

ANALISIS SENTIMEN OPINI MASYARAKAT TWITTER TERHADAP KASUS DIABETES DI INDONESIA MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER

Unsa Ahdannia Susilo^{1*}, Betha Nurina Sari², Adhi Rizal³

^{1,2,3}Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361; Telp. (0267) 64177

Keywords:

Diabetes Melitus; Analisis Sentimen; Twitter; Naïve Bayes

Correspondent Email:

2110631170114@student.unsika.ac.id

Abstrak. Penyakit diabetes melitus merupakan salah satu penyakit kronis yang prevalensinya terus meningkat di Indonesia, terutama akibat gaya hidup tidak sehat, termasuk konsumsi minuman kemasan tinggi gula. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap penyakit diabetes melalui opini yang diunggah pada platform media sosial Twitter. Metodologi penelitian yang digunakan adalah *Knowledge Discovery in Database* (KDD). Algoritma yang digunakan adalah *Naive Bayes Classifier* dengan dua kategori sentimen, yaitu positif dan negatif. Data dikumpulkan menggunakan Twitter API dan diproses melalui tahapan *preprocessing*, termasuk pembersihan teks, tokenisasi, dan *stemming*. Data yang semula berjumlah 4802 berkurang setelah dibersihkan menjadi 4369 data yang terdiri dari 2373 tweet label negatif dan 1996 tweet label positif. Pengujian dilakukan berdasarkan empat skenario pembagian data latih dan data uji, yaitu 90:10, 80:20, 70:30, dan 60:40. Hasil penelitian menunjukkan bahwa algoritma *Naive Bayes* mampu mengklasifikasikan sentimen secara efektif dengan akurasi sebesar 81%, precision 80%, recall 81%, dan *f-score* 80% pada skenario eksperimen 90:10. Temuan ini mengindikasikan bahwa sebagian besar opini masyarakat di Twitter sentimen negatif terhadap diabetes, mencerminkan tingginya kekhawatiran publik terhadap penyakit tersebut. Penelitian ini diharapkan dapat menjadi referensi bagi pihak terkait dalam merancang strategi edukasi dan pencegahan yang lebih tepat sasaran.

Abstract. *Diabetes mellitus is a chronic disease whose prevalence continues to rise in Indonesia, primarily due to unhealthy lifestyles, including the consumption of high-sugar packaged beverages. This study aims to analyze the sentiments of the Indonesian public toward diabetes based on opinions posted on the social media platform Twitter. The research methodology employed is Knowledge Discovery in Databases (KDD). The algorithm used is the Naive Bayes Classifier, with two sentiment categories: positive and negative. Data was collected using the Twitter API and processed through several preprocessing stages, including text cleaning, tokenization, and stemming. The original dataset of 4,802 tweets was reduced to 4,369 after cleaning, consisting of 2,373 tweets labeled as negative and 1,996 labeled as positive. The model was tested under four different training and testing data split scenarios: 90:10, 80:20, 70:30, and 60:40. The results show that the Naive Bayes algorithm effectively classified sentiments, achieving an accuracy of 81%, precision of 80%, recall of 81%, and F-score of 80% in the 90:10 experimental scenario. These findings indicate that most public opinions on Twitter express negative sentiments toward diabetes, reflecting a*

high level of public concern about the disease. This research is expected to serve as a reference for relevant stakeholders in designing more targeted education and prevention strategies.

1. PENDAHULUAN

Penyakit diabetes melitus merupakan penyebab kematian keenam di dunia dengan peningkatan prevalensi signifikan di Indonesia. Menurut IDF, pada 2021 jumlah penderita diabetes di Indonesia mencapai 19,47 juta, menjadikannya negara dengan kasus kelima terbanyak di dunia[1]. Fenomena ini memunculkan kekhawatiran masyarakat yang diekspresikan melalui Twitter, sebuah media sosial real-time yang digunakan luas sebagai sarana penyebaran opini publik [2].

Kasus meningkatnya konsumsi minuman berpemanis berkemasan yang dikaitkan dengan lonjakan penderita diabetes pada anak mendorong diskusi publik[3], terutama setelah viralnya isu gagal ginjal akut[4]. Respons masyarakat terhadap isu ini tercermin dalam opini dan pernyataan di Twitter. Hal ini relevan untuk dianalisis guna memahami kecenderungan sentimen publik terhadap diabetes.

Penelitian sebelumnya telah menerapkan Naive Bayes Classifier untuk analisis sentimen isu-isu kesehatan seperti HIV/AIDS dan stunting, menunjukkan performa yang cukup baik. Dengan demikian, penelitian ini akan menganalisis opini publik mengenai diabetes di Indonesia menggunakan data Twitter yang tujuannya untuk menerapkan algoritma Naive Bayes Classifier dalam mengklasifikasikan opini masyarakat Twitter menjadi sentimen positif dan negatif, serta mengevaluasi performa algoritma tersebut dalam menganalisis sentimen terhadap isu diabetes. Dengan pendekatan ini, diharapkan hasil penelitian dapat memberikan gambaran umum mengenai sikap masyarakat terhadap diabetes dan menjadi bahan pertimbangan dalam upaya peningkatan kesadaran akan bahaya konsumsi gula berlebih.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan penelitian berbasis komputer mengenai pendapat, perasaan, dan emosi yang terkandung

dalam suatu teks. Ulasan dan informasi tentang suatu produk disimpan dalam format teks, sehingga diperlukan metode tertentu untuk mengekstrak informasi tersebut. Salah satu metode dalam proses pengambilan informasi adalah Analisis Sentimen. Metode ini berfungsi untuk mengekstraksi opini, memahami, serta memproses data teks secara otomatis guna mengidentifikasi sentimen yang terkandung dalam suatu opini [5]. Analisis sentimen diperlukan sebagai alat untuk mengkategorikan informasi dalam bentuk teks ke dalam kelas sentimen positif dan negatif. Metode ini digunakan untuk mengidentifikasi kecenderungan suatu sentimen, apakah bersifat positif atau negatif [6].

2.2 Algoritma Naïve Bayes Classifier

Naïve Bayes Classifier adalah metode klasifikasi yang didasarkan pada teorema Bayes yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Naive Bayes menghitung probabilitas dari setiap kelas keputusan dengan asumsi bahwa kelas tersebut benar, berdasarkan vektor informasi objek. Algoritma ini beranggapan bahwa setiap atribut pada objek saling bebas dan tidak saling bergantung. Probabilitas yang digunakan untuk mendapatkan estimasi akhir dihitung dari total frekuensi pada tabel keputusan "master". Naive Bayes Classifier menunjukkan kinerja yang sangat baik dibandingkan dengan model klasifikasi lainnya [7]. Berikut persamaan Teorema Bayes sebagai berikut:

$$P(X) = \frac{P(X|H)P(H)}{P(X)}$$

Keterangan:

- X : Sample data yang label atau kategorinya belum diketahui
- H : Hipotesis bahwa data X termasuk ke dalam label tertentu
- P(X) : Probabilitas data sample

- $P(H)$: Probabilitas hipotesa H (priorprobability)
- $P(H|X)$: Probabilitas dari hipotesa H setelah diketahui data X (posteriori probability)
- $P(X|H)$: Probabilitas kemunculan data X jika diketahui hipotesis H benar

2.3 Pre-processing

Pre-processing merupakan tahapan penting dalam analisis data dan *machine learning* yang mengubah data mentah menjadi format terstruktur. Tujuannya adalah meningkatkan kualitas dan akurasi data agar siap digunakan dalam model pembelajaran mesin, sehingga menghasilkan dataset yang dapat diandalkan untuk analisis dan memastikan efektivitas model yang dikembangkan[8]. Tahapan pre-processing meliputi *Cleaning*, *Case folding*, *Tokenizing*, *Stopword Removal*, dan *Stemming*.

2.4 Term Frequency – Inverse Document Frequency (TF – IDF)

TF-IDF merupakan salah satu metrik yang efektif untuk menilai tingkat kepentingan suatu kata dalam sebuah teks atau kumpulan dokumen (korpus). TF-IDF disebut juga sebagai metode pembobotan yang memberikan nilai pada setiap kata berdasarkan frekuensi kemunculan kata dalam satu dokumen (tf) dan frekuensi kata yang jarang muncul (idf). Kata-kata yang memiliki skor bobot lebih tinggi dianggap lebih signifikan. Berikut merupakan rumus perhitungan TF-IDF[9].

2.5 Confusion Matrix

Confusion matrix adalah salah satu metode untuk mengevaluasi performa dari model klasifikasi yang telah dibangun. Confusion matrix terdiri dari empat komponen yang memuat nilai prediksi dan nilai aktual, diantaranya True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) yang merepresentasikan hasil klasifikasi [10]. Dibawah ini merupakan contoh dari *Confusion Matrix*.

Tabel 1. *Confusion Matrix*

<i>Confusion Matrix</i>		Nilai Prediksi	
		Positif	Negatif
Nilai Aktual	Positif	TP	FN
	Negatif	FP	TN

Keterangan :

TP (True Positive) : Jumlah data positif dengan prediksi positif

TN (True Negative): Jumlah data negatif dengan prediksi negatif

FP (False Positive) : Jumlah data negatif dengan prediksi positif

FN (False Negative): Jumlah data positif dengan prediksi negatif

2.6 Twitter

Twitter adalah salah satu platform media sosial yang masih banyak digunakan oleh masyarakat Indonesia untuk berbagi informasi dan mengungkapkan pendapat mengenai topik yang sedang tren di laman pribadi dengan batas maksimal 280 karakter[11]. Kemudahan dalam menganalisis data serta mendapatkan opini dari Twitter sangat membantu peneliti dalam melakukan penelitian tentang analisis sentimen, di mana data tersebut akan digunakan sebagai sumber informasi bagi mereka yang membutuhkan [12].

2.7 Python

Python merupakan bahasa pemrograman yang menggunakan sintaksis sederhana sehingga bahasa ini relatif mudah dipelajari dan dipahami saat dibaca. Dengan demikian, Python menempati posisi pertama dalam kategori "paling ingin dipelajari" selama lima tahun berturut-turut walaupun menurun di tahun 2021 dalam hasil survei besar pengembang stackoverflow[13]. Sebagai salah satu bahasa yang banyak dipilih oleh ilmuwan data, Python memiliki banyak pustaka yang menyederhanakan proses kerja dengan data sekaligus memperluas peluang. Visualisasi data adalah salah satu aspek penting dalam ilmu data yang digunakan untuk

mendapatkan ringkasan visual dari data. Dengan Python, memperoleh wawasan berharga dan memahami data tidak hanya mudah, tetapi juga sangat fleksibel, interaktif, dan dapat disesuaikan [14].

2.8 Diabetes Melitus

Diabetes melitus merupakan suatu penyakit metabolik kronis yang ditandai oleh tingginya kadar glukosa dalam darah (hiperglikemia), yang disebabkan oleh gangguan pada fungsi pankreas dalam memproduksi hormon insulin atau karena tubuh tidak dapat menggunakan insulin secara efektif (resistensi insulin). Diagnosis penyakit ini dapat dilakukan melalui pengukuran kadar glukosa dalam darah [15].

2.9 Google Collaboratory

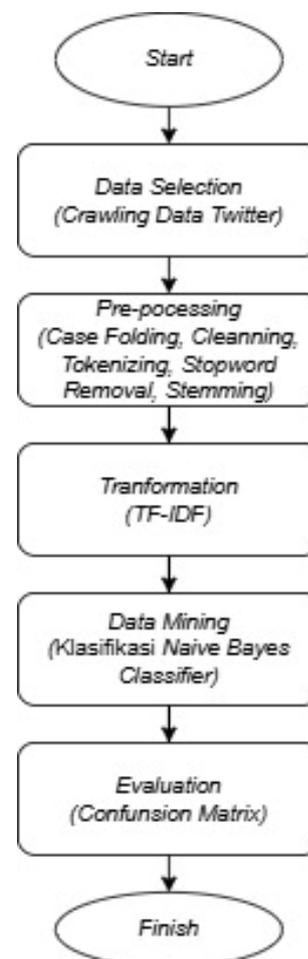
Google Colab adalah platform cloud computing yang disediakan oleh Google untuk mendukung pengembangan dan penelitian dalam bidang ilmiah. Dikenal sebagai Collaboratory atau "Colab," produk ini berasal dari Google Research. Colab memberikan kemudahan bagi pengguna untuk menulis dan menjalankan kode Python secara langsung melalui browser, menjadikannya ideal untuk machine learning, analisis data, serta pendidikan. Selain itu, Google Colab juga mendukung kerja sama antar pengembang aplikasi, sehingga sangat menguntungkan bagi kolaborasi tim. Meskipun demikian, penggunaan Google Colab masih belum luas dikalangan masyarakat umum, karena lebih banyak digunakan oleh individu yang membutuhkan lingkungan pemrograman, seperti para developer atau programmer [16].

3. METODE PENELITIAN

Penelitian ini bertujuan untuk menganalisis sentimen opini masyarakat Twitter terhadap kasus diabetes di Indonesia dengan menggunakan algoritma *Naive Bayes Classifier*. Objek penelitian berupa data teks dari pengguna media sosial Twitter pada Februari hingga Desember tahun 2024 menggunakan kata kunci "DiabetesMelitus", "PenyakitDiabetes", dan "Diabetes" dalam bahasa Indonesia. Data dikumpulkan melalui

proses crawling menggunakan token API Twitter dan bahasa pemrograman Python. Tweet yang diperoleh kemudian dilabeli secara manual dan dibantu oleh validator ahli Bahasa Indonesia ke dalam dua kategori sentimen, yaitu positif dan negatif.

Metode penelitian mengacu pada tahapan *Knowledge Discovery in Database (KDD)* yang terdiri dari lima langkah utama, yakni data selection, pre-processing, transformation, data mining, dan evaluation. Berikut rancangan penelitian yang disajikan pada gambar 1.



Gambar 1. Rancangan Penelitian

3.1 Data Selection

Tahap pertama adalah data selection, yaitu pengambilan data berdasarkan kata kunci yang telah ditentukan dan pelabelan data berdasarkan kategori sentimen.

3.2 Pre-processing

dilakukan untuk membersihkan dan mempersiapkan data sebelum dianalisis.

Proses pre-processing meliputi cleaning untuk menghapus unsur tidak relevan seperti username, hashtag, URL, emotikon, dan simbol; case folding untuk mengubah semua teks menjadi huruf kecil; tokenizing untuk memisahkan kalimat menjadi kata-kata tunggal; stopword removal untuk menghapus kata-kata umum yang tidak bermakna signifikan; serta stemming untuk mengubah kata menjadi bentuk dasarnya.

3.3 Transformation

Pemberian bobot pada setiap kata dalam dokumen menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Bobot ini membantu mengukur pentingnya sebuah kata dalam suatu dokumen dibandingkan dengan keseluruhan korpus data.

3.4 Data Mining

algoritma Naive Bayes digunakan untuk mengklasifikasikan data teks ke dalam dua kelas sentimen. Proses klasifikasi ini menggunakan pembagian data latih dan data uji dengan beberapa skenario perbandingan, yakni 90:10, 80:20, 70:30, dan 60:40, guna menguji kestabilan model.

3.5 Evaluation

Evaluasi model dilakukan dengan menggunakan confusion matrix. Evaluasi ini menghasilkan beberapa metrik performa, yaitu

1. Accuracy digunakan untuk menilai proporsi prediksi y

$$\frac{TP+TN}{TP+TN+FP+FN} 100\% \tag{1}$$

2. Precision

$$\frac{TP}{TP+FP} \tag{2}$$

3. Recall

$$\frac{TP}{TP+FN} \tag{3}$$

4. F-Score

$$2 * \frac{Precision * Recall}{Precision+Recall} \tag{4}$$

4. HASIL DAN PEMBAHASAN

Hasil dan pembahasan untuk analisis sentimen opini masyarakat twitter terhadap kasus diabetes di Indonesia menggunakan algoritma naïve bayes classifier yang diolah peneliti menggunakan google collaboratory.

Berikut hasil *crawling* data, preprocessing dan nilai Accuracy :

4.1 Pengumpulan Data

Hasil crawling dari data Twitter melalui authtoken sebanyak 4802 yang berkurang menjadi 4369 data setelah dilakukan penghapusan pada data duplikat. Berikut sample hasil dari proses crawling data.

Username	Fulltext
@fauziwong	WASPADA Terhadap GULA TERSEMBUNYI Potensi memicu Diabetes https://t.co/PaGeS4nf sE
@withseungcheols	ka cherrrr apakah kamu tidak terkena diabetes menulis seanrisya??? HAHAHAA yaampun nggak soalnya yang sekarang kebanyakan pahitnya .. #tellyonym
@jevdraz	1. Peningkatan Risiko Diabetes Konsumsi minuman manis dapat meningkatkan risiko diabetes tipe 2. Gula berlebih memicu resistensi insulin yang dapat menyebabkan gangguan metabolisme. (Sumber: American Diabetes Association

Setelah data terkumpul, dilakukan proses pelabelan secara manual dengan perolehan 2373 label negatif dan 1996 label positif.

4.2 Preprocessing

Setelah data terkumpul dan telah dilakukan pelabelan, selanjutnya adalah tahap preprocessing yang meliputi tahap cleaning, case folding, tokenizing, stopworld removal, dan stemming.

Fulltext	ï, □ Tanda Pankreas Lemah: Jika pankreas tidak berfungsi dengan baik kesehatan anda mungkin dalam bahaya
----------	--

	Risiko komplikasi serius seperti diabetes dan masalah pencernaan semakin meningkat. https://t.co/ILz3OyEY Mj
Cleaning	Tanda Pankreas Lemah Jika pankreas tidak berfungsi dengan baik kesehatan anda mungkin dalam bahaya Risiko komplikasi serius seperti diabetes dan masalah pencernaan semakin meningkat
Case Folding	tanda pankreas lemah jika pankreas tidak berfungsi dengan baik kesehatan anda mungkin dalam bahaya risiko komplikasi serius seperti diabetes dan masalah pencernaan semakin meningkat
Tokenizing	'tanda', 'pankreas', 'lemah', 'jika', 'pankreas', 'tidak', 'berfungsi', 'dengan', 'baik', 'kesehatan', 'anda', 'mungkin', 'dalam', 'bahaya', 'risiko', 'komplikasi', 'serius', 'seperti', 'diabetes', 'dan', 'masalah', 'pencernaan', 'semakin', 'meningkat'
Stopword Removal	'pankreas', 'lemah', 'pankreas', 'berfungsi', 'kesehatan', 'bahaya', 'risiko', 'komplikasi', 'serius', 'diabetes', 'masalah', 'pencernaan', 'semakin', 'meningkat'
Stemming	'pankreas', 'lemah', 'pankreas', 'fungsi', 'kesehatan', 'bahaya', 'risiko', 'komplikasi', 'serius', 'diabetes',

'masalah', 'cerna', 'makin', 'tingkat'
--

4.3

Hasil Accuracy

Dari keempat skenario pengujian, masing-masing melakukan pengujian berdasarkan confusion matrix. Perbandingan 90 : 10 menunjukkan hasil terbaik dengan jumlah *True Positive* (TP) sebesar 148, *True Negatif* (TN) sebesar 204, *False Positive* (FP) sebesar 53, dan *False Negatif* (FN) sebesar 32. Berdasarkan hasil confusion matrix tersebut, diperoleh hasil dengan nilai *accuracy* 81%, *precision* 80%, *recall* 81%, dan *f-score* 80%.

```

◆ Rasio Data Latih : Uji = 90 : 10
Akurasi: 0.8054919908466819
Confusion Matrix:
[[204 53]
 [ 32 148]]

```

	precision	recall	f1-score	support
Negatif	0.86	0.79	0.83	257
Positif	0.74	0.82	0.78	180
accuracy			0.81	437
macro avg	0.80	0.81	0.80	437
weighted avg	0.81	0.81	0.81	437

```

Multinomial NB Accuracy : 0.8054919908466819
Multinomial NB Precision : 0.800362593810608
Multinomial NB Recall : 0.8079982706441851
Multinomial NB F1-Score : 0.8022445470178297

```

Hal ini menunjukkan sentimen negatif lebih dominan daripada sentimen positif dengan tingkat akurasi 81%. Setelah diperoleh hasil perhitungan, dilakukan identifikasi terhadap 10 kata yang paling sering muncul beserta frekuensi terhadap sentimen negatif. Untuk memberikan gambaran yang lebih jelas, hasil analisis frekuensi kata yang sering muncul divisualisasikan ke dalam bentuk *wordcloud*.



5. KESIMPULAN

- a. Penelitian ini menerapkan algoritma Naive Bayes Classifier dalam metode penelitian KDD untuk menganalisis sentimen tweet

tentang diabetes. Data sebanyak 4369 tweet (dari 4802 tweet awal setelah duplikasi dihapus) dikumpulkan dari Twitter pada Februari-Desember 2024 menggunakan kata kunci "DiabetesMelitus", "PenyakitDiabetes", dan "Diabetes". Tweet tersebut kemudian dilabeli secara manual menjadi 1996 positif dan 2373 negatif. Data diprapemrosesan (pembersihan, case folding, tokenizing, stopword removal, stemming) dan ditransformasikan ke bentuk numerik menggunakan TF-IDF. Klasifikasi dilakukan dengan Naive Bayes Classifier menggunakan empat skenario pembagian data latih dan uji (90:10, 80:20, 70:30, 60:40).

- b. Kinerja Algoritma Naive Bayes Classifier dalam analisis sentimen opini Twitter tentang diabetes dievaluasi menggunakan confusion matrix, serta nilai akurasi, presisi, recall, dan F1-score. Skenario pembagian data 90:10 menunjukkan performa terbaik dengan akurasi 81%, presisi 80%, recall 81%, dan F1-score 80%. Hasil ini mengindikasikan model klasifikasi sentimen yang baik dan dapat diandalkan. Lebih lanjut, sentimen negatif ditemukan lebih dominan dibandingkan positif, mencerminkan kekhawatiran publik akibat unggahan keluhan atau peringatan tanpa solusi. Oleh karena itu, diperlukan peningkatan konten edukatif positif di media sosial untuk menyeimbangkan persepsi ini.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] M. Ardila, D. T. W. S. Humolungo, D. P. Amukti, and A. Akrom, "Promosi Kesehatan Pencegahan dan Pengendalian Diabetes

Melitus Pada Remaja," *J. Abdimas Indones.*, vol. 4, no. 2, pp. 534–540, 2024, doi: 10.53769/jai.v4i2.729.

- [2] T. Juniardi and C. A. Sugianto, "ANALISIS SENTIMEN TIM NASIONAL SEPAK BOLA INDONESIA DI TURNAMEN PIALA DUNIA U-17 INDONESIA PADA TWITTER (X) MENGGUNAKAN," vol. 12, no. 3, 2024.
- [3] A. R. Hidayat, H. Hanipah, A. Nurjanah, and R. Farizki, "Upaya untuk Mencegah Penyakit Diabetes pada Usia Dini," *J. Forum Kesehatan. Media Publ. Kesehat. Ilm.*, vol. 11, no. 2, pp. 63–69, 2022, doi: 10.52263/jfk.v11i2.229.
- [4] A. Syafrudin, D. Valentino, D. Arya, S. Pratama, and R. Pramudya, "Penerapan Metode Content Based Filtering Untuk Rekomendasi Minuman Dengan Kadar Gula Di Supermarket," pp. 18–21, 2024.
- [5] A. Hendra and F. Fitriyani, "Analisis Sentimen Review Halodoc Menggunakan Naive Bayes Classifier," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 2, pp. 78–89, 2021, doi: 10.14421/jiska.2021.6.2.78-89.
- [6] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor," *UNNES J. Math.*, vol. 10, no. 2, pp. 22–27, 2021, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [7] Rayuwati, Husna Gemasih, and Irma Nizar, "IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID," *Jural Ris. Rumpun Ilmu Tek.*, vol. 1, no. 1, pp. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.
- [8] R. Merdiansah and A. Ali Ridha, "Sentiment Analysis of Indonesian X Users Regarding Electric Vehicles Using IndoBERT," *J. Ilmu Komput. dan Sist. Inf. (JIKOMSI)*, vol. 7, no. 1, pp. 221–228, 2024.
- [9] I. Widaningrum, D. Mustikasari, R. Arifin, S. L. Tsaqila, and D. Fatmawati, "Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan K-Means Clustering Untuk Menentukan Kategori Dokumen," *Pros. Semin. Nas. Sist. Inf. dan Teknol.*, pp. 145–149, 2022.
- [10] D. Normawati and S. A. Prayogi, "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [11] C. Jovanica, D. D. Rahmintangrum, H. A. Nuradni, and A. Salsabila, "Analisis Pengaruh Aktor Pada Tagar #Roketchina Di Media Sosial Twitter Menggunakan Social Network

- Analysis (Sna),” *J. Ilm. Komun. Makna*, vol. 10, no. 1, p. 43, 2022, doi: 10.30659/jikm.v10i1.15644.
- [12] D. Alita, “Multiclass SVM Algorithm for Sarcasm Text in Twitter,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 118–128, 2021, doi: 10.35957/jatisi.v8i1.646.
- [13] G. Syam, S. Lal, and T. Chen, “Empirical Study of the Evolution of Python Questions on Stack Overflow,” *E-Informatica Softw. Eng. J.*, vol. 17, no. 1, pp. 1–44, 2023, doi: 10.37190/e-Inf230107.
- [14] V. Khisanova, “Python libraries and their usage in visualizing meteorological data,” 2022.
- [15] L. I. Umayya and I. S. Wardani, “Hubungan Antara Diabetes Melitus Dengan Glaukoma,” *J. Med. Hutama*, vol. 04, no. 01, pp. 3280–3291, 2023.
- [16] Febby Wilyani, Qonaah Nuryan Arif, and Fitri Aslimar, “Pengenalan Dasar Pemrograman Python Dengan Google Colaboratory,” *J. Pelayanan dan Pengabd. Masy. Indones.*, vol. 3, no. 1, pp. 08–14, 2024, doi: 10.55606/jppmi.v3i1.1087.