

Sentiment Analysis of Instagram Comments on the BPS Province X Account Using the Naive Bayes Algorithm Based on Machine Learning

Jessika^{1*}, Al Khaidar², Nurdin³, Syarifah Muliana⁴

^{1,2,3,4} Program Studi Magister Teknologi Informasi, Fakultas Teknik, Universitas Malikussaleh, Lhokseumawe, 24355, Indonesia

Informasi Artikel

Diterima : 1 September 2025
Revisi : 25 September 2025
Publikasi : 30 September 2025

Kata Kunci:

Analisis Sentimen
Naïve Bayes
Instagram
Machine Learning

ABSTRAK

Analisis sentimen merupakan salah satu pendekatan dalam pemrosesan bahasa alami yang bertujuan untuk mengidentifikasi dan mengkategorikan opini atau sikap pengguna terhadap suatu entitas berdasarkan data teks. Data yang digunakan terdiri atas 500 caption unggahan terakhir yang diperoleh melalui alat bantu Phantombuster. Tahapan analisis meliputi proses crawling data, preprocessing (penghapusan data duplikat dan kosong, tokenization, stopword removal, dan case folding), pemodelan menggunakan algoritma Naïve Bayes, serta visualisasi hasil klasifikasi. Berdasarkan hasil pengolahan, ditemukan bahwa sebagian besar data tergolong dalam kategori netral (97,65%), sementara sisanya terbagi dalam kategori positif (1,57%) dan negatif (0,78%), dengan akurasi model sebesar 94%. Meskipun akurasi model tergolong tinggi, dominasi kelas netral menunjukkan adanya ketidakseimbangan distribusi data (imbalanced data) yang dapat memengaruhi kualitas generalisasi model.

ABSTRACT

Sentiment analysis is an approach in natural language processing that aims to identify and categorize user opinions or attitudes towards an entity based on text data. The data used consists of the last 500 uploaded captions obtained through the Phantombuster tool. The analysis stages include data crawling, preprocessing (removal of duplicate and empty data, tokenization, stopword removal, and case folding), printing using the Naïve Bayes algorithm, and visualization of the classification results. Based on the processing results, it was found that the majority of the data was classified as neutral (97.65%), while the rest was divided into positive (1.57%) and negative (0.78%) categories, with a model accuracy of 94%. Although the model accuracy is relatively high, the dominance of the neutral class indicates an imbalance in data distribution (imbalanced data) which can affect the quality of the generalization model.

This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



*Penulis Koresponden

Email: jesjessikaa@gmail.com

Cara sitasi IEEE:

Jessika, A. Khaidar, Nurdin, & S. Muliana, "Sentiment Analysis of Instagram Comments on the BPS Province X Account Using the Naive Bayes Algorithm Based on Machine Learning," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 3, pp. 1231-1237, September 2025, doi: 10.30811/jaise.v5i3.7815

1. PENDAHULUAN

Perkembangan pesat teknologi informasi dan komunikasi (TIK) dalam dua dekade terakhir telah mengubah secara fundamental cara masyarakat dalam mengakses dan menyebarkan informasi [1]-[2]. Media sosial sebagai salah satu inovasi TIK yang paling dominan menjadi medium utama interaksi sosial di era digital saat ini [3]-[4]. Platform-platform seperti Instagram tidak hanya menjadi ruang berbagi konten visual, tetapi juga menjadi sumber data berharga yang dapat dimanfaatkan untuk berbagai analisis, salah satunya adalah analisis sentimen. Analisis sentimen merupakan cabang dari pengolahan bahasa alami (Natural Language Processing - NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini, perasaan, atau sikap pengguna terhadap suatu objek atau topik tertentu berdasarkan teks yang mereka hasilkan [5]-[6].

Instagram dipilih dalam konteks penelitian ini karena popularitasnya yang terus meningkat di kalangan masyarakat lintas usia dan latar belakang sosial, menjadikannya sumber data real-time yang kaya dan dinamis [7]. Di Indonesia sendiri, penggunaan Instagram oleh berbagai lembaga, termasuk instansi pemerintah, semakin meningkat sebagai saluran komunikasi yang efektif dalam menyampaikan informasi kepada masyarakat luas [8]. Badan Pusat Statistik (BPS), sebagai lembaga pemerintah yang bertanggung jawab menyediakan data dan informasi statistik resmi, juga memanfaatkan Instagram untuk menyebarkan hasil survei, data demografi, dan informasi statistik penting lainnya secara cepat dan interaktif [9].

Keberhasilan penyebaran informasi bukan hanya soal kuantitas unggahan, melainkan bagaimana informasi tersebut diterima dan dipersepsikan oleh masyarakat [10]-[11]. Oleh karena itu, analisis sentimen terhadap konten yang dipublikasikan oleh akun resmi BPS Provinsi X sangat penting dilakukan untuk mengukur opini publik dan mengidentifikasi sentimen yang dominan terhadap informasi yang disajikan. Dengan demikian, hasil analisis dapat memberikan umpan balik yang berharga bagi BPS dalam memperbaiki strategi komunikasi dan penyajian data agar lebih diterima dan dipahami oleh masyarakat.

Teknik analisis sentimen telah banyak digunakan dalam berbagai bidang untuk mengekstrak opini dari teks di media sosial, ulasan produk, dan forum diskusi daring. Pendekatan ini mengelompokkan data teks ke dalam kategori sentimen seperti positif, negatif, dan netral berdasarkan pola bahasa yang digunakan [12]-[13]. Penelitian ini caption dari 500 unggahan terakhir akun Instagram BPS Provinsi X menjadi fokus utama untuk dianalisis, karena caption merupakan teks yang mengiringi setiap unggahan dan berisi pesan utama yang ingin disampaikan oleh akun tersebut.

Metode Naïve Bayes dipilih sebagai algoritma utama dalam proses klasifikasi sentimen karena beberapa alasan. Pertama, algoritma ini memiliki prinsip probabilistik yang sederhana namun efektif, mampu bekerja baik pada dataset berukuran besar dengan dimensi tinggi seperti data teks [14]. Kedua, Naïve Bayes relatif mudah diimplementasikan dan memiliki performa yang cukup baik dalam klasifikasi teks, terutama dalam tugas-tugas seperti spam filtering dan analisis sentimen [15]. Meskipun demikian, tantangan utama yang sering muncul dalam penerapan Naïve Bayes adalah ketidakseimbangan kelas data, di mana distribusi kategori sentimen tidak merata sehingga dapat memengaruhi akurasi prediksi [16]-[17].

Penelitian ini akan melewati beberapa tahapan utama, yaitu pengambilan data (crawling) menggunakan alat bantu Phantombuster, proses pra-proses data (preprocessing) seperti penghilangan data duplikat dan kosong, tokenisasi, penghapusan kata-kata umum (stopword removal), serta normalisasi teks (case folding). Setelah itu, dilakukan pemodelan dengan algoritma Naïve Bayes untuk mengklasifikasikan sentimen caption ke dalam kategori positif, negatif, dan netral. Tahap akhir berupa visualisasi hasil untuk mempermudah interpretasi data serta menarik kesimpulan terkait pola sentimen masyarakat terhadap konten BPS Provinsi X.

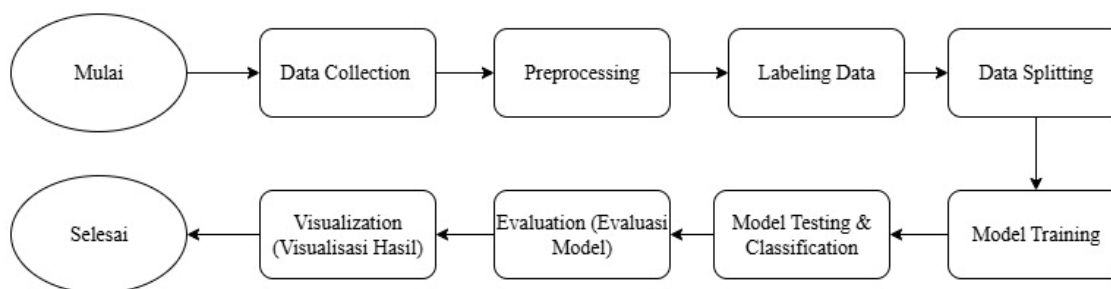
Meskipun analisis sentimen dengan Naïve Bayes telah banyak diterapkan, termasuk pada platform Instagram, penelitian ini menonjolkan aspek novelty melalui fokus pada konten resmi lembaga pemerintah dalam hal ini BPS Provinsi X dan pengolahan caption yang bersifat informatif serta terkait data statistik publik. Pendekatan ini diharapkan dapat memberikan gambaran yang lebih spesifik tentang persepsi masyarakat terhadap penyampaian informasi statistik, sekaligus menjadi bahan pertimbangan dalam meningkatkan kualitas komunikasi data publik melalui media sosial, khususnya dalam konteks penyebaran informasi statistik yang akurat, relevan, dan mudah dipahami.

2. METODE

2.1 Perancangan Analisis Sentimen

Perancangan analisis sentimen pada penelitian ini dilakukan untuk memberikan gambaran yang jelas dan sistematis mengenai tahapan yang dilalui dalam mengolah data komentar Instagram pada akun BPS Provinsi X. Perancangan ini bertujuan agar proses penelitian dapat tersusun secara terstruktur mulai dari pengumpulan data, pengolahan awal data, hingga pada tahap klasifikasi dan evaluasi hasil. Dengan adanya perancangan ini, alur penerapan algoritma Naïve Bayes berbasis machine learning dapat dipahami secara lebih

komprehensif, baik dari sisi teknis maupun metodologis, sehingga penelitian tidak hanya menghasilkan output berupa klasifikasi. Adapun perancangan analisis sentimen dapat dilihat pada Gambar 1 berikut.



Gambar 1. Perancangan Analisis Sentimen

Gambar 1 menggambarkan alur perancangan analisis sentimen komentar Instagram pada akun BPS Provinsi X menggunakan algoritma Naïve Bayes berbasis machine learning. Proses dimulai dari tahap data collection untuk memperoleh data komentar, kemudian dilanjutkan dengan preprocessing agar data bersih dan siap diolah melalui tahapan case folding, cleansing, tokenizing, stopword removal, dan stemming. Selanjutnya dilakukan labeling data untuk memberikan kategori sentimen positif, negatif, atau netral, kemudian data dibagi menjadi data latih dan data uji melalui tahap data splitting. Data latih digunakan dalam proses model training, sedangkan data uji digunakan pada tahap model testing & classification untuk mengukur kemampuan model dalam melakukan klasifikasi. Hasil klasifikasi selanjutnya dievaluasi melalui tahap evaluation dengan menggunakan metrik kinerja seperti akurasi, precision, recall, dan F1-score. Tahap akhir berupa visualization bertujuan untuk menyajikan distribusi sentimen dalam bentuk grafik agar hasil klasifikasi lebih mudah dipahami, kemudian proses dinyatakan selesai.

2.1 Analisis Sentimen

Analisis sentimen merupakan proses sistematis untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini subjektif dari teks guna memahami sikap atau penilaian seseorang terhadap suatu [18]. Pendekatan ini banyak dimanfaatkan dalam pemasaran, politik, dan sektor publik untuk mengevaluasi persepsi masyarakat. Dalam konteks media sosial, seperti Instagram, Twitter, dan Facebook, analisis sentimen menjadi relevan karena platform ini memungkinkan pengguna mengekspresikan opini secara terbuka dan real-time, sehingga dapat memberikan wawasan berharga bagi lembaga pemerintahan, termasuk Badan Pusat Statistik (BPS), dalam memahami respon publik terhadap konten dan layanan yang disampaikan.

2.2 Media Sosial Instagram

Instagram merupakan salah satu media sosial berbasis visual yang banyak digunakan di Indonesia, dengan pengguna aktif yang terus meningkat setiap tahunnya [19]. Fitur caption pada setiap unggahan memungkinkan pengguna untuk menyampaikan informasi atau opini secara tekstual, sehingga dapat dimanfaatkan sebagai sumber data dalam penelitian analisis sentimen. Data yang diambil dari Instagram bersifat unstructured dan memiliki karakteristik khusus, seperti singkat, padat, dan seringkali menggunakan bahasa informal atau slang. Oleh karena itu, diperlukan teknik pemrosesan bahasa alami (NLP) untuk mengubah data mentah menjadi informasi yang dapat dianalisis secara komputasional.

2.3 Metode Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilistik yang banyak digunakan dalam analisis teks karena kesederhanaannya dan kemampuannya dalam menangani data berdimensi tinggi [20]. Algoritma ini menggunakan Teorema Bayes untuk menghitung probabilitas suatu teks masuk dalam kategori tertentu berdasarkan fitur-fitur (kata-kata) yang terkandung dalam teks tersebut. Dalam konteks analisis sentimen, Naïve Bayes mampu melakukan klasifikasi terhadap teks yang telah dilabeli ke dalam kategori sentimen, seperti positif, negatif, dan netral. Beberapa penelitian sebelumnya menunjukkan bahwa algoritma ini efektif dan efisien dalam menganalisis data teks pendek seperti tweet dan [21].

3. HASIL DAN PEMBAHASAN

3.1 Klasifikasi Sentimen

Setelah dilakukan tahapan crawling, preprocessing, pelabelan data, dan klasifikasi menggunakan algoritma Naïve Bayes, diperoleh hasil distribusi sentimen dari 500 data caption Instagram akun resmi BPS Provinsi X. Proses splitting data dilakukan dengan pembagian 80% data latih (training data) dan 20% data uji (testing data). Model menghasilkan tingkat akurasi sebesar 94%, yang menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan data dengan tingkat ketepatan yang tinggi. Tabel 1 berikut menunjukkan hasil distribusi sentimen:

Tabel 1. Hasil Distribusi Sentimen

Kategori	Jumlah Data	Persentase (%)
Netral	488	97.65%
Positif	8	1.57%
Negatif	4	0.78%
Total	500	100%

Berdasarkan tabel tersebut, diketahui bahwa mayoritas berada dalam kategori netral, yaitu sebanyak 97,65%, sedangkan kategori positif dan negatif hanya menyumbang masing-masing sebesar 1,57% dan 0,78%. Hal ini menunjukkan bahwa konten yang diunggah oleh BPS Provinsi X cenderung bersifat informatif dan netral, tanpa memuat opini atau ekspresi emosional yang kuat.

3.2 Confusion Matrix

Mengevaluasi kinerja model klasifikasi sentimen, dilakukan pengujian menggunakan Confusion Matrix. Confusion Matrix merupakan alat evaluasi yang menampilkan perbandingan antara label sebenarnya (actual) dengan prediksi model (predicted), sehingga dapat diketahui sejauh mana model mampu mengklasifikasikan setiap kategori sentimen secara akurat. Adapun hasil pengujian Analisis Sentimen Komentar Instagram pada Akun BPS Provinsi X menggunakan algoritma Naive Bayes berbasis machine learning dapat dilihat pada Gambar 2 berikut.

Classification Report:

	precision	recall	f1-score	support
Netral	1.00	0.94	0.97	488
Positif	0.23	0.75	0.35	8
Negatif	0.33	1.00	0.50	4
accuracy			0.94	500
macro avg	0.52	0.90	0.61	500
weighted avg	0.98	0.94	0.95	500

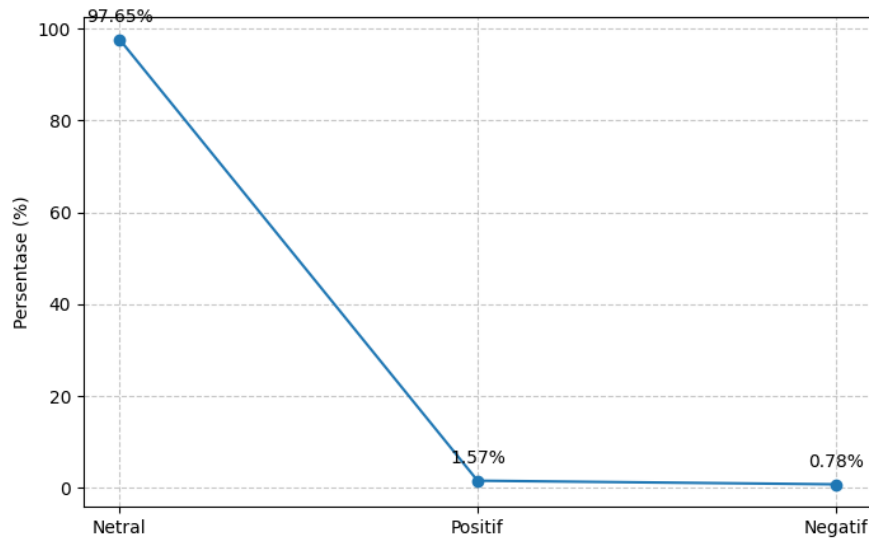
Gambar 2. Classification Report

Gambar 2 menunjukkan hasil evaluasi model klasifikasi dengan tiga kelas, yaitu Netral, Positif, dan Negatif. Model memiliki performa sangat baik pada kelas Netral dengan precision 1.00, recall 0.94, dan f1-score 0.97 karena jumlah data yang dominan. Sebaliknya, performa pada kelas Positif dan Negatif masih rendah, terlihat dari nilai precision yang kecil (0.23 dan 0.33) serta f1-score masing-masing hanya 0.35 dan 0.50, meskipun nilai recall cukup tinggi. Secara keseluruhan, akurasi model mencapai 94%, namun nilai macro average f1-score sebesar 0.61 mengindikasikan adanya ketidakseimbangan performa antar kelas akibat distribusi data yang tidak seimbang.

3.3 Pengujian ROC Curve Multi Class

Pengujian ROC (Receiver Operating Characteristic) Curve dilakukan untuk mengevaluasi kinerja model klasifikasi sentimen secara lebih mendalam, khususnya dalam membedakan masing-masing kelas. Karena data sentimen terdiri dari tiga kategori, yaitu Netral, Positif, dan Negatif, ROC Curve dibuat menggunakan pendekatan One-vs-Rest (OvR) agar dapat menampilkan kemampuan model dalam memprediksi setiap kelas secara terpisah. Dengan pengujian ini, peneliti dapat mengidentifikasi seberapa baik

dibandingkan hanya menampilkan data numerik. Melalui pendekatan visual ini, peneliti dapat menunjukkan ketimpangan jumlah data antar kategori sentimen sehingga pola persepsi publik terhadap akun BPS Provinsi X dapat teridentifikasi secara lebih intuitif. Adapun visualisasi sentimen dapat dilihat pada Gambar 2.



Gambar 2. Visualisasi Distribusi Sentimen

Gambar 1 menunjukkan distribusi hasil klasifikasi sentimen komentar Instagram pada akun BPS Provinsi X. Berdasarkan visualisasi tersebut, dapat dilihat bahwa mayoritas komentar termasuk ke dalam kategori netral dengan proporsi sebesar 97,65%, sedangkan komentar dengan sentimen positif hanya mencapai 1,57%, dan sentimen negatif berada pada angka 0,78%. Kondisi ini memperlihatkan adanya ketidakseimbangan (imbalanced data) yang cukup signifikan antar kategori sentimen. Ketimpangan distribusi ini menjadi tantangan dalam proses klasifikasi karena model cenderung memberikan prediksi bias terhadap kelas mayoritas, yakni sentimen netral. Meskipun demikian, temuan ini juga memberikan gambaran bahwa interaksi pengguna pada akun BPS Provinsi X cenderung bersifat informatif atau tidak mengandung muatan opini yang kuat, baik dalam bentuk positif maupun negatif.

3.3 Interpretasi dan Evaluasi Model

Nilai akurasi yang diperoleh sebesar 94% pada hasil pengujian model memang secara sekilas menunjukkan performa yang tinggi dalam melakukan klasifikasi. Namun, apabila dianalisis lebih mendalam, capaian akurasi tersebut tidak serta merta mencerminkan kualitas klasifikasi yang sebenarnya. Hal ini disebabkan oleh dominasi kategori sentimen netral yang mencapai 97,65%, sehingga terdapat kecenderungan model untuk secara default memberikan prediksi pada kelas mayoritas. Kondisi ini mengakibatkan nilai akurasi terlihat tinggi karena sebagian besar data uji memang berasal dari kelas netral, padahal kemampuan model dalam membedakan kelas lain (positif dan negatif) relatif rendah.

Hasil evaluasi terhadap data uji juga memperlihatkan adanya indikasi overfitting terhadap kelas mayoritas. Model mampu mengklasifikasikan data pada kategori netral dengan sangat baik, namun menunjukkan kelemahan dalam mengidentifikasi data pada kategori positif maupun negatif. Gejala ini menandakan bahwa model lebih terlatih untuk mengenali pola dominan yang terdapat pada data netral, tetapi gagal menggeneralisasi dengan baik terhadap data minoritas. Dengan demikian, meskipun nilai akurasi tampak tinggi, kualitas klasifikasi tidak merata antar kelas, sehingga diperlukan evaluasi dengan metrik lain seperti precision, recall, dan F1-score untuk memberikan gambaran yang lebih komprehensif mengenai performa model dalam menangani distribusi data yang tidak seimbang (imbalanced).

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Naïve Bayes berhasil diterapkan dalam klasifikasi sentimen pada 500 caption Instagram akun resmi BPS Provinsi X dengan capaian akurasi sebesar 94%, yang secara umum menunjukkan performa model berada pada kategori tinggi. Proses pelatihan dilakukan dengan pembagian data sebesar 80% untuk data latih dan 20% untuk data uji, sehingga model memiliki dasar yang cukup dalam mempelajari pola distribusi data. Hasil klasifikasi

memperlihatkan dominasi yang sangat signifikan pada kategori sentimen netral sebesar 97,65%, sedangkan sentimen positif dan negatif hanya mencapai masing-masing 1,57% dan 0,78%.

Ketimpangan distribusi ini mengindikasikan adanya masalah imbalanced data yang berpotensi menimbulkan bias pada model. Akurasi yang tinggi tidak sepenuhnya mencerminkan kinerja klasifikasi yang adil terhadap semua kategori sentimen, karena model cenderung lebih akurat dalam mengidentifikasi kelas mayoritas, yakni sentimen netral, sementara kemampuan dalam mengenali sentimen positif dan negatif relatif rendah. Oleh karena itu, evaluasi lebih lanjut menggunakan metrik tambahan seperti precision, recall, dan F1-score sangat diperlukan untuk memberikan gambaran performa yang lebih komprehensif, sekaligus memastikan bahwa model dapat bekerja secara seimbang dalam mengklasifikasikan seluruh kategori sentimen.

REFERENSI

- [1] A. Khaidar, M. Arhami, And M. Abdi, "Application Of The Random Forest Method For Ukt Classification At Politeknik Negeri Lhokseumawe," *J. Artif. Intell. Softw. Eng.*, Vol. 4, No. 2, Pp. 94–103, 2024.
- [2] S. Kurnia And A. Khaidar, "Perbandingan Metode Machine Learning Menggunakan Metode Support Vector Machine Dan Artificial Neural Network Dalam Memprediksi Serangan Jantung," *J. Inform. Kaputama (Jik)*, Vol. 9, No. 2, Pp. 87–94, 2025.
- [3] A. S. Das And S. A. G. G. Singh, "A Survey Of Sentiment Analysis In Social Media: Techniques And Challenges," *J. Data Intell.*, Vol. 3, No. 2, Pp. 211–235, 2022.
- [4] M. A. Hasan, M. A. Al-Ghamdi, And M. A. Hossain, "Sentiment Analysis Of Social Media Data Using Machine Learning: A Comprehensive Review," *Ieee Access*, Vol. 9, Pp. 88754–88772, 2021.
- [5] R. Sharma, S. Sharma, And R. Kumar, "Performance Analysis Of Naïve Bayes And Svm In Sentiment Classification Of Twitter Data," *Int. J. Inf. Technol.*, Vol. 14, No. 1, Pp. 245–253, 2022.
- [6] P. Singh, N. Kumar, And A. Singh, "Text Preprocessing Techniques For Sentiment Analysis: A Comparative Study," *J. King Saud Univ. - Comput. Inf. Sci.*, Vol. 34, No. 8, Pp. 5867–5878, 2022.
- [7] H. Kaur, S. Singh, And M. Kumar, "A Deep Learning-Based Model For Sentiment Analysis Of Instagram Captions," *Multimed. Tools Appl.*, Vol. 81, No. 1, Pp. 1123–1142, 2022.
- [8] D. Supriyanto, N. Nuryanti, And A. Noegroho, "Media Sosial Dalam Efektivitas Komunikasi Pelayanan Publik Instansi Pemerintah Daerah," *Innov. J. Soc. Sci. Res.*, Vol. 4, No. 4, Pp. 7796–7810, 2024.
- [9] Y. E. Putri, F. M. Elita, I. Gemiharto, And F. I. Komunikasi, "Pengaruh Media Sosial Instagram@ Bps_Statistics Terhadap Ekuitas Merek Badan Pusat Statistik," *Ekspresi Dan Persepsi J. Ilmu Komun.*, Vol. 6, No. 1, Pp. 17–31, 2023.
- [10] A. N. Ahmed, M. F. A. Fauzi, And M. A. A. M. Ali, "Sentiment Analysis Visualization Tools For Social Media: A Review," *J. Phys. Conf. Ser.*, Vol. 2163, No. 1, P. 12024, 2022.
- [11] T. H. Nguyen, H. T. Nguyen, And D. T. Le, "A Hybrid Approach For Sentiment Analysis On Imbalanced Social Media Data," *Expert Syst. Appl.*, Vol. 194, P. 116524, 2022.
- [12] S. Al-Sallab, A. H. A. Al-Zoubi, And A. Al-Soud, "Challenges Of Arabic Sentiment Analysis: A Systematic Literature Review," *Acm Trans. Asian Low-Resource Lang. Inf. Process.*, Vol. 21, No. 4, Pp. 1–28, 2022.
- [13] S. B. K. Kumar, S. S. Kumar, And R. Kumar, "Ethical Considerations In Social Media Data Crawling For Academic Research," *J. Information, Commun. Ethics Soc.*, Vol. 20, No. 3, Pp. 411–427, 2022.
- [14] Z. Li, Q. Wang, And Y. Liu, "Improving Naïve Bayes For Text Classification By Feature Weighting," *J. Ambient Intell. Humaniz. Comput.*, Vol. 13, No. 5, Pp. 2651–2662, 2022.
- [15] F. Saberi, M. A. Cheraghlikhani, And A. H. Ghorbani, "A Novel Framework For Sentiment Analysis On Instagram Comments Using Deep Learning," *J. Big Data*, Vol. 8, No. 1, P. 112, 2021.
- [16] L. Deng, Y. Wang, And J. Li, "Public Opinion Analysis On Government Social Media During Public Health Crises: A Case Study Of Covid-19," *Inf. Process. Manag.*, Vol. 58, No. 5, P. 102652, 2021.
- [17] J. A. P. Jiménez, M. A. R. Díaz, And D. G. García, "Sentiment Analysis Of Public Services User Comments On Social Networks Using Nlp Techniques," *Ieee Lat. Am. Trans.*, Vol. 20, No. 7, Pp. 1655–1663, 2022.
- [18] A. Zakira, M. Arhami, M. I. Abdi, And S. Safriadi, "Text-Based Emotion Sentiment Analysis On Social Media Using Nlp And Lexicon Approach (Case Study: Gaza Conflict)," *J. Informatics Eng. Softw. Appl.*, Vol. 1, No. 1, Pp. 139–149, 2025.
- [19] R. R. Armyani, L. C. Tambunan, R. M. Siregar, N. R. Lubis, And A. Azahra, "Analisis Peran Media Sosial Instagram Dalam Meningkatkan Penjualan Online," *J. Pendidik. Tambusai*, Vol. 5, No. 3, Pp. 8920–8928, 2021.
- [20] L. L. Larasati, "Analisis Sentimen Terhadap Opini Warganet Tentang Wisata Di Yogyakarta Pada Platform Instagram Menggunakan Naïve Bayes Classifier (Studi Kasus: Data Opini Warganet Di Instagram Selama Satu Tahun)," 2024.
- [21] S. D. Parameswari, M. Lubis, S. Suakanto, Y. Z. Ramadhan, R. N. Amanah, And R. A. Dila, "Studi Perbandingan Naïve Bayes Dan Support Vector Machine (Svm) Dalam Analisis Sentimen Pengguna Metaverse," *J. Teknol. Dan Manaj. Ind. Terap.*, Vol. 4, No. 3, Pp. 1059–1065, 2025.