

Analisis Sentimen Komentar Warga Twitter terhadap Calon Wakil Presiden Nomor Urut 02 saat Debat Cawapres Menggunakan Naive Bayes

Nawal Aulia¹, Akmal Zidan Fahreza², Muhammad Ridho Habibie³, Anggraini Puspita Sari⁴
^{1,2,3} Universitas Pembangunan Nasional “Veteran” Jawa Timur, Surabaya

Email : 22081010097@student.upnjatim.ac.id¹, 22081010163@student.upnjatim.ac.id²,
22081010246@student.upnjatim.ac.id³

*Corresponding author email : anggraini.puspita.if@gmail.com

Abstrak

Debat calon wakil presiden Indonesia yang dilaksanakan pada 21 Januari 2024, menuai banyak komentar positif maupun negatif dari masyarakat, yang menimbulkan pro-kontra di kalangan masyarakat pada media sosial Twitter. Tujuan penelitian ini adalah untuk mengolah dan menganalisis hasil sentimen dari komentar masyarakat Indonesia terkait debat calon wakil presiden. Penelitian ini mengambil dataset sebanyak 367 data: 181 data positif, 52 data negatif, dan 135 data netral yang berasal dari komentar warga Twitter dan berfokus pada calon wakil presiden nomor urut 02 yaitu Gibran Rakabuming Raka. Metode yang digunakan untuk mengolah data yakni algoritma Naïve Bayes. Langkah-langkah yang dilakukan dalam penelitian ini yakni *cleaning data*, *Preprocessing*, *Translating*, *Labeling*, *Visualization*, dan *data training*. Untuk melakukan analisis sentiment, dataset dibagi menjadi dua bagian dengan perbandingan 70:30 untuk pelatihan dan pengujian. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 83.70%.

Kata kunci: Debat, Komentar, Naive Bayes

Abstract

The debate for Indonesia's vice-presidential candidates held on January 21, 2024, garnered a significant amount of both positive and negative comments from the public, leading to a pro and contra situation on the social media platform Twitter. The aim of this research is to process and analyze the sentiment results from the Indonesian public's comments regarding the vice-presidential debate. This study collected a dataset comprising 367 entries: 181 positive comments, 52 negative comments, and 135 neutral comments, sourced from Twitter users' remarks and focused on the vice-presidential candidate number 02, Gibran Rakabuming Raka. The method used for data processing is the Naïve Bayes algorithm. The steps undertaken in this research include data cleaning, preprocessing, translating, labeling, visualization, and data training. For sentiment analysis, the dataset is divided into two parts with a 70:30 ratio for training and testing. The evaluation results indicate that the model has an accuracy of 83.70%.

Keywords: Debate, Comments, Naive Bayes

PENDAHULUAN

Pemilihan umum (Pemilu) merupakan cara Indonesia menjalankan demokrasi. Demokrasi Indonesia terwujud melalui Pemilu yang digunakan dalam menentukan perwakilan rakyat dan pejabat publik lainnya (Subiyanto, 2020). Awal tahun 2024 ini Indonesia akan mengadakan pemilihan umum untuk memilih presiden serta wakil presiden. Pemilu adalah momen bagi rakyat Indonesia untuk memilih pemimpin. Pada pemilu kali ini, ada tiga pasangan calon presiden dan wakil presiden yang akan bersaing. Karena itu, Komisi Pemilihan Umum (KPU) memberi kesempatan bagi semua kandidat untuk menyampaikan visi dan misi mereka didalam debat antar calon. Debat ini dimulai pada bulan Desember 2023 hingga bulan Februari 2024, dengan kesempatan debat sebanyak lima kali yang membahas tema yang berbeda-beda, dimana tiga debat dilakukan oleh antar calon presiden dan dua debat lainnya dilakukan oleh antar calon wakil presiden. Hasil dari perdebatan ini, selalu menuai komentar dari masyarakat, baik komentar negatif maupun komentar positif, sehingga selalu menjadi topik hangat di sosial media.

Media sosial merupakan tempat yang memungkinkan orang-orang untuk berbicara dan berbagi informasi, termasuk blog, wiki, media sosial, forum, dan dunia virtual (Cahyono, 2016). Media sosial memiliki banyak manfaat, seperti mendapat kenalan baru, menemukan beragam informasi serta isu politik di Indonesia. Pemilu tahun 2024 di Indonesia, sebagai pesta demokrasi terbesar, menjadi topik menarik bagi pengguna media sosial untuk berbagi pendapat.

Twitter merupakan media sosial yang digunakan sebagian besar masyarakat

untuk berbagi informasi dan opini, aplikasi ini menyediakan berbagai fitur seperti mengambil gambar serta video, juga fitur utama yaitu saling balas tweet guna mempermudah interaksi antar pengguna. Twitter telah menjadi bagian penting dari kehidupan sehari-hari (Aggrawal, 2018). Dengan Twitter, kita bisa menyampaikan opini objektif tentang berbagai topik, sehingga menjadikannya salah satu media sosial yang populer (Colleta, Ponti, Hruschka, Acharya, & Ghosh, 2019). Debat calon wakil presiden adalah salah satu topik yang ramai dibicarakan oleh masyarakat di media sosial Twitter. Pasalnya salah satu calon wakil Presiden menarik perhatian didalam debat tersebut. Karena banyaknya Tweet yang ada, memerlukan durasi yang cukup lama untuk mengklasifikasikannya. Oleh sebab itu, diperlukan suatu metode untuk mengelompokkan data Twitter tersebut agar bisa dianalisis dengan mengelompokkannya ke dalam beberapa kategori.

Analisis sentimen adalah bidang studi yang mengkaji sikap, pendapat, perasaan, penilaian, dan emosi seseorang dalam tulisan (Vindua & Zailani, 2023). Penelitian ini akan dibagi menjadi tiga kategori: positif, netral, dan negatif. Karena Indonesia adalah negara demokrasi, warganya bebas beropini, yang bisa menghasilkan dampak positif maupun negatif, seperti keributan politik. Tidak semua pengguna Twitter berhati-hati dalam memilih kata-kata dalam postingannya, dan banyak dari mereka yang mengekspresikan diri dengan kata-kata kasar dan ofensif (Hidayatullah, Fadila, Juwairi, & Nayoan, 2019). Dokumen-dokumen tersebut bisa dikelompokkan menggunakan metode klasifikasi, dan dalam penelitian ini kami menggunakan metode Naïve Bayes. Metode

Naive Bayes sering dipilih oleh peneliti karena cukup efektif dan bisa menghasilkan hasil yang akurat (Imam Kurniawan, 2019). Terdapat beberapa penelitian yang telah dilakukan tentang

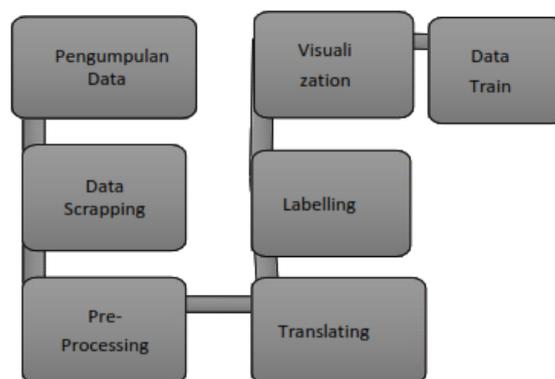
analisis sentimen menggunakan metode klasifikasi Naïve Bayes. Salah satunya adalah penelitian yang menggabungkan metode Naive Bayes dan K-Means untuk menganalisis opini masyarakat terhadap pencalonan presiden dan wakil presiden 2019 (Imam Kurniawan, 2019). Penelitian tersebut menghasilkan tingkat akurasi yang tinggi yaitu sebesar 93,35%. Berdasarkan penjelasan tersebut, dilakukan penelitian berjudul "Analisis Sentimen Pengguna Twitter terhadap calon Wakil Presiden nomor urut 02 pada Debat Calon Wakil Presiden menggunakan Naïve Bayes." Penelitian ini bertujuan untuk mengklasifikasikan opini masyarakat tentang debat calon wakil presiden menjadi tiga kategori: negatif, positif, atau netral.

METODE

Naive bayes banyak digunakan karena lebih efektif jika dibandingkan dengan metode lain dalam hal klasifikasi (Dina Agustina, 2022). Metode ini sering digunakan dalam lingkup analisis sentimen karena berbagai faktor, salah satunya yaitu naive bayes menyuguhkan algoritma perhitungan yang cepat dan menghasilkan akurasi yang cukup tinggi sehingga dapat membantu suatu sistem bekerja lebih optimal (Harjanta, 2015). Pada Naïve bayes, kita dapat menggunakan rumus sebagai berikut:

$$\text{Akurasi Data} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data}}$$

Alur metode dapat dilihat pada alur



Gambar 1. Alur metode

A. Pengumpulan Data

Kami menggunakan *google colab* sebagai tools dan bahasa pemrograman *python*, selain itu kami menggunakan beberapa library tambahan seperti *pandas*, *seaborn*, *matplotlib*, *sastrawi* dan lain sebagainya. Data yang digunakan merupakan hasil scrapping data dari aktifitas sosial media *twitter*.

B. Data Scrapping

Dataset yang terkumpul didapatkan dari aplikasi twitter. kami menggunakan metode data scrapping dengan modul *Tweet-Harvest*, kami menambahkan keyword "gibran" saat melakukan pengumpulan data dengan tujuan agar data- data yang kami dapat lebih berfokus pada topik kami. Setelah data terkumpul, kemudian dataset dibagi menjadi dua bagian, yaitu data train (70%) dan data uji (30%).

C. Pre-Processing

Proses preprocessing mengeliminasi data yang berlebihan dan tidak relevan sebelum dataset diproses (Rasenda, 2020). Pada tahap ini data akan disiapkan menjadi data yang siap untuk diproses oleh model. tahapan Preprocessing terdiri dari:

1. Data Cleaning, adalah proses untuk membersihkan karakter yang tidak diperlukan seperti angka, tanda baca, icon, username (@username), link url, dan lain – lain. Tujuannya adalah untuk

mengurangi noise pada teks (Subagio, Sari, & Sihananto, 2024).

2. Case Folding merupakan proses memodifikasi kalimat dalam teks menjadi satu bentuk saja, biasanya semua huruf besar (*uppercase*) atau semua huruf kecil (*lowercase*) (Thariq Muhammad Firdausy, 2022).
3. Normalisasi data adalah proses mengubah kata-kata tidak resmi dalam dataset menjadi kata-kata resmi, misalnya mengubah "gugel" menjadi "google".
4. Tokenisasi adalah proses memecah kalimat atau paragraf menjadi kata-kata individual yang berdiri sendiri.
5. Stemming adalah proses mengubah kata berimbuhan menjadi bentukdasar.

D. Translating

Pada tahap translating, komentar yang sudah melalui tahap *Pre-processing* akan diubah menjadi Bahasa Inggris agar mempermudah pengolahan data pada kata.

E. Labelling

Labeling dalam Naive Bayes merujuk pada memberikan kelas atau label kepada setiap contoh dalam dataset. Hal ini penting karena dalam melatih suatu model memerlukan data yang telah diberi label. Dengan menggunakan informasi tersebut, model dapat memahami distribusi fitur-fitur dalam setiap kelas dan menghitung probabilitas kelas untuk contoh baru.

F. Visualization

Visualization digunakan untuk memberikan gambaran kata apa saja yang sering muncul pada dataset komentar. Semakin besar visualisasinya, semakin sering kata tersebut muncul

G. Data Train

Data train pada Naive Bayes adalah kumpulan data berlabel yang digunakan untuk melatih model dalam menghitung

probabilitas apriori dan probabilitas kondisional fitur terhadap kelas. Melalui data ini, model mempelajari hubungan antara fitur dan kelas sehingga dapat membuat prediksi untuk data baru berdasarkan aturan Bayes.

HASIL DAN PEMBAHASAN

A. Data Scrapping

Data scrapping menggunakan Twitter-Harvest dan menerapkan filter untuk hanya mengambil komentar berbahasa Indonesia, metode ini menghasilkan sebanyak 368 data. Dan mendapatkan hasil sebagai berikut.

Gambar 3. Dataset lengkap

Pada Gambar 3, menunjukkan data diperoleh sejumlah 368 data dan kemudian dilanjutkan dengan pemilihan atribut, pemilihan atribut bertujuan untuk hanya menampilkan atribut yang akan digunakan nanti. Hasil ditunjukkan pada Gambar 4.

Gambar 4. Dataset yang telah dikurangi atributnya

A. Pre-Processing

Tahap ini data akan disiapkan melalui beberapa proses hingga data siap untuk diolah oleh model

1. Data Cleaning

Data dibersihkan dari karakter spesial seperti dataset akan dibersihkan

dari karakter spesial seperti “@”, “RT”, dan “”HTTPS”. Hasil diperlihatkan pada gambar 5.

	full_text	username	created_at
0	Gibran bikin Solo eksis cak Imin mungkin masih...	jaenabet261330	Mon Dec 25 11:29:06 +0000 2023
1	Gibran berhasil bikin Solo eksis cak Imin mung...	AsepSusant	Mon Dec 25 11:29:02 +0000 2023
2	Gibran bikin Solo jadi pusat perhatian cak Imi...	nawang_nugroho	Mon Dec 25 11:28:59 +0000 2023
3	Urang Jabar naha ikhlas disebut keun bakal nga...	h68052	Mon Dec 25 11:28:57 +0000 2023
4	Gibran gesit cak Imin agak lambat kirakira har...	elisuwarsih	Mon Dec 25 11:28:57 +0000 2023
...
363	lahhh katanya silahkan cek di gugel udah banya...	zidnsprta	Mon Dec 25 11:06:24 +0000 2023
364	Yang penting bukan petugas partai apalagi bokeper	Gr4sssss	Mon Dec 25 11:06:23 +0000 2023
365	Solo membuktikan bahwa Prabowo Gibran punya re...	soemardihasan	Mon Dec 25 11:05:49 +0000 2023
366	tapi fans nya yg muslim pada gaterima beliau d...	nilrofficial	Mon Dec 25 11:05:46 +0000 2023
367	Solo yang sukses di bawah Gibran adalah bukti ...	MarlianaSarii	Mon Dec 25 11:05:43 +0000 2023

Gambar 5. Data Cleaning

Jika gambar 4 dan 5 dibandingkan, akan terlihat bahwa karakter spesial yang terlihat pada gambar 2 telah hilang di gambar

2. Case Folding

Ditahap ini akan dilakukan konversi kalimat yang awalnya terdapat huruf besar akan diubah menjadi seragam berhuruf kecil seperti pada gambar 6.

	full_text	username	created_at
0	gibran bikin solo eksis cak imin mungkin masih...	jaenabet261330	Mon Dec 25 11:29:06 +0000 2023
1	gibran berhasil bikin solo eksis cak imin mung...	AsepSusant	Mon Dec 25 11:29:02 +0000 2023
2	gibran bikin solo jadi pusat perhatian cak imi...	nawang_nugroho	Mon Dec 25 11:28:59 +0000 2023
3	urang jabar naha ikhlas disebut keun bakal nga...	h68052	Mon Dec 25 11:28:57 +0000 2023
4	gibran gesit cak imin agak lambat kirakira har...	elisuwarsih	Mon Dec 25 11:28:57 +0000 2023
...
363	lahhh katanya silahkan cek di gugel udah banya...	zidnsprta	Mon Dec 25 11:06:24 +0000 2023
364	yang penting bukan petugas partai apalagi bokeper	Gr4sssss	Mon Dec 25 11:06:23 +0000 2023
365	solo membuktikan bahwa prabowo gibran punya re...	soemardihasan	Mon Dec 25 11:05:49 +0000 2023
366	tapi fans nya yg muslim pada gaterima beliau d...	nilrofficial	Mon Dec 25 11:05:46 +0000 2023
367	solo yang sukses di bawah gibran adalah bukti ...	MarlianaSarii	Mon Dec 25 11:05:43 +0000 2023

Gambar 6. Case Folding

Terlihat pada gambar 6, huruf besar yang ada pada awal kalimat pada gambar sebelumnya telah hilang.

3. Normalisasi

Setelah melalui proses Case Folding, data akan dinormalisasi dengan mengubah kata yang disingkat dan kata yang tidak baku menjadi bentuk yang benar. Hasil dari normalisasi ini ditunjukkan pada gambar 7.

	full_text	username	created_at
0	gibran bikin solo eksis cak mungkin masih menc...	jaenabet261330	Mon Dec 25 11:29:06 +0000 2023
1	gibran berhasil bikin solo eksis cak mungkin m...	AsepSusant	Mon Dec 25 11:29:02 +0000 2023
2	gibran bikin solo jadi pusat perhatian cak nam...	nawang_nugroho	Mon Dec 25 11:28:59 +0000 2023
3	urang jabar naha ikhlas disebut keun bakal nga...	h68052	Mon Dec 25 11:28:57 +0000 2023
4	gibran gesit cak agak lambat kirakira harusnya...	elisuwarsih	Mon Dec 25 11:28:57 +0000 2023
...
363	lahhh katanya silahkan cek di google udah bany...	zidnsprta	Mon Dec 25 11:06:24 +0000 2023
364	yang penting bukan petugas partai apalagi bokeper	Gr4sssss	Mon Dec 25 11:06:23 +0000 2023
365	solo membuktikan bahwa gibran punya rencana ek...	soemardihasan	Mon Dec 25 11:05:49 +0000 2023
366	tapi fans nya yang muslim pada gaterima beliau...	nilrofficial	Mon Dec 25 11:05:46 +0000 2023
367	solo yang sukses di bawah gibran adalah bukti ...	MarlianaSarii	Mon Dec 25 11:05:43 +0000 2023

Gambar 7. Normalisasi

Terlihat pada baris ke 363 pada gambar 6 yang menampilkan kata “gugel”, diubah pada gambar 7 yang telah dilakukan normalisasi menjadi kata “google”.

4. Tokenisasi

Proses memisahkan atau memecah sebuah kalimat menjadi kata yang berdiri sendiri. Hasil dari pemisahan, terdapat pada gambar 8.

```
[ ] tokenized = df['full_text'].apply(lambda x:x.split())
tokenized
0      [gibran, bikin, solo, eksis, cak, mungkin, mas...
1      [gibran, berhasil, bikin, solo, eksis, cak, mu...
2      [gibran, bikin, solo, jadi, pusat, perhatian, ...
3      [urang, jabar, naha, ikhlas, disebut, keun, ba...
4      [gibran, gesit, cak, agak, lambat, kirakira, h...
...
363     [lahhh, katanya, silahkan, cek, di, google, ud...
364     [yang, penting, bukan, petugas, partai, apalag...
365     [solo, membuktikan, bahwa, gibran, punya, renc...
366     [tapi, fans, nya, yang, muslim, pada, gaterima...
367     [solo, yang, sukses, di, bawah, gibran, adalah...
Name: full_text, Length: 368, dtype: object
```

Gambar 8. Tokenisasi

5. Stemming

Stemming adalah proses menghapus imbuhan dari kata, seperti me-, -kan, di-, -an, dan lainnya. Contoh hasil stemming bisa dilihat pada gambar 9.

media asing label gibran rakabundung nepo baby apa itu teliti bukti nah lhong sampai nas gibranongotau gimana lah umur bawah 40 apa punya hak denda 2014 dan terus ya bisa nolak buat anak sino solo tarut soya pasang dan taut sama gibran kena me unggala pake narasi gantung kami i di dalam amplop gibran bukti cerdas solo gonilang jakarta dan jeteng kacau di bawah mahasin gibran bikin solo hebat cak wawawawaw masih sialah nyari taruk di partai politik yg aneh itu elo gibran babulu di kuly nasional sial ye sori ye dedak masi go lah naha dia pabal hlu wakan pake sate kaling teran kawang loh kalau pengembili bijak ubah ke arah lha baik kan bagus lah usung pres Jokowi saja ubah contoh ikn gak masuk rpijan tapi mah jd prior berutili cak fakta dia makanya gubillah itu yang di video slapa itu yang sangkem slapa jadi buzer emg harus boga sih tp ya boga jangen di mb spt negara milik jokowi kang cemo-cemoit pada gibran bikin solo hebat cak kayak masih sialah nyari hati di dunia politik dlu nggak nyambung hahah gaweset gibran nih maskkeba urang mau nanya klo ada tino ngumpanya di hal ato di lombo itu gema y nongkep nyam k stik mau di tangkep klo udh balik terap bawahan presiden indonesia telah jinnah wawaw jadi jinnah realisasi semua dan akhir dgn indah gak gagal prabowo gibran buruh jepura kila solid satu suara dekad08 wawaka sekandung la moyet ajak pilin dan masuk kandang wawakak outat gibran bikin waki anak mada tani unil prapomp konstitusi go usah bawo anak mada deh greli caupre nomor urut 2 gibran rakabundung aku baru tahu saat debat dan wawingwawing calm boleh bawo kontak hal yang makud gibran adal omong ngaduraw panti tahu sora tempo urus dagan murabah ato bangwet padahal sdi topang modal milyaran cetah dari mana uang tni ma soya orang dr sekian banyak duhak orbi ragu mampu dan trah debat kwarin pa mahi gak percaya yang tampil itu anda solo jadi pambawa jakarta malah jadi lombo sendiri cak boga dong tunjuk gibran berangg yang animo kening dia sendiriitu yang berapangwawing gibran solo kece cak jakarta malah kacau gibran bukti bawaw solo kila lebih sialer daripada jakarta dengan prestasi yang luar biasa kirain ayah gibran gibran di solo oloosoi lonjak somentara cak di mana ya gibran sukses baw solo ke puncak somentara jakarta dan jeteng belm tahu arah di bawah mahasin apresiasi debat caupre: manan pandawa lino gibran memu saling solo jawa gibran pialawa dengan cerdas sayang cak gugal alir prestasi mungkin lebih cocok jadi tonton gibran hasil solo jadi pusat perhati cak mungkin perlu introspeksi diri untuk kejar tingal gibran bukti pialaw solo sialer lebih dari jakarta dan jeteng cak ungerit barwa lebih wawaw dalam kerja cc apakah anda benar bohong cak data yang lengkap baru jadi caupre: aja sdi bohong ke rakyat gibran bikin solo eks cak kayak masih sialah nyari kibur di dunia politik solo wawaw: nawa gibran somentara jakarta dan jeteng perlu baik bawaw

Gambar 9. Stemming

Tabel 2. Hasil Prediksi Benar

No.	Teks	Prediksi	Label sebenarnya
1	gibran make solo ex cak maybe still looking for the right political stage	Positif	Positif
2	urang jabar naha ikhlas said keun would support gibran ulah poho 2019 dibobodo prabowo	Netral	Netral
...
...
307	but his Muslim fans on Gaterima he baptized so yohanes	Netral	Netral
308	a successful solo under the gibran is proof that positive change can be achieved	Positif	Positif

Tabel 3. Hasil Prediksi Salah

No	Teks	Prediksi	Label Sebenarnya
1	gibran results make solo ex cak may still teach how to appear on the political screen	Positif	Netral
2	gibran makes solo the center of attention cak seems to be still busy looking for fans in the world of politics	Positif	Netral
...
...
...
59	samsul ngibulnya talent comes down from mukidi	Netral	Negatif
60	lahhh said silah check on google there are many private companies that invest in clowns	Netral	Positif

SIMPULAN (PENUTUP)

Analisis sentimen masyarakat terhadap calon wakil presiden nomor urut 02 di Twitter menggunakan metode Naïve Bayes dilakukan dengan data sebagai berikut: 181 data positif, 52 data negatif, dan 135 data netral. Dataset ini dibagi menjadi dua bagian dengan perbandingan 70:30 untuk pelatihan dan pengujian. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 83.70%. Dari hasil ini, dapat disimpulkan bahwa model memiliki akurasi yang baik dengan performa yang cukup memadai.

DAFTAR PUSTAKA

- Aggrawal, N. (2018). Detection of Offensive Tweets: A Comparative Study. *Computer Reviews Jurnal*, Vol 1 No 1 (2018) ISSN: 2581-6640. Bhatia, S., Sharma, M., & Bhatia, K. K. (2017). Sentiment Analysis and Mining of Opinions. *Stud. Big Data*, Vol. 30, Pp. 503-523.
- Cahyono, A. S. (2016). Pengaruh Media Sosial Terhadap Perubahan Sosial Masyarakat di Indonesia. *Publiciana*, Hal. 140-157.
- Colleta, L. F., Ponti, M., Hruschka, E. R., Acharya, A., & Ghosh, J. (2019). Combining clustering and active learning for the detection and learning of new image classes. *Neurocomputing*, Vol. 358, Pages 150-165.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL. *Jurnal Teknokompak*, Vol 15, No 1 (2021).
- Dina Agustina, F. R. (2022). Analisis Sentimen pada Sosial Media Twitter terhadap MRT Jakarta Menggunakan Machine Learning. *Information System Research Jurnal*, 2022, Vol. II, Nomor 1.
- Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, Vol18, No 1 (2020).
- Harjanta, A. T. (2015). Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Jurnal Informatika Upgris*, Vol.1, No 1.
- Hidayatullah, A. F., Fadila, A. A., Juwairi, K. P., & Nayoan, R. A. (2019). Identifikasi Konten Kasar Pada Tweet Bahasa Indonesia. *Jurnal Linguistik Komputasional*, Vol. 2 No.1.
- Imam Kurniawan, A. S. (2019). Implementasi Metode K-Means dan Naïve Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019. *JURNAL EKSPLORA INFORMATIKA*.
- Rasenda, H. L. (2020). Implementasi K-NN Dalam Analisa Sentimen Riba Pada Bunga Bank. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, IV, Nomor II, 369-376.
- Subagio, A. W., Sari, A. P., & Sihananto A.N. (2024). Klasifikasi Lexicon-Based Sentiment Analysis Tragedi Kanjuruhan pada Twitter Menggunakan Algoritma Convolutional Neural Network. *JUISIK*, Vol. 4 No. 1 (2024).
- Subiyanto, A. E. (2020). Pemilihan Umum Serentak yang Berintegritas sebagai Pembaruan Demokrasi Indonesia. *Jurnal Konstitusi*, Vol. 17, No. 2, 2020.
- Thariq Muhammad Firdausy, P. P. (2022). Deteksi Iklan pada Twit menggunakan Metode Naive Bayes. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, VI No. I, 204-209.
- Vindua, R., & Zailani, A. U. (2023). Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python. *JURIKOM : Jurnal Riset Komputer*, Vol. 10, No.