

ANALISIS SENTIMEN TERHADAP WACANA KEBIJAKAN SUBSIDI KRL BERBASIS NIK MENGUNAKAN ALGORITMA NAIVE BAYES

Jeffrey Alexander ¹⁾ Tri Sutrisno ²⁾ Irvan Lewenusa ³⁾

^{1) 2) 3)} Sistem Informasi Universitas Tarumanagara

Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia

¹⁾ email : jeffrey.825210120@stu.untar.ac.id, ²⁾ email : tris@fti.untar.ac.id, ³⁾ email : irvanl@fti.untar.ac.id

ABSTRACT

Kereta Rel Listrik (KRL) is a public transportation that uses rail-based electric power. Although KRL is the main choice of many people because of its efficiency, it has various challenges to face, especially regarding the discourse on the NIK-based KRL subsidy policy. This can be seen from various comments from KRL users on social media that contain comments on the policy discourse. Based on the Financial Memorandum Book document of the Draft State Budget (APBN) 2025, the general KRL fee subsidy will be changed based on the population's Family Identification Number (NIK). As a result, the KRL fee will change from Rp3,000 for the first 25km to approximately Rp25,000 for the first 25km. Complaints and grievances of people today tend to prefer to share posts (uploads) in the form of stories of their daily life through social media. One of the social media that is often used to argue is X (Twitter). One topic that is often discussed is public transportation, usually shared through social media posts or quotes as a form of service user satisfaction. To improve the quality of KRL services, it is necessary to know the level of satisfaction of KRL public transportation users in Indonesia. Sentiment analysis is used in this research to determine the sentiment of KRL service user comments in Indonesia. The Software Development Life Cycle (SDLC) method used is the Agile method. Public sentiment is collected using X social media, and sentiment analysis stages will be carried out such as crawling, labeling, preprocessing, naive bayes classification, accuracy testing, and visualization using Power BI software. The accuracy test results obtained using confusion matrix with 80% training data division and 20% test data obtained 76% accuracy results.

Key words

Sentiment Analysis, KRL, Naive Bayes, Visualization

1. Pendahuluan

Indonesia telah mengalami perkembangan yang sangat cepat di berbagai sektor, termasuk infrastruktur dan teknologi [1]. Perkembangan ini tidak hanya

membawa dampak positif seperti peningkatan kualitas hidup, tetapi juga menimbulkan masalah baru, terutama di kota-kota besar. Dengan populasi yang terus meningkat, kebutuhan dan penggunaan transportasi juga ikut meningkat.

Transportasi umum menjadi salah satu solusi dalam mengatasi masalah mobilitas di perkotaan [2]. Transportasi umum meliputi berbagai jenis, seperti angkutan kota (angkot), bus, taksi, Light Rail Transit (LRT), Mass Rapid Transit (MRT), hingga Kereta Rel Listrik (KRL) [3]. Berbagai jenis transportasi umum memiliki kelebihan dan kekurangan masing-masing, namun semuanya bertujuan untuk memberikan akses mobilitas yang terjangkau dan efisien bagi masyarakat. Transportasi umum berperan penting dalam mengurangi penggunaan kendaraan pribadi yang berlebihan, yang sering kali menjadi penyebab utama kemacetan lalu lintas [4]. Selain itu, transportasi umum juga dianggap lebih ramah lingkungan karena dapat mengangkut banyak masyarakat dalam sekali perjalanan, sehingga emisi karbon per kapita menjadi berkurang [5].

Banyaknya masyarakat yang menggunakan transportasi umum disebabkan oleh berbagai faktor. Salah satu faktor utama adalah biaya yang lebih terjangkau dibandingkan dengan menggunakan kendaraan pribadi. Selain itu, kemudahan akses dan kenyamanan juga menjadi alasan penting mengapa transportasi umum banyak diminati, terutama di kota-kota besar [6]. Transportasi umum juga sering kali menawarkan rute yang mencakup area yang luas, sehingga memudahkan masyarakat untuk mencapai tujuan mereka tanpa perlu berganti kendaraan. Akibatnya, pemerintah harus berinovasi dan meningkatkan kualitas pada transportasi umum untuk memenuhi kebutuhan masyarakat.

Transportasi umum yang sering digunakan di Indonesia adalah Kereta Rel Listrik (KRL). KRL adalah transportasi umum yang menggunakan tenaga listrik berbasis rel [7]. Meskipun KRL menjadi pilihan utama banyak masyarakat karena efisiensinya, KRL memiliki berbagai tantangan yang harus dihadapi, terutama terkait wacana kebijakan subsidi KRL berbasis NIK. Hal ini dapat dilihat dari berbagai komentar pengguna KRL di media sosial yang berisi komentar terhadap wacana

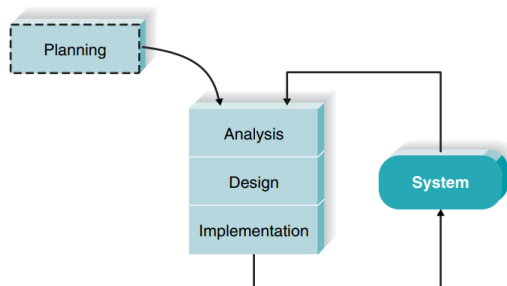
kebijakan tersebut. Berdasarkan dokumen Buku Nota Keuangan Rancangan Anggaran Pendapatan dan Belanja Negara (APBN) 2025, subsidi biaya KRL yang selama ini bersifat umum akan diubah berdasarkan Nomor Induk Keluarga (NIK) penduduk [8]. Hal ini mengakibatkan biaya KRL berubah dari Rp3.000 untuk 25km pertama dapat berubah menjadi kurang lebih Rp25.000 untuk 25km pertama [9]. Hal ini memicu pro dan kontra dari masyarakat. Keluh dan kesah masyarakat pada zaman sekarang cenderung lebih suka membagikan Post (unggahan) berupa cerita perjalanan kehidupannya sehari-hari melalui media sosial. Media sosial merupakan media komunikasi berbasis internet yang memungkinkan penggunaannya mengekspresikan diri dan berinteraksi, berkolaborasi, berbagi, berkomunikasi dengan pengguna lain, dan menjalin hubungan sosial secara virtual [10].

Salah satu media sosial yang sering digunakan untuk berpendapat adalah X (Twitter). Komentar biasanya dibagikan melalui Post atau Quote media sosial sebagai wujud kepuasan pengguna layanan. Pada tanggal 28 Agustus 2024, akun X @detikcom mengunggah Post yang menyertakan link berita yang membahas perubahan kebijakan subsidi KRL berbasis NIK. Melalui Post tersebut masyarakat turut memberikan pendapat mereka terkait wacana tersebut melalui akun X mereka masing-masing.

Maka dari itu, untuk menentukan pilihan yang tepat dalam meresmikan kebijakan ini, perlu mengetahui pendapat mayoritas masyarakat terkait wacana kebijakan tersebut. Menentukan pendapat mayoritas dapat dilakukan dengan analisis sentimen. Analisis sentimen adalah proses untuk mengidentifikasi isi dari dataset yang berupa teks dokumen, kalimat, ataupun paragraf yang memiliki sentimen positif, negatif, atau netral [11]. Analisis sentimen digunakan pada penelitian ini bertujuan untuk mengetahui sentimen dari komentar pengguna layanan KRL di Indonesia.

2. Metode Penelitian

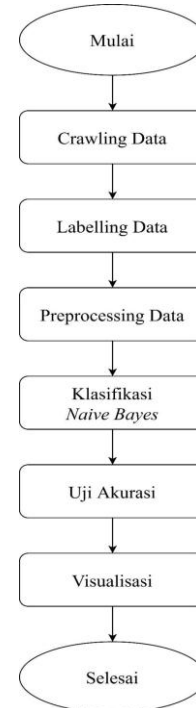
Metode Software Development Life Cycle (SDLC) yang digunakan dalam penelitian ini adalah metode Agile yang dapat dilihat pada **Gambar 1**. Metode Agile merupakan salah satu metode dalam SDLC berbasis iteratif dan inkremental yang memiliki fleksibilitas dalam perubahan pada persyaratan ataupun setelah perencanaan awal selesai [12].



Gambar 1 Metode SDLC Agile [13].

2.1 Alur Penelitian

Perancangan alur penelitian bertujuan untuk memberikan gambaran umum proses analisis dan perancangan yang dilakukan dalam penelitian yang dapat dilihat pada **Gambar 2**.



Gambar 2 Alur Penelitian.
(Sumber: Dokumentasi Pribadi)

2.2 Analisis Sentimen

Analisis sentimen adalah gabungan dari data mining dan *text mining*, yang digunakan untuk mengolah berbagai opini yang diberikan oleh pengguna melalui berbagai media tentang produk, jasa, atau instansi [14]. Analisis sentimen merupakan metode yang digunakan untuk memahami, mengekstrak, dan mengolah opini secara otomatis dari teks untuk mengetahui sentimen yang terdapat pada suatu opini [15]. Analisis Sentimen berkaitan dengan ulasan pengguna, informasi yang diberikan mencakup dasar pemikiran dalam bentuk kalimat yang bisa bersifat subjektif (seperti opini) atau objektif (seperti fakta) [16]. Ulasan subjektif didasarkan pada pendapat pribadi, perasaan, keyakinan, serta penilaian terhadap sesuatu. Ulasan objektif berdasarkan pada fakta, bukti, dan observasi yang dapat diukur.

Post dan Quote di media sosial sering kali mengungkapkan berbagai emosi seperti kegembiraan, kemarahan, kekecewaan, kesenangan, dan perasaan lainnya. Dalam analisis sentimen, terdapat tiga jenis opini yaitu opini positif, opini negatif, dan opini netral. Melalui analisis sentimen perusahaan terkait sehingga dapat memahami kepuasan masyarakat terhadap suatu layanan atau produk melalui respons dari masyarakat atau para ahli [17]. Analisis sentimen dalam penelitian ini bertujuan untuk memahami persepsi dan opini pengguna terhadap pelayanan KRL di Indonesia. Dengan mengkategorikan

sentimen pengguna menjadi positif, netral, atau negatif, penelitian ini dapat memberikan wawasan mengenai bagaimana pelayanan KRL diterima oleh masyarakat.

2.3 Confusion Matrix

Confusion matrix merupakan metode pengujian untuk mengevaluasi kinerja metode klasifikasi terhadap suatu data klasifikasi [18]. *Confusion matrix* membandingkan klasifikasi data hasil prediksi dengan klasifikasi data aktual dalam bentuk tabel.

Tabel 1 Tabel *Confusion Matrix*.
(Sumber: Dokumentasi Pribadi)

Aktual	Prediksi	
	Positif	Negatif
Positif	TP (True Positive)	FN (False Negative)
Negatif	FP (False Positive)	TN (True Negative)

Pada Tabel 1 evaluasi kinerja *confusion matrix*, terdapat 4 istilah sebagai representasi hasil proses klasifikasi. berikut merupakan penjelasan 4 istilah tersebut:

1. TP (True Positive) merupakan jumlah data klasifikasi positif dan diprediksi sebagai positif oleh metode klasifikasi.
2. TN (True Negative) merupakan jumlah data klasifikasi negatif dan diprediksi sebagai negatif oleh metode klasifikasi.
3. FP (False Positive) merupakan jumlah data klasifikasi negatif tetapi diprediksi sebagai positif oleh metode klasifikasi.
4. FN (False Negative) merupakan jumlah data klasifikasi positif tetapi diprediksi sebagai negatif oleh metode klasifikasi.

2.4 Crawling

Crawling merupakan proses pengumpulan data dari sumber yang tersedia di internet dengan menggunakan program atau script tertentu [19]. Proses ini dilakukan oleh perangkat lunak yang dikenal sebagai *crawler* atau spider [20]. *Crawler* dirancang untuk menjelajahi web dengan menggunakan API (Application Programming Interface) untuk mendapatkan informasi yang diinginkan, Informasi akan disimpan dalam bentuk data yang dapat dianalisis lebih lanjut [21]. Penelitian ini menggunakan metode *crawling* untuk mengumpulkan data berupa Post dan Quote yang berkaitan dengan KRL pada aplikasi X.

2.5 Naive Bayes

Naive Bayes Classifier adalah metode yang digunakan untuk mencari nilai probabilitas tertinggi dalam mengklasifikasikan data uji ke dalam kategori yang paling sesuai, dan menggunakan perhitungan probabilitas dalam metode machine learning [22]. Metode Naive Bayes

merupakan klasifikasi untuk menghitung probabilitas yang berdasarkan frekuensi dan kombinasi nilai dalam dataset yang tersedia [23]. Naive bayes memerlukan data pelatihan yang relatif kecil untuk memperkirakan parameter yang diperlukan dalam proses klasifikasi [24]. Metode ini mengklasifikasikan kelas berdasarkan probabilitas sederhana, setiap atribut diasumsikan saling terpisah satu sama lain [25]. Secara matematis, perhitungan Naive Bayes dapat dilihat pada persamaan (1).

$$P(x|y) = \frac{P(y|x) \cdot P(x)}{P(y)} \quad (1)$$

Pada nilai x merupakan kumpulan atribut yang terdiri dari kata-kata yang mencerminkan tiga klasifikasi pendapat yaitu positif, negatif, dan netral, Sementara itu pada nilai y merupakan klasifikasi pendapat tersebut, yang menentukan suatu atribut atau kata termasuk dalam kategori positif, negatif, atau netral [26].

2.6 Preprocessing

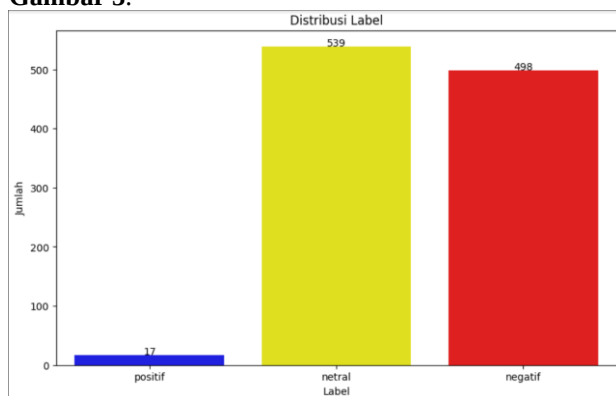
Preprocessing merupakan teknik pemrosesan data yang tidak terstruktur menjadi data yang dapat digunakan untuk analisis atau pemodelan lebih lanjut [27]. Penggunaan *preprocessing* harus dilakukan karena data yang diperoleh melalui *crawling* masih tidak terstruktur, sehingga diperlukan *preprocessing* untuk memastikan data yang digunakan relevan dan berkualitas. Terdapat beberapa tahap untuk melakukan *preprocessing*:

1. Cleaning Data
Membersihkan data dari elemen yang tidak digunakan, seperti menghapus nilai duplikat, tanda baca, simbol, spasi berlebihan, angka, dan karakter khusus yang tidak memiliki makna.
2. Normalization
Mengubah singkatan menjadi kata lengkap dan menyelaraskan bentuk teks dengan mengubah semua huruf menjadi huruf kecil, sehingga data menjadi konsisten tanpa adanya penggunaan huruf kapital yang dapat mempengaruhi hasil analisis.
3. Stopwords Removal
Menghapus kata-kata umum yang tidak memiliki bobot seperti "dan", "di", "atau", dan sejenisnya yang tidak memiliki nilai informatif signifikan. Stopwords biasanya diabaikan dalam analisis teks karena tidak memberikan makna kontekstual yang penting.
4. Tokenize
Memecah teks menjadi unit-unit yang lebih kecil yang disebut tokens, biasanya berupa kata atau frasa. Proses ini memungkinkan teks yang dianalisis menjadi lebih mendetail.
5. Stemming
Proses yang digunakan untuk mengubah kata imbuhan menjadi kata dasar, seperti kata "berlari" menjadi "lari", atau "bermain" menjadi "main". Stemming membantu mengurangi variasi kata yang berasal dari bentuk yang sama, sehingga meningkatkan konsistensi data dan mempermudah analisis.

Stemming sering kali menghasilkan kata-kata yang tidak sempurna, tetapi proses ini cukup efektif dalam mengurangi dimensi data dan menjaga relevansi konten.

3. Hasil dan Pembahasan

Analisis sentimen dilakukan dengan menggunakan metode klasifikasi Naive Bayes terhadap sentimen masyarakat terkait wacana kebijakan subsidi KRL berbasis NIK. Data yang diperoleh melalui tahap *crawling data*, dilakukan *labelling manual* yang dapat dilihat pada **Gambar 3**.



Gambar 3 Distribusi *Labelling Manual*.
(Sumber: Dokumentasi Pribadi)

Tahap yang dilakukan setelah *labelling manual* adalah *preprocessing* yang melalui tahap *cleaning data*, *normalization*, *stopwords removal*, *tokenize*, dan *stemming*. Hasil yang diperoleh setelah dilakukan *preprocessing* dapat dilihat pada **Gambar 4**.

	date	likes	text	label
0	Mon Sep 30 12:24:19 +0000 2024	0	ngerasa krl makin padat praduga salah kata ban...	netral
1	Mon Sep 30 11:31:00 +0000 2024	2	wacana terap subsidi tarif krl dasar nomor ind...	netral
2	Sun Sep 29 15:26:36 +0000 2024	0	subsidi krl sempit dasar nik warga kian jepit	negatif
3	Sun Sep 29 12:50:51 +0000 2024	0	yakin analis trubus subsidi krl menteri bumh m...	netral
4	Sun Sep 29 12:33:14 +0000 2024	25	menteri bumh mau ubah badan amat sebut pso anc...	netral
...	...	...	...	...
1049	Wed Aug 28 10:36:47 +0000 2024	16	apbn darah darah ketemu cara benar cabut bagi ...	positif
1050	Wed Aug 28 10:30:22 +0000 2024	2	bbm gas krl busway segala subsidi model bas ni...	negatif
1051	Wed Aug 28 10:21:44 +0000 2024	0	perintah ubah skema subsidi krl jabodetabek ba...	netral
1052	Wed Aug 28 10:16:00 +0000 2024	2656	perintah ubah skema subsidi krl jabodetabek ba...	netral
1053	Wed Aug 28 07:53:15 +0000 2024	22	salur subsidi tiket krl jabodetabek bas nik mulai	netral

Gambar 4 Data Setelah *Preprocessing*.
(Sumber: Dokumentasi Pribadi)

Klasifikasi Naive Bayes dilakukan dengan membagi data latih sebesar 80% dan data uji sebesar 20%. Pembagian data dapat dilihat pada **Gambar 5**.

```
# Memisahkan data menjadi data latih dan data uji
x = df['text']
y = df['label']
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)

# Mencetak jumlah data latih dan data uji
print("Jumlah Data Latih:", len(x_train))
print("Jumlah Data Uji:", len(x_test))

Jumlah Data Latih: 843
Jumlah Data Uji: 211
```

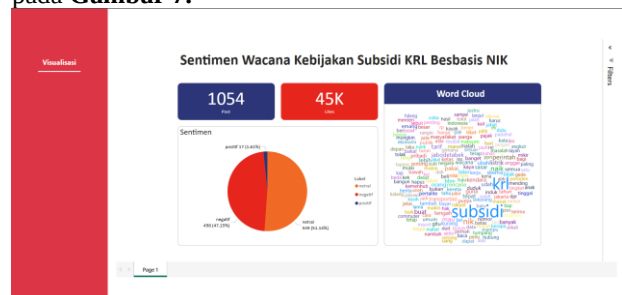
Gambar 5 Pembagian Data Latih dan Data Uji.
(Sumber: Dokumentasi Pribadi)

Hasil uji akurasi yang diperoleh menggunakan *confusion matrix* dengan pembagian data latih 80% dan data uji 20% memperoleh hasil akurasi 76%. Hasil uji akurasi menggunakan *confusion matrix* dapat dilihat pada **Gambar 6**.

Akurasi: 0.7630331753554502				
Uji Akurasi:	precision	recall	f1-score	support
negatif	0.74	0.83	0.78	109
netral	0.80	0.69	0.74	101
positif	0.00	0.00	0.00	1
accuracy			0.76	211
macro avg	0.51	0.51	0.51	211
weighted avg	0.76	0.76	0.76	211
Confusion Matrix:				
[[91 18 0]				
[31 70 0]				
[1 0 0]]				

Gambar 6 Uji Akurasi.
(Sumber: Dokumentasi Pribadi)

Visualisasi sentimen masyarakat terkait wacana kebijakan subsidi KRL berbasis NIK dibuat menggunakan software Power BI dan hasil visualisasi akan ditampilkan dalam bentuk website sederhana dengan bahasa pemrograman PHP. Hasil visualisasi dapat dilihat pada **Gambar 7**.



Gambar 7 Visualisasi.
(Sumber: Dokumentasi Pribadi)

4. Kesimpulan

Setelah dilakukan analisis sentimen terkait wacana kebijakan subsidi KRL berbasis NIK diperoleh kesimpulan:

1. Hasil uji akurasi yang diperoleh menggunakan *confusion matrix* dengan pembagian data latih 80% dan data uji 20% memperoleh hasil akurasi 76%.
2. Visualisasi dibuat menggunakan software Power BI dan hasil visualisasi yang ditampilkan dalam bentuk website sederhana dengan bahasa pemrograman PHP,

bertujuan untuk mempermudah pemahaman terhadap sentimen masyarakat.

5. Saran

Setelah dilakukan analisis sentimen terkait wacana kebijakan subsidi KRL berbasis NIK diperoleh Saran:

1. Membuat website dengan penambahan fitur untuk melakukan analisis sentimen.
2. Menggunakan metode algoritma selain Naive Bayes seperti *K-Means Clustering*.

REFERENSI

- [1] Prof. Dr. Sri Adiningsih, S.E., M.Sc, Transformasi Ekonomi Berbasis Digital di Indonesia: Lahirnya Tren Baru Teknologi, Bisnis, Ekonomi, dan Kebijakan di Indonesia. Gramedia Pustaka Utama, 2019.
- [2] Diah Pramesti, J. Andini, Dyah Aurum Kianthi Raharjo, and Arif Devi Dwipayana, "Efektivitas Penggunaan Moda Transportasi Umum Dengan Kendaraan Pribadi," Indonesian Journal of Multidisciplinary on Social and Technology, vol. 2, no. 1, pp. 6–16, Jan. 2024, doi: <https://doi.org/10.31004/ijmst.v2i1.246>.
- [3] W. Nabila Putri, F. Nuraida Rezeki, Z. Dianita Zahra, L. Anisa Surya, A. Muhammad Rahmat, and S. Yuliani, "Transportasi Publik Meningkatkan Ekonomi Hijau Secara Berkelanjutan di Jakarta," Jurnal Multidisiplin Ilmu Akademik, vol. 1, no. 3, Jun. 2024.
- [4] Hani Subagio, Dwinanta Nugroho, Muhamad Ridwan, Hari Rachmadi, and Ajeng Tri Kadesti, "Mengurai Kemacetan dalam Konteks Tata Ruang Perkotaan Yogyakarta," Jurnal Politik dan Pemerintahan Daerah, vol. 5, no. 2, pp. 23–31, Dec. 2023, doi: <https://doi.org/10.36355/jppd.v5i2.124>.
- [5] M. Benny, Investasi dan Reformasi Transportasi Kota. Syiah Kuala University Press, 2024.
- [6] A. Rosa and A. Widad, "Faktor-Faktor yang Membedakan Keputusan Konsumen Menggunakan Jasa Transportasi Dalam Kota (Studi Kasus Angkutan Umum Konvensional dan Berbasis Online)," Jurnal Manajemen dan Bisnis Sriwijaya, vol. 15, no. 3, pp. 164–172, 2017.
- [7] A. Ayu, None Aulia Rofika Dewi, D. Panca, None Raisya Kanahaya Achmad, and I. Viska, "Analisis Pengaruh Manajemen Mutu Terhadap Kinerja, Efektivitas, dan Keandalan pada Transportasi Kereta Rel Listrik (KRL)," Jurnal Riset Manajemen dan Ekonomi (JRIME), vol. 2, no. 3, pp. 97–107, Jun. 2024, doi: <https://doi.org/10.54066/jrime-itb.v2i3.1996>.
- [8] A. Hikam, "Subsidi KRL Jadi Berbasis NIK Mulai 2025, Kemenhub: Agar Tepat Sasaran," detikfinance, Aug. 28, 2024. https://finance.detik.com/berita-ekonomi-bisnis/d-7512817/subsidi-krl-jadi-berbasis-nik-mulai-2025-kemenhub-agar-tepat-sasaran?utm_campaign=detikcomsocmed&utm_medium=oa&utm_source=twitter&utm_content=detikfinance.
- [9] I. G. Oswaldo, "Bukan Rp 3.000, Ternyata Segini Tarif Asli Naik KRL Jika Tak Disubsidi," detikfinance, Sep. 13, 2024. <https://finance.detik.com/berita-ekonomi-bisnis/d-7539523/aturan-subsidi-mau-diubah-ternyata-segini-tarif-asli-naik-krl>.
- [10] T. Liedfray, F. J. Waani, and J. J. Lasut, "Peran Media Sosial Dalam Mempererat Interaksi Antar Keluarga Di Desa Esandom Kecamatan Tombatu Timur Kabupaten Minahasa Tenggara," Jurnal Ilmiah Society, vol. 2, no. 1, Jan. 2022.
- [11] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," Expert Systems with Applications, vol. 40, no. 10, pp. 4065–4074, Aug. 2013, doi: <https://doi.org/10.1016/j.eswa.2013.01.001>.
- [12] A. A. F. Amarta and I. G. Anugrah, "Implementasi Agile Scrum Dengan Menggunakan Trello Sebagai Manajemen Proyek Di PT Andromedia," Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI), vol. 4, no. 6, pp. 528–534, Dec. 2021, doi: <https://doi.org/10.32672/jnkti.v4i6.3702>.
- [13] A. Dennis, Barbara Haley Wixom, and R. M. Roth, Systems analysis and design. Hoboken, NJ: Wiley, 2015.
- [14] Evasaria Magdalena Sipayung, Herastia Maharani, and I. Zefanya, "Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier," JSI: Jurnal Sistem Informasi (E-Journal), vol. 8, no. 1, 2016, doi: <https://doi.org/10.18495/jsi.v8i1.3250>.
- [15] D. Ikasari and Widiastuti Widiastuti, "Sentiment Analysis Review Novel 'Goodreads' Berbahasa Indonesia Menggunakan Naive Bayes Classifier," Semnas Ristek (Seminar Nasional Riset dan Inovasi Teknologi), vol. 5, no. 1, 2021, doi: <https://doi.org/10.30998/semnasristek.v5i1.5040>.
- [16] A. R. Alaci, S. Becken, and B. Stantic, "Sentiment Analysis in Tourism: Capitalizing on Big Data," Journal of Travel Research, vol. 58, no. 2, pp. 175–191, Dec. 2019, doi: <https://doi.org/10.1177/0047287517747753>.
- [17] A. Halim and None Andri Safuwani, "Analisis Sentimen Opini Warganet Twitter Terhadap Tes Screening Genose Pendeteksi Virus Covid-19 Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization," Jurnal Informatika Teknologi dan Sains (Jinteks), vol. 5, no. 1, pp. 170–178, Feb. 2023, doi: <https://doi.org/10.51401/jinteks.v5i1.2229>.
- [18] G. P. Wiratama and A. Rusli, "Sentiment Analysis of Application User Feedback in Bahasa Indonesia Using Multinomial Naive Bayes," 2019 5th International Conference on New Media Studies (CONMEDIA), Oct. 2019, doi: <https://doi.org/10.1109/conmedia46929.2019.8981850>.
- [19] P. Y. Saputra, D. H. Subhi, and F. Z. A. Winatama, "Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah di Youtube Menggunakan Algoritma Naive Bayes," Jurnal Informatika Polinema, vol. 5, no. 4, pp. 209–213, Aug. 2019, doi: <https://doi.org/10.33795/jip.v5i4.259>.
- [20] S. Sarwosri, A. H. Basori, and W. B. Surastyo, "Aplikasi Web Crawler Untuk Web Content Pada Mobile Phone," JUTI: Jurnal Ilmiah Teknologi Informasi, vol. 7, no. 3, p. 127, Jan. 2009, doi: <https://doi.org/10.12962/j24068535.v7i3.a79>.
- [21] Iksan Ramadhan and Husni Sastramihardja, "Pemanfaatan Web Crawler Dalam Mengumpulkan Informasi Melalui Internet," Konferensi Nasional Sistem Informasi (KNSI) 2018.
- [22] Nurman Satya Marga, "Sentimen Analisis Tentang Kebijakan Pemerintah Terhadap Kasus Corona Menggunakan Metode Naive Bayes," Jurnal Informatika dan Rekayasa Perangkat Lunak, vol. 2, no. 4, pp. 453–463, Feb. 2022, doi: <https://doi.org/10.33365/jatika.v2i4.1602>.

- [23] H. Zhang and D. Li, "Naïve Bayes Text Classifier," *IEEE Xplore*, Nov. 01, 2007, doi:<https://doi.org/10.1109/GrC.2007.40>.
- [24] H. Apriyani and K. Kurniati, "Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus," *Journal of Information Technology Ampera*, vol. 1, no. 3, pp. 133–143, Dec. 2020, doi:<https://doi.org/10.51519/journalita.volume1.issue3.year2020.page133-143>.
- [25] J. Azhar and Widya Syaharani, "Prakiraan Cuaca Dengan Menggunakan Metode Naïve Bayes Classifier," *Jurnal Media Teknik Elektro dan Komputer*, vol. 1, no. 1, pp. 11–18, 2024.
- [26] D. Jurafsky and J. H. Martin, *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 2023.
- [27] Farkhan Nuruz Zaman, Mohamad Agung Fadhillah, Masy Ari Ulinuha, and Khothibul Umam, "Menganalisis Respons Netizen Twitter Terhadap Program Makan Siang Gratis Menerapkan NLP Metode Naïve Bayes," *Just IT : Jurnal Sistem Informasi, Teknologi Informasi dan Komputer*, vol. 14, no. 3, pp. 201–208, 2024, doi:<https://doi.org/10.24853/justit.14.3.201-208>.