METODE NAÏVE BAYES UNTUK KLASIFIKASI SENTIMEN *TWEET* PEMAIN NATURALISASI TIM NASIONAL SENIOR SEPAK BOLA INDONESIA

Dito Dang Kurniawan, Asriyanik, Didik Indrayana
Prodi Teknik Informatika, Universitas Muhammadiyah Sukabumi
Jl. R. Syamsudin, SH. No. 50 Kota Sukabumi Jawa Barat Indonesia 43113
ditodangkurniawan03@ummi.ac.id

ABSTRAK

Pemain naturalisasi di Timnas Senior Sepak Bola Indonesia sering menjadi perbincangan di media sosial, khususnya di X Permasalahan yang muncul berkaitan dengan beragamnya persepsi di kalangan penggemar, yang sering kali memunculkan perbedaan pendapat dan meningkatkan polarisasi opini. Beragam opini muncul, baik positif, negatif, maupun netral, sehingga diperlukan metode otomatis untuk mengklasifikasi sentimen secara efektif. Dalam penelitian ini, model *Naïve Bayes* diterapkan untuk mengklasifikasikan sentimen tweet terkait pemain naturalisasi. Metode yang digunakan adalah SEMMA (*Sample, Explore, Modify, Model, dan Assess*), yang mencakup tahapan pengambilan sampel, eksplorasi data, modifikasi, pembangunan model, serta evaluasi. Model ini mampu mengategorikan sentimen menjadi tiga jenis dengan akurasi 85%. Hasil evaluasi menunjukkan bahwa untuk sentimen negatif, *precision* mencapai 85%, *recall* 77%, dan *F1-score* 81%. Pada sentimen netral, *precision* sebesar 100%, *recall* 65%, dan *F1-score* 79%. Sementara itu, sentimen positif memiliki *precision* 82%, *recall* 98%, dan *F1-score* 89%. Hasil ini menunjukkan bahwa model memiliki performa yang baik dalam mendeteksi sentimen positif, negatif dan netral.

Kata kunci : *Klasifikasi Sentimen, Naive Bayes, SEMMA, Tweet*

1. PENDAHULUAN

Sepak bola merupakan olahraga dengan basis penggemar terbesar di Indonesia, dengan 69% responden mengidentifikasi diri sebagai penggemar berdasarkan survei Ipsos 2022, menjadikannya sebagai negara dengan proporsi penggemar sepak bola terbesar di dunia. Salah satu langkah untuk meningkatkan kualitas tim nasional adalah dengan mengadopsi kebijakan naturalisasi pemain asing. Namun, keputusan ini sering menimbulkan perdebatan di kalangan penggemar yang diekspresikan melalui platform media sosial seperti X.

Kehadiran pemain naturalisasi di Tim Nasional Indonesia sering menjadi topik perbincangan yang menimbulkan beragam reaksi dari publik, baik positif, negatif, maupun netral. Permasalahan utama yang muncul dari kebijakan naturalisasi adalah perbedaan persepsi di kalangan penggemar dan dampaknya terhadap hubungan antara tim nasional dan pendukungnya. Sebagian mendukung langkah ini sebagai solusi cepat untuk meningkatkan performa tim nasional, sementara yang lain menolak karena dianggap menghambat perkembangan pemain lokal.

Dengan semakin meluasnya penggunaan media sosial, X menjadi salah satu saluran utama bagi masyarakat untuk mengekspresikan opini terkait isu ini. Klasifikasi sentimen adalah proses untuk menilai keadaan suatu data berdasarkan perasaan atau opini yang terkandung di dalamnya [1].

Klasifikasi sentimen terhadap tweet yang berkaitan dengan pemain naturalisasi dapat memberikan gambaran bagaimana persepsi publik terhadap kebijakan tersebut berkembang seiring waktu, serta mengidentifikasi potensi faktor-faktor yang memengaruhi sentimen masyarakat. Oleh karena itu, pemahaman tentang sentimen publik sangat penting untuk menilai dampak dari kebijakan naturalisasi pemain terhadap hubungan antara penggemar dan tim nasional.

Model naive bayes terbukti dapat mengklasifikasi sentimen apalagi data yang berupa sebuah text. Hal itu dibuktikan dengan penelitian sebelumnya. Sebagai contoh, penelitian [2] menunjukkan bahwa Naïve Bayes memiliki akurasi sebesar 86%, lebih tinggi dibandingkan SVM yang mencapai 84%.

Penelitian lain [3] membandingkan Naïve Bayes dengan metode KNN dan Decision Tree, dengan hasil akurasi masing-masing sebesar 100%, 98.25%, dan 62.28%. Berdasarkan uraian tersebut, penulis memilih model naive bayes sebagai metode utama pada penelitian ini. Penelitian ini bertujuan untuk mengklasifikasikan sentimen publik terkait pemain naturalisasi Tim Nasional Indonesia melalui platform X. Dengan menggunakan dataset tweet yang menggambarkan berbagai reaksi masyarakat, mengklasifikasi ini akan mengelompokkan data ke dalam kategori sentimen positif, netral dan negatif. Melalui metode penelitian SEMMA (*Sample, Explore, Modify, Modell, dan Assess*). SEMMA adalah pendekatan yang mudah dimengerti dan digunakan sebagai panduan dalam proyek data mining [4].

Penelitian ini diharapkan dapat memberikan pemahamanyang lebih mendalam mengenai pandangan penggemar terhadap kebijakan naturalisasi dan dampaknya terhadap opini publik serta mengidentifikasi pola-pola sentimen yang dominan dalam diskusi di platform X.

2. TINJAUAN PUSTAKA

2.1. Klasifikasi Sentimen

Klasifikasi sentimen adalah proses penggunaan teknik pemrosesan bahasa alami (NLP), analisis teks, dan kecerdasan buatan untuk mengidentifikasi, mengekstrak, dan memahami opini atau emosi yang terkandung dalam suatu teks. Tujuan utama klasifikasi sentimen adalah menentukan apakah suatu opini bersifat positif, negatif, atau netral [5].

2.2. Naïve Bayes

Naïve Bayes adalah sebuah algoritma klasifikasi yang didasarkan pada Teorema Bayes, dengan asumsi bahwa setiap fitur bersifat independen satu sama lain [5].

Algoritma ini sering digunakan dalam berbagai aplikasi seperti klasifikasi teks, deteksi spam, analisis sentimen, dan lainnya karena kesederhanaannya serta efisiensinya dalam menangani dataset besar [6]. Naïve bayes dapat dijelaskan sebagai berikut [7]:

$$P(A|B)\frac{P(B|A).P(A)}{P(B)} \qquad (1)$$

Di mana:

a. P(A|B) adalah probabilitas hipotesis A terjadi jika diketahui kejadian B.
b. P(B|A) adalah probabilitas kejadian B jika diketahui A terjadi.
c. P(A) adalah probabilitas awal dari hipotesis A tanpa memperhatikan B.
d. P(B) adalah probabilitas dari kejadian BBB secara keseluruhan.

2.3. SEMMA

SEMMA adalah suatu proses yang digunakan dalam data mining untuk mengeksplorasi, memodelkan, dan menganalisis data guna menemukan pola yang bermakna [4].

SEMMA merupakan singkatan dari lima tahapan utama dalam proses ini, yaitu:

a. *Sample*
   Proses pengumpulan data untuk di analisis.
b. *Explore*
   Menganalisis data secara eksploratif untuk memahami pola dalam data.
c. *Modify*
   Memproses dan menyiapkan data, seperti transformasi dan pembersihan data.
d. *Modell*
   Menerapkan algoritma statistik atau pembelajaran mesin untuk membangun model prediktif atau deskriptif.
e. *Assess*
   Mengevaluasi kinerja model menggunakan metrik tertentu untuk memastikan akurasi dan keandalannya.

2.4. Tweet

Tweet adalah pesan singkat yang diposting di platform media sosial X (sebelumnya dikenal sebagai Twitter). Tweet dapat berisi teks hingga 280 karakter, gambar, video, gif, atau tautan. Pengguna dapat membalas, menyukai, membagikan (retweet), atau mengutip tweet orang lain. Tweet sering digunakan untuk berbagi pemikiran, berita, opini, atau informasi secara cepat dan luas [8].

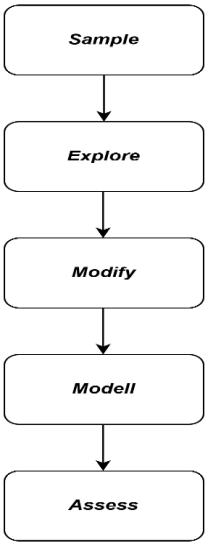
2.5. TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF). *Term Frequency* (TF) mengacu pada jumlah kemunculan suatu kata dalam dokumen [9].

Semakin sering kata tersebut muncul, semakin tinggi nilai TF-nya. Sedangkan *Inverse Document Frequency* (IDF) mengukur sejauh mana kelangkaan suatu kata dalam kumpulan dokumen. [10].

3. METODE PENELITIAN

Berikut metode penelitian yang digunakan penulis:



Gambar 1. Metode Penelitian

3.1. Sample

Merupakan tahap dalam pengumpulan data *tweet*. Pengumpulan data menggunakan *API X* dengan kata kunci "Pemain Naturalisasi", "#TimnasIndonesia", dan "#PemainNaturalisasi". Rentang waktu yang yaitu dari bulan januari hingga bulan oktober 2024.

3.2. Explore

Di tahap ini, dilakukan pelabelan manual dari setiap data tweet yang sudah didapatkan dengan label positif, netral dan negatif.

3.3. Modify

Pada tahap ini, dilakukan modifikasi data, agar dataset siap digunakan untuk pelatihan model. Tahap ini terdiri dari beberapa proses yaitu:

a. Cleaning
   Cleaning merupakan proses pembersihan data dari simbol-simbol, angka, ataupun tanda baca.

b. Case Folding
Mengubah semua data atau seluruh kalimat menjadi huruf kecil.
c. Stopword Removal
Membersihkan data dari kata yang tidak diperlukan seperti tanda penghubung.
d. Stemming
Merubah seluruh kata kedalam bentuk dasarnya seperti menghilangkan imbuhan.
e. Tokenize
Tokenize yaitu memisahkan kalimat menjadi kata per kata (token), yang biasanya dipisahkan dengan spasi dan tanda baca.

3.4. Modell

Di tahap ini, dilakukan beberapa proses yaitu:

a. Pembagian data
Data dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan perbandingan 80:20. Pembagian data dengan rasio 80:20 dipilih untuk menjaga keseimbangan antara pelatihan dan pengujian model. Dengan 80% data digunakan untuk pelatihan, model dapat mempelajari pola dengan lebih baik, sementara 20% data yang dialokasikan untuk pengujian cukup untuk menilai kinerjanya tanpa mengurangi keakuratan evaluasi [11].
b. Ekstraksi Fitur
Proses ini memanfaatkan *Term Frequency-Inverse Document Frequency* (TF-IDF).
c. Implementasi Model
Dilakukan pelatihan model naïve bayes dengan data latih dan selanjutnya model akan di uji dengan data uji.

3.5. Assess

Kinerja dari model. Untuk menilai kemampuan model dalam melakukan klasifikasi, dilakukan proses evaluasi. Confusion matrix dimanfaat kan dalam evaluasi ini, confusion matrix ini terdiri dari; akurasi, presisi, dan recall dari model yang ada. Berikut persamaannya [8]:

$$Akurasi = \frac{TP + TN}{TP + TP + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TPositif + FNegatif} \quad (4)$$

$$F1\ Score = 2 \times \frac{Recall \times Precision}{Precision + Recall} \quad (5)$$

4. HASIL DAN PEMBAHASAN

4.1. Sample

Sebanyak 870 data tweet yang telah dikumpulkan. Data yang telah dikumpulkan disimpan kedalam file dengan format *csv*.

Tabel 1. Sample

No	Tweet
1	Pemain naturalisasi untuk indonesia. Maju terus king indo dan terus menang #naturalisasi #TimnasIndonesia #garuda #Indonesia
2	Pokoknya timnas lebih hidup sejak ada pemain naturalisasi #timnasindonesia
3	Pemain lokal atau naturalisasi, selama mainnya oke ya nggak masalah. #TimnasIndonesia #naturalisasi #lokal

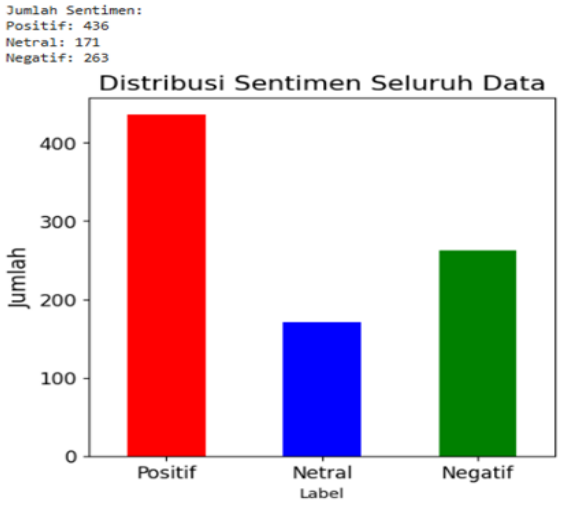
4.2. Explore

Selanjutnya, data yang telah dikumpulkan akan melalui tahap pelabelan manual dengan kategori positif, netral, dan negatif.

Tabel 2. Pelabelan manual

No	Tweet	Label
1	Pemain naturalisasi untuk indonesia. Maju terus king indo dan terus menang #naturalisasi #TimnasIndonesia #garuda #Indonesia	Positif
2	Pokoknya timnas lebih hidup sejak ada pemain naturalisasi #timnasindonesia	Negatif
3	Mending lokal boss, karna ada regenerasi #TimnasIndonesia #pemainnaturalisasi	Netral

Dari hasil pelabelan manual ditemukan sentimen positif sebanyak 436, netral 171, dan 263. Berikut visualisasi data dengan diagram batang:



Gambar 2. Visualisasi data

4.3. Cleaning

Tabel 3. Contoh hasil cleaning

Sebelum	Sesudah
Pemain naturalisasi untuk indonesia. Maju terus king indo dan terus menang #naturalisasi #TimnasIndonesia #garuda #Indonesia	Pemain naturalisasi untuk indonesia Maju terus king indo danterus menang

Dari tabel 3 diatas dapat dilihat seluruh tagar dan tanda bacanya dihilangkan.

4.4. Case Folding

Tabel 4. Contoh hasil case folding

Sebelum	Sesudah
Pemain naturalisasi untuk indonesia Maju terus king indo	pemain naturalisasi untuk indonesia maju terus king indo dan terus menang

Pada tabel 4 diatas huruf yang masih kapital diubah menjadi huruf kecil.

4.5. Stopword Removal

Tabel 5. Contoh hasil stopword removal

Sebelum	Sesudah
pemain naturalisasi untuk indonesia maju terus king indo menang	pemain naturalisasi untuk indonesia maju king indo terus menang

Tabel 5 bisa dilihat menghilangkan kata yang tidak dibutuhkan.

4.6. Stemming

Tabel 6. Contoh hasil stemming

Sebelum	Sesudah
pemain naturalisasi untuk indonesia maju king indo menang	main naturalisasi untuk indonesia maju king indo menang

Tabel 6 diatas bisa dilihat merubah kata kedalam bentuk dasarnya.

4.7. Tokenize

Tabel 7. Contoh hasil tokenize

Sebelum	Sesudah
main naturalisasi untuk indonesia maju king indo menang	['main', 'naturalisasi', 'untuk', 'indonesia', 'maju', 'king', 'indo', 'menang']

Dari tabel 7 bisa dilihat setiap kata di pisah kedalam bentuk kata perkata.

4.8. Pembagian Data

Berikut tabel distribusi sentimen dari data seluruh data yang sudah dibagi dengan rasio data latih 80% dan data uji 20%.

Tabel 8. Distribusi sentimen

Label Sentimen	Data Latih	Data Uji
Positif	349	87
Netral	137	34
Negatif	210	53
Total	**696**	**174**

4.9. Ekstraksi Fitur

Berikut contoh hasil ekstraksi fitur menggunakan TF-Idf.

Tabel 9. Sebelum ekstraksi fitur

Dokumen	Tweet
D1	['pemain', 'naturalisasi', 'indonesia', 'maju', 'king', 'indo']
D2	['main','naturalisasi','ngebawa', 'dampak', 'positif','timnas']
D3	['timnas', 'isi', 'main', 'naturalisasi', 'lokal', 'semua','tetap','wakil','indonesia','lapang']

Tabel 10. Sesudah ekstraksi fitur

Fitur	*D1*	*D2*	*D3*
dampak	0.0	0.484	0.0
indonesia	0.0	0.0	0.357
isi	0.0	0.0	0.357
kalau	0.435	0.0	0.0
kapan	0.435	0.0	0.0
kembang	0.435	0.0	0.0
lapang	0.0	0.0	0.357
lokal	0.331	0.0	0.271
main	0.257	0.286	0.211
naturalisasi	0.257	0.286	0.211
ngebawa	0.0	0.484	0.0
positif	0.0	0.484	0.0
semua	0.0	0.0	0.357
terus	0.435	0.0	0.0
tetap	0.0	0.0	0.357
timnas	0.0	0.368	0.271
wakil	0.0	0.0	0.357

4.10. Implementasi Model

Proses pelatihan dilakukan menggunakan data latih yang telah melalui tahap preprocessing dan pembobotan. Model Naïve Bayes menghitung probabilitas setiap kata dalam tweet terhadap masing-masing kelas sentimen positif, netral, dan negatif. Setelah probabilitas setiap kata dalam tweet dihitung, model Naïve Bayes membangun tabel probabilitas untuk setiap kelas sentimen. Setelah model selesai dilatih, pengujian dilakukan menggunakan data uji. Model memprediksi label sentimen berdasarkan probabilitas tertinggi dari masing-masing kelas sentimen. Berikut hasilnya.

Tabel 11. Hasil pengujian

Tweet	P Positif	P Netral	P Negatif	Hasil Label
D1	0.344	0.105	0.5502	Negatif
D2	0.772	0.061	0.1654	Positif
D3	0.257	0.477	0.2649	Netral

Dari Tabel 11 di atas, setiap tweet mendapatkan label sentimen berdasarkan nilai probabilitas tertinggi yang dihitung oleh model Naïve Bayes. Misalnya, untuk tweet D1, probabilitas terbesar adalah pada kelas negatif (0.5502), sehingga model memprediksi label sentimen D1 sebagai negatif.

4.11. Assess

Pada tahap assess dilakukan evaluasi kinerja model naive bayes. Evaluasi ini dilakukan dengan menghitung akurasi, precision, recall, dan F1-score.

```
Hasil Evaluasi:

Akurasi: 0.8505747126436781

Laporan Klasifikasi:
              precision    recall  f1-score   support

     Negatif       0.85      0.77      0.81        53
      Netral       1.00      0.65      0.79        34
     Positif       0.82      0.98      0.89        87

    accuracy                           0.85       174
   macro avg       0.89      0.80      0.83       174
weighted avg       0.86      0.85      0.85       174
```

Gambar 3. Hasil evaluasi

Hasil evaluasi model klasifikasi sentimen menunjukkan bahwa model memiliki akurasi sebesar 85.06%, yang berarti sekitar 85% data uji diklasifikasikan dengan benar. Untuk kelas Negatif, precision mencapai 0.85, recall 0.77, dan f1-score 0.81 dengan support 53, menunjukkan bahwa model cukup baik dalam mengenali sentimen negatif meskipun masih ada beberapa kesalahan dalam recall. Kelas Netral memiliki precision sempurna sebesar 1.00, tetapi recall yang lebih rendah (0.65), yang berarti model cenderung kurang mampu mengenali data yang sebenarnya netral, dengan f1-score sebesar 0.79 dari 34 data uji. Sementara itu, kelas Positif memiliki precision 0.82 dan recall yang tinggi sebesar 0.98, menghasilkan f1-score 0.89 dari 87 data uji, menunjukkan bahwa model sangat baik dalam mengenali sentimen positif. Secara keseluruhan, nilai macro average untuk f1-score adalah 0.83, sedangkan weighted average mencapai 0.85, yang mengindikasikan keseimbangan performa model terhadap distribusi data dalam dataset uji.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penerapan model Naïve Bayes untuk menklasifikasi sentimen tweet terkait pemain naturalisasi Timnas Senior Sepak Bola Indonesia, dapat disimpulkan bahwa model ini cukup efektif dalam mengklasifikasikan sentimen tweet. Dengan tingkat akurasi 85%, model berhasil dengan baik dalam mengidentifikasi sentimen positif, negatif, dan netral dari tweet yang berhubungan dengan pemain naturalisasi. Evaluasi menunjukkan bahwa model memiliki kemampuan baik dalam mendeteksi sentimen negatif dengan tingkat akurasi (precision) 85% dan kemampuan mendeteksi sentimen negatif yang benar (recall) 77%. Model juga mampu mendeteksi sentimen netral dengan akurasi sempurna (100%), meskipun recall-nya relatif rendah (65%).

Untuk sentimen positif, model memiliki kemampuan yang sangat baik dengan recall 98% dan precision 82%, yang menunjukkan efektivitasnya dalam mengenali tweet dengan sentimen positif. Secara keseluruhan, model Naïve Bayes menunjukkan kemampuan yang baik dalam mengklasifikasi sentimen tweet yang berkaitan dengan pemain naturalisasi Timnas Indonesia, meskipun masih ada kesempatan untuk meningkatkan performanya, terutama dalam mendeteksi sentimen netral. Disarankan untuk mencoba algoritma lain selain Naïve Bayes, seperti *Support Vector Machine* (SVM), *Random Forest*, atau deep learning. Selain itu, penggunaan dataset yang lebih besar dan beragam juga penting untuk meningkatkan performa model. Dengan data yang lebih banyak dan variasinya lebih luas, model dapat lebih baik dalam mengenali berbagai bentuk bahasa yang digunakan dalam tweet dan menangani beragam sentimen yang ada, sehingga model dapat menjadi lebih akurat dan dapat diterapkan pada berbagai kasus.

DAFTAR PUSTAKA

[1] T. Walasary, “Survey Paper tentang Analisis Sentimen,” *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 201–206, 2022, doi: 10.24002/konstelasi.v2i1.5378.

[2] C. F. Hasri and D. Alita, “Penerapan Metode Naã• Ve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter,” *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022, doi: 10.33365/jatika.v3i2.2026.

[3] M. K. Anam, B. N. Pikir, and M. B. Firdaus, “Penerapan Na¨ıve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen danPemeritah,” *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 1, pp. 139–150, 2021, doi: 10.30812/matrik.v21i1.1092.

[4] Y. A. Suwitono and F. J. Kaunang, “Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras,” *J. Komtika (Komputasi dan Inform.*, vol. 6, no. 2, pp. 109–121, 2022, doi: 10.31603/komtika.v6i2.8054.

[5] Heliyanti Susana, “Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet,” *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.

[6] E. B. Susanto, Paminto Agung Christianto, Mohammad Reza Maulana, and Sattriedi Wahyu Binabar, “Analisis Kinerja Algoritma Naïve Bayes Pada Dataset Sentimen Masyarakat Aplikasi NEWSAKPOLE Samsat Jawa Tengah,” *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 3, no. 3, pp. 234–241, 2022, doi: 10.37859/coscitech.v3i3.4343.

[7] H. Taufiqqurrahman, F. Tri Anggraeny, and M. Muharrom Al Haromainy, “Perbandingan

Algoritma Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Sentimen Ulasan Aplikasi Mypertamina," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3934–3939, 2024, doi: 10.36040/jati.v7i6.7801.

[8] N. S. Marga, "Sentimen Analisis Tentang Kebijakan Pemerintah Terhadap Kasus Corona Menggunakan Metode Naive Bayes," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 4, pp. 453–463, 2022, doi: 10.33365/jatika.v2i4.1602.

[9] K. Dalimunthe and B. H. Hayadi, "Information Text Retrieval Untuk Pencarian Data Penilaian Mengacu Pada Saran Dari Pengunjung Menggunakan Vector Space Modelimplementasi," *J. Comput. Sci. Inf. Technol. Progr. Stud. Teknol. Inf.*, vol. 3, no. 2, pp. 73–79, 2022, [Online]. Available: http://jurnal.ulb.ac.id/index.php/JCoInT/index

[10] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 6, no. 2, p. 129, 2021, doi: 10.30998/string.v6i2.10133.

[11] T. Menggunakan, M. T. Dan, and N. Bayes, "Jurnal Informatika Terpadu TWITTER MENGGUNAKAN METODE TF-IDF DAN NAIVE BAYES," vol. 10, no. 2, pp. 93–100, 2024.