy, recall, precision* dan *F1-score*. Metrik-metrik tersebut dipilih karena mampu memberikan gambaran yang komprehensif mengenai performa model, khususnya dalam membandingkan keseimbangan antara prediksi negatif dan positif. Studi ini juga menggunakan *confusion matrix* untuk memvisualisasikan hasil prediksi model terhadap data uji. *Confusion matrix* menyajikan informasi rinci tentang jumlah prediksi yang tepat dan keliru pada setiap kelas, sehingga dapat membantu dalam menganalisis pola kesalahan klasifikasi serta menilai konsistensi model dalam mengklasifikasikan data.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang terkumpul merupakan komentar dari video YouTube yang khusus membahas isu kenaikan PPN 12%. Hasil dari web scraping didapatkan data sebanyak 3832 yang disimpan dalam format *excel.* Tabel 1 menampilkan hasil dan atribut pada data yang telah dikumpulkan.

Tabel 1. Hasil *Web Scraping*

Author	***Comment***	***LikeCount***	***PublishedAt***
@MuhammadAryaZen	udah diulangin juga masih salah bangðŸ˜† lanjuuut	0	2025-01-07T01:05:15Z
@bujank93	Ga usah beli tiketnya guys, ingat 2025 bakal suram hemat2ðŸ˜,ðŸ˜,ðŸ˜,ðŸ˜,	0	2025-01-06T13:17:35Z
@andrichannel10	aman boss... PPN 12% hanya untuk Barang mewah	1	2025-01-01T07:41:01Z
@adnansuryanaputra8607	saya sih setuju ppn naik 12%, lho kok bisa?	0	2024-12-31T07:47:32Z
@twinslightart5274	setelah PPN 12 dijalankan ternyata fail ujungnya PRIVATISASI	0	2024-12-29T07:22:06Z

Tabel 1 menunjukkan 5 contoh hasil pengumpulan data menggunakan *web scraping*. Dapat dilihat secara *default* terdapat 4 atribut data yaitu *author, comment, likeCount,* dan *publishedAt.* Data yang dikumpulkan ini khusus berbahasa Indonesia yang memiliki rentang waktu dari 26 November 2024 sampai 8 Januari 2025.

3.2. Pre-processing

Sebelum data dianalisis lebih lanjut, dilakukan tahap preprocessing untuk memastikan kualitas dan konsistensi data komentar yang diperoleh. Tabel 2 berikut adalah hasil dari masing masing tahap preprocessing yang diambil dari salah satu contoh data dari Tabel 1.

Tabel 2. Hasil *Preprocessing*

Sebelum *Preprocessing*	**Setelah *Preprocessing***
aman boss... PPN 12% hanya untuk Barang mewah	['safe', 'boss', 'goods', 'luxury']

Tabel 2 menunjukkan data awal sebelum dilaukan tahap *preprocessing* pada kolom pertama dan hasil nya dapat dilihat pada kolom kedua.

3.2.1. Case Folding

Seluruh huruf pada data komentar diubah menjadi huruf kecil. Ini dilakukan untuk menyeragamkan teks sehingga tidak ada perbedaan antara kata yang sama namun berbeda penulisan huruf besar maupun kecil.

Tabel 3. *Case Folding*

Comment	***Case Folding***
aman boss... PPN 12% hanya untuk Barang mewah	aman boss... ppn 12% hanya untuk barang mewah

Dari Tabel 3 dapat dilihat bahwa kata "PPN" berhasil dirubah menjadi huruf kecil. Hal ini penring karena perbedaan huruf kapital dapat dianggap berbeda oleh sistem walaupun memiliki kata yang sama.

3.2.2. Cleaning

Tahap cleaning dilakukan dengan menghapus karakter yang tidak relevan seperti angka, tanda baca, URL, emotikon, maupun simbol tertentu. Proses ini membantu mengurangi noise dalam data sehingga hanya menyisakan kata-kata yang bermakna.

Tabel 4. Hasil *Cleaning*

Case Folding	***Cleaning***
aman boss... ppn 12% hanya untuk barang mewah	aman boss ppn hanya untuk barang mewah

Tabel 4 menunjukkan kata "…" dan "12%" sudah dihilangkan, ini dilakukan agar data lebih bersih dan tidak mengganggu proses analisis.

3.2.3. Tokenisasi

Pada tahap ini, data teks akan dipecah menjadi unit-unit kata atau token.

Tabel 5. Hasil Tokenisasi

Case Folding	**Tokenisasi**
aman boss ppn hanya untuk barang mewah	['aman', 'boss', 'ppn', 'hanya', 'untuk', 'barang', 'mewah']

Tabel 5 menunjukkan kata yang sebelumnya masih dalam bentuk kalimat diubah menjadi potongan kata yang dipisah dengan koma dan tanda petik.

3.2.4. Normalisasi

Pada tahap ini, dilakukan normalisasi untuk menyamakan kata tidak baku atau singkatan kembali ke bentuk baku.

Tabel 6. Hasil Normalisasi

Tokenisasi	**Normalisasi**
['aman', 'boss', 'ppn', 'hanya', 'untuk', 'barang', 'mewah']	['aman', 'bos', 'ppn', 'hanya', 'untuk', 'barang', 'mewah']

Tabel 6 menunjukkan ada sedikit perubahan pada bagian normalisasi yaitu pada kata "boss" yang diubah menjadi "bos".

3.2.5. Stopword

Tahap ini melakukan penghapusan kata yang dianggap tidak bermakna tapi sering muncul untuk menjaga kualitas data.

Tabel 7. Hasil *Stopword*

Normalisasi	***Stopword***
['aman', 'bos', 'ppn', 'hanya', 'untuk', 'barang', 'mewah']	['aman', 'boss', 'barang', 'mewah']

Tabel 7 menunjukkan beberapa kata seperti "ppn", "hanya", dan "untuk" dihilangkan dari dataset karena dianggap tidak berbobot.

3.2.6. Stemming

Dalam *stemming*, kata diubah ke bentuk dasarnya dengan menghilangkan imbuhan misalnya sisipan, awalan, dan akhiran. Ini dilakukan agar menyatukan kata dengan bentuk berbeda tapi asalnya dari akar yang sama, dengan demikian analisis dapat dilakukan lebih efektif.

Tabel 8. Hasil *Stemming*

Stopword	***Stemming***
['aman', 'bos', 'barang', 'mewah']	['aman', 'bos', 'barang', 'mewah']

Dari Tabel 8, sekilas tidak nampak perubahan pada kolom *stemming* karena data yang sudah di normalisasi kedua sudah bersih dan dalam bentuk yang baku, namun *stemming* tetap dilakukan karena pada data yang lain masih terdapat kata yang perlu di *stem*.

3.2.7. Translate

Tahap akhir adalah penerjemahan atau *translate* teks ke dalam bahasa Inggris. Proses ini diperlukan karena kamus NRC Lexicon yang digunakan dalam penelitian ini hanya tersedia dalam bahasa Inggris, sehingga data komentar perlu disesuaikan dengan bahasa kamus tersebut.

Tabel 9. Hasil *Translate*

Stemming	***Translate***
['aman', 'bos', 'barang', 'mewah']	['safe', 'boss', 'goods', 'luxury']

Tabel 9 menunjukkan hasil kata pada kolom stemming yang diterjemahkan kedalam Bahasa Inggris demi menyesuaikan kamus NRC Lexicon yang kamu nya juga berbahasa Inggris.

3.3. Pelabelan

Pada tahap pelabelan, penelitian ini menggunakan kamus emosi NRC Emotion Lexicon yang dapat diakses secara *open source* di *github*. NRC Lexicon merupakan kamus emosi berbahasa Inggris yang memetakan ribuan kata ke dalam kategori emosi dasar seperti *anger, fear, sadness, joy, trust, anticipation, surprise, disgust*, serta kategori *positive*, *negative* dan *neutral*.

Tabel 10. Hasil Pelabelan

Translate	***Emotion***
['safe', 'boss', 'goods', 'luxury']	positive

Tabel 10 menampilkan hasil pada salah satu data yang diberi label pada kolom *Emotion*. Proses pelabelan ini bertujuan untuk mengklasifikasikan setiap data ke dalam kelas tertentu sehingga dapat digunakan pada tahap pemodelan selanjutnya.

3.4. TF-IDF

Setelah data diberi label, tahap selanjutnya yaitu merepresentasikan teks ke dalam format numerik dengan metode TF-IDF. Metode ini menghitung bobot setiap *term* berdasarkan frekuensi kemunculannya pada suatu dokumen dibandingkan dengan keseluruhan koleksi dokumen. Tabel 11 berikut menampilkan hasil representasi data menggunakan *TF-IDF* dari lima dokumen komentar (D1–D5).

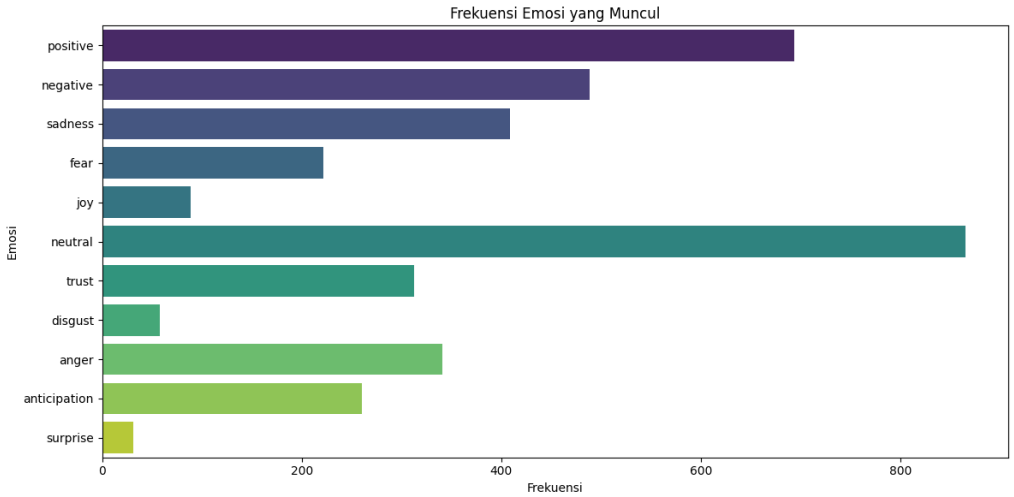
Tabel 11. Hasil TF IDF

Kata	D1	D2	D3	D4	D5
Already	0	0	0	0	0,2
Repeated	0	0	0	0	0,2
Wrong	0	0	0	0	0,2
Continued	0	0	0	0	0,2
Buy	0	0	0	0,175	0
Ticket	0	0	0	0,175	0
Gloomy	0	0	0	0,175	0
Saving	0	0	0	0,175	0
Safe	0	0	0,2	0	0
Boss	0	0	0,2	0	0
Goods	0	0	0,2	0	0
Luxury	0	0	0,2	0	0
Agree	0	0,233	0	0	0
You	0	0,233	0	0	0
Know	0	0,233	0	0	0
Run	0,175	0	0	0	0
Fail	0,175	0	0	0	0
Ends	0,175	0	0	0	0
Privatization	0,175	0	0	0	0

Tabel 11 menunjukkan bobot *TF-IDF* untuk setiap *term* pada lima dokumen komentar. Nilai bobot yang lebih tinggi menunjukkan bahwa *term* tersebut lebih penting atau lebih representatif pada dokumen tertentu dibandingkan dokumen lainny

3.5. Pemodelan

Setelah data sudah siap untuk dianalisis, tahap berikut nya adalah pemodelan menggunakan metode *Random Forest* dan *Decision Tree*. Pemodelan memiliki beberapa tahapan yaitu *split data*, *train test*, dan pembuatan model. Data yang akan di split sebanyak 3.772 dengan dengan rasio 80:20 untuk *training* dan *testing*. Berikut gambaran distribusi data yang akan digunakan untuk pembuatan model.



Gambar 2. Distribusi Data

Gambar 2 menunjukkan distribusi emosi pada dataset yang sudah diberi label, terlihat bahwa netral menjadi emosi yang paling dominan sebanyak 800 lebih, diikuti oleh positif, negatif dan *sadness*. Sementara *surprise* menjadi emosi yang paling sedikit sebanyak kurang dari 50. Untuk *Random Forest*, parameter yang digunakan adalah *n_estimators=450 criterion=entropy,* dan *random_state=42*. Sedangkan *Decision Tree* menggunakan parameter *cv=5, n_jobs=-1, verbose=1, max_depth: [None, 3, 5, 7, 10], min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 4]*.

3.6. Evaluasi

Tahap ini mengevaluasi model agar dapat mengukur kinerja klasifikasi sentimen yang telah dibangun. Pada penelitian ini, pengujian dilakukan berdasarkan empat metrik evaluasi, yaitu *accuracy, precision, recall, dan f1-score*. Berikut adalah *classification report* untuk model *Decision Tree*.

Tabel 12. Evaluasi *Decision Tree*

	Precision	***Recall***	***F1-Score***	***Support***
anger	0,494118	0,512195	0,502994	82
anticipation	0,583333	0,428571	0,494118	49
disgust	0,714286	0,555556	0,625000	9
fear	0,617647	0,437500	0,512195	48
joy	0,583333	0,437500	0,500000	16
negative	0,557895	0,504762	0,530000	105
neutral	0,622318	0,852941	0,719603	170
positive	0,533898	0,492188	0,512195	128
sadness	0,694915	0,561644	0,621212	73
surprise	0,333333	0,250000	0,285714	4
trust	0,465753	0,478873	0,472222	71
accuracy	0,573510	0,573510	0,573510	0,7351
macro avg	0,563712	0,501066	0,525023	755
weighted avg	0,572652	0,573510	0,565868	755

Tabel di atas menunjukkan hasil evaluasi kinerja model *Decision Tree* pada klasifikasi emosi menggunakan metrik *precision, recall, f1-score,* dan *support*. Secara umum, kategori *neutral* memiliki performa tertinggi dengan nilai *recall* 0,85 dan *f1-score* 0,71, yang menunjukkan model lebih mampu mengenali kelas ini dibanding kelas lainnya. Sebaliknya, kelas dengan jumlah data kecil seperti *surprise* menunjukkan nilai *f1-score* yang rendah, yaitu 0,28. Secara keseluruhan, model memperoleh skor akurasi sebesar 57,35% dengan nilai *weighted average f1-score* sebesar 0,56. Tabel 13 berikut menampilkan evaluasi utnuk *Random Forest* yang hasilnya sedikit berbeda dari model sebelumnya.

Tabel 13. Evaluasi *Random Forest*

	Precision	***Recall***	***F1-Score***	***Support***
anger	0,597015	0,487805	0,536913	82
anticipation	0,705882	0,489796	0,578313	49
disgust	0,833333	0,555556	0,666667	9
fear	0,718750	0,479167	0,575000	48
joy	1,000000	0,500000	0,666667	16
negative	0,765625	0,466667	0,579882	105
neutral	0,565657	0,988235	0,719486	170
positive	0,711712	0,617188	0,661088	128
sadness	0,680851	0,876712	0,766467	73
surprise	0,000000	0,000000	0,000000	4
trust	0,809524	0,478873	0,601770	71
Accuracy	0,654305	0,654305	0,654305	755
Macro Avg	0,671668	0,540000	0,577477	755
Weighted Avg	0,683938	0,654305	0,639905	755

Tabel 13 menunjukkan hasil evaluasi kinerja model *Random Forest*. Sama seperti sebelumnya, kelas *neutral* juga memiliki performa paling tinggi dengan nilai *recall* mencapai 0,98 dan *f1-score* sebesar 0,71, menandakan bahwa kedua model lebih mudah mengenali kelas ini dibanding emosi

lainnya. Selain itu, kelas *sadness* juga menunjukkan performa baik dengan *f1-score* sebesar 0,76. Sebaliknya, kelas *surprise* tidak berhasil dikenali oleh model dengan nilai *f1-score* 0,00. Secara keseluruhan, model *Random Forest* mendapatkan skor akurasi sebesar 65,43% dengan nilai *weighted average f1-score* sebesar 0,63.

4. KESIMPULAN

Jika dibandingkan dengan *Decision Tree*, performa yang sedikit lebih tinggi didapatkan model *Random Forest* dengan akurasi 65,43% dan *weighted average f1-score* sebesar 0,63, sedangkan *Decision Tree* hanya mencapai akurasi 57,35% dengan *weighted average f1-score* sebesar 0,56. Hal ini menandakan jika *Random Forest* lebih mampu mengklasifikasikan data secara konsisten, terutama pada kelas *neutral* dan *sadness* yang memiliki nilai *f1-score* relatif tinggi. Namun, kelemahan *Random Forest* terlihat pada kelas *surprise* yang sama sekali tidak terklasifikasi, berbeda dengan *Decision Tree* yang masih dapat mengenali sebagian kecil kelas tersebut. Dengan demikian, meskipun *Random Forest* memiliki keunggulan dalam akurasi keseluruhan, *Decision Tree* relatif lebih baik dalam mendistribusikan prediksi ke kelas kelas minor.

DAFTAR PUSTAKA

[1] E. Sihombing, M. Halmi Dar, and F. Aini Nasution, "Comparison of Machine Learning Algorithms in Public Sentiment Analysis of TAPERA Policy," *Int. J. Sci. Technol. Manag.*, vol. 5, no. 5, pp. 1089–1098, 2024, doi: 10.46729/ijstm.v5i5.1164.

[2] N. S. I. Al-agele, "A Social Media Sentiment Analysis Using Machine Learning Approaches," vol. 30, no. Ml, pp. 70–82, 2025.

[3] K. Munger, J. Phillips, and T. George, "Pressing Play on Politics : Quantitative Description of YouTube James Bisbee Omer Yalcin University of Massachusetts Amherst , USA Cardiff University , Wales Matthew Hindman," vol. 5, pp. 1–37, 2025.

[4] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.

[5] M. Taufiq Anwar, D. Riandhita Arief Permana, P. STMI Jakarta, P. Sistem Informasi Industri Otomotif, J. Letjen Suprapto No, and J. Pusat, "Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER," *Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 783–792, 2023, [Online]. Available: https://jurnal.mdp.ac.id/index.php/jatisi/article/view/3406/1173

[6] W. Apriliah, I. Kurniawan, M. Baydhowi, and T. Haryati, "Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest," *Sistemasi*, vol. 10, no. 1, p. 163, 2021, doi: 10.32520/stmsi.v10i1.1129.

[7] A. S. Aribowo, H. Basiron, N. F. A. Yusof, and S. Khomsah, "Cross-domain sentiment analysis model on indonesian youtube comment," *Int. J. Adv. Intell. Informatics*, vol. 7, no. 1, pp. 12–25, 2021, doi: 10.26555/ijain.v7i1.554.

[8] C. Slamet, R. Andrian, D. S. Maylawati, Suhendar, W. Darmalaksana, and M. A. Ramdhani, "Web Scraping and Naïve Bayes Classification for Job Search Engine," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, no. 1, 2018, doi: 10.1088/1757-899X/288/1/012038.

[9] Ф. Котлер *et al.*, "Analisis Struktur Kovarians Indikator Terkait Kesehatan pada Lansia yang Tinggal di Masyarakat, dengan Fokus pada Persepsi Kesehatan SubjektifTitle," *Accid. Anal. Prev.*, vol. 183, no. 2, pp. 153–164, 2023.

[10] Q. Alvi, S. F. Ali, S. B. Ahmed, N. A. Khan, M. Javed, and H. Nobanee, "On the frontiers of Twitter data and sentiment analysis in election prediction: a review," *PeerJ Comput. Sci.*, vol. 9, pp. 1–25, 2023, doi: 10.7717/peerj-cs.1517.

[11] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," *TrAC - Trends Anal. Chem.*, vol. 132, p. 116045, 2020, doi: 10.1016/j.trac.2020.116045.

[12] F. H. Nesa and A. C. Siregar, "Implementasi Data Mining untuk Klasifikasi Komentar Hate

Speech Menggunakan Algoritma Support Vector Machine (SVM) Implementation of Data Mining for Hate Speech Comment Classification Using Support Vector Machine (SVM) Algorithm," vol. 14, no. 105, 2025.

[13] A. Chamekh, M. Mahfoudh, and G. Forestier, "Sentiment Analysis Based on Deep Learning in E-Commerce," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13369 LNAI, pp. 498–507, 2022, doi: 10.1007/978-3-031-10986-7_40.

[14] S. Astuti, A. Erfina, and U. N. Putra, "Analisis Sentimen Pandangan Publik Terhadap Kenaikan Pajak 12 % dari Twitter Menggunakan Indonesian Roberta Base Classifier Sentiment Analysis of Public Views on 12 % Tax Increase From Twitter Using English Roberta Base Classifier," vol. 14, no. 105, pp. 698–706, 2025.

[15] B. Khemani and A. Adgaonkar, "A Review on Reddit News Headlines with NLTK tool," *SSRN Electron. J.*, pp. 1–5, 2021, doi: 10.2139/ssrn.3834240.

[16] A. A. Aliero, B. S. Adebayo, H. O. Aliyu, A. G. Tafida, B. U. Kangiwa, and N. M. Dankolo, "Systematic Review on Text Normalization Techniques and its Approach to Non-Standard Words," *Int. J. Comput. Appl.*, vol. 185, no. 33, pp. 44–55, 2023, doi: 10.5120/ijca2023923106.

[17] S. Sarica and J. Luo, "Stopwords in technical language processing," *PLoS One*, vol. 16, no. 8 August, pp. 1–13, 2021, doi: 10.1371/journal.pone.0254937.

[18] S. Perveen *et al.*, "Unsupervised fake news detection on social media using hybrid Gaussian Mixture Model," *PLoS One*, vol. 20, no. 8 August, pp. 1–26, 2025, doi: 10.1371/journal.pone.0330421.

[19] D. Ataman, A. Birch, N. Habash, M. Federico, and P. Koehn, "Machine Translation in the Era of Large Language Models : A Survey of Historical and Emerging Problems," pp. 1–36, 2025.

[20] A. Segura Navarrete, C. Martinez-Araneda, C. Vidal-Castro, and C. Rubio-Manzano, "A novel approach to the creation of a labelling lexicon for improving emotion analysis in text," *Electron. Libr.*, vol. 39, no. 1, pp. 118–136, 2021, doi: 10.1108/EL-04-2020-0110.

[21] A. S. Ahmed, A. A. A. Haddad, R. S. Hameed, and M. S. Taha, "An Accurate Model for Text Document Classification Using Machine Learning Techniques," *Ing. des Syst. d'Information*, vol. 30, no. 4, pp. 913–921, 2025, doi: 10.18280/isi.300408.

[22] H. Luzern, *Cyber Security Cyber Security*, no. March. 2017. doi: 10.1007/978-981-16-9229-1.