

ANALISIS SENTIMEN MASYARAKAT TERHADAP PROGRAM MAKAN SIANG GRATIS PADA MEDIA SOSIAL X MENGGUNAKAN ALGORITMA NAÏVE BAYES

Altolyto Sitanggang¹, Yuyun Umaidah², Riza Ibnu Adam³

^{1,2,3}Universitas Singaperbangsa Karawang; Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur Karawang, Indonesia; (0267) 641177

Received: 12 Juli 2024

Accepted: 31 Juli 2024

Published: 7 Agustus 2024

Keywords:

Naïve Bayes, KDD, Makan Siang Gratis, X, Sentimen.

Correspondent Email:

altolyto@gmail.com

Abstrak. Dalam era digital, media sosial seperti X, Facebook, dan Instagram telah menjadi bagian penting dari kehidupan modern, memungkinkan individu untuk berbagi pandangan dan opini dengan cepat. Salah satu topik hangat di X adalah program makan siang gratis dari pasangan calon presiden nomor urut 02, yang bertujuan meningkatkan gizi anak dan ibu hamil, mencakup 82,9 juta orang. Program ini memicu beragam tanggapan masyarakat. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap program tersebut menggunakan algoritma Naïve Bayes dan metode *Knowledge Discovery in Database* (KDD). Data dikumpulkan melalui *crawling* pada media sosial X, menghasilkan 2.211 *tweet* yang kemudian diseleksi dan diberi label sentimen positif dan negatif. Algoritma Naïve Bayes diuji dengan tiga skenario pembagian data *training* dan *testing*, dan dievaluasi menggunakan *confusion matrix*. Hasil evaluasi menunjukkan model mencapai hasil terbaik pada rasio data 60:40 dengan akurasi 72,2%, presisi 63,2%, *recall* 66,1%, dan *F1-Score* 64%.

Abstract. In the digital age, social media such as X, Facebook and Instagram have become an essential part of modern life, allowing individuals to share views and opinions quickly. One of the hot topics on X is the free lunch program of presidential candidate number 02, which aims to improve the nutrition of children and pregnant women, covering 82.9 million people. This program triggered a variety of public responses. This research aims to analyze public sentiment towards the program using Naïve Bayes algorithm and Knowledge Discovery in Database (KDD) method. Data was collected through crawling on X social media, resulting in 2,211 tweets which were then selected and labeled with positive and negative sentiments. The Naïve Bayes algorithm was tested with three scenarios of training and testing data sharing, and evaluated using confusion matrix. The evaluation results showed that the model achieved the best results at a data ratio of 60:40 with 72.2% accuracy, 63.2% precision, 66.1% recall, and 64% F1-Score.

1. PENDAHULUAN

Ketika kita terus menyaksikan pesatnya pertumbuhan era digital, media sosial telah muncul sebagai elemen penting dalam kehidupan sehari-hari individu dalam masyarakat modern. *Platform-platform* seperti X, Facebook, dan Instagram memungkinkan individu untuk berbagi pandangan, pengalaman, dan opini mereka dengan cepat dan mudah kepada cakupan yang lebih luas. Informasi terbaru bisa didapatkan dari media sosial apapun, termasuk aplikasi X. X telah menjadi salah satu *platform* yang membuat masyarakat dapat berinteraksi dan berbagi pendapat tentang berbagai topik, termasuk isu-isu politik, sosial, dan ekonomi. Pembahasan isu politik terkait visi misi pasangan calon presiden telah menjadi topik yang hangat diperbincangkan.

Dari ketiga pasangan calon presiden, terdapat program dari salah satu pasangan calon presiden yang menarik perhatian masyarakat, yaitu program makan siang gratis dari pasangan calon presiden nomor urut 02. Program makan siang gratis merupakan program yang memberi makan siang dan susu di sekolah maupun di pesantren secara gratis, untuk membantu gizi anak balita dan ibu hamil (<https://prabowogibran2.id/>). Program makan siang gratis ini direncanakan akan diberikan kepada 82,9 juta orang, termasuk anak balita, murid sekolah, dan ibu hamil, dengan tujuan membangun generasi yang pintar, sehat, dan unggul dalam daya saing, serta berkontribusi pada produktivitas ekonomi demi mencapai visi Indonesia Emas 2045. Namun, seperti halnya program-program besar lainnya, program makan siang gratis juga menuai beragam tanggapan dan pendapat dari berbagai kalangan masyarakat.

Dalam konteks ini, analisis sentimen masyarakat terhadap program makan siang gratis memiliki nilai penting untuk memahami persepsi dan sikap masyarakat terhadap program ini. Pada analisis sentimen data subjektif, khususnya opini, adalah data yang digunakan sebagai dasar, tanpa nilai objektif apa pun. Selain itu, data ini berasal dari individu, yang masing-masing memiliki cara uniknya sendiri dalam mengungkapkan sudut pandang dan preferensinya [1]. Analisis sentimen melibatkan prinsip dasar pengelompokan polaritas yang terdapat dalam

teks, yang kemudian diidentifikasi menjadi kategori-kategori positif, negatif, dan netral[2].

Terdapat lebih dari satu algoritma yang dapat dipakai untuk melakukan analisis sentimen, seperti misalnya algoritma Naïve Bayes. Naïve Bayes adalah sebuah algoritma klasifikasi yang mengandalkan konsep probabilitas dengan merujuk pada Teorema Bayes sebagai dasar konsepnya[3]. Keuntungan penggunaan Naïve Bayes adalah bahwa metode ini memerlukan jumlah data pelatihan (*Training Data*) yang relatif kecil untuk menentukan estimasi parameter yang dibutuhkan dalam proses klasifikasi. Algoritma Naïve Bayes adalah algoritma yang cocok untuk melakukan analisis sentimen, karena menggunakan pendekatan di mana kata-kata dalam sebuah dokumen dianggap saling berguna tanpa memperhitungkan susunan kata atau informasi yang terdapat dalam kalimat atau dokumen itu[4].

Berdasarkan penelitian yang sudah pernah dilakukan oleh [5] dengan judul Komparasi Algoritma Naïve Bayes Dan *Support Vector Machine* (SVM) Pada Analisis Sentimen Spotify, dalam analisis sentimen terhadap ulasan Spotify menggunakan 1500 data, algoritma Naïve Bayes berhasil memberikan hasil terbaik dengan akurasi mencapai 86,4%. Sebaliknya, algoritma *Support Vector Machine* (SVM) hanya mencapai akurasi sebesar 84%. Adapun juga penelitian yang telah dilakukan oleh [6] yang membandingkan kinerja algoritma Naïve Bayes, decision tree, dan random forest dalam analisis sentimen terhadap kenaikan biaya haji 2023 di *platform* media sosial YouTube. Hasil penelitian menunjukkan tingkat akurasi Naïve Bayes sebesar 90%, decision tree sebesar 83%, dan random forest sebesar 87%. Dari data tersebut, dapat disimpulkan bahwa algoritma Naïve Bayes menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan decision tree dan random forest.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah suatu proses yang menghasilkan informasi baru dengan cara memahami dan mengolah data teks secara otomatis[7]. Analisis sentimen melibatkan pengumpulan pendapat dari masyarakat umum melalui jejaring sosial yang mencakup layanan

umum dan isu-isu terkini[8]. Adapun tujuan dari analisis sentimen yaitu untuk mengidentifikasi dan mengklasifikasikan sentimen ke dalam bentuk yang negatif atau positif. Analisis sentimen dapat digunakan pada opini berbagai bidang seperti politik, ekonomi, sosial, dan hukum[9].

2.2 Program Makan Siang Gratis

Program makan siang gratis adalah salah satu program yang dijanjikan pada saat masa kampanye oleh pasangan calon presiden nomor urut 2 pada pemilu 2024. Program makan siang gratis merupakan program yang memberi makan siang dan susu di sekolah maupun di pesantren secara gratis, untuk membantu gizi anak balita dan ibu hamil. Program makan siang gratis ini direncanakan akan diberikan kepada 82,9 juta orang, termasuk anak balita, murid sekolah, dan ibu hamil, dengan tujuan membangun generasi yang pintar, sehat, dan unggul dalam daya saing, serta berkontribusi pada produktivitas ekonomi demi mencapai visi Indonesia Emas 2045 (<https://prabowogibran2.id/>).

2.3 Media Sosial X

X atau yang sebelumnya dikenal sebagai twitter ialah wadah media sosial yang memberikan kesempatan pada penggunanya untuk menulis, membaca, dan berpartisipasi dalam diskusi mengenai topik terkini serta isu-isu yang sedang berkembang. X banyak digunakan menjadi *platform* bagi pengguna untuk mengekspresikan perasaan terkait dengan berbagai hal, baik itu pujian, kritik maupun celaan [10]. Pujian, kritik maupun celaan itu bisa dijadikan data untuk menganalisis tren atau topik tertentu. Menurut [11] X merupakan sumber data yang cukup baik karena tingkat keakuratan dari opini yang ada di *platform* tersebut dianggap tinggi, sehingga dapat digunakan untuk memahami pandangan masyarakat terhadap suatu topik.

2.4 Text Mining

Text mining merupakan salah satu teknik dalam pengolahan informasi yang berguna untuk menganalisis data dalam bentuk teks, seperti dokumen atau tulisan, baik itu untuk tujuan klasifikasi, penentuan, maupun pengelompokan[12]. Menurut [13] *Text mining* ialah proses yang membuat teks yang tadinya

tidak terstruktur menjadi teks yang terstruktur dengan tujuan mendapatkan pola untuk memperoleh wawasan ataupun pengetahuan yang baru. *Text mining* memiliki kesamaan dengan *data mining* dalam ukuran data, tetapi memiliki perbedaan dalam pola dan struktur sumber data yang digunakan.

2.5 Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) adalah suatu metode untuk mendapatkan pola yang belum diketahui, dapat dimengerti, dan bermanfaat dari suatu *dataset* yang besar[14]. Proses *Knowledge Discovery in Database* (KDD) melibatkan pengguna secara aktif dalam setiap langkahnya, dan dapat dilakukan ulang terus-menerus untuk memperoleh hasil yang terbaik.

2.6 Text Preprocessing

Text Preprocessing merupakan serangkaian aktivitas untuk merapikan dan membersihkan data agar data siap untuk diproses ke tahap berikutnya[15]. Menurut [16] Tahapan dalam *text preprocessing* yaitu:

1. *Cleaning*, Pada tahap ini, dilakukan tindakan menghilangkan karakter dan tanda baca yang tidak relevan dari teks dilakukan untuk mengurangi *noise* pada *dataset*.
2. *Casefolding*, Pada tahap ini, dilakukan suatu proses untuk menjadikan semua huruf dalam teks diubah menjadi huruf kecil (*lowercase*).
3. *Stopword Removal*, Pada tahap ini, dilakukan proses menghilangkan kata-kata sambung dari teks karena dianggap tidak memiliki makna dan mengganggu analisis teks.
4. *Tokenizing*, Pada langkah ini, dilakukan pemisahan dari sebuah kalimat menjadi beberapa kata.
5. *Stemming*, pada tahap ini, dilakukan proses penyederhanaan kata yang memiliki awalan atau akhiran dan akan menjadi bentuk dasarnya.

2.7 TF-IDF

TF-IDF adalah metode pembobotan kata yang menghitung nilai *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). *Term Frequency-Inverse Document Frequency* (TF-IDF) adalah suatu proses konversi data teks

menjadi vektor untuk memperhitungkan urutan kata yang signifikan. *Term Frequency* mengukur seberapa sering kemunculan sebuah kata dalam sebuah dokumen, sementara *Inverse Document Frequency* mengukur seberapa jarang kemunculan kata tersebut dalam seluruh koleksi dokumen. Hasil pembobotan kata merupakan hasil perkalian antara nilai *Term Frequency* dan *Inverse Document Frequency*, di mana kata yang paling sering muncul akan memperoleh nilai atau bobot lebih rendah, sementara kata paling jarang muncul akan memperoleh bobot lebih tinggi[17].

2.8 Algoritma Naïve Bayes

Naïve Bayes adalah sekelompok algoritma pada metode klasifikasi yang mengikuti konsep yang serupa berdasarkan Teorema Bayes. Algoritma ini merupakan pendekatan berbasis probabilitas yang umumnya digunakan untuk masalah klasifikasi[18]. Algoritma Naïve Bayes digunakan untuk membuat suatu keputusan dengan memprediksi suatu kasus ataupun peristiwa berdasarkan klasifikasi yang telah dilakukan. Menurut [19] Karakteristik yang kuat dari algoritma Naïve Bayes adalah dugaan yang begitu kuat terhadap kemandirian antara setiap kejadian ataupun kondisi. Kelebihan dari penggunaan algoritma naïve bayes adalah hanya memerlukan sedikit *data train* untuk menetapkan ukuran yang dibutuhkan dalam proses klasifikasi[20].

2.9 Confusion Matrix

Dalam bidang *data mining*, salah satu cara untuk mengevaluasi kinerja suatu model yang didapatkan adalah dengan mempergunakan *confusion matrix*. *Confusion matrix* adalah tabel yang secara khusus menggambarkan kinerja model atau algoritma, di mana setiap baris mencerminkan kelas aktual dari data, dan setiap kolom mencerminkan kelas prediksi dari data (ataupun sebaliknya)[21]. Hasil dari evaluasi menggunakan *confusion matrix* merupakan akurasi, *precision*, dan *recall*. Akurasi mengevaluasi keberhasilan model dalam membuat prediksi yang benar terhadap seluruh data uji, *precision* mengukur seberapa baik model dalam membuat prediksi yang benar untuk kelas positif dari seluruh prediksi positif, sementara *recall* mengevaluasi

seberapa baik model dalam mengidentifikasi kelas positif dengan benar[22].

2.10 Confusion Matrix

Python adalah bahasa pemrograman yang sering dipilih untuk pengembangan aplikasi analisis data, karena menawarkan beragam *library* dan modul yang mendukung penerapan algoritma klasifikasi[23]. Python banyak digunakan karena penggunaannya yang lebih mudah. Python cenderung lebih menyerupai bahasa manusia daripada bahasa mesin karena python merupakan bahasa pemrograman tingkat tinggi [24]. Karena Python dilengkapi dengan beragam *library* yang dapat digunakan secara fleksibel untuk berbagai jenis aplikasi, mulai dari analisis statistik hingga kecerdasan buatan, dan ditambah dengan sintaksis yang mudah dipahami, menjadikannya sebagai pilihan bahasa pemrograman yang sangat ideal [25].

2.11 Google Colab

Google Colab adalah layanan *cloud computing* gratis yang diberikan oleh Google yang memungkinkan seseorang untuk menulis, menjalankan, dan berbagi kode Python, serta melakukan pemrosesan data dan pembelajaran mesin[26]. Google Colab adalah sebuah *platform* yang dirancang untuk menyederhanakan tugas-tugas yang terkait dengan ilmu data, dengan tujuan memudahkan pekerjaan yang ada pada bidang *data science*[27]. Google colab digunakan banyak orang karena penggunaannya yang mudah. Google Colab memberi fasilitas kepada pengguna untuk mengeksekusi kode Python tanpa perlu melakukan instalasi dan penyiapan yang rumit[28]. Kemudahan penggunaan, fleksibilitas, dan ketersediaan sumber daya yang melimpah membuat Google Colab menjadi salah satu pilihan utama bagi pengguna yang bekerja pada bidang data.

2.12 Wordcloud

Wordcloud adalah visualisasi teks yang menampilkan kata-kata dengan ukuran dan warna yang berbeda sesuai dengan seberapa sering kata tersebut muncul dalam suatu dokumen[29]. Wordcloud banyak digunakan karena representasi visual yang efektif dalam menyajikan informasi teks secara singkat dan mudah dimengerti[30]. Dengan


```
df['tokenized_text'] = df['stopword_removed_text'].apply(lambda x:x.split())
df[['stopword removed text','tokenized text']]
```

Gambar 5 Fungsi *Tokenizing*

- ```

3 from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()

def stemmed_wrapper(term):
 return stemmer.stem(term)

term_base = {}

for document in df['tokenized_text']:
 for term in document:
 if term not in term_base:
 term_base[term] = ""

for term in term_base:
 term_base[term] = stemmed_wrapper(term)

def get_stemmed_term(document):
 return [term_base[term] for term in document]

df['stemmed_text'] = df['tokenized_text'].apply(get_stemmed_term)
df[['tokenized_text', 'stemmed_text']]

```

Gambar 6 Fungsi *Stemming*

Pada tahap ini akan dilakukan perubahan pada teks *tweet* yang sudah menjalani proses *data preprocessing*, menjadi representasi data numerik. Tahap ini dilakukan dengan menerapkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Pembobotan ini bertujuan untuk memberikan nilai pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen atau seberapa umum penggunaannya.

[illegible]

Gambar 7 Tahap *Data Transformation* TF-IDF

Pada tahap ini, akan ditentukan data latih dan data uji. Proses ini dilakukan dengan menggunakan operator *split data* yang membagi data dalam tiga pengujian yang berbeda, yaitu perbandingan 80:20, 70:30, dan 60:40 secara otomatis sesuai dengan perbandingan dari *dataset* yang sudah bersih. Pada tahap ini data akan diproses menggunakan Algoritma Multinomial Naïve Bayes.

```

from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB

Pemisahan data dengan rasio 80:20
X_train_80, X_test_20, y_train_80, y_test_20 = train_test_split(text_data['label'], text_data['text'], test_size=0.2, random_state=0)

Pemisahan data dengan rasio 70:30
X_train_70, X_test_30, y_train_70, y_test_30 = train_test_split(text_data['label'], text_data['text'], test_size=0.3, random_state=0)

Pemisahan data dengan rasio 60:40
X_train_60, X_test_40, y_train_60, y_test_40 = train_test_split(text_data['label'], text_data['text'], test_size=0.4, random_state=0)

Membuat model Multinomial Naive Bayes
model = MultinomialNB()

```

Gambar 8 Tahap *Data Mining*

Pada tahap ini, evaluasi terhadap klasifikasi sentimen dilakukan menggunakan algoritma Naïve Bayes. *Confusion matrix* digunakan untuk menganalisis performa model, serta menghitung nilai akurasi, presisi, *recall*, dan *F1-Score* guna memberikan gambaran tentang efektivitasnya. Setiap pembagian data dengan rasio 80:20, 70:30, dan 60:40 dievaluasi secara terpisah untuk memberikan pemahaman yang mendalam tentang kinerja model ini.

```
Rasio 80:20 - Akurasi: 0.6667, Presisi: 0.5890, Recall: 0.6254, F1-Score: 0.5869
Confusion Matrix 80:20:
[[73 32]
 [12 15]]

Rasio 70:30 - Akurasi: 0.6954, Presisi: 0.6083, Recall: 0.6416, F1-Score: 0.6128
Confusion Matrix 70:30:
[[114 41]
 [19 23]]

Rasio 60:40 - Akurasi: 0.7224, Presisi: 0.6319, Recall: 0.6612, F1-Score: 0.6397
Confusion Matrix 60:40:
[[158 47]
 [26 32]]
```

### Gambar 9 Hasil Evaluasi Performa Model Naïve Bayes

Berdasarkan gambar diatas maka didapatkan hasil seperti tabel dibawah ini.

Tabel 1. Hasil Evaluasi

| <b>Rasio</b> | <b>Akurasi</b> | <b>Presisi</b> | <b>Recall</b> | <b>F1-Score</b> |
|--------------|----------------|----------------|---------------|-----------------|
| 80:20        | 66,7%          | 58,9%          | 62.5%         | 58,7%           |
| 70:30        | 69,5%          | 60,8%          | 64.2%         | 61,3%           |
| 60:40        | 72.2%          | 63.2%          | 66.1%         | 64%             |

Dapat dilihat bahwa pada rasio pemisahan 80:20, model mencapai akurasi sebesar 66,7%, dengan presisi 58,9%, *recall* 62,5%, dan *F1-Score* 58,7%. Dari rasio 70:30, model menunjukkan peningkatan dengan akurasi

69,5%, presisi 60,8%, *recall* 64,2%, dan *F1-Score* 61,3%. Sementara itu, dari rasio 60:40, model mencapai performa terbaiknya dengan akurasi sebesar 72,2%, presisi 63,2%, *recall* 66,1%, dan *F1-Score* 64%.

## 5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan beberapa poin utama, diantaranya:

1. Penelitian ini mengevaluasi sentimen masyarakat terhadap program makan siang gratis dari aplikasi X menggunakan metodologi *Knowledge Discovery in Database* (KDD). *Tweet* yang berbentuk data nominal diubah menjadi data numerik menggunakan TF-IDF, yang kemudian diproses menggunakan algoritma Naïve Bayes untuk *data mining* dan dievaluasi menggunakan *Confusion Matrix*.
2. Dari hasil evaluasi model menggunakan algoritma Naïve Bayes dalam tiga skenario pemisahan data yang berbeda, model menunjukkan kinerja terbaik pada skenario pemisahan 60:40 dengan akurasi sebesar 72,2%, presisi 63,2%, *recall* 66,1%, dan *F1-Score* 64%.

Sebagai rekomendasi, dalam pengembangan selanjutnya, dapat dilakukan:

1. Dapat ditingkatkan jumlah data dalam *dataset* pada penelitian selanjutnya untuk meningkatkan hasil dari klasifikasi.
2. Mengimplementasikan algoritma yang berbeda untuk mengetahui algoritma yang lebih baik dalam melakukan analisis sentimen.
3. Melakukan berbagai skenario pemisahan data untuk memperoleh hasil optimal dari model yang diterapkan menggunakan algoritma yang dipilih.

## UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada kedua orang tua yang selalu memberikan *support* dan doa restu dalam setiap langkah penelitian ini, serta kepada dosen pembimbing yang telah membantu dalam penelitian yang dilakukan.

## DAFTAR PUSTAKA

- [1] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. S. T. Ayu, and W. F. Dicka, "Dataset Indonesia untuk Analisis Sentimen," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, p. 334, 2019, doi: 10.22146/jnteti.v8i4.533.
- [2] S. Suprianto, "Perbandingan Metode Naïve Bayes Classifier Dan Holistic Lexicon Based Dalam Analisis Sentimen Angket Mahasiswa," *JSI J. Sist. Inf.*, vol. 11, no. 2, 2019, doi: 10.36706/jsi.v11i2.9140.
- [3] H. Z. Muflih, A. R. Abdillah, and F. N. Hasan, "Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes," vol. 4, no. 3, pp. 1613–1621, 2023, doi: 10.30865/klik.v4i3.1303.
- [4] R. Yunita and M. Kamayani, "Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Kebijakan Penghapusan Kewajiban Skripsi," vol. 12, no. 1, pp. 2879–2890, 2023.
- [5] A. S. Rahayu and A. Fauzi, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine ( SVM ) Pada Analisis Sentimen Spotify," vol. 4, pp. 349–354, 2022, doi: 10.30865/json.v4i2.5398.
- [6] M. Yasir And R. Suraji, "Perbandingan Metode Klasifikasi Naïve Bayes , Decision Tree , Random Forest Terhadap Analisis Sentimen Kenaikan Biaya Haji 2023 Pada Media Sosial Youtube," 2023.
- [7] C. F. Hasri And D. Alita, "Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," Vol. 3, No. 2, Pp. 145–160, 2022.
- [8] M. Syarifuddin, "Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn," Vol. 15, No. 1, Pp. 23–28, 2020.
- [9] P. Arsi And R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine ( Svm ) Sentiment Analysis On The Discussion Of Relocating I Ndonesia ' S Capital City Using The Support Vector Machine ( Svm )," Vol. 8, No. 1, Pp. 147–156, 2021, Doi: 10.25126/Jtiik.202183944.
- [10] W. Gafatia, Desta and N. Hadinata, "Analisis Pro Kontra Vaksin Covid 19 Menggunakan Sentiment Analysis Sumber Media Sosial Twitter," vol. 2, no. 1, pp. 34–42, 2021.
- [11] M. Rifqi, "Content Analysis on Twitter Users Interaction within First 100 Days of Jakarta ' s New Government by Using Text Mining Analisis Konten Interaksi Pengguna Twitter pada Masa 100 Hari Pertama Pemerintahan Baru DKI Jakarta Menggunakan Text Mining

- Content Anal,” no. October 2018, 2019, doi: 10.30818/jpkm.2018.2030203.
- [12] C. J. E. Munthe, N. A. Hasibuan, and H. Hutabarat, “Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace,” vol. 2, no. 3, pp. 110–115, 2022.
- [13] N. Silalahi and G. Ginting, Leonarde, “Rekomendasi Berita Berkaitan dengan Menerapkan Algoritma Text Mining dan TF-IDF,” vol. 3, no. 4, pp. 276–282, 2023, doi: 10.47065/bulletincsr.v3i4.266.
- [14] J. Wandana, S. Defit, and S. Sumijan, “Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means,” *J. Inf. dan Teknol.*, vol. 2, pp. 4–9, 2020, doi: 10.37034/jidt.v2i4.73.
- [15] S. Chohan, A. Nugroho, A. Maezar, B. Aji, and W. Gata, “Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naïve Bayes dan Synthetic Minority Over Sampling Technique,” vol. 22, no. 2, 2020.
- [16] M. I. Ghozali, W. H. Sugiharto, and A. F. Iskandar, “Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes,” vol. 3, no. 6, pp. 1340–1348, 2023, doi: 10.30865/klik.v3i6.936.
- [17] M. H. Mahendra, D. T. Murdiansyah, and K. M. Lhaksana, “Analisis Sentimen Tweet COVID-19 Menggunakan Metode K-Nearest Neighbors dengan Ekstraksi Fitur TF-IDF dan CountVectorizer Dike: Jurnal Ilmu Multidisiplin,” vol. 1, pp. 37–43, 2023.
- [18] B. Ghogh et al., “Feature Selection and Feature Extraction in Pattern Analysis: A Literature Review”.
- [19] Sulindawaty, E. Laia, and M. Yamin, “Penerapan Algoritma Naïve Bayes dalam Menganalisis Sentimen pada Review Pengguna E-Commerce,” vol. 4, no. 1, pp. 305–316, 2023, doi: 10.30865/klik.v4i1.1186.
- [20] R. Syahputra, G. J. Yanris, and D. Irmayani, “SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter,” vol. 6, no. 2, pp. 671–678, 2022.
- [21] M. Hidayatullah, S. Alam, and I. Jaelani, “Sentiment Analysis of Police Performance On Twitter Users Using Naïve Bayes Method,” *RISTEC Res. Inf. Syst. Technol.*, vol. 2, no. 2, pp. 86–97, 2021, doi: 10.31980/ristec.v2i2.1945.
- [22] M. Luthfi Bangun Permadi and R. Gumilang, “Penerapan Algoritma CNN (Convolutional Neural Network) Untuk Deteksi Dan Klasifikasi Target Militer Berdasarkan Citra Satelit,” *J. Sos. Teknol.*, vol. 4, no. 2, pp. 134–143, 2024, doi: 10.59188/jurnalsostech.v4i2.1138.
- [23] P. S. Zakaria, R. Julianto, and R. S. Bernada, “Implementasi Naive Bayes Menggunakan Python dalam Klasifikasi Data,” *BIKMA Bul. Ilm. Ilmu Komput. dan Multimed.*, vol. 1, no. 2, pp. 126–131, 2023.
- [24] N. Cahyono and Dewi Setiyawati, “Analisis Sentimen Pengguna Sosial Media Twitter Terhadap Perokok Di Indonesia,” *Indones. J. Comput. Sci.*, vol. 12, no. 1, pp. 262–272, 2023, doi: 10.33022/ijcs.v12i1.3154.
- [25] S. Junaidi, M. Devegi, and H. Kurniawan, “Pelatihan Pengolahan dan Visualisasi Data Penduduk menggunakan Python,” *ADMA J. Pengabd. dan Pemberdaya. Masy.*, vol. 4, no. 1, pp. 151–162, 2023, doi: 10.30812/adma.v4i1.2963.
- [26] R. Syahrial, T. Sukmawati, and E. N. Dewi, “Face Mask Detection Menggunakan Python Dan Opencv Untuk Mendeteksi Pelanggaran Protokol Kesehatan Covid-19,” *JEIS J. Elektro dan Inform. Swadharma*, vol. 3, no. 1, pp. 77–86, 2023, doi: 10.56486/jeis.vol3no1.304.
- [27] Rangga Gelar Guntara, “Pelatihan Sains Data Bagi Pelaku UMKM di Kota Tasikmalaya Menggunakan Google Colab,” *Joong-Ki J. Pengabd. Masy.*, vol. 2, no. 2, pp. 245–251, 2023, doi: 10.56799/joongki.v2i2.1572.
- [28] M. Arkan et al., “Webinar & Workshop Peran Artificial Intelligence Dalam Analisis Big Data Untuk Mendukung Sektor Bisnis E-Commerce,” *J. Pengabd. Kolaborasi dan Inov. IPTEKS*, vol. 2, no. 1, pp. 228–236, 2024, doi: 10.59407/jpki2.v2i1.490.
- [29] W. Saefudin, A. Komarudin, and R. Ilyas, “Visualisasi Kumpulan Berita Dalam Bentuk Peta Digital Dengan Metode Term Frequency-Inverse Document Frequency dan Gazetteer,” *Semin. Nas. Sains dan Teknol. Inf.*, vol. 2, no. 1, pp. 665–670, 2019.
- [30] M. Akmal Hakim, P. Gunawan, R. Pratama, and F. Kurniawam, “Implementasi Algoritma K-Nearest Neighbors untuk Menganalisis Pendapat Pakar AI tentang Kemajuan Kecerdasan Buatan,” *J. Comput. Inf. Syst. Ampera*, vol. 5, no. 2, pp. 2775–2496, 2024, [Online]. Available: <https://journal-computing.org/index.php/journal-cisa/index>