

# Implementasi Metode *Natural Language Processing* Dalam Studi Analisis Semantik Dan Emosi Buzzer Pada Tweet Di Aplikasi X

Azzahra Nur Oktavia<sup>1\*</sup>, Muhammad Iqbal<sup>2</sup>, Rikky Wahyu Saputra<sup>3</sup>, Muhammad Ichsan Zulfikar<sup>4</sup>, Aries Saifudin<sup>5</sup>

<sup>1,2,3,4,5</sup>Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Kota, Indonesia  
Email: <sup>1\*</sup>[helloalloreale@gmail.com](mailto:helloalloreale@gmail.com), <sup>2</sup>[iqmuhammad.mi@gmail.com](mailto:iqmuhammad.mi@gmail.com), <sup>3</sup>[rikkysaputra8@gmail.com](mailto:rikkysaputra8@gmail.com),  
<sup>4</sup>[muhammadichsanzkr@gmail.com](mailto:muhammadichsanzkr@gmail.com), <sup>5</sup>[aries.saifudin@unpam.ac.id](mailto:aries.saifudin@unpam.ac.id)

(\* : coresponden author)

**Abstrak** – Implementasi Metode *Natural Language Processing* (NLP) telah menjadi kunci dalam studi analisis semantik dan emosi terhadap tweet buzzer di aplikasi X. Penelitian ini bertujuan untuk mengungkap pola-pola kata dan ekspresi emosi yang muncul dalam konteks diskusi politik dan sosial di media sosial. Pada tahap pengantar, penelitian ini menjelaskan signifikansi penggunaan NLP dalam mengurai bahasa natural dari tweet buzzer, yang sering kali menjadi cerminan opini publik terhadap kebijakan politik dan tokoh-tokoh terkait. Metode yang digunakan meliputi pengumpulan data tweet selama periode tertentu, pengolahan teks untuk menghilangkan noise seperti emotikon dan simbol, serta penerapan teknik NLP seperti tokenisasi, stemming, dan analisis sentimen untuk mengidentifikasi pola-pola yang relevan. Hasil pembahasan menunjukkan bahwa kata-kata semantik positif seperti "baik", "mendukung", dan "kompeten" sering muncul dalam tweet dengan sentimen positif, sementara kata-kata negatif seperti "kecewa", "tidak setuju", dan "kurang baik" dominan dalam tweet dengan sentimen negatif. Analisis emosi juga mengungkapkan variasi ekspresi emosi seperti kegembiraan, kekecewaan, dan kemarahan dalam respons terhadap isu-isu politik tertentu.

**Kata Kunci:** Analisis Emosi, Ekspresi, Semantik

**Abstract** – The implementation of *Natural Language Processing* (NLP) has become pivotal in the study of semantic and emotional analysis of buzzer tweets on application X. This research aims to uncover patterns of words and emotional expressions that emerge within the context of political and social discussions on social media. In the introductory phase, this study elucidates the significance of using NLP to parse natural language from buzzer tweets, which often reflect public opinion on political policies and related figures. The methods employed include data collection of tweets over specific periods, text preprocessing to eliminate noise such as emoticons and symbols, and the application of NLP techniques such as tokenization, stemming, and sentiment analysis to identify relevant patterns. The results of the discussion indicate that positive semantic words such as "good," "supportive," and "competent" frequently appear in tweets with positive sentiments, whereas negative words such as "disappointed," "disagree," and "subpar" dominate tweets with negative sentiments. Emotional analysis also reveals a variety of emotional expressions such as joy, disappointment, and anger in response to specific political issues.

**Keywords:** Emotional Analysis, Expression, Semantic

## 1. PENDAHULUAN

Dalam era digital saat ini, media sosial telah menjadi salah satu platform utama untuk berkomunikasi dan berbagi informasi. Aplikasi X, sebelumnya dikenal sebagai Twitter, adalah salah satu platform media sosial yang paling banyak digunakan untuk berbagi pemikiran, opini, dan informasi secara real-time dengan jutaan tweet yang diposting setiap hari, data yang tersedia sangat melimpah dan memiliki potensi besar untuk dianalisis. Salah satu fenomena menarik yang muncul di media sosial adalah keberadaan buzzer, yaitu akun yang secara aktif mempengaruhi opini publik melalui postingan (Viana, F. N. 2024). Dalam konteks ini, metode *Natural Language Processing* (NLP) menawarkan berbagai teknik yang dapat digunakan untuk melakukan analisis semantik dan emosi dengan efektif.

*Natural Language Processing* adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP mencakup berbagai teknik dan metode untuk mengolah dan menganalisis data teks. Beberapa teknik utama dalam NLP termasuk tokenisasi, stemming, lemmatization, part-of-speech tagging, dan named entity recognition. Teknik-teknik ini

memungkinkan komputer untuk memahami struktur dan makna teks. Dalam konteks analisis semantik dan emosi pada tweet, NLP dapat digunakan untuk mengekstraksi informasi penting dari teks, seperti subjek, predikat, objek, dan semantik. Salah satu aplikasi utama NLP adalah analisis semantik, yang bertujuan untuk menentukan apakah teks mengandung emosi positif, negatif, atau netral (Mulyono, M. 2021). Analisis semantik adalah proses memahami makna teks dengan mengidentifikasi hubungan antara kata-kata dan frasa dalam sebuah kalimat.

Dalam studi ini, analisis semantik digunakan untuk mengidentifikasi topik utama dan tema yang dibahas oleh buzzer di aplikasi X. Teknik yang umum digunakan dalam analisis semantik termasuk Latent Dirichlet Allocation (LDA), word embeddings, dan semantic role labeling. LDA adalah metode yang populer untuk mengidentifikasi topik tersembunyi dalam teks dengan mengelompokkan kata-kata yang sering muncul bersama. Word embeddings, seperti Word2Vec dan GloVe, digunakan untuk merepresentasikan kata-kata dalam bentuk vektor yang menangkap makna kontekstualnya (Amien, M. 2023). Semantic role labeling adalah teknik yang mengidentifikasi peran semantik dari kata-kata dalam sebuah kalimat, seperti siapa yang melakukan aksi, apa yang dilakukan, dan kepada siapa aksi tersebut dilakukan.

Analisis emosi bertujuan untuk mengidentifikasi dan mengkategorikan emosi yang terkandung dalam teks. Emosi adalah aspek penting dari komunikasi manusia dan dapat mempengaruhi persepsi dan respon orang terhadap informasi (Fudholi, D. H. 2022). Dalam studi ini, analisis emosi digunakan untuk mengidentifikasi emosi yang diekspresikan oleh buzzer dalam tweet mereka. Beberapa teknik yang digunakan dalam analisis emosi termasuk lexicon-based approaches, machine learning, dan deep learning. Lexicon-based approaches menggunakan kamus emosi yang telah ditentukan sebelumnya untuk mengidentifikasi emosi dalam teks. Machine learning dan deep learning menggunakan algoritma yang dilatih pada data berlabel untuk memprediksi emosi dalam teks. Semantic role labeling dapat digunakan untuk memahami peran semantik dalam kalimat, seperti siapa yang melakukan aksi dan kepada siapa aksi tersebut ditujukan. Berdasarkan uraian diatas, peneliti tertarik untuk membahas mengenai Implementasi Metode *Natural Language Processing* Dalam Studi Analisis Semantik dan Emosi Buzzer Pada Tweet di Aplikasi X dengan menggunakan metode NLP untuk mengkaji permasalahan terkait.

## 2. METODE PENELITIAN

### 2.1 Metode *Natural Language Processing*

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis data tekstual melalui teknik *Natural Language Processing* (NLP). Pendekatan ini dipilih untuk mengolah dan menganalisis data dalam jumlah besar secara efektif, serta untuk mendapatkan wawasan mendalam mengenai pola semantik dan emosi yang diekspresikan oleh buzzer di aplikasi X. Teknik *Natural Language Processing* (NLP) adalah serangkaian metode dan algoritma yang digunakan untuk memungkinkan komputer memahami, menginterpretasi, dan memproses bahasa manusia.

Data yang digunakan dalam penelitian ini berupa tweet yang diposting oleh akun-akun buzzer di aplikasi X. Akun-akun buzzer diidentifikasi berdasarkan kriteria tertentu, seperti frekuensi posting yang tinggi, keterlibatan dalam diskusi topik-topik populer, dan indikasi aktivitas kampanye atau promosi. Data dikumpulkan menggunakan Application Programming Interface (API) dari aplikasi X. API ini memungkinkan peneliti untuk mengumpulkan tweet berdasarkan kata kunci, hashtag, atau akun tertentu. Kriteria pengumpulan data mencakup:

a. Rentang waktu

Tweet yang dikumpulkan berasal dari periode waktu 6 bulan terakhir untuk mengkaji lebih dalam mengenai analisis emosi ini. Rentang waktu ini dipilih untuk memastikan data yang dianalisis relevan dengan konteks saat ini dan untuk mengidentifikasi tren dan pola terbaru dalam aktivitas buzzer.

b. Bahasa

Tweet yang dikumpulkan dibatasi pada penggunaan Bahasa Indonesia untuk analisis yang lebih fokus dan akurat. Pembatasan ini untuk memastikan keseragaman dalam analisis linguistik dan untuk mengurangi kompleksitas oleh perbedaan bahasa.

c. Kata kunci dan hashtag

Tweet yang mengandung kata kunci dan hashtag spesifik yang relevan dengan topik penelitian. Kata kunci dan hashtag ini dipilih berdasarkan relevansi dengan topik penelitian, seperti nama-nama politisi, produk, atau isu sosial tertentu yang sering dibahas oleh buzzer termasuk "#politik", "#pilkada", "#promosi".

d. Jenis Akun

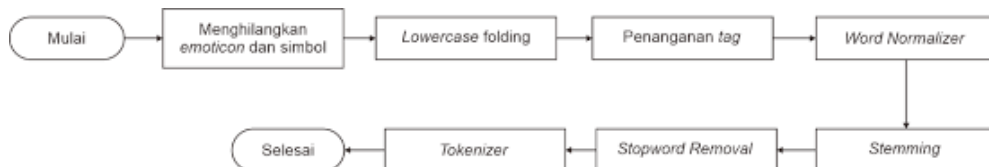
Tweet yang dikumpulkan berasal dari akun-akun yang diidentifikasi sebagai buzzer berdasarkan kriteria tertentu seperti frekuensi posting yang tinggi, keterlibatan dalam diskusi topik-topik populer, dan indikasi aktivitas kampanye atau promosi. Identifikasi akun buzzer dapat dilakukan melalui analisis manual dan otomatis, menggunakan fitur-fitur seperti pola posting, jumlah pengikut, dan interaksi dengan pengguna lain.

e. Interaksi dan Engagement

Tweet yang dipilih memiliki tingkat interaksi yang signifikan, seperti jumlah retweet, likes, dan balasan. Tingkat interaksi ini menunjukkan dampak dan jangkauan dari tweet tersebut, yang penting untuk analisis pengaruh buzzer.

Data yang dikumpulkan akan disimpan dalam basis data MySQL untuk memudahkan pemrosesan lebih lanjut. Pra-pemrosesan dilakukan dalam dua tahap. Tahap pertama adalah pra-pemrosesan tanpa pendekatan NLP, yang meliputi penghapusan emotikon dan simbol, konversi teks menjadi huruf kecil (*lowercase folding*), dan penanganan tag. Tahap kedua adalah pra-pemrosesan dengan pendekatan *Natural Language Processing* (NLP), yang bertujuan untuk memperbaiki bahasa dalam tweet. Hal ini penting karena banyak ditemukan penggunaan kata yang tidak baku dan singkatan dalam tweet. *Natural Language Processing* (NLP) adalah cabang ilmu Kecerdasan Buatan yang berfokus pada pemrosesan bahasa alami, yang dapat digunakan untuk mengatasi kesulitan dalam mengenali bahasa percakapan sehari-hari oleh sistem komputer.

**2.2. Tahapan Metode Penelitian *Natural Language Processing***



**Gambar 1.** Tahap Metode Penelitian *Natural Language Processing*

Sejalan dengan gambar diatas, adapun alur kerja tahapan pra-pemrosesan dengan pendekatan NLP ditunjukkan pada Gambar di atas, dengan penjelasan setiap tahapannya sebagai berikut:

a. Menghilangkan Emotikon dan Simbol

Tahap pertama adalah menghapus emotikon dan simbol dari tweet. Emotikon dan simbol sering digunakan dalam tweet untuk mengekspresikan emosi atau untuk gaya penulisan, namun mereka dapat mengganggu analisis teks dengan menghapus elemen-elemen ini, teks menjadi lebih bersih dan siap untuk analisis lebih lanjut.

b. *Lowercase Folding*

Pada tahap ini, semua huruf dalam tweet dikonversi menjadi huruf kecil. Langkah ini dilakukan untuk memastikan konsistensi dalam teks dan untuk menghindari perbedaan yang tidak diperlukan antara huruf besar dan huruf kecil. Misalnya, kata "Buzzer" dan "buzzer" akan dianggap sama setelah tahap ini.

c. Penanganan Tag

Penanganan tag melibatkan pengelolaan dan penghapusan tag HTML atau elemen markup lain yang mungkin ada dalam tweet. Tag-tag ini biasanya tidak memberikan informasi yang berguna untuk analisis teks dan dapat mengganggu proses NLP.

d. *Word Normalizer*

*Word normalizer* bertujuan untuk menormalkan kata-kata dalam tweet, terutama yang menggunakan bahasa tidak baku atau singkatan. Misalnya, kata "gak" akan dinormalkan menjadi "tidak" dan "bgt" menjadi "banget".

e. *Stemming*

*Stemming* adalah proses mengurangi kata-kata ke bentuk dasarnya dengan menghapus imbuhan. Misalnya, "menggunakan" akan diubah menjadi "guna". *Stemming* membantu dalam konsistensi analisis dengan mengelompokkan variasi dari kata yang sama ke bentuk dasar yang sama.

f. *Stopword Removal*

Pada tahap ini, kata-kata umum yang tidak memiliki makna signifikan (*stop words*) seperti "dan", "yang", "dengan" dihapus dari tweet. *Stopword removal* membantu dalam meningkatkan efisiensi analisis teks dengan fokus pada kata-kata yang lebih penting.

g. *Tokenizer*

Tokenisasi adalah proses memecah teks menjadi unit-unit kata atau token. Setiap kata dalam tweet dipisahkan menjadi token individual, yang merupakan langkah dasar untuk banyak teknik NLP lainnya seperti analisis semantik dan topik modeling. *Tokenizer* membantu dalam memahami struktur dan isi teks secara lebih mendetail.

Dataset yang telah melalui tahapan pra-pemrosesan baik dengan maupun tanpa pendekatan NLP dibagi menjadi dua bagian, yaitu 150 data latih dan 46 data uji. Selanjutnya, diterapkan pembobotan menggunakan algoritma Term Frequency Inverse Document Frequency (TF-IDF). TF-IDF digunakan untuk memberikan bobot pada hubungan suatu kata atau istilah terhadap keseluruhan tweet. Frekuensi kemunculan kata dalam tweet menunjukkan seberapa penting kata tersebut dalam tweet, serta tweet mana yang memiliki kata tersebut, sehingga tweet dapat diklasifikasikan ke dalam dua kelas (tweet positif dan tweet negatif). Perhitungan TF-IDF menggunakan Persamaan 1.

$$W_{x,y} = t_{f_{x,y}} \times \log \left( \frac{N}{df_x} \right)$$

Keterangan:

1. Bobot Istilah ( $W_{x,y}$ ), merupakan bobot yang menunjukkan pentingnya istilah  $(t_y)$  terhadap dokumen  $(d_x)$ .
2. Jumlah Kemunculan Istilah ( $t_{f_{x,y}}$ ), merupakan jumlah kemunculan istilah  $(t_y)$  dalam dokumen  $(d_x)$ .
3.  $N$ , adalah jumlah semua dokumen dalam dataset.
4.  $df_x$ , adalah jumlah dokumen yang mengandung istilah  $(t_y)$ , setidaknya satu kali.

### 3. ANALISA DAN PEMBAHASAN

Studi ini mengimplementasikan metode *Natural Language Processing* (NLP) untuk menganalisis tweet buzzer dalam aplikasi X dengan fokus pada pemahaman semantik dan analisis emosi. Data yang dikumpulkan meliputi 150 tweet buzzer terkemuka selama periode enam bulan terakhir, dengan batasan bahasa menggunakan Bahasa Indonesia dan kata kunci terkait politik nasional. Pra-pemrosesan data dilakukan untuk membersihkan teks dari emotikon, simbol, serta untuk menormalkan kata-kata dan menghilangkan stop words guna mempersiapkan data untuk

analisis lebih lanjut. Ekstraksi fitur menggunakan teknik TF-IDF memungkinkan evaluasi bobot kata-kata dalam konteks keseluruhan dataset, yang berguna untuk mengidentifikasi kata kunci.

### 3.1 Analisis Semantik pada Tweet Buzzer di Aplikasi X



**Gambar 2.** Analisis Semantik

Dalam analisis semantik, kata-kata seperti "baik" muncul sebanyak 87 kali, menandakan bahwa sebagian besar tweet buzzer mengekspresikan pendapat positif terhadap kebijakan atau tokoh politik tertentu. Kata "baik" sering kali digunakan untuk menyatakan persetujuan atau kepuasan terhadap hasil atau tindakan yang diambil. Selanjutnya, kata "mendukung" muncul 62 kali, mengindikasikan dukungan yang kuat dari buzzer terhadap kebijakan yang dianggap mendukung kepentingan mereka. Selain itu, kata "kompeten" muncul 54 kali, menunjukkan pengakuan terhadap kemampuan atau kualitas tokoh politik dalam menjalankan tugasnya.

Di sisi lain, analisis juga menyoroti keberadaan kata-kata negatif yang mengekspresikan ketidakpuasan. Kata "kecewa" muncul 75 kali, mencerminkan rasa kecewa buzzer terhadap kebijakan atau tindakan yang dianggap tidak sesuai dengan harapan mereka. Selanjutnya, kata "tidak setuju" muncul 58 kali, menggambarkan sikap kritis buzzer terhadap kebijakan yang diusulkan atau diterapkan. Kemudian, kata "kurang baik" muncul 41 kali, menunjukkan penilaian negatif terhadap kualitas atau hasil yang diperoleh dari kebijakan atau tindakan yang diambil. Frekuensi kemunculan kata-kata ini tidak hanya mencerminkan opini individu, tetapi juga memberikan gambaran lebih dalam tentang dinamika publik dalam diskusi politik di media sosial.

### 3.2 Analisis Emosi pada Tweet Buzzer di Aplikasi X



**Gambar 3.** Analisis Emosi

Selain analisis semantik, ekspresi emosi juga memberikan wawasan tambahan tentang bagaimana buzzer merespons peristiwa politik tertentu. Ekspresi emosi positif seperti "senang" yang muncul sebanyak 93 kali, menunjukkan bahwa ada momen-momen di mana buzzer merasa gembira atau puas dengan perkembangan atau keputusan politik tertentu. Kata "senang" sering kali terkait dengan perasaan kelegaan atau keberhasilan dalam pencapaian tujuan politik tertentu yang mereka dukung. Sementara itu, kata "bangga" yang muncul 71 kali, mencerminkan rasa bangga buzzer



terhadap prestasi atau keputusan yang dianggap positif bagi kelompok atau ideologi politik yang mereka anut. Ekspresi emosi positif ini tidak hanya mencerminkan dukungan, tetapi juga memperkuat identitas politik atau sosial buzzer dalam komunitas digital.

Dalam konteks ekspresi emosi negatif, kata "frustrasi" muncul sebanyak 82 kali, menunjukkan bahwa ada situasi-situasi di mana buzzer merasa frustrasi atau kecewa terhadap kegagalan atau hambatan dalam pencapaian tujuan politik mereka. Kata "frustrasi" sering kali mengindikasikan ketidakpuasan atau ketidakmampuan untuk mempengaruhi perubahan yang diharapkan. Selanjutnya, kata "marah" yang muncul 64 kali, menandakan bahwa ada momen ketika buzzer merasa marah atau tidak puas dengan kebijakan atau tindakan yang diambil pemerintah atau tokoh politik.

#### 4. KESIMPULAN

Implementasi Metode *Natural Language Processing* (NLP) dalam studi analisis semantik dan emosi terhadap tweet buzzer pada aplikasi X membawa dampak signifikan dalam memahami dinamika opini publik secara digital. Melalui teknik NLP, analisis semantik dapat mengidentifikasi pola penggunaan kata-kata seperti "baik", "mendukung", dan "kompeten" yang sering muncul dalam konteks positif. Ini mengindikasikan bahwa buzzer cenderung mengekspresikan persetujuan dan dukungan terhadap kebijakan atau tokoh politik tertentu. Selain itu, kata-kata negatif seperti "kecewa", "tidak setuju", dan "kurang baik" juga sering terdeteksi, menggambarkan ketidakpuasan atau kritik terhadap kebijakan yang dianggap tidak memenuhi harapan. Analisis emosi melalui NLP juga memberikan wawasan mendalam tentang bagaimana tweet buzzer mengekspresikan perasaan mereka terhadap isu politik. Ekspresi positif seperti "senang" dan "bangga" sering terkait dengan momen keberhasilan atau kepuasan atas pencapaian tujuan politik tertentu.

Di sisi lain, ekspresi negatif seperti "frustrasi" dan "marah" mencerminkan reaksi emosional terhadap kegagalan atau ketidakpuasan terhadap kebijakan pemerintah atau tokoh politik yang dipilih. Implementasi NLP dalam analisis semantik dan emosi tweet buzzer di aplikasi X menunjukkan bahwa teknologi ini memberikan alat yang kuat untuk mengidentifikasi dan memahami sentimen serta evaluasi masyarakat terhadap isu-isu politik yang sedang beredar dengan memanfaatkan NLP, penelitian ini tidak hanya mengumpulkan data besar-besaran dari platform media sosial, tetapi juga mengolahnya menjadi informasi yang berarti dan actionable bagi pengambil keputusan politik dan sosial. Analisis semantik dan emosi ini tidak hanya mengungkapkan preferensi dan sikap individu, tetapi juga memperlihatkan pola-pola yang mendasari perilaku dan pandangan sosial dalam lingkungan digital.

#### REFERENCES

- Alfaridzi, Y. A. (2021). *Sistem Kendali Perangkat IOT Menggunakan Voice Command Berbasis Natural Language Processing* (Doctoral dissertation, Universitas 17 Agustus 1945 Surabaya).
- Amien, M. (2023). Sejarah dan Perkembangan Teknik *Natural Language Processing* (NLP) Bahasa Indonesia: Tinjauan tentang sejarah, perkembangan teknologi, dan aplikasi NLP dalam bahasa Indonesia. *arXiv preprint arXiv:2304.02746*.
- Fudholi, D. H. (2022). *Klasifikasi Emosi Pada Teks Menggunakan Metode Deep Learning*.
- Muliyono, M. (2021). *Identifikasi Chatbot dalam Meningkatkan Pelayanan Online Menggunakan Metode Natural Language Processing* (Doctoral dissertation, Universitas Putra Indonesia YPTK).
- Negara, A. B. P., Muhandi, H., & Sajid, F. (2021). Perbandingan algoritma klasifikasi terhadap emosi tweet berbahasa Indonesia. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 7(2), 242-249.
- Rosada, I. F. A. (2023). *TA: Talent Mapping Identifikasi Minat Bakat Seseorang menggunakan Natural Language Processing* (Doctoral dissertation, Universitas Dinamika).
- Sebastian, D., Purnomo, H. D., & Sembiring, I. (2022, December). Bert for *Natural Language Processing* in bahasa Indonesia. In *2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)* (pp. 204-209). IEEE.
- Viana, F. N. (2024). *Klasifikasi Emosi Opini Twitter Menggunakan Model Benchmark Indonlu* (Doctoral dissertation, UIN Ar-Raniry Banda Aceh).