

Penerapan Naive Bayes Terhadap Sentimen Analisis Media Sosial Twitter Pengguna Kereta Cepat Jakarta-Bandung (*Whoosh*)

Tifani Agustiranti¹, Aulia Khalfani Izzati Kurdiana², Bilal Al Ghiffari³, Elza Dwi Juniar⁴, Diki Gita Purnama⁵

^{1,2,3,4,5}Teknik Informatika, Universitas Paramadina, Jakarta, Indonesia

Email: ¹tifani.agustiranti@students.paramadina.ac.id, ²aulia.kurdiana@students.paramadina.ac.id,

³bilal.ghiffari@students.paramadina.ac.id, ⁴elza.juniar@students.paramadina.ac.id, ⁵diki.purnama@paramadina.ac.id

Email Penulis Korespondensi: tifani.agustiranti@students.paramadina.ac.id

Abstrak—Dalam era digital saat ini, media sosial telah menjadi platform utama bagi pengguna untuk berbagi pengalaman dan pendapat mereka terkait berbagai layanan publik, termasuk layanan transportasi. Kereta Cepat Jakarta-Bandung (*whoosh*) adalah salah satu proyek infrastruktur yang penting di Indonesia, menawarkan alternatif transportasi yang efisien antara dua kota metropolitan utama. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap penggunaan kereta cepat Jakarta-Bandung (*whoosh*) melalui data ulasan yang dikumpulkan dari media sosial Twitter. Metode penelitian melibatkan pengumpulan data melalui *crawling* Twitter, pelabelan data manual, pra-pemrosesan teks, pembobotan *TF-IDF*, dan penerapan algoritma *Naive Bayes Classifier*. Hasil penelitian menunjukkan bahwa tanggapan masyarakat cenderung positif terhadap pengguna kereta cepat tersebut dengan tingkat akurasi klasifikasi mencapai 88%. Meskipun demikian, terdapat sejumlah data yang belum terklasifikasi dengan baik, menunjukkan tantangan dalam mendeteksi sentimen netral dengan tepat. Implikasi dari penelitian ini adalah pentingnya analisis sentimen dalam memahami persepsi publik terhadap pengguna layanan transportasi dan potensi pengembangan metode klasifikasi yang lebih canggih untuk meningkatkan akurasi klasifikasi sentimen pada data yang kompleks.

Kata Kunci: Sentimen Analisis, Kereta Cepat, Twitter, *Naive Bayes Classifier*, *whoosh*

Abstract— During the digital era, social media has emerged as a key platform for people to share their experiences and opinions regarding various public services, including transport services. The Jakarta-Bandung High Speed Rail is one of the most important infrastructure sectors in Indonesia, offering an efficient transport alternative between the two major metropolitan cities. This study aims to analyze public sentiment towards the use of the Jakarta-Bandung Whoosh high-speed train through review data collected from Twitter social media. The research method involves data collection through Twitter crawling, manual labeling of data, text pre-processing, TF-IDF weighting, and application of Naive Bayes Classifier algorithm. The results show that the public response tends to be positive towards the fast train service, with the classification accuracy rate reaching 88%. Nonetheless, there are a number of data that have not been classified well, demonstrating the challenge in detecting neutral sentiments appropriately. The implications of this research are the importance of sentiment analysis in understanding public perception of transportation services and the potential development of more sophisticated classification methods to improve sentiment classification accuracy on complex data.

Keywords: Sentiment Analysis, High Speed Train, Twitter, Naive Bayes Classifier, whoosh

1. PENDAHULUAN

Transportasi adalah salah satu sarana yang bisa dipergunakan untuk mengangkut orang atau barang dari lokasi awal ke tujuan tertentu. Saat ini, Indonesia telah memiliki berbagai jenis transportasi, termasuk transportasi darat, laut, udara, dan transportasi yang menggunakan rel. Transportasi berbasis rel menarik minat masyarakat dengan kapasitas penumpang yang tinggi dan keunggulan tidak terkena kemacetan karena memiliki jalur khusus. Oleh karena itu, kereta menjadi pilihan transportasi yang sangat efektif untuk mobilitas massal di wilayah perkotaan yang padat

Kereta cepat adalah kereta yang mampu melaju dengan kecepatan lebih dari 350 km/jam. Penetapan rute Kereta Cepat Jakarta-Bandung bukanlah tanpa pertimbangan yang jelas. Jakarta sebagai pusat kegiatan industri, dan Bandung sebagai ibukota Provinsi Jawa Barat dan merupakan kota destinasi wisata yang merupakan faktor utama dalam keputusan ini. Pembangunan Kereta Api Cepat Jakarta-Bandung ini bertujuan untuk mengembangkan efektivitas dan efisiensi perpindahan transport publik, sambil menjunjung aspek ramah lingkungan[1].

Dalam era digital ini, media sosial, terutama twitter, menjadi saluran utama bagi masyarakat untuk berbagi pengetahuan dan pandangan mereka terkait beragam hal, termasuk pengalaman perjalanan dengan kereta cepat. Hal ini menciptakan peluang untuk melakukan analisis sentimen terhadap komentar dan ulasan pengguna mengenai layanan kereta cepat Jakarta-Bandung.

Analisis sentimen dapat memberikan gambaran keseluruhan tentang persepsi masyarakat melalui mengelompokkan jenis opini yang muncul ke dalam kategori positif, negatif, dan netral. Analisis sentimen merupakan studi komputasional mengenai pendapat, perilaku, dan emosi individu terhadap entitas tertentu, yang dapat mencakup individu, peristiwa, atau topik tertentu. Analisis sentimen melibatkan proses mengelompokkan polaritas teks dalam sebuah dokumen, kalimat, atau aspek tertentu, dengan tujuan menentukan apakah pendapat yang disampaikan dalam dokumen atau kalimat tersebut bersifat positif, negatif, atau netral. Dalam konteks yang lebih kompleks, analisis sentimen

dapat mengidentifikasi nuansa emosional seperti kesedihan, kegembiraan, atau kemarahan berdasarkan teks yang terdapat dalam dokumen.

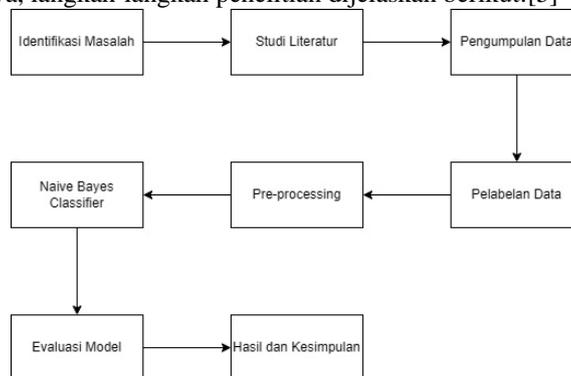
Dalam penelitian terdahulu, beberapa artikel digunakan untuk menganalisis sentimen publik pada tweet yang memuat pembangunan kereta cepat Jakarta-Bandung. Salah satunya Algoritma *Naive Bayes* yang digunakan untuk menghasilkan hasil yang sangat akurat dengan skor sentimen akurasi 73%, ketepatan 89%, dan recall 89% [2]. Studi analisis sentiment tentang kereta cepat Jakarta-Bandung lainnya membahas tentang “Perbandingan Model Algoritma Klasifikasi Pada Analisis Sentimen Opini Masyarakat Terhadap Layanan Kereta Cepat Jakarta-Bandung (The Whoosh)” dengan membandingkan algoritma *Support Vector Machine*, *Naive Bayes*, dan algoritma *K-NN* sebagai metode perbandingan pembobotan kata yang memiliki tingkat akurasi lebih dari 70% dengan kinerja yang baik [3]

Penelitian ini bertujuan untuk menjelajahi aplikasi *Naive Bayes* dalam analisis sentimen terhadap pengguna Kereta Cepat Jakarta-Bandung melalui data yang dikumpulkan dari media sosial twitter. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam memahami persepsi masyarakat terhadap layanan transportasi yang krusial ini serta memberikan dasar bagi perbaikan dan pengembangan layanan yang lebih baik di masa [4].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian dengan menggunakan algoritma *naive bayes classifier* untuk melakukan analisis sentimen masyarakat di Twitter. Dimulai dengan identifikasi, membaca literatur, mengumpulkan data dari Twitter, melabelkan data positif, negatif dan netral, melakukan *pre-processing*, dan kemudian menguji algoritma *naive bayes classifier* untuk menemukan hasil dan kesimpulan. Selanjutnya, langkah-langkah penelitian dijelaskan berikut: [5]



Gambar 1. Langkah-Langkah Tahapan Dalam Penelitian

2.2 Metode Pengumpulan Data

Twitter (X) adalah platform media sosial mikroblog tempat pengguna dapat berbagi berita blog. Namun platform membatasi jumlah karakter yang dapat diposting pengguna di setiap halaman profil pengguna maksimal 140 karakter [6]. Penelitian ini menggunakan media sosial Twitter dengan mengumpulkan 714 komentar atau data postingan dan mengumpulkan data dari *crawling*. Opini penggunaan kereta cepat di twitter dapat dijadikan sebagai sumber data untuk penelitian dengan menggunakan teknik *crawling*.

2.3 Pelabelan Data

Setelah proses pengumpulan data, langkah selanjutnya adalah pemberian pelabelan data yang diubah menjadi sentimen positif, negatif dan netral. Proses ini dilakukan manual.

2.4 Pre-processing Data

Proses preprocessing merupakan langkah penting untuk meningkatkan kualitas data mentah agar dapat diolah. Data yang masih mengandung simbol, hashtag, dan kata-kata yang tidak bermakna akan diperbaiki pada tahap ini. Pra-pemrosesan pada penelitian ini meliputi *cleansing*, *case folding*, *tokenisasi*, *stopword*, dan *stemming*. Pra-pemrosesan ini melakukan beberapa aktivitas transformasi data, yaitu [7]:

- Cleansing* : proses penghapusan karakter atau tanda baca, mention, hashtag, URL, dan karakter lain selain karakter.
- Case folding* : proses mengubah semua data teks menjadi huruf kecil.

- c. *Tokenizing* : proses pemotongan atau pemisahan kata yang dipisahkan dengan *whitespace* atau spasi.
- d. *Stopword removal* : proses menghilangkan kata-kata yang sering muncul pada data teks namun tidak mempunyai arti penting.
- e. *Stemming* : proses mengubah seluruh kata pada data teks menjadi kata dasar dengan menghilangkan awalan, sisipan, dan akhiran

2.5 Pembobotan TF-IDF

TF-IDF adalah singkatan dari *Term Frekuensi Invers Document Frequency* yaitu nilai numerik yang digunakan dalam pemrosesan bahasa alami untuk menunjukkan arti kata dalam suatu dokumen dibandingkan dengan sekumpulan dokumen lainnya[8]. Ada beberapa tahapan klasifikasi melakukan TF-IDF yaitu menjalankan klasifikasi *Naive Bayes*. TF-IDF bertujuan untuk menghitung bobot setiap kata yang digunakan sebagai fitur[9]. Semakin banyak dokumen yang diproses, semakin banyak pula fitur yang dikumpulkan.

2.6 Naive Bayes Classifier

Metode klasifikasi *naive bayes classifier* digunakan untuk menganalisis kelompok data dengan menggunakan perhitungan probabilitas. Prinsip inti *naive bayes* mencakup penerapan teorema *bayes*, sebuah konsep statistik untuk estimasi probabilitas. Pengklasifikasian *naive bayes* menghitung probabilitas setiap kelas berdasarkan atribut yang tersedia dan menentukan kelas yang optimal. Fungsinya untuk menghitung dan mencari nilai probabilitas tertinggi untuk pengklasifikasian data uji ke dalam kategori yang sesuai[10].

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

Keterangan :

X = Data dari class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

P(H) = Probabilitas hipotesis H (prior prob.)

P(X|H) = Probabilitas X berdasarkan kondisi tersebut

2.7 Evaluasi Model

Setelah melatih model *naive bayes*, penulis mengevaluasi kinerjanya menggunakan metrik-metrik evaluasi standar seperti akurasi, presisi, *recall*, dan *F1-score*. Klasifikasi ini mencakup melakukan penilaian yang terdiri dari matriks konfusi dan visualisasi dashboard. Matriks konfusi adalah tabel yang membantu menentukan seberapa akurat proses klasifikasi dilakukan[11]. Visualisasi dashboard dilakukan untuk menampilkan data yang diolah dalam bentuk grafik dan *cloud* kerja.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berdasarkan ulasan twitter terhadap postingan kereta cepat dengan mengumpulkan 714 data komentar. Data yang digunakan menggunakan *scraping* dengan bahasa *python* dan menggunakan perpustakaan *node.js*. Data twitter dikumpulkan pada tahun 2024[12]. Proses *scraping data* dilakukan dengan *query data whoosh* dan data tersebut diekstraksi ke dalam *excel* (lihat Gambar 2).

```
# Crawl Data

filename = 'whoosh.csv'
search_keyword = 'whoosh lang:id'
limit = 1500

!npx --yes tweet-harvest@2.2.8 -o "{filename}" -s "{search_keyword}" -l {limit} --token {twitter_auth_token}
```

Gambar 2. Proses *Crawling Data*

c. **Tokenizing**

Tokenizing proses pemecahan kata pada kalimat berdasarkan spasi.

Tabel 3. Proses *tokenizing*

Tiket	Whoosh	Maret	Dibelinya	Sih	kemaren
Nanya	Admin	Kai	Suruh	Check	berkala

Sumber : data peneliti

d. **Stopword**

Stopword merupakan proses menghilangkan kata yang tidak diperlukan atau tidak bermakna pada saat proses pengolahan data[17].

Tabel 4. Hasil *Stopword*

Text	Hasil <i>Stopword</i>
dulu banyak yang ngecengin saat pembangunan terowongan whoosh setelah jadi ya fomo nyobain begitu juga nanti kalau IKN udah kelar kalau ada yang yang nyerah terus ngatain itu biasa ada juga yang berjuang mencari cara bagaimana menyelesaikan masalah belajar memang mahal	ngecengin bangun terowongan whoosh ya fomo nyobain ikn udah kelar nyerah ngatain juang cari selesai ajar mahal

Sumber : data peneliti

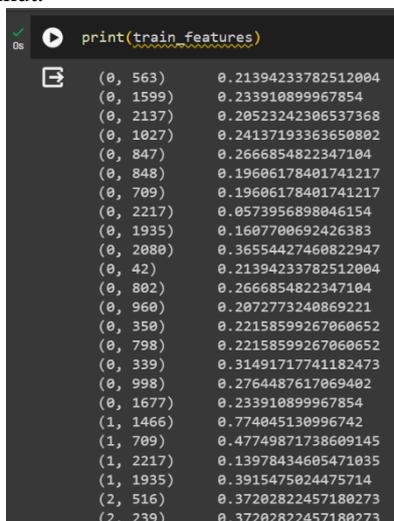
e. **Stemming**

Stemming merupakan proses mengubah kata yang berimbuhan menjadi kata dasar[18]

3.4 Implementasi *Naïve Bayes Classifier*

1. **Pembobotan *TF-IDF*.**

Hasil *preprocessing* teks diubah menjadi representasi vektor dari setiap dokumen dalam koleksi. Setiap dimensi vektor mewakili bobot relatif kata dalam dokumen[19]. Hasil pembobotan menggunakan *TF-IDF* yang ditunjukkan pada Gambar 4 berikut.



Gambar 4. Proses *TF-IDF*

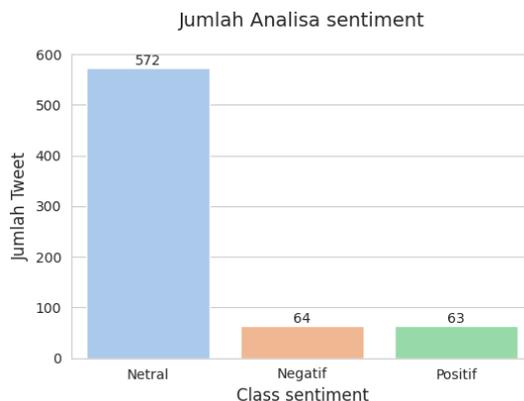
2. ***Splitting Data* dan *Pemodelan Klasifikasi***

Seluruh data teks dilakukan analisis dan melalui proses komputasi untuk menentukan skor setiap label atau nilai sentimen. Sentimen dengan skor tertinggi dijadikan hasil prediksi dengan menerapkan klasifikasi menggunakan algoritma *Naive Bayes*[20]. Visualisasi hasil klasifikasi sentimen menggunakan algoritma *Naive Bayes* dapat

dilihat pada gambar 5. Visualisasi dilakukan dengan menggunakan *widget vadersentiment* dengan menghitung *compound score stemming data* yang sudah melalui tahap *preprocessing*.

Tabel 5. Hasil *Scoring Naïve Bayes Teknik Vadersentiment*

Label	Review	Compound_Score	Naïve Bayes
Positif	oalah pantas nyoba cari jadwal argo parahyangan pilih jadwal dikit udah gitu mahal karna konsisten argo parahyangan naik harga goodbye ajah haha untung alternatif whoosh	0.4588	Positif
Negatif	whoosh target market nya tibatiba argo parahyangan middle to low class kena	-0.2732	Negatif
Neutral	transportasi lanjut stasiun whoosh tegalluar kota bandung kcic sedia area tunggu	0	Neutral



Gambar 5. Distribusi Klasifikasi *Naive Bayes*

Hasil analisis sentimen tersebut menampilkan data mayoritas tweet masuk dalam kategori sentimen netral, disusul sentimen positif dan negatif. Hasil tersebut menunjukkan bahwa mayoritas pengguna Twitter mempunyai opini netral terhadap layanan kereta cepat Jakarta-Bandung, namun ada juga yang memberikan opini positif dan negatif.

Selanjutnya, data yang dimiliki dibagi lagi untuk dilakukan *splitting stemming data* untuk dilakukan pengujian menjadi data latih dan data uji dengan perbandingan 80 : 20 yaitu 80% data latih dan 20% data uji. Dengan jumlah data masing-masing yaitu data latih 559 baris data dan data uji sebanyak 140 baris data

```
[85] X_train, X_test, y_train, y_test = train_test_split
      (df['stemming_data'], df['sentiment'], test_size=0.2, random_state=42)

print(f'Jumlah data latih : {len(X_train)}')
print(f'Jumlah data uji : {len(X_test)}')
```

Gambar 6. Proses *Splitting data*

Pada pengujian ini, *Algoritma Naive Bayes* menggunakan *parameter alpha=1* dan *fit_prior* kemudian diatur menjadi *False*. Model ini menghasilkan 120 prediksi benar positif dan 3 prediksi benar negatif. Lihat pada gambar 7 di bawah ini. Tahap ini menguji untuk menunjukkan bahwa pengguna kereta cepat Jakarta-Bandung pada media sosial Twitter setelah dilakukan *splitting* cenderung memberikan sentimen positif terhadap penggunaan tersebut. Hal ini dapat disebabkan oleh beberapa faktor, seperti mayoritas tweet data yang menyebutkan kata mengenai kenyamanan dan kecepatan layanan kereta cepat Jakarta-Bandung yang memuaskan.



Gambar 7. Confusion Matrix Naïve Bayes

Untuk menghitung hasil tingkat akurasi tersebut apakah sudah sesuai, digunakan persamaan berikut:

$$Accuracy\ whoosh = \frac{TP+TN}{P}$$

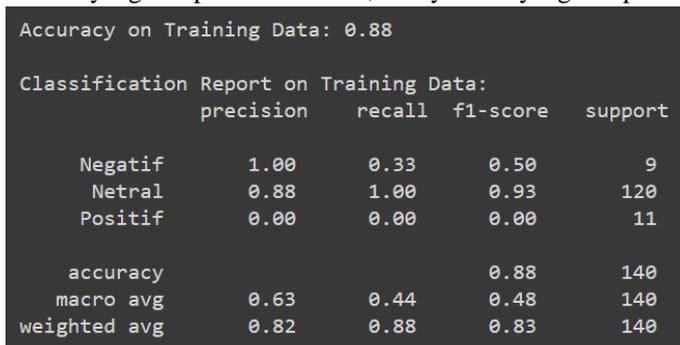
$$Accuracy\ whoosh = \frac{120+3}{140}$$

$$Accuracy\ whoosh = \frac{123}{140} = 0,878$$

Berdasarkan sample perhitungan kata sentimen diatas, kita ketahui hasil perhitungan probabilitas mendapat kesimpulan nilai probabilitas tertinggi yaitu positif.

3. Evaluasi Model dan Visualisasi

Setelah melatih model *Naive Bayes* selanjutnya dilakukan evaluasi performa menggunakan metrik evaluasi standar seperti akurasi, presisi, recall, dan skor F1. Akurasi mengukur seberapa baik model dapat mengklasifikasikan tweet dengan benar ke dalam kategori sentimen yang benar. *Presisi* mengukur proporsi tweet yang diklasifikasi dengan benar sebagai sentimen tertentu dari total tweet yang diklasifikasikan ke dalam kategori tersebut. *Recall* mengukur proporsi tweet dengan benar dari jumlah total tweet yang benar-benar termasuk dalam kategori tersebut. Skor F1 adalah rata-rata harmonik antara presisi dan *recall*. Pada gambar 8 di bawah ini, dapat dilihat hasil sentimen yang digunakan dalam perhitungan algoritma *Naive Bayes Classifier* menggunakan *Python*. Akurasi yang dicapai sebesar 88%, artinya hasil yang dicapai sangat baik.



Gambar 7. Hasil Evaluasi Model Naive bayes

Selanjutnya, hasil visualisasi dari analisis sentimen ini disebut *Word Cloud* yaitu visualisasi yang memuat sebagian besar kata kunci atau kata yang sering muncul pada data yang dihasilkan, seperti “Whoosh”, “Cepat”, “Stasiun”, “Kereta”, “Bandung” dan “KCIC”.



Gambar 8. Hasil *Word Cloud*

4. KESIMPULAN

Setelah melalui pemrosesan dengan teknik *vadersentiment* dan *splitting stemming* data, hasil sentimen negatif sebanyak 64, hasil sentimen positif sebanyak 63, dan hasil sentimen netral sebanyak 572 dapat disimpulkan bahwa hasil dari tanggapan masyarakat mengenai penggunaan Kereta Cepat Jakarta Bandung setelah dilakukan tahap *splitting* data *compound score* tergolong positif dan bekerja dengan baik dengan menghasilkan nilai akurasi cukup tinggi yaitu 88%, presisi 82%, dan *recall* 88% dengan melalui proses normalisasi kata yang sudah diubah sesuai dengan pengujian kamus kata baku. Meskipun begitu, masih banyak kata data yang tidak terklasifikasi dengan baik atau dapat dikatakan sebagai data yang salah klasifikasi. Terdapat total 396 data yang belum terklasifikasi secara lengkap pada data berlabel netral yang dianggap sulit dideteksi. Dengan demikian, penelitian ini memberikan kontribusi penting dalam memahami sentimen pengguna terhadap penggunaan Kereta Cepat Jakarta-Bandung melalui analisis data yang dikumpulkan dari media sosial Twitter, serta menyoroti pentingnya penerapan metode klasifikasi *Naive Bayes* dalam menganalisis sentimen secara efektif.

REFERENCES

- [1] O. Dina *et al.*, “Jurnal Ilmiah ‘Neo Politea’ FISIP Universitas Al-Ghifari PERSPEKTIF EKOLOGI ADMINISTRASI: PEMBANGUNAN INFRASTRUKTUR KERETA API CEPAT JAKARTA-BANDUNG,” 2021.
- [2] J. Florensus Sianipar, Y. R. Ramadhan, and I. Jaelani, “Analisis Sentimen Pembangunan Kereta Cepat Jakarta-Bandung di Media Sosial Twitter Menggunakan Metode Naive Bayes,” *Media Online*, vol. 4, no. 1, pp. 360–367, 2023, doi: 10.30865/klik.v4i1.1033.
- [3] S. Wahyu, “Konferensi Nasional Ilmu Komputer (KONIK) 2023 Perbandingan Model Algoritma Klasifikasi Pada Analisis Sentimen Opini Masyarakat Terhadap Layanan Kereta Cepat Jakarta Bandung (The Whoosh)”.
- [4] I. Iwandini, A. Triayudi, and G. Soepriyono, “Analisa Sentimen Pengguna Transportasi Jakarta Terhadap Transjakarta Menggunakan Metode Naives Bayes dan K-Nearest Neighbor,” *Journal of Information System Research (JOSH)*, vol. 4, no. 2, pp. 543–550, Jan. 2023, doi: 10.47065/josh.v4i2.2937.
- [5] A. Asroviana Putri and R. Febrian Umbara, “Analisis Sentimen terhadap Opini Mahasiswa pada Survei Kepuasan terhadap Fasilitas dan Layanan Universitas Telkom dengan menggunakan Algoritma Naïve Bayes.”
- [6] N. Tri Romadloni, I. Santoso, S. Budilaksono, and M. Ilmu Komputer STMIK Nusa Mandiri Jakarta, “PERBANDINGAN METODE NAIVE BAYES, KNN DAN DECISION TREE TERHADAP ANALISIS SENTIMEN TRANSPORTASI KRL COMMUTER LINE.”
- [7] I. Bagus Nyoman Wijana Manuaba *et al.*, “Volume 5 Nomor 2 Maret Tahun 2022 Analisis Sentimen Data Provider Layanan Internet Pada Twitter Menggunakan Support Vector Machine (SVM) Dengan Penambahan Algoritma Levenshtein Distance.”
- [8] R. T. Wahyuni, D. Prastiyanto, and D. E. Suprpto, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi.”

-
- [9] A. Apriani, H. Zakiyudin, and K. Marzuki, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta," *Jurnal Bumigora Information Technology (BITe)*, vol. 3, no. 1, pp. 19–27, Jul. 2021, doi: 10.30812/bite.v3i1.1110.
- [10] R. Fajar, S. Program, P. Rekayasa, N. Lunak, and R. Bengkalis, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," vol. 3, no. 1.
- [11] K. Yan, D. Arisandi, and) Tony, "Jurnal Ilmu Komputer dan Sistem Informasi ANALISIS SENTIMEN KOMENTAR NETIZEN TWITTER TERHADAP KESEHATAN MENTAL MASYARAKAT INDONESIA."
- [12] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.
- [13] S. Fransiska, "ANALISIS SENTIMEN TWITTER UNTUK REVIEW FILM MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER (NBC) PADA SENTIMEN R PROGRAMMING," *Jurnal Siliwangi*, vol. 5, no. 2, 2019.
- [14] P. S. M. Suryani, L. Linawati, and K. O. Saputra, "Penggunaan Metode Naive Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia," *Majalah Ilmiah Teknologi Elektro*, vol. 18, no. 1, p. 145, May 2019, doi: 10.24843/mite.2019.v18i01.p22.
- [15] S. Hikmawan *et al.*, "Halaman: 167-176 Terakreditasi Peringkat 5 (SINTA 5) sesuai SK RISTEKDIKTI Nomor," 2020. [Online]. Available: <http://ejournal.ubharajaya.ac.id/index.php/JKI>
- [16] M. Hudha, E. Supriyati, and T. Listyorini, "ANALISIS SENTIMEN PENGGUNA YOUTUBE TERHADAP TAYANGAN #MATANAJWAMENANTITERAWAN DENGAN METODE NAIVE BAYES CLASSIFIER," *Jurnal Informatika dan Komputer) Akreditasi KEMENRISTEKDIKTI*, vol. 5, no. 1, pp. 2614–8897, 2022, doi: 10.33387/jiko.
- [17] A. H. Tri Jaka, "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining."
- [18] N. M. A. J. Astari, Dewa Gede Hendra Divayana, and Gede Indrawan, "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier," *Jurnal Sistem dan Informatika (JSI)*, vol. 15, no. 1, pp. 27–29, Nov. 2020, doi: 10.30864/jsi.v15i1.332.
- [19] R. P. Sidiq, B. A. Dermawan, and Y. Umaidah, "Sentimen Analisis Komentar Toxic pada Grup Facebook Game Online Menggunakan Klasifikasi Naive Bayes," *Jurnal Informatika Universitas Pamulang*, vol. 5, no. 3, p. 356, Sep. 2020, doi: 10.32493/informatika.v5i3.6571.
- [20] A. Z. Amrullah, A. Sofyan Anas, M. Adrian, and J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *Jurnal*, vol. 2, no. 1, 2020, doi: 10.30812/bite.v2i1.804.