

Analysis of Public Sentiment towards Corruption Based on Tweets Using Naive Bayes Classifier

Alivia Zulzila, Latifah Jayatri Febiola, Dodi Vionanda*

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: dodi_vionanda@fmipa.unp.ac.id

Submitted : 10 Februari 2025

Revised : 12 Februari 2025

Accepted : 28 Februari 2025

ABSTRACT

Corruption is one of the big problems faced in Indonesia. The high level of corruption can damage the integrity of government, hamper economic growth, and reduce public trust in public institutions. Even though the government has made efforts to eradicate corruption, such as the formation of the Corruption Eradication Commission (KPK), these big challenges remain. Social media, especially Twitter, has become an important platform for people to voice opinions and criticize corruption issues. Sentiment analysis is the process of identifying and classifying opinions or emotions from a text into positive, negative, and neutral categories. The textual classification algorithm used in this research is Naive Bayes. This research aims to determine public sentiment towards corruption in Indonesia in positive, negative and neutral categories. This is done by data preprocessing, data labeling, and classification. The results of sentiment classification using the Naive Bayes method obtained positive sentiment of 11, negative sentiment of 14, and neutral sentiment of 1485. So it can be concluded that neutral sentiment reflects that the public tends to be cautious in expressing their views, especially on sensitive issues such as corruption.

Keywords: *Corruption, Naive Bayes Classifier, Sentiment, Twitter.*



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Korupsi adalah salah satu masalah besar yang dihadapi di Indonesia. Tingginya angka korupsi dapat merusak integritas pemerintahan, menghambat pertumbuhan ekonomi, menurunkan kepercayaan masyarakat terhadap institusi publik dan menciptakan ketidakadilan dalam distribusi sumber daya (Kajian et al., 2020). Meskipun pemerintah telah berupaya memberantas korupsi, termasuk dengan membentuk Komisi Pemberantas Korupsi (KPK), tantangan besar dalam menghadapi masalah ini masih tetap berlangsung.

Kemajuan pesat dari perkembangan teknologi dan informasi yang terjadi pada saat ini, salah satu penyebabnya adalah adanya pengaruh internet (Mubarok, 2021). Informasi sangat mudah didapatkan dan disebarkan pada orang lain melalui sosial media salah satunya twitter. Twitter merupakan situs microblogging yang penggunanya dapat membuat pesan singkat yang biasa disebut "tweet" (Nugroho, 2018). Twitter dianggap menjadi platform yang populer bagi penggunanya dalam mengekspresikan gagasan dan pandangan mereka secara bebas. Hal ini disebabkan oleh kemudahan akses, jumlah karakter yang singkat yaitu hanya 280 karakter, yang mendorong pengguna untuk menyampaikan maksud dan tujuannya secara singkat, jelas dan padat pada inti permasalahan yang ingin disampaikan (Fitriyyah et al., 2019). Kondisi ini juga didukung dengan banyaknya pengguna telepon seluler di kalangan masyarakat, yang memungkinkan masyarakat untuk terhubung dengan internet dan memanfaatkan platform tersebut dengan lebih mudah (Mas Pintoko & Muslim, 2018). Sehingga dengan adanya twitter sebagai media, dapat memberikan kemudahan bagi para pengguna terutama masyarakat dalam menyuarkan pendapat dan mengkritik isu-isu korupsi.

Analisis sentimen merupakan proses menilai dan mengolompokkan opini, sikap serta evaluasi seseorang terhadap suatu isu tertentu untuk menentukan apakah sentimen yang dihasilkan bersifat positif atau negatif (Harun et al., 2023). Analisis sentimen juga dapat dimanfaatkan untuk menggali pendapat, karena berfokus pada pendapat, sikap, atau emosi yang mencerminkan pandangan seseorang pada suatu kejadian atau kasus tertentu (Natasuwarna, 2019).

Analisis ini dapat memberikan pemahaman tentang persepsi publik terhadap korupsi dan upaya yang dapat dilakukan dalam memberantas korupsi tersebut, dengan mengumpulkan *tweet* yang relevan dengan kata kunci atau tagar terkait korupsi, kemudian mengolahnya untuk mengidentifikasi sentimen positif, negatif, atau netral, analisis ini dapat menggambarkan tentang bagaimana masyarakat merespon berbagai peristiwa dan kebijakan terkait korupsi. Sebelum

melakukan analisis sentimen, langkah pertama yang dilakukan adalah preprocessing data untuk mempersiapkan data agar dapat dianalisis. Tahapan preprocessing ini meliputi pembersihan data (*cleaning*), konversi huruf menjadi kecil (*case folding*), pemecahan teks menjadi kata-kata (*tokenizing*), normalisasi kata (*normalization*), penghapusan kata yang tidak memiliki makna signifikan (*stopword removal*), serta proses mengkonversi kata ke bentuk dasar (*stemming*) (Indriyani et al., 2022).

Untuk mengklasifikasikan sentimen tersebut digunakan metode *Naïve Bayes Classifier* (NBC). Metode tersebut merupakan salah satu metode sederhana, namun memiliki tingkat akurasi tinggi dalam mengklasifikasikan data teks (Hussein, 2018).

Penelitian sebelumnya oleh Aprilia & Isnain (2024) tentang analisis sentimen terhadap media sosial twitter dengan kasus kampanye anti korupsi di Indonesia menggunakan metode Naïve Bayes dan mengklasifikasikan ke dalam dua kategori positif dan negatif. Hasil penelitiannya menunjukkan bahwa sebagian besar *tweet* adalah sentimen positif. Sementara itu, Darwis et al. (2020) tentang analisis sentimen Komisi Pemberantasan Korupsi di Twitter menggunakan metode *Support Vector Machine* (SVM). Dalam penelitian tersebut kinerja KPK diklasifikasikan dalam tiga kategori, yaitu positif, negatif, dan netral, penelitian ini menunjukkan persepsi publik terhadap kinerja KPK adalah sentimen negatif. Penelitian lain oleh Pramudiya et al. (2024) tentang analisis sentimen opini publik terhadap kasus korupsi timah di youtube menggunakan metode *oversampling* dan algoritma *decision tree*.

Berdasarkan penelitian sebelumnya telah melakukan analisis sentiment terhadap isu korupsi, namun umumnya memiliki fokus yang terbatas. Penelitian ini berbeda karena bertujuan menggambarkan persepsi publik secara umum terhadap isu korupsi di Indonesia melalui Twitter, tanpa membatasi peristiwa tertentu. Selain itu penelitian ini mengimplementasikan metode Naïve Bayes dan mengklasifikasikannya ke dalam kategori positif, negatif, dan netral.

Diharapkan ini dapat bermanfaat dalam menganalisis opini masyarakat berdasarkan data *tweet* yang membahas korupsi serta mengidentifikasi sentimen positif, negatif, dan netral yang terkandung di dalamnya dengan menggunakan metode klasifikasi *Naïve Bayes*.

II. METODE PENELITIAN

A. Sumber Data dan Variabel Penelitian

Penelitian ini menggunakan data sekunder yang dihasilkan dengan melakukan proses *crawling* atau pengambilan data dari Twitter, menggunakan *Google Colaboratory*. Data yang dikumpulkan berupa *tweet* tentang korupsi yang diperoleh dari masyarakat Indonesia. Variabel peneliti yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X	Jumlah kemunculan kata ke-i dalam suatu objek (sentimen)
Y	Kelas data menggunakan sentiment positif, negatif dan netral

Dalam penelitian ini, objek merujuk pada topik utama dalam *tweet*, yaitu korupsi. Sementara itu, kelas adalah kategori sentimen dari *tweet* yang dianalisis, yang dibagi menjadi tiga kategori yaitu positif, negatif, dan netral.

B. Tahapan Analisis Data

Tahapan yang dilakukan dalam analisis data adalah sebagai berikut.

1. Tahap awal dimulai dengan proses pengumpulan data (*crawling data*)
2. *Preprocessing Data*, tahapan ini bertujuan untuk membersihkan data dengan tujuan mengatasi masalah pada dataset dan mempertahankan konsistensi data yang digunakan. Berikut merupakan langkah-langkah dalam *preprocessing data*.
 - a. *Cleaning* merupakan tahap di mana karakter atau tanda baca yang tidak digunakan dan dihapus dari ulasan untuk membersihkan data. Contoh karakter yang dihapuskan seperti URL, tag (#), titik (.), koma (,) maupun tanda baca lainnya.
 - b. *Tokenisasi* merupakan proses memecah teks menjadi unit kata yang umum digunakan disebut sebagai token. Pada penelitian ini, kata-kata akan dipisahkan berdasarkan unigram.
 - c. *Stopword*, untuk membuang kata-kata yang tidak relevan serta menghapus imbuhan dari kata, kata hubung, atau kata depan.

- d. *Stemming* digunakan untuk mengubah kata-kata dalam data ulasan menjadi bentuk dasarnya dengan bantuan library *sastrawi*. Cara kerja proses ini adalah dengan menghapus imbuhan atau afiks yang melekat pada kata seperti *me-*, *ber-*, *di-*, *-kan*, dan *-an*.
3. Melakukan pelabelan data
Langkah selanjutnya yaitu adalah melakukan pelabelan. Pada penelitian dilakukan pelabelan otomatis dengan *Python TextBlob* (Azhar et al., 2022). *TextBlob* adalah Pustaka *Python* yang berfungsi dalam memproses data teks. Pustaka ini menyediakan API untuk berbagai tugas pengolahan bahasa alami (NLP), seperti ekstraksi frasa kata benda, analisis sentimen, klasifikasi, penerjemahan, dan lainnya (Parlika et al., 2020). Proses pelabelan ini dilakukan dengan mengidentifikasi teks positif, negatif, dan netral.
4. Melakukan klasifikasi
Dalam melakukan klasifikasi digunakan metode *Naïve Bayes Classifier* (NBC). Metode NBC merupakan metode didasarkan pada Teorema Bayes, dengan menghitung peluang dalam memperkirakan kemungkinan di masa yang akan datang berdasarkan data sebelumnya (Tanggaraeni & Sitokdana, 2022). Nilai peluang NBC ditentukan melalui persamaan (2) berikut.

$$P(V_j|a_i) = \frac{P(a_j|V_j)P(V_j)}{P(a_i)} \quad (2)$$

Keterangan:

$P(V_j|a_i)$ = peluang kelas ke-*j* ketika kata ke-*i* muncul.

$P(a_j|V_j)$ = peluang kata ke-*i* terdapat dalam kelas ke-*j*

$P(V_j)$ = peluang jumlah kejadian kelas ke-*j*,

$P(a_i)$ = peluang jumlah munculnya kata ke-*i*.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dataset yang digunakan berasal dari *tweet* masyarakat pada aplikasi twitter dengan kata kunci “korupsi”. Data tersebut diperoleh dengan menggunakan teknik *crawling* menggunakan *Google Colaboratory*. Data yang digunakan adalah data ulasan dari tanggal 28 April 2024 pukul 14:00 WIB hingga data ulasan tanggal 29 April 2024 pukul 16:00 WIB. Jumlah data yang berhasil diambil sebanyak 1511 data. Setelah data dari twitter diperoleh, langkah selanjutnya adalah mengubahnya ke dalam bentuk *dataframe*, di mana hanya kolom *full text* yang dipilih untuk dianalisis. Kemudian, *dataframe* yang telah disusun tersebut disimpan dalam format *csv* agar dapat diolah pada tahap analisis selanjutnya.

B. Hasil Crawling Data

Berikut adalah hasil dari analisis *crawling* data. Data tersebut terdiri dari 15 kolom dan 1511 baris komentar. Kolom yang digunakan untuk proses analisis berikutnya adalah kolom *full_text*, *username*, dan *created_at*

conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	lang	location	quote_count	reply_count	retweet_count	tweet_url	user_id_str	username
0	1784364512681205476 Mon Apr 29 15:14:37 +0000 2024	0	SALAH LAYU ya one dia ditarik ke atas belum jua...	1784364512681205476	NaN	NaN	in	she/her	0	0	0	https://twitter.com/dotyento/status/1784364512681205476	1188012363758333002	dotyento
1	1784364434291348565 Mon Apr 29 15:14:18 +0000 2024	0	Bla nu korupsi baneas ajeura kar naonnya ? Gena...	1784364434291348565	NaN	NaN	in	NaN	0	0	0	https://twitter.com/snapfno/status/1784364434291348565	168291341425848321	snappfno
2	1784316446319178170 Mon Apr 29 15:13:43 +0000 2024	0	@mazzni_gop Si anjg'ri kig ga dihukum mat...	1784316446319178170	NaN	mazzni_gop	in	bandung west java,Indonesia	0	0	0	https://twitter.com/VIDIGARIBALDI/status/1784316446319178170	59968462	VIDIGARIBALDI1
3	17843177324770050 Mon Apr 29 15:13:27 +0000 2024	0	@ykw002 KORUPSI BESAR ANURRRR KAYANNYA GUE ...	17843177324770050	NaN	ykw002	in	kla, rdyf	0	0	0	https://twitter.com/vuanara/status/17843177324770050	13916875458170785	vuanara
4	17843394887705189 Mon Apr 29 15:13:11 +0000 2024	0	@Anak_Og @okowi @KajaksarRI Da mah korupsi...	17843394887705189	NaN	Anak_Og	in	NaN	0	0	0	https://twitter.com/Veeveskyosh/status/17843394887705189	170541041689004768	Veeveskyosh
...
1506	1784345385596244271 Sun Apr 28 14:04:55 +0000 2024	1	Tim penyidik Komisi Pemberantasan Korupsi (PKP...	1784345385596244271	https://pbs.twimg.com/media/GMqNvN4AAQd.jpg	NaN	in	Sulawesi Tengah, Indonesia	0	0	0	https://twitter.com/Karabawesid/status/1784345385596244271	77882299510779004	Karabawesid
1507	178433028054070320 Sun Apr 28 14:04:17 +0000 2024	1	@Jibay_45back Korupsi gaya baru satisnya ba...	178433028054070320	NaN	Jibay_45back	in	NaN	0	0	0	https://twitter.com/DedenWyo5193/status/178433028054070320	16884392222371872	DedenWyo5193
1508	1784403107429426564 Sun Apr 28 14:01:24 +0000 2024	2	@radhidayat Big boss korupsi da mencapai baka...	1784403107429426564	NaN	radhidayat	in	NaN	0	0	0	https://twitter.com/winter_lost/status/1784403107429426564	10590205437042689	winter_is
1509	1784322784758174008 Sun Apr 28 14:00:09 +0000 2024	6	@sai_idAnti korupsi...sark kelatunya usah di L...	1784322784758174008	NaN	sai_id	in	NaN	0	1	1	https://twitter.com/DSarasam31774/status/1784322784758174008	1743461704074514412	DSarasam31774
1510	178438337787702472 Sun Apr 28 14:00:07 +0000 2024	0	Nitp 1. Tugas pengantar komputer 2. LTS stat...	178438337787702472	NaN	NaN	in	kerak bumi	0	1	0	https://twitter.com/mingricadanza/status/178438337787702472	13026038078795942	mingricadanza

1511 rows x 15 columns

Gambar 1. Hasil Crawling data Twitter

Data pada hasil diatas masih berisikan kalimat-kalimat yang tidak dibutuhkan pada analisis berikutnya, seperti tanda baca, emoji, kalimat duplikat, nama pengguna, dan tautan. Semua itu bisa mempengaruhi hasil dari klasifikasi, sehingga sebelum digunakan dalam proses klasifikasi maka dilakukan proses *preprocessing* terlebih dahulu.

C. Preprocessing Data

Tahapan awal setelah diperoleh data hasil *crawling* adalah *text preprocessing*, tahapan yang digunakan, yaitu *cleaning*, *stopword*, *tokenize*, dan *stemming data*. Adapun hasil *text preprocessing* yang dilakukan dapat dilihat pada Tabel 2.

Tabel 2. Hasil dari *Text Preprocessing*

Proses	Output
Dataset	SALAH LAH! ya oke dia daftar kipk pas belum punya semuanya cuma kalo udah punya semuanya bahkan sampe
<i>Cleaning</i>	SALAH LAH ya oke dia daftar kipk pas belum punya semuanya cuma kalo udah punya semuanya bahkan sampe income nya dua digit ya ngundurin diri dong kan udah mampu masa masih mau minta hak orang kurang mampu korupsi itu anjing
<i>Tokonize</i>	"['salah', 'lah', 'ya', 'oke', 'dia', 'daftar', 'kipk', 'pas', 'belum', 'punya', 'semuanya', 'cuma', 'kalo', 'udah', 'punya', 'semuanya', 'bahkan', 'sampe', 'income', 'nya', 'dua', 'digit', 'ya', 'ngundurin', 'diri', 'dong', 'kan', 'udah', 'mampu', 'masa', 'masih', 'mau', 'minta', 'hak', 'orang', 'kurang', 'mampu', 'korupsi', 'itu', 'anjing']"
<i>Stopword</i>	"['salah', 'ya', 'oke', 'daftar', 'kipk', 'pas', 'kalo', 'udah', 'sampe', 'income', 'nya', 'digit', 'ya', 'ngundurin', 'udah', 'hak', 'orang', 'korupsi', 'anjing']"
<i>Stemming data</i>	salah ya oke daftar kipk pas kalo udah sampe income nya digit ya ngundurin udah hak orang korupsi anjing

Pada saat melakukan *text preprocessing* terdapat beberapa data yang dihapuskan, berupa data duplikat atau ulasannya. Jumlah dataset yang tersisa setelah melewati proses *text preprocessing* adalah sebanyak 1492 data dari data awal yaitu 1511 data. Selanjutnya dilakukan proses pelabelan data untuk mengklasifikasikan kalimat berdasarkan fitur yang terdapat didalamnya.

D. Pelabelan Data

Langkah berikutnya yaitu pemberian label pada data ulasan yang sudah dipersiapkan. Proses penentuan label dilakukan dengan menggunakan fungsi *TextBlob* pada pemrograman. Hasil pelabelan data menggunakan fnsgi *TextBlob*.

	full_text	klasifikasi
0	salah lah oke daftar kipk pas punya semua cuma...	Netral
1	eta nu korupsi bansos ayena ker naonnya genah ...	Netral
2	si anjig klo hukum mati udah anehmiskinkan huk...	Netral
3	korupsi besar anjirrr kayak gue banyak baca b...	Netral
4	mah korupsi genetik kompak se sodara2	Netral
...
1505	tim sidik komisi berantas korupsi kpk sita bua...	Netral
1506	korupsi gaya baru selisih 2 dgn broker amp sim...	Netral
1507	big boss korup ae aku beacukai sarang korupsi ...	Netral
1508	anti korupsi wonk ketua laporin kpk tnggl di t...	Netral
1509	nitip 1 tugas antar komputer 2 uts statistika ...	Netral

1510 rows x 2 columns

Gambar 2. Gambar Pelabelan Data

Hasil proses pelabelan data dapat ditemukan dalam Tabel 3. Total dari pelabelan data tersebut di klasifikasikan menjadi kelas sentimen positif, negatif dan netral.

Tabel 3. Hasil Pelabelan Data

Kelas Sentimen	Jumlah
Positif	68
Negatif	84
Netral	1358

Berdasarkan Tabel 3, dari 1510 sentimen yang ada, diperoleh sentimen positif sebanyak 68 sentimen, sentimen negatif sebanyak 84 sentimen, dan sentimen netral sebanyak 1358 sentimen.

E. Klasifikasi Menggunakan Naïve Bayes

Setelah dilakukan pelabelan data, maka dilakukan klasifikasi menggunakan metode *Naïve Bayes*. Hasil klasifikasi dapat dilihat pada tabel 4 berikut

Tabel 4. Hasil klasifikasi *Naïve Bayes*

<i>full text</i>	<i>Klasifikasi Naïve Bayes</i>
miskin pelaksanaanya butuh manusia 12 malaekat butuh uuyang kuat atau good will presiden keluar perpu hukum mati relatif lebih murah tega hukum anti korupsi akan lebih efektif beri effect jerih	Positif
bayangin rp40000 x 10000 orang bahkan 100000 orang udah bayar failed activation good job masuk ranah korupsi nih undangundangnya lho	Positif
mreka crab mentalilty yang ngajuin kpk padhal layak dapet nama mental korupsi	Negatif
salah satu mental punya malu jabat laku koruptor bisa bicara politik negeri olah mahluk suci mantan pidana korupsi berani tampil dipublikseolah tindak korupsi biasa dan rakyat permisif atas laku	Negatif
kpk skrg cuma singkat kantor polisi kuningan sejak ketua pegang polisi pdhl dulu kpk buat krn tak becus polisi tangan korupsi kpk badan independen luar polisi lha skrg pegang polisi samjugbo sama boong	Netral
kalau kasus gila anggap yang biasa fix bangsa ini sedang sakit kronis korupsi tunggu kapan indonesia terap hukum mati koruptor	Netral
nama negera indonesia jadi jangan heran memang semua sdm rata rata suka korupsi	Netral
hadir kpk komisi berantas korupsi bukti hasil berantas praktik korupsi indonesia	Netral
halah indonesia apa gak korupsi	Netral

Tabel 4. Jumlah Sentimen *Naïve Bayes*

Kelas Sentimen	Jumlah
Positif	11
Negatif	14
Netral	1485

Pada tabel 4, berdasarkan analisis yang telah dilakukan, jumlah sentimen diklasifikasikan menjadi kelas sentimen, yaitu positif, negatif, dan netral mengenai korupsi. Terdapat 11 sentimen yang termasuk kedalam kelas sentimen positif, 14 sentimen yang termasuk kedalam kelas sentimen negatif, dan 1485 sentimen yang masuk kedalam kelas sentimen netral.

IV. KESIMPULAN

Berdasarkan hasil analisis, dapat disimpulkan bahwa analisis sentimen diterapkan menggunakan metode *Naïve Bayes* dapat digunakan untuk melihat sentimen atau pandangan masyarakat Indonesia di Twitter tentang korupsi. Hal ini bertujuan untuk mengetahui jumlah masyarakat yang memiliki sentimen positif, negatif, dan netral terhadap masalah korupsi yang terjadi di Indonesia.

Dan dari penelitian ini, disimpulkan bahwa terdapat 11 sentimen positif, 14 sentimen negatif, dan 1485 sentimen netral. Terlihat bahwa sentimen netral dari masyarakat terhadap korupsi lebih banyak dibandingkan sentimen positif dan negatif. Hal ini menunjukkan bahwa masyarakat cenderung berhati-hati dalam menyampaikan pandangan, terutama pada isu yang sensitif seperti korupsi. Saran untuk penelitian selanjutnya adalah perlunya penggunaan referensi khusus berbahasa Indonesia untuk membantu mengidentifikasi kata-kata yang tidak relevan pada tahap *text preprocessing*. Terutama pada kalimat yang tidak baku yang umum digunakan pada media sosial seperti Twitter. Sehingga hasil sentimen yang diperoleh lebih akurat.

DAFTAR PUSTAKA

- Aprilia, N. W. A. S., & Isnain, A. R. (2024). Analisis Sentimen Terhadap Media Sosial Twitter dengan Kasus Kampanye Anti-Korupsi di Indonesia Menggunakan Naive Bayes. *Jurnal Media Informatika Budidarma*, 8(2), 695. <https://doi.org/10.30865/mib.v8i2.7582>
- Azhar, R., Surahman, A., & Juliane, C. (2022). Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 267–281.
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Eduatic - Scientific Journal of Informatics Education*, 7(1), 1–11. <https://doi.org/10.21107/edutic.v7i1.8779>
- Fitriyiah, S. N. J., Safriadi, N., & Pratama, E. E. (2019). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes. *Jepin*, 5(3), 279–285. <http://dev.twitter.com>.
- Harun, R., Ishak, R., & Panna, S. (2023). Analisis Sentimen Opini Publik Pengguna Twitter Terhadap Kenaikan Harga BBM Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, 2(1), 26–33. <https://doi.org/10.37195/balok.v2i1.414>
- Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University - Engineering Sciences*, 30(4), 330–338. <https://doi.org/10.1016/j.jksues.2016.04.002>
- Indriyani, E. R., Paradise, P., & Wibowo, M. (2022). Perbandingan Metode Naïve Bayes dan Support Vector Machine Untuk Analisis Sentimen Terhadap Vaksin Astrazeneca di Twitter. *Jurnal Media Informatika Budidarma*, 6(3), 1545. <https://doi.org/10.30865/mib.v6i3.4220>

- Kajian, P., Publik, P., & Bumi, A. (2020). *Pengawasan Publik terhadap Pejabat Publik yang Melakukan Tindakan Korupsi : Perspektif Hukum Administrasi*. 4(1), 80–102. <https://doi.org/10.25072/jwy.v4i1.316>
- Mas Pintoko, B., & Muslim, K. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier. *E-Proceeding of Engineering*, 5(3), 8121–8130.
- Mubarok, R. (2021). Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Sosial Berskala Besar (Psb) Dengan Metode *Jurnal Siliwangi Seri Sains Dan Teknologi*, 7(1), 19–24. <http://jurnal.unsil.ac.id/index.php/jssainstek/article/view/3726>
- Natasuwarna, A. P. (2019). Analisis Sentimen Keputusan Pemindahan Ibukota Negara Menggunakan Klasifikasi Naive Bayes. *Sensitif (Seminar Nasional Sistem Informasi Dan Teknik Informatika)*, 47–53.
- Nugroho, A. (2018). Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstraksi Fitur N-Gram. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 2(2), 200. <https://doi.org/10.30645/j-sakti.v2i2.83>
- Parlika, R., Pradika, S. I., Hakim, A. M., & N M, K. R. (2020). Analisis Sentimen Twitter Terhadap Bitcoin Dan Cryptocurrency Berbasis Python Textblob. *Jurnal Ilmiah Teknologi Informasi Dan Robotika*, 2(2), 33–37. <https://doi.org/10.33005/jifti.v2i2.22>
- Pramudiya, R., Kadafi, A., & Udjulawa, D. (2024). Analisis Sentimen Opini Publik terhadap Kasus Korupsi Timah di Youtube Menggunakan Metode Oversampling dan Algoritma Decision Tree. *Arcitech: Journal of Computer Science and Artificial Intelligence*, 4(1), 1. <https://doi.org/10.29240/arcitech.v4i1.10472>
- Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naive Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 785–795. <https://doi.org/10.35957/jatisi.v9i2.1835>