

Analisis Sentimen Terhadap Kontroversi Putusan MK Mengenai Usia Capres-Cawapres Menggunakan Multi-Layer Perceptron Dengan Teknik SMOTE

Sasmita¹, Reski Nurul Jariah S. Intam², Dewi Fatmarani Surianto³, Muhammad Fajar B^{4*}

^{1,2,3,4}Prodi Teknik Komputer, Jurusan Teknik Informatika dan Komputer, Universitas Negeri Makassar, Indonesia

Article Info

Article history:

Received Feb 22, 2024

Revised Apr 30, 2024

Accepted Jun 08, 2024

Keywords:

Analisis Sentimen

MLP

SMOTE

Putusan MK

ABSTRACT

In October 2023, the Constitutional Court's decision on age limit requirements for presidential and vice-presidential candidates stirred controversy, perceived as favoring a specific vice-presidential candidate. Public reactions flooded social media platforms, particularly on Najwa Shihab's YouTube channel, where sentiment analysis was conducted on 505 comments under the video titled "Putusan MK: Publik memang Seharusnya Marah" (Constitutional Court Decision: The Public Should Indeed Be Angry). The comments were categorized into three sentiment classes: 425 negative, 42 neutral, and 38 positive. The study employed Multi-Layer Perceptron (MLP) models tested on imbalanced and balanced data using the SMOTE oversampling technique. Two feature extraction methods, TF-IDF weighting and countvectorizer, were applied. Results showed that combining TF-IDF with balanced data yielded the most effective classification model, boasting a remarkable accuracy, precision, recall, and F1-score, each at 99%.

Copyright © 2024 Universitas Indraprasta PGRI.

All rights reserved.

Corresponding Author:

Muhammad Fajar B,

Teknik Komputer, Fakultas Teknik

Universitas Negeri Makassar

Jl. Jalan Daeng Tata Raya Parang Tambung, Mannuruki

Email: fajarb@unm.ac.id

1. PENDAHULUAN

Mahkamah Konstitusi (MK) merupakan lembaga penegak hukum tertinggi di negara Republik Indonesia. MK sendiri memiliki peran penting dalam menafsirkan dan menguji konstitusi serta perundang-undangan yang berlaku [1]. Pada tanggal 16 Oktober 2023, MK mengeluarkan putusan Nomor 90/PUU-XXI/2023 menyangkut persyaratan batas usia untuk menjadi calon presiden dan wakil presiden. Dalam hal ini, MK memutuskan bahwa seseorang yang berusia di bawah 40 tahun tetapi pernah menjabat menjadi kepala daerah, bisa mencalonkan sebagai presiden atau wakil presiden [2]. Keputusan ini berasal dari tuntutan yang diajukan oleh seorang mahasiswa dari Universitas Surakarta terkait masalah usia minimum untuk calon presiden dan calon wakil presiden, sebagaimana diatur dalam Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum [3].

Dari hasil putusan MK tersebut, Gibran Rakabuming Raka yang baru berusia 36 tahun dapat maju sebagai bakal calon wakil presiden. Beliau merupakan putra sulung dari Presiden sekaligus seorang Wali Kota Surakarta pada periode 2021. Namun, putusan ini dianggap kontroversial karena beberapa hal, salah satunya yaitu persepsi terkait adanya hubungan keluarga antara ketua MK dengan presiden. Sehingga, hal ini memunculkan persepsi bahwa MK melakukan pelanggaran kode etik hakim konstitusi dan dianggap terlibat konflik kepentingan dalam perkara pengeluaran putusan ini [4]. Kontroversial tersebut menjadi bahan pembicaraan di kalangan masyarakat baik secara langsung maupun melalui media sosial di berbagai platform.

Mengingat penggunaan media sosial semakin meningkat, dimana di Indonesia jumlahnya sebanyak 167 juta orang, dihitung per Januari 2023 [5]. Hal ini membuat masyarakat dapat dengan mudah mendapatkan berbagai macam informasi seperti isu-isu politik yang terjadi, salah satunya kontroversi keputusan MK yang sudah dijabarkan sebelumnya. Terkait isu tersebut, muncul berbagai macam opini masyarakat yang tertuang

dalam bentuk komentar di semua media sosial salah satunya adalah Youtube. Sejalan dengan itu, maka perlu dilakukan analisis sentimen terhadap opini masyarakat pada kolom komentar youtube mengenai isu tersebut.

Penerapan teknik *text mining* dalam melakukan analisis sentimen, data opini berupa teks dapat diolah menjadi sebuah informasi sesuai isu yang diangkat [6]. Dalam lingkup penelitian, telah banyak kajian sebelumnya yang secara menyeluruh membahas berbagai teknik pengolahan teks dan algoritma yang digunakan untuk mengklasifikasikan sentimen. Dimana pengolahan teks (*text mining*) merupakan penambangan data yang mengacu pada proses pengambilan informasi berkualitas tinggi dalam bentuk teks melalui penciptaan pola [7][8]. *Text mining* setara dengan analisis teks yang melibatkan pencarian informasi, analisis leksikal, distribusi frekuensi kata, pengenalan pola, ekstraksi informasi [9].

Salah satu model yang digunakan dalam melakukan klasifikasi sentimen adalah model *Artificial Neural Network* (ANN). Kelebihan dari model ini yaitu memberikan hasil klasifikasi yang lebih baik dibandingkan dengan teknik *machine learning* lainnya [10]. ANN dibagi menjadi dua kategori utama, yakni *single-layer* ANN yang melibatkan satu lapisan tersembunyi, dan *multi-layer* ANN yang melibatkan lebih dari satu lapisan tersembunyi [11].

Multi-Layer Perceptron (MLP) merupakan salah satu jenis algoritma ANN yang terkenal karena keakuratannya yang tinggi dalam melakukan klasifikasi data [12]. Keunggulan ini didukung oleh keberadaan lebih dari satu hidden layer dalam arsitekturnya, yang secara efektif meningkatkan kemampuan model dibandingkan dengan model *single-layer perceptron*. Pemanfaatan fungsi aktivasi pada algoritma ini memungkinkan proses iterasi penjumlahan bobot dan bias dari input, dengan tujuan untuk mengurangi kesalahan dalam langkah klasifikasi, sebelum mengeluarkan lapisan *output* [13].

Pada penelitian sebelumnya terdapat dua penelitian yang menggunakan algoritma *Backpropagation* menggunakan *learning rate* masing-masing bernilai 0,000002 dan 0,01 untuk melakukan pengklasifikasian terhadap analisis sentimen. Penelitian pertama yaitu yang dilakukan oleh [14], menghasilkan akurasi sebesar 71,630%. Sedangkan penelitian kedua yang dilakukan oleh [10], menghasilkan akurasi 86,4%.

Adapun penelitian berikutnya dengan menggunakan algoritma *Support Vector Machine* dalam mengklasifikasikan sentimen. Pada penelitian yang dilakukan oleh [15], mendapatkan akurasi sebesar 85,54%. Berikutnya, penelitian yang dilakukan oleh [16], mendapatkan akurasi sebesar 53,88% serta nilai *recall* 49,69%, *precision* 48,77%, *classification error* 46,12% dan *f-measure* 49,23%. Penelitian selanjutnya yang dilakukan oleh [17], dengan menggunakan algoritma SVM dan *Lexicon Based*, mendapatkan akurasi sebesar 60% dengan menggunakan *lexicon based* dan 48% tanpa menggunakan *lexicon based*.

Pada penelitian lainnya yang dilakukan oleh [18], dengan mengkombinasi antara algoritma *K-Means* dan *Naive bayes* mendapatkan akurasi sebesar 76,46%. Sedangkan pada penelitian yang dilakukan oleh [19], dengan membandingkan antara tiga algoritma dimana dengan menggunakan *Naive Bayes* mendapatkan akurasi sebesar 91,85%, SVM sebesar 91,82%, dan *Logistic Regression* sebesar 92,4%. Penelitian lain menggunakan *Long Short-Term Memory Networks* (LSTM) sebagai model klasifikasi serta penggunaan metode *word embedding* menghasilkan akurasi sebesar 81% [20]. Adapun penelitian lain dengan menggunakan model klasifikasi *Soft Voting* dalam melakukan analisis sentimen covid19 mendapatkan akurasi sebesar 91% [21].

Terdapat juga penelitian yang memanfaatkan algoritma *Naive Bayes* serta Teknik *Optimasi Information Gain* dan SMOTE. Hasil penelitian menunjukkan bahwa tingkat keakuratan mencapai 87,20%. Penerapan kedua teknik ini berhasil meningkatkan akurasi, *recall*, dan nilai *F1-score*. Secara rata-rata, terjadi peningkatan sebesar 6,25% untuk tingkat keakuratan, 23,9% untuk *recall*, dan 25,44% untuk nilai *F1-score* [22]. Penelitian lain menggunakan teknik SMOTE yang dilakukan oleh [23], dengan dua penggunaan model klasifikasi yaitu *Support Vector Machine* dan *Naive Bayes*. Penggunaan SMOTE pada penelitian ini menghasilkan keakuratan untuk *Naive Bayes* sebesar 88,54% dan *Support Vector Machine* sebesar 69,7%.

Sedangkan tiga penelitian sebelumnya yang menggunakan algoritma *Multilayer Perceptron*. Penelitian pertama yang dilakukan oleh [24], untuk melakukan analisis sentimen dengan membandingkan dua algoritma optimasi yaitu *Adam solver* dan *SGD solver*. Penggunaan parameter *Adam solver* menghasilkan tingkat akurasi sebesar 86%, sementara penggunaan parameter *SGD solver* menghasilkan tingkat akurasi sebesar 72%. Kemudian pada penelitian kedua yang dilakukan oleh [25], dengan membandingkan tingkat akurasi antara *MultiLayer Perceptron* dan *Random Forest*. Penggunaan algoritma *Multi-Layer Perceptron Classifier* menghasilkan akurasi 94,44%, sedangkan *Random Forest* menghasilkan akurasi 96,42%. Sedangkan penelitian ketiga yang dilakukan oleh [26], mendapatkan hasil akurasi 93,26% dalam mengklasifikasikan opini publik di *twitter* terhadap presiden Jokowi.

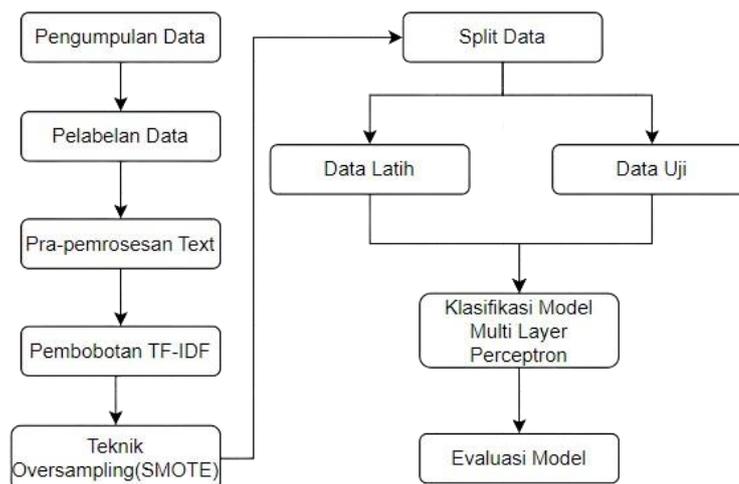
Namun dari beberapa penelitian sebelumnya, tingkat akurasi dari metode yang digunakan untuk klasifikasi masih perlu ditingkatkan. Kemudian jika terdapat data yang tidak seimbang pada dataset yang digunakan, maka perlu penambahan suatu teknik untuk mengatasi masalah tersebut sehingga hasil dan tingkat akurasi terhadap proses klasifikasi akan jauh lebih baik.

Oleh karena itu, penelitian ini menggunakan algoritma *Multi-Layer Perceptron* (MLP) dimana dinilai cukup baik dalam melakukan klasifikasi serta menghasilkan tingkat akurasi yang paling baik dibandingkan

dengan algoritma *machine learning* lainnya. Kemudian penelitian ini juga menggunakan teknik SMOTE untuk mengatasi data yang tidak seimbang pada dataset. Adapun tujuan dilakukan penelitian ini yaitu untuk mengetahui kinerja dari model MLP dengan melakukan pembobotan kata serta penggunaan teknik oversampling dalam menganalisis sentimen masyarakat terkait isu ini.

2. METODE

Pada Penelitian ini dilakukan beberapa tahap seperti yang ditunjukkan dalam alur penelitian di bawah ini. Ini mencakup pengumpulan data, pelabelan data, prapemrosesan, hingga sampai pada tahap analisis sentimen menggunakan model klasifikasi MLP. Langkah-langkah penelitian analisis sentimen ini ditunjukkan dalam gambar 1.



Gambar 1. Alur Tahapan Penelitian Analisis Sentimen

2.1. Pengumpulan Data

Tahap ini merupakan tahapan yang bertujuan untuk mengumpulkan data yang akan digunakan di dalam penelitian ini. data yang dikumpulkan merupakan komentar yang berasal dari video di kanal YouTube milik Najwa Shihab dengan judul "Putusan MK: Publik Memang Seharusnya Marah". Proses Pengumpulan data dilakukan dengan menggunakan *Website Scraper* yaitu netlytic.org [27]. Data yang terkumpul kemudian disimpan dalam format excel yang dimana akan dilakukan pengambilan kembali data komentar yang diperlukan secara manual. Sehingga, data yang berhasil dikumpulkan sebanyak 505 data. Adapun sampel data yang dikumpulkan dapat dilihat pada tabel 1.

Tabel 1. Sampel Data yang digunakan

No	Komentar
1	GIBRAN....CONTOH ANAK MUDA....YG RAKUS MANFAATKAN AYAH N PAMAN
2	Gak perlu marah, toh palunya sudah di Ketok, demo sekalipun gak berguna, mending pastikan 2024, Menangkan Akal Sehat
3	MAHKAMAH KASUR AMAT MENJIJIKKAN 😡😡😡👊
4	Publik mana yg marah !? Publik PDI-P kah !!! Biarkan Prabowo-Gibran maju agr bertarung di pemilu 2024 !!! Ahli2 hukum tolol !!! 😡
5	Takut hasil pemilu diobok obok lagi 😡,, kalo kamu berjiwa besar untuk kepentingan NKRI bran,, walaupun itu "LEGAL" kamu tidak akan maju bran,,

2.2. Pelabelan Data

Tahap ini merupakan tahap dilakukannya proses pelabelan data yang telah terkumpul dengan cara manual. Proses pelabelan ini dilakukan oleh tim peneliti yang merupakan dosen di Universitas Negeri Makassar dengan bidang keilmuan NLP (Natural Language Processing) dan bahasa. Terdapat 3 label yang digunakan yaitu label negatif, netral dan positif, dimana ketentuan yang digunakan pada masing-masing label yaitu, ketentuan label negatif adalah komentar yang berisi kata-kata yang mencaci maki keluarga presiden dan ketua MK, marah akan keputusan MK, ketentuan label netral adalah komentar yang berisi kata-kata netral akan keputusan MK, sedangkan label positif adalah komentar yang setuju akan keputusan MK. Dataset yang

digunakan pada penelitian ini terdiri dari 38 data dengan kelas positif, 42 data pada kelas netral, dan 425 data lainnya dengan kelas negatif . Berikut contoh pelabelan data komentar dapat dilihat pada tabel 2.

Tabel 2. Contoh Pelabelan Data

Label	Komentar
Negatif	Inilah KERUSAKAN HUKUM, KETATANEGARAAN. SUNGGUH MENGERIKAN !!! Saya tdk bisa membayangkan, setelah perhitungan suara pilpres nanti. Tuhan selamatkan INDONESIA
Netral	Yang marah cukup yang punya kepentingan aja,kalo gak punya kepentingan gak usah ikut ² an,mubazir gak dapat duit..
Positif	Kenapa harus menolak pemimpin muda yang berkualitas sudah saat nya negri ini dipimpin oleh orang-orang muda yg energik dan pintar seperti Gibran

2.3. Prapemrosesan Teks

Pada Tahap ini dilakukan tahap prapemrosesan teks terhadap data yang telah terkumpul, tujuan dari tahap ini yaitu untuk meminimalisir kesalahan yang terdapat pada data. Adapun tahapan prapemrosesan teks yang dilakukan dalam penelitian ini yaitu *Normalization-Cleansing-Case Folding-Stopword-Stemming*. *Normalization* adalah tahap mengubah kata yang semulanya tidak baku menjadi baku. *Cleaning* adalah tahap menghilangkan karakter khusus seperti titik (.), koma (,), tanda tanya (?), tanda seru (!), angka numerik (0-9), karakter non-abjad (\$, %, *), simbol "@" atau simbol "#", *Uniform Resource Locator* (URL), dan emoji. Semua karakter ini dihilangkan karena tidak memiliki pengaruh besar pada penentuan label. *Case-folding* adalah proses menyeragamkan semua huruf menjadi huruf kecil. *Stopword removal* adalah proses dimana kata sambung dan kata yang dianggap tidak memiliki makna dihapus. *Stemming* adalah proses dimana kata dengan imbuhan diubah menjadi kata dasar. Adapun hasil setiap tahap prapemrosesan teks dapat dilihat pada tabel 3.

Tabel 3. Contoh Proses Prapemrosesan Teks

Tahap Pra-Pemrosesan	Komentar
Data Awal	Bubarkan saja MK, terbukti Mahkamah Keponakan
Cleaning dan Normalisasi	Bubarkan saja MK terbukti Mahkamah Keponakan
Case Folding (lower)	bubarkan saja mk terbukti mahkamah keponakan
Tokenization	['bubarkan', 'saja', 'mk', 'terbukti', 'mahkamah', 'keponakan']
Stopword Removal	['bubarkan', 'mk', 'terbukti', 'mahkamah', 'keponakan']
Stemming	['bubar', 'mk', 'bukti', 'mahkamah', 'keponakan']
Hasil Bersih	bubar mk bukti mahkamah keponakan

2.4. Ekstraksi Fitur

Pada penelitian ini menggunakan dua ekstraksi fitur data yang telah dilabeli diantaranya penggunaan *countvectorizer* dan pembobotan TF-IDF. *Countvectorizer* digunakan untuk menghitung frekuensi kemunculan kata dalam dokumen. Sedangkan metode TF-IDF, perhitungan dilakukan untuk menentukan seberapa sering kata muncul dalam dokumen tertentu, serta mempertimbangkan seberapa umum kata tersebut muncul dalam dokumen [28]. Adapun rumus yang digunakan untuk menghitung TF-IDF yaitu :

$$TF\text{-}IDF_{t,d} = TF_{t,d} \times IDF_t \tag{1}$$

keterangan :

- TF-IDF_{t,d} = bobot dari kata t dalam satu dokumen
- TF_{t,d} = frekuensi kemunculan kata t dalam dokumen d
- IDF_t = *inverse document frequency*, dimana

$$IDF_t = \log\left(\frac{N}{DF_{t+1}}\right) \tag{2}$$

keterangan :

- N = banyaknya dokumen
- DF_t = banyaknya dokumen yang mengandung kata t

Berikut beberapa sampel dokumen yang diambil dari dataset untuk melakukan perhitungan bobot TF-IDF dapat dilihat pada tabel 4.

Tabel 4. Sampel Komentar untuk Ekstraksi Fitur TF-IDF

Dokumen	Komentar
D1	mahkamah keluarga mahkamah komedi malu masyarakat kecewa

D2	kentara mahkamah keluarga
D3	mahkamah konoha makelar

Setelah menentukan dokumen yang akan digunakan, maka langkah selanjutnya adalah melakukan proses perhitungan TF-IDF, dimana hasilnya dapat dilihat pada tabel 5.

Tabel 5. Hasil Perhitungan TF-IDF

Term	TF1	TF2	TF3	DF	IDF+1	TF-IDF 1	TF-IDF 2	TF-IDF 3
Mahkamah	0,29	0,33	0,33	0,29	1,00	0,29	0,33	0,33
Keluarga	0,14	0,33	0	0,14	1,18	0,17	0,39	0
Komedi	0,14	0	0	0,14	1,48	0,21	0	0
Malu	0,14	0	0	0,14	1,48	0,21	0	0
Masyarakat	0,14	0	0	0,14	1,48	0,21	0	0
Kecewa	0,14	0	0	0,14	1,48	0,21	0	0
Kentara	0	0,33	0	0	1,48	0	0,49	0
Konoha	0	0	0,33	0	1,48	0	0	0,49
Makelar	0	0	0,33	0	1,48	0	0	0,49

2.5. Teknik Oversampling (SMOTE)

Synthetic Minority Oversampling Technique (SMOTE) merupakan teknik yang digunakan dalam mengatasi permasalahan data yang tidak seimbang. Teknik ini akan melakukan peningkatan jumlah sampel data dari kelas minoritas dalam rasio kelas mayoritas. Dari proses tersebut akan menghasilkan lebih banyak fitur dari data latih sehingga membantu meningkatkan akurasi model[29].

Algoritma dari SMOTE ini pertama-tama akan memilih secara acak sebuah instansi kelas minoritas dari set pelatihan dan menemukan nilai K tetangga terdekatnya. Akhirnya, N (jumlah *oversampling* yang dibutuhkan) dari K instansi ini dipilih secara acak untuk menghitung instansi baru melalui interpolasi. Dalam melakukannya, selisih antara vektor fitur (sampel dan masing-masing tetangga yang dipilih diambil, kemudian selisih ini dikalikan dengan angka acak yang ditarik antara 0 dan 1, lalu ditambahkan ke vektor fitur sebelumnya. Hal ini menyebabkan pemilihan titik acak sepanjang "segmen garis" antara fitur. Dalam hal atribut nominal, salah satu dari dua nilai dipilih secara acak.

2.6. Pembagian Data

Pada tahap ini, terjadi pemisahan data menjadi dua kategori, yakni data latih dan data uji. Dimana proporsi 90% untuk data latih dan 10% untuk data uji. Penetapan data latih dan data uji 90:10 dipilih berdasarkan studi sebelumnya yang memperoleh tingkat akurasi tertinggi dengan menggunakan porsi data latih dan data uji tersebut [30]. Penelitian tersebut berhasil mengklasifikasi penderita penyakit hati dari 583 data dengan perolehan akurasi tertinggi yaitu 78.40% dibandingkan dengan perbandingan data yang lainnya. Selain itu, di sisi lain, pelatihan model MLP bisa menjadi komputasi yang mahal secara komputasi dan waktu, sehingga dengan menggunakan proporsi yang lebih kecil untuk data pengujian dapat menghemat waktu dan sumber daya komputasi. Hasil yang diperoleh data latih berfungsi sebagai materi pelatihan untuk mengajarkan algoritma yang telah dikembangkan, sementara data uji digunakan untuk mengukur sejauh mana akurasi algoritma yang telah dibuat.

2.7. Klasifikasi Data

Pada tahap ini dilakukan tahap pengimplementasian algoritma MLP, yang bertujuan untuk melakukan pengklasifikasian terhadap data. Dengan menggunakan metode ini, penelitian ini dapat menentukan model klasifikasi data komentar ke dalam salah satu dari ketiga kelas yang tersedia, berdasarkan ciri-ciri dari data komentar dari data latih. Penelitian ini dilakukan 2 skenario pengujian yaitu pengujian tanpa menggunakan SMOTE dan pengujian dengan menggunakan SMOTE. Hal ini dilakukan untuk membandingkan kinerja dari kedua skenario tersebut terhadap model yang digunakan. Adapun parameter yang digunakan dalam pengklasifikasian menggunakan MLP yaitu [24][26]:

1. Fungsi Aktivasi = Relu
2. *Hidden Layer* = 30
3. Maksimum Iterasi = 100

4. *Random State* = 42

2.8. Model Evaluasi

Evaluasi dilakukan untuk menentukan metode pembagian dataset dan klasifikasi yang dapat menghasilkan tingkat akurasi tertinggi. Dalam penelitian ini, model dievaluasi dengan mempertimbangkan *confussion matrix*. *Confussion matrix* merupakan salah satu metrik kinerja yang paling informatif dan dapat diandalkan untuk menilai sistem pembelajaran dengan berbagai kelas [31]. Adapun bentuk *confussion matrix* dapat dilihat pada tabel 6.

Tabel 6. Tabel *Confussion Matrix*

Data Aktual	Data Prediksi	
	TN FN	FP TP

Keterangan :

True Negative (TN) : banyaknya dokumen negatif yang diklasifikasi benar sebagai negatif

False Positive (FP) : banyaknya dokumen negatif yang diklasifikasi salah sebagai positif

False Negative (FN) : banyaknya dokumen positif yang diklasifikasi salah sebagai negatif

True Positive (TP) : banyaknya dokumen positif yang diklasifikasi benar sebagai positif.

Dari tabel *confussion matrix* ini, akan dievaluasi hasilnya dengan menghitung nilai akurasi, presisi, *recall*, dan *F1-score*. Akurasi mengukur sejauh mana model dapat memprediksi dengan benar, presisi mengukur ketepatan dari prediksi positif, *recall* mengukur sejauh mana model mampu mengenali kasus positif, dan *F1-score* adalah pengukuran gabungan antara presisi dan *recall*.

$$Akurasi = \frac{TN+TP}{TN+TP+FN+FP} \times 100\% \tag{3}$$

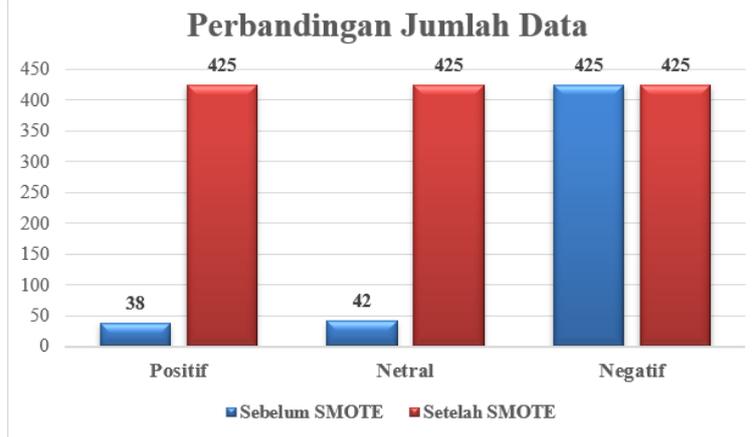
$$Presisi = \frac{TP}{TP+FP} \times 100\% \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{5}$$

$$F1 - Score = \frac{2}{1} \times \frac{presisi \times recall}{Presisi + recall} \times 100\% \tag{6}$$

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, dilakukan klasifikasi dengan menggunakan model *Multi Layer Perceptron (MLP)*. Adapun pada tahap klasifikasi dilakukan dua skenario pengujian yaitu pengujian tanpa menggunakan SMOTE dan pengujian dengan menggunakan SMOTE. Teknik SMOTE di sini bertujuan untuk menyelaraskan data yang sebelumnya tidak seimbang menjadi seimbang. Adapun perbandingan jumlah data sebelum dan sesudah proses SMOTE dapat dilihat pada gambar 2.



Gambar 2. Perbandingan Jumlah Data

Gambar 2 merupakan visualisasi perbandingan jumlah data sebelum dan setelah menggunakan SMOTE. Penggunaan SMOTE berhasil melakukan penyeimbangan jumlah data, sehingga seluruh kelas yang digunakan memiliki jumlah data yang sama yakni 425. Dimana total data sebelum menggunakan SMOTE sebanyak 505 data, kemudian menjadi 1275 data setelah menggunakan SMOTE.

Adapun sebelum masuk ke pengujian, dilakukan tahap ekstraksi fitur dan split data pada masing-masing pengujian. Ekstraksi fitur pada penelitian ini menggunakan dua pendekatan yaitu fitur *countvectorizer* dan fitur TF-IDF. Dua pendekatan ini digunakan untuk mengetahui seberapa besar pengaruh dari kedua fitur ini terhadap hasil klasifikasi. Kemudian untuk split data digunakan perbandingan 90:10 artinya 90% data training dan 10% data testing dari keseluruhan data yang ada.

3.1 Skenario pengujian tanpa menggunakan SMOTE

Pengujian ini menggunakan metode MLP dan dilakukan dengan menggunakan kumpulan data asli yang tidak seimbang. Proses ekstraksi fitur dilakukan terpisah menggunakan *Countvectorizer* dan TF-IDF. Adapun jumlah data testing yang digunakan pada pengujian ini sebanyak 57 data dari hasil pembagian. Kemudian data ini digunakan pada tahap evaluasi model klasifikasi dimana hasilnya berupa bentuk *confussion matrix*. Adapun hasil *confussion matrix* pada pengujian ini dapat dilihat pada tabel 7 dan tabel 8.

Tabel 7. *Confussion Matrix* menggunakan *Countvectorizer*

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	37	1	0
Netral	4	2	1
Positif	5	0	1

Tabel 7 menunjukkan hasil eksperimen dengan menggunakan ekstraksi fitur *countvectorizer*. Terdapat 38 data pada kelas negatif, di mana 37 data yang teridentifikasi dengan benar sebagai kelas negatif, dan 1 data lainnya keliru teridentifikasi sebagai kelas netral. Pada kelas netral yang memiliki 7 data uji, di antaranya hanya 2 data yang berhasil terdeteksi sebagai kelas netral, kemudian 4 data terdeteksi sebagai kelas negatif, dan 1 data lainnya terdeteksi sebagai kelas positif. Sementara pada kelas positif yang terdiri dari 6 data, hanya 1 data yang berhasil terdeteksi sebagai kelas positif, sedangkan 5 data lainnya terdeteksi sebagai kelas negatif.

Tabel 8. *Confussion Matrix*s menggunakan TF-IDF

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	37	1	0
Netral	3	3	1
Positif	5	0	1

Tabel 8 menunjukkan hasil eksperimen dengan menggunakan ekstraksi fitur TF-IDF. Hasil prediksi klasifikasinya yang hampir sama dengan hasil ekstraksi fitur *countvectorizer*, namun perbedaannya hanya di bagian kelas netral. Dimana data di kelas netral yang berjumlah 7 sebagai data uji, hanya 3 yang terprediksi benar diklasifikasi sebagai kelas netral, 3 data yang terdeteksi sebagai kelas negatif, dan 1 data yang terdeteksi sebagai kelas positif. Hal tersebut mengindikasikan bahwa prediksi benar dari penggunaan ekstraksi fitur TF-IDF lebih banyak dibandingkan dengan *countvectorizer* di bagian kelas netral.

Kedua tabel tersebut menunjukkan bahwa model mengalami *overfitting* pada data kelas mayoritas dibandingkan dengan kelas minoritas. Akibatnya, jumlah kesalahan prediksi pada kelas minoritas lebih tinggi dibandingkan kelas mayoritas. Tabel di bawah merupakan hasil evaluasi kinerja matrik dari masing-masing *confussion matrix* di atas yang menghasilkan nilai akurasi, presisi, recall, dan f1-score pada masing-masing kelas. Adapun hasilnya dapat dilihat pada tabel 9.

Tabel 9. Evaluasi Kinerja Matrik Pengujian Tanpa SMOTE

Ekstraksi Fitur	Kelas Sentimen	Akurasi	Presisi	Recall	f1-Score
Countvectorizer	Negatif	78%	80%	97%	88%
	Netral		67%	29%	40%
	Positif		50%	17%	25%
TF-IDF	Negatif	80%	82%	97%	89%
	Netral		75%	43%	55%
	Positif		50%	17%	25%

Tabel 9 merupakan hasil evaluasi kinerja dari masing-masing klasifikasi model MLP dengan penggunaan dua ekstraksi fitur. Berdasarkan nilai akurasinya, penggunaan ekstraksi fitur TF-IDF lebih tinggi dibandingkan dengan fitur *countvectorizer*. Namun, keduanya sama-sama mengalami *overfitting* dikarenakan adanya ketidakseimbangan data. Hal tersebut mengakibatkan model cenderung lebih baik dalam memprediksi kelas mayoritas karena ada lebih banyak contoh dari kelas tersebut dalam data pelatihan. Sehingga kinerja yang dihasilkan kelas negatif lebih tinggi karena merupakan kelas mayoritas dibandingkan dengan kelas netral dan positif.

3.2 Skenario Pengujian dengan menggunakan SMOTE

Pengujian selanjutnya menggunakan data yang seimbang dari penggunaan teknik SMOTE dengan ekstraksi fitur *countvectorizer* dan TF-IDF. Berikut merupakan hasil evaluasi dalam bentuk *confussion matrix* dari masing-masing ekstraksi fitur dapat dilihat pada tabel 10 dan 11.

Tabel 10. *Confusion Matrix* menggunakan *Countvectorizer*

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	32	3	7
Netral	2	43	3
Positif	0	12	26

Tabel 10 menunjukkan bahwa hasil klasifikasi dengan kombinasi SMOTE dan *Countvectorizer* pada model MLP ini tidak memberikan hasil yang optimal untuk setiap dataset. Masih terdapat banyak kesalahan prediksi sehingga hasil klasifikasinya kurang efektif.

Tabel 11. *Confussion Matrix* menggunakan TF-IDF

Data Aktual	Data Prediksi		
	Negatif	Netral	Positif
Negatif	41	1	0
Netral	0	48	0
Positif	0	0	38

Tabel 11 menunjukkan bahwa kombinasi SMOTE dan TF-IDF menghasilkan klasifikasi data yang akurat dimana hanya 1 data saja salah diklasifikasikan sebagai kelas negatif. Adapun hasil evaluasi matrik pada pengujian ini dapat dilihat pada tabel 12.

Tabel 12. Evaluasi kinerja Matrik Pengujian dengan SMOTE

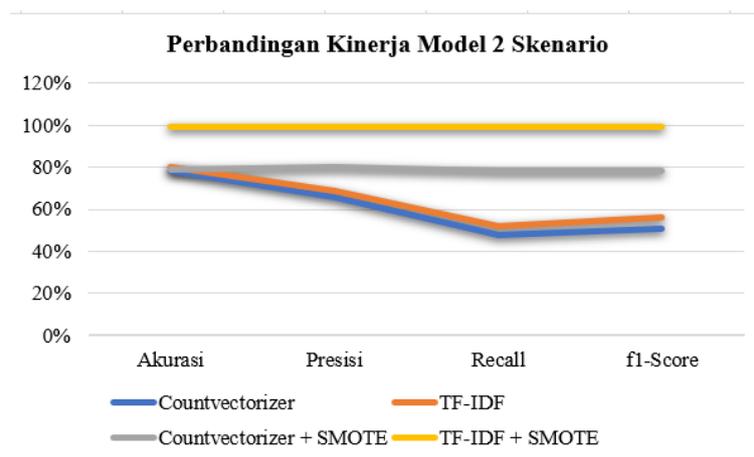
Ekstraksi Fitur	Kelas Sentimen	Akurasi	Presisi	Recall	f1-Score
<i>Countvectorizer</i>	Negatif		94%	76%	84%
	Netral	79%	74%	90%	81%
	Positif		72%	68%	70%
TF-IDF	Negatif		100%	98%	99%
	Netral	99%	98%	100%	99%
	Positif		100%	100%	100%

Pada tabel 12 membuktikan bahwa penggunaan kombinasi SMOTE dengan TF-IDF memiliki kinerja yang lebih baik dibandingkan kombinasi SMOTE dengan *countvectorizer*. Dimana dengan penggunaan TF-IDF menghasilkan nilai akurasi tertinggi yaitu 0,99. Hal ini membuktikan bahwa menggunakan TF-IDF memberikan hasil yang lebih baik dibandingkan dengan *countvectorizer* karena TF-IDF dapat memberikan bobot lebih besar pada kata-kata yang lebih informatif, sedangkan *countvectorizer* hanya menghitung frekuensi kata-kata. Dalam beberapa kasus, TF-IDF dapat menangkap informasi yang lebih penting. Adapun rata-rata hasil evaluasi kinerja model MLP dari dua pengujian sebelumnya dapat dilihat pada tabel 13.

Tabel 13. Hasil rata-rata evaluasi model MLP

Klasifikasi	Ekstraksi Fitur	Akurasi	Presisi	Recall	f1-Score
Tanpa SMOTE	<i>Countvectorizer</i>	78%	66%	48%	51%
	TF-IDF	80%	69%	52%	56%
Dengan SMOTE	<i>Countvectorizer</i>	79%	80%	78%	78%
	TF-IDF	99%	99%	99%	99%

Berdasarkan kinerja rata-rata hasil klasifikasi model MLP dengan menggunakan teknik SMOTE lebih tinggi dibandingkan tanpa menggunakan SMOTE, hal disebabkan karena teknik SMOTE berfungsi dalam menyeimbangkan data yang tidak seimbang. Dalam hal ini kelas netral dan positif yang sebelumnya merupakan kelas minoritas akan seimbang datanya dengan kelas mayoritas yaitu kelas negatif, sehingga bisa meningkatkan akurasi model. Adapun dua penggunaan ekstraksi fitur yang digunakan pada masing-masing pengujian, dapat dilihat bahwa penggunaan TF-IDF lebih unggul dalam meningkatkan kinerja dari model MLP di sini. Pada gambar 3 merupakan hasil visualisasi dari perbandingan kinerja dari dua pengujian yang telah dilakukan.



Gambar 3. Perbandingan Kinerja Model

Dari gambar 3 dapat dilihat bahwa grafik dengan pengujian tanpa smote garisnya dari seakan menurun dari nilai akurasi yang tinggi namun menurun pada nilai presisi, recall, dan f1-score. Hal ini biasa terjadi dikarenakan terdapat dataset yang tidak seimbang, sehingga model cenderung memprediksi kelas mayoritas untuk mencapai akurasi tinggi. Model dapat dengan mudah mencapai akurasi tinggi dengan memprediksi kelas mayoritas sebagian besar waktu, tetapi ini tidak mencerminkan kemampuan model untuk membedakan antara kelas yang berbeda. Berbeda dengan grafik dengan pengujian dengan bantuan SMOTE, visualisasi yang dihasilkan sangat stabil. Hal ini dikarenakan data yang digunakan sudah seimbang sehingga model dapat memprediksi dengan baik untuk setiap kelas setelah adanya *oversampling* menggunakan SMOTE.

Dari hasil evaluasi matrik yang diperoleh, dapat disimpulkan bahwa kinerja dari model MLP dengan melakukan pembobotan TF-IDF serta penggunaan teknik *oversampling* SMOTE dapat mengklasifikasi sentimen terkait isu ini dengan sangat baik. Di mana eksperimen ini mampu menghasilkan akurasi sebesar 99% pada dataset terkait. MLP digunakan sebagai model klasifikasi, sedangkan teknik *oversampling* SMOTE dan pembobotan TF-IDF mampu memaksimalkan pengklasifikasian yang diperoleh oleh model MLP sehingga mampu menghasilkan kinerja yang maksimal.

4. PENUTUP

Penelitian ini menganalisis kinerja dari model *Multi Layer Perceptron* (MLP) untuk melakukan klasifikasi sentimen masyarakat dari komentar youtube dari channel Najwa Shihab dengan judul “Putusan MK : Publik memang Seharusnya Marah”. Hasil yang diperoleh menunjukkan bahwa kinerja model pada data yang seimbang lebih baik dibandingkan dengan data tidak seimbang. Selain itu, penggunaan ekstraksi fitur TF-IDF cenderung meningkatkan akurasi klasifikasi dibandingkan dengan menggunakan perhitungan *countvectorizer* saja. Dimana pada penelitian ini, kombinasi teknik SMOTE dengan fitur TF-IDF menghasilkan kinerja terbaik dengan nilai akurasi mencapai 99%, presisi 99%, recall 99%, dan f1-score 99%.

Untuk penelitian selanjutnya, disarankan untuk menggunakan model klasifikasi, penggunaan ekstraksi fitur, serta penambahan komponen lainnya untuk membandingkan performa model klasifikasi. Tidak hanya itu, sumber data dari berbagai media sosial lain seperti Tiktok, Instagram, Facebook, maupun halaman berita dapat digunakan sebagai variasi penelitian selanjutnya. Adapun untuk mengurangi kemungkinan model *overfitting* pada saat pelatihan, maka disarankan untuk menggunakan data yang seimbang.

DAFTAR PUSTAKA

- [1] S. Anggara, “Buku Sistem Politik Indonesia,” in *Sistem Politik Indonesia*, Bandung: CV Pustaka Setia, 2013.
- [2] “MK Putuskan Syarat Usia Capres-Cawapres 40 Tahun Inkonstitusional Bersyarat,” *Kompas.com*, 2023. <https://nasional.kompas.com/read/2023/10/16/15300701/mk-putuskan-syarat-usia-capres-cawapres-40-tahun-inkonstitusional-bersyarat> (accessed Oct. 25, 2023).
- [3] Mahkamah Konstitusi Republik Indonesia, “Pengujian Materiil Undang-Undang Nomor 7 Tahun 2017 tentang Pemilihan Umum.” Jln. Medan Merdeka Barat No. 6, Jakarta Pusat, 2023.
- [4] F. Pratama, “Beragam Tanggapan soal Putusan MK Mengenai Batas Usia Capres/Cawapres,” *detikNews*, 2023. <https://news.detik.com/berita/d-6987142/beragam-tanggapan-soal-putusan-mk-mengenai-batas-usia-capres-cawapres> (accessed Oct. 25, 2023).

- [5] Hermila, S. Ayu Ashari, R. R. Taufik Bau, and S. Suhada, "Eksplorasi Intensitas Penggunaan Sosial Media (Studi Deskriptif Pada Mahasiswa Teknik Informatika UNG)," *Invert. J. Inf. Technol. Educ.*, vol. 3, no. 2, 2023, [Online]. Available: <http://ejurnal.ung.ac.id/index.php/inverted>
- [6] F. A. Wenando, R. Hayami, and A. J. Anggrawan, "Analisis Sentimen Pada Pemerintahan Terpilih Pada Pilpres 2019 Ditwitter Menggunakan Algoritme Naïvebayes," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 7, no. 1, pp. 101–106, 2020, doi: 10.33330/jurteksiv7i1.851.
- [7] M. R. Shinde and P. C. Gill, "Pattern Discovery Techniques for the Text Mining and its Applications," *Int. J. Sci. Res. ISSN*, vol. 3, no. 5, pp. 1660–1664, 2012, [Online]. Available: www.ijsr.net
- [8] A. Macanovic, "Text mining for social science – The state and the future of computational text analysis in sociology," *Soc. Sci. Res.*, vol. 108, no. August, p. 102784, 2022, doi: 10.1016/j.ssrresearch.2022.102784.
- [9] E. M. O. N. Haryanto, A. K. A. Estetikha, and R. A. Setiawan, "Implementasi Smote Untuk Mengatasi Imbalanced Data Pada Sentimen Analisis Sentimen Hotel Di Nusa Tenggara Barat Dengan Menggunakan Algoritma Svm," *J. Inf. Interaktif*, vol. 7, no. 1, pp. 16–20, 2022.
- [10] Niharika and S. Malhotra, "Analyzing the Sentiments with Neural Network," *J. Sci. Res.*, vol. 65, no. 01, pp. 266–272, 2021, doi: 10.37398/jsr.2021.650134.
- [11] S. Sarosa, *Analisis Data Penelitian Kualitatif*. PT Kanisius, 2021. doi: 10.1080/10916466.2018.1425717.
- [12] A. F. Hardiyanti and D. Fitriana, "Perbandingan Algoritma C4.5 dan Multilayer Perceptron untuk Klasifikasi Kelas Rumah Sakit di DKI Jakarta," *J. Telekomun. dan Komput.*, vol. 11, no. 3, p. 198, 2021, doi: 10.22441/incomtech.v11i3.10632.
- [13] J. T. Samudra and B. H. Hayadi, "Comparison of Adam's Optimization Function and Stochastic Gradient Descent on Bad Credit Classification of Savings and Loan Cooperatives Using Multilayer Perceptron," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 7, no. 2, p. 435, 2022, doi: 10.24114/cess.v7i2.35210.
- [14] I. K. A. G. Wiguna, P. Sugiartawan, I. G. I. Sudipa, and I. P. Y. Pratama, "Sentiment Analysis Using Backpropagation Method to Recognize the Public Opinion," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 16, no. 4, p. 423, 2022, doi: 10.22146/ijccs.78664.
- [15] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J. Inform.*, vol. 8, no. 2, pp. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.
- [16] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.35957/jatisi.v7i1.196.
- [17] N. E. Oktaviana, Y. A. Sari, and I. Indriati, "Analisis Sentimen terhadap Kebijakan Kuliah Daring Selama Pandemi Menggunakan Pendekatan Lexicon Based Features dan Support Vector Machine," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, pp. 357–362, 2022, doi: 10.25126/jtiik.2022925625.
- [18] V. R. Prasetyo, G. Erlangga, and D. A. Prima, "Analisis Sentimen untuk Identifikasi Bantuan Korban Bencana Alam berdasarkan Data di Twitter Menggunakan Metode K-Means dan Naive Bayes," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 5, pp. 1055–1062, 2023, doi: 10.25126/jtiik.20231057077.
- [19] K. W. Gusti, "Klasifikasi Bencana Alam Pada Twitter Menggunakan Naïve Bayes, Support Vector Machine Dan Logistic Regression," *Technol. J. Ilm.*, vol. 14, no. 4, p. 349, 2023, doi: 10.31602/tji.v14i4.11614.
- [20] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [21] N. Agustina and C. N. Ihsan, "Pendekatan Ensemble Untuk Analisis Sentimen Covid19 Menggunakan Pengklasifikasi Soft Voting," vol. 10, no. 2, pp. 263–270, 2023, doi: 10.25126/jtiik.2023106215.
- [22] H. Hidayatullah, P. Purwantoro, and Y. Umaidah, "Penerapan Naïve Bayes Dengan Optimalisasi Information Gain Dan Smote Untuk Analisis Sentimen Pengguna Aplikasi Chatgpt," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, 2023.
- [23] D. Rosiyadi, "Klasifikasi Komentar Instagram Untuk Identifikasi Keluhan Pelanggan Jasa Pengiriman Barang Dengan Teknik SMOTE," vol. 12, no. 4, pp. 280–290, 2019, doi: 10.30998/faktorexacta.v12i4.4981.
- [24] M. E. Johan and S. A. Azka, "Prediction of Alleged Stress Symptoms based on Indonesian Sentiment Lexicon using Multilayer Perceptron," *G-Tech J. Teknol. Terap.*, vol. 7, no. 3, pp. 958–966, 2023.
- [25] J. Asian, M. Dholah Rosita, and T. Mantoro, "Sentiment Analysis for the Brazilian Anesthesiologist Using Multi-Layer Perceptron Classifier and Random Forest Methods," *J. Online Inform.*, vol. 7, no. 1, p. 132, 2022, doi: 10.15575/join.v7i1.900.

- [26] N. Munasatya and S. Novianto, "Natural Language Processing untuk Sentimen Analisis Presiden Jokowi Menggunakan Multi Layer Perceptron," *Techno.Com*, vol. 19, no. 3, pp. 237–244, 2020, doi: 10.33633/tc.v19i3.3630.
- [27] Netlytic, "Netlytic - social media text and social networks analyzer," 2016. <https://netlytic.org/index.php?logon> (accessed Oct. 30, 2023).
- [28] N. Silalahi and G. Leonarde Ginting, "Rekomendasi Berita Berkaitan dengan Menerapkan Algoritma Text Mining dan TF-IDF," *Bull. Comput. Sci. Res.*, vol. 3, no. 4, pp. 276–282, 2023, doi: 10.47065/bulletincsr.v3i4.266.
- [29] V. Rupapara, F. Rustam, H. F. Shahzad, A. Mehmood, I. Ashraf, and G. Y. U. S. Choi, "Impact of SMOTE on Imbalanced Text Features for Toxic Comments Classification Using RVVC Model," vol. 9, pp. 78621–78634, 2021, doi: 10.1109/ACCESS.2021.3083638.
- [30] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, "Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.
- [31] S. Koço and C. Capponi, "On multi-class classification through the minimization of the confusion matrix norm," *J. Mach. Learn. Res.*, vol. 29, pp. 277–292, 2013.