

## Public Opinion on The MBG Program: Comparative Evaluation of InSet and VADER Lexicon Labeling Using SVM on Platform X

Na'ilah Puti Zakiyah <sup>1\*</sup>, Khothibul Umam <sup>2\*</sup>, Adzhal Arwani Mahfudh <sup>3\*</sup>

\* Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang

[2208096102@student.walisongo.ac.id](mailto:2208096102@student.walisongo.ac.id) <sup>1</sup>, [khothibul\\_umam@walisongo.ac.id](mailto:khothibul_umam@walisongo.ac.id) <sup>2</sup>, [adzhal@walisongo.ac.id](mailto:adzhal@walisongo.ac.id) <sup>3</sup>

### Article Info

#### Article history:

Received 2025-06-27

Revised 2025-11-15

Accepted 2025-12-10

#### Keyword:

*Analisis Sentimen,  
InSet Lexicon, MBG,  
Support Vector Machine,  
VADER Lexicon,*

### ABSTRACT

This study aims to examine public opinion regarding the MBG program on platform X by utilizing the Support Vector Machine (SVM) algorithm using two sentiment labeling methods, namely InSet Lexicon and VADER Lexicon. The data was then divided into 70% for training and 30% for testing, and extracted using Term Frequency–Inverse Document Frequency (TF-IDF) to convert the text into numerical representations. The SVM model was trained on both labeled data sets to compare their performance based on evaluation metrics such as accuracy, precision, recall, and F1 score. The results show that labeling with VADER produces a more dominant number of neutral sentiments, while InSet Lexicon produces a more balanced distribution between positive, negative, and neutral sentiments. At the modeling stage, SVM with InSet labels achieved an accuracy of 80.10%, with precision of 0.81, recall of 0.80, and an F1 score of 0.79. Meanwhile, SVM with VADER labels achieved an accuracy of 93.83%, precision of 0.94, recall of 0.94, and an F1 score of 0.93. Although VADER showed higher accuracy values, InSet Lexicon is considered more efficient and relevant for sentiment analysis in Indonesia because it is capable of producing more balanced and contextual classifications.



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

### I. PENDAHULUAN

Pemerintah Indonesia meluncurkan program MBG (Makan Bergizi Gratis) sebagai salah satu upaya untuk meningkatkan status gizi dan kesehatan masyarakat, terutama bagi kelompok rentan. Namun, implementasi program ini kerap menuai sorotan publik di media sosial, termasuk platform X. Persepsi masyarakat yang beragam—positif, netral, atau negatif—terlihat dari cuitan, komentar, dan diskusi daring. Misalnya, dalam studi mengenai komunikasi kebijakan MBG disebut bahwa opini publik di media sosial menunjukkan perasaan ambigu atau kritik terhadap percepatan pelaksanaan program dan ketidakjelasan regulasi.

Analisis konten di platform sosial seperti Twitter atau X telah terbukti menjadi metode yang efektif untuk menilai pandangan publik tentang berbagai topik kebijakan, produk, atau isu sosial, karena mampu mengidentifikasi cara pandang masyarakat terhadap opini yang ada di publik [1]. Dari segi teknis, analisis sentimen berada di antara pemrosesan bahasa alami, penggalan teks, dan linguistik berbasis komputer.

Metode ini menggunakan teknik seperti pendekatan berbasis kamus atau metode machine learning untuk mengklasifikasikan teks [2]. Dalam konteks ini, proses pelabelan data menjadi langkah krusial yang mempengaruhi kualitas model klasifikasi. Terdapat sejumlah teknik pelabelan, terutama penggunaan kamus sentimen seperti Indonesia Sentiment Lexicon atau pendekatan VADER yang banyak digunakan di platform media sosial.

Pendekatan berbasis lexicon menggunakan kamus kata yang sudah diberikan nilai sentimen, misalnya InSet Lexicon yang khusus dikembangkan untuk bahasa Indonesia, serta VADER yang dirancang untuk teks media sosial dengan pendekatan aturan dan lexicon. InSet Lexicon dipilih karena kustomisasinya yang fokus pada bahasa Indonesia sehingga lebih relevan dalam konteks lokal [3]. Sedangkan VADER dipilih karena kemampuannya yang dirancang khusus untuk teks media sosial dengan memperhitungkan intensitas kata, penggunaan huruf kapital, tanda baca, dan emotikon sehingga bisa menangkap nuansa sentimen lebih baik di media sosial [4]. Pengujian kedua metode pelabelan ini bertujuan

mengetahui mana yang memberikan representasi sentimen terbaik pada opini program MBG.

Algoritma SVM banyak digunakan dalam analisis sentimen karena kemampuannya melakukan klasifikasi dengan margin maksimal dan hasil yang akurat. Fitur-fitur data teks, seperti kata-kata atau istilah, diubah menjadi representasi numerik (misalnya TF-IDF, word embedding) agar dapat digunakan oleh SVM. Studi terkini menunjukkan kombinasi pelabelan berbasis lexicon dengan SVM menghasilkan performa baik dalam klasifikasi sentimen aplikasi dan produk digital.

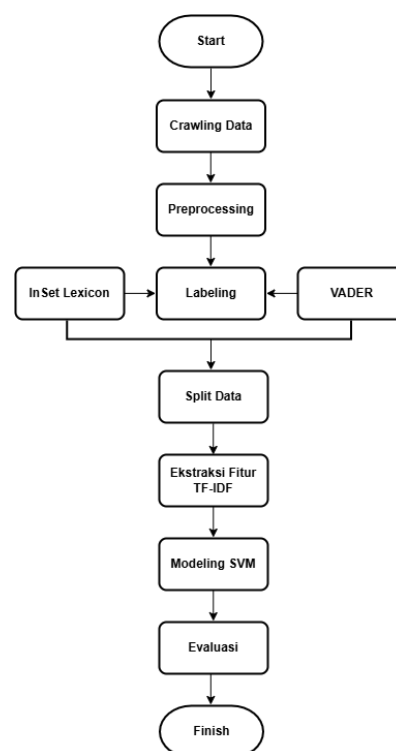
Berbagai penelitian telah membahas penggunaan kedua pelabelan ini dalam analisis sentiment. Sebagai contoh, studi [3] membandingkan pelabelan menggunakan rating-based dan InSet Lexicon pada data ulasan aplikasi GoBiz dan menguji dengan SVM, menemukan InSet Lexicon memberikan akurasi hingga 89,7% yang lebih baik daripada rating-based. Penelitian lain oleh [5] membandingkan pelabelan InSet Lexicon dan VADER pada analisis sentimen isu Rohingya di media sosial, menemukan InSet Lexicon memiliki akurasi rata-rata 85,8% dan VADER 82,65% saat diuji dengan SVM. Studi lain [6] mengevaluasi berbagai pendekatan pelabelan termasuk InSet Lexicon dan menemukan InSet efektif untuk sentiment classification di bahasa Indonesia. Penelitian lain [7] menggunakan VADER untuk analisis sentimen pada forum web dan menunjukkan bahwa VADER efektif dalam penyajian opini dengan catatan keterbatasan dalam kosakata dan konteks sarkasme.

Tujuan utama penelitian ini adalah membandingkan performa pelabelan InSet Lexicon dan VADER dalam mengklasifikasikan opini publik tentang program MBG menggunakan SVM, serta menentukan metode pelabelan mana yang memberikan hasil klasifikasi sentimen paling akurat dan dapat diandalkan. Hasil penelitian ini diharapkan dapat memberikan wawasan mengenai metode pelabelan yang lebih sesuai untuk data Indonesia dan kebijakan public, memberi pemerintah atau pembuat kebijakan alat untuk mengukur persepsi publik tentang MBG secara otomatis dan lebih akurat, serta membantu dalam monitoring dan evaluasi kebijakan berbasis opini publik daring, sehingga pemerintah dapat merespon proaktif terhadap kritik atau masalah persepsi.

## II. METODE

Proses diawali dengan pengumpulan informasi (crawling) menggunakan API, dan kemudian dilanjutkan dengan tahap preprocessing untuk membersihkan serta mempersiapkan data teks. Setelah itu, data dilabeli secara bersamaan dengan memanfaatkan dua metode pelabelan, yakni InSet Lexicon dan VADER, untuk membandingkan hasil yang diperoleh. Data yang sudah dilabeli selanjutnya dibagi menjadi set latihan dan set uji, lalu dilakukan ekstraksi fitur dengan menggunakan TF-IDF untuk merepresentasikan teks dalam format numerik. Model SVM dibangun menggunakan data latih tersebut, dan akhirnya dilakukan evaluasi performa

model berdasarkan hasil klasifikasi. Prosedur penelitian secara keseluruhan digambarkan dalam Gambar 1.



Gambar 1. Diagram Alur Penelitian

### A. Pengumpulan Data

Pengambilan data dilakukan menggunakan Application Programming Interface (API) serta memanfaatkan pustaka Python dengan teknik crawling yang dijalankan melalui platform Google Colab. Data yang dikumpulkan berasal dari komentar dan pendapat publik pada platform X (sebelumnya Twitter), yang diambil dengan kata kunci relevan dengan program MBG. Proses crawling ini memungkinkan peneliti untuk mendapatkan data dalam jumlah besar secara sistematis, serta dapat disesuaikan parameter pencarian seperti waktu, kata kunci, dan filter lain sesuai kebutuhan analisis sentimen.

### B. Preprocessing Data

Langkah awal yang penting dalam pengolahan data adalah *preprocessing*, yang dilakukan untuk memastikan teks bersih dan siap diolah. Proses ini meliputi case folding, penghapusan URL, mention dan hashtag, normalisasi, tokenization, stopword removal, dan stemming menggunakan library Sastrawi yang sesuai dengan karakteristik Bahasa Indonesia. Langkah-langkah ini penting untuk mengurangi noise pada data dan menghasilkan fitur yang representatif sebelum proses pembelajaran mesin dilakukan. Menurut penelitian terbaru, pra-pemrosesan teks memiliki pengaruh signifikan

terhadap peningkatan akurasi dalam klasifikasi sentimen berbasis TF-IDF dan SVM [8].

### C. Pelabelan Sentimen

Pelabelan adalah suatu proses pengelompokan untuk mengevaluasi emosi dari setiap komentar, yang dibagi menjadi dua kategori yakni emosi positif dan negatif. Metode penandaan ini mengimplementasikan strategi yang berakar pada leksikon, artinya memanfaatkan kumpulan kata yang telah dikelompokkan berdasarkan emosi tertentu sebagai referensi untuk menentukan polaritas dari setiap ulasan. [9]. Setelah teks dibersihkan, dilakukan pelabelan otomatis (automatic labeling) menggunakan dua pendekatan lexicon-based, yaitu InSet Lexicon dan VADER. InSet (Indonesia Sentiment Lexicon) merupakan kamus kata sentimen Bahasa Indonesia yang dikembangkan untuk mengukur polaritas opini dalam teks lokal dengan menghitung bobot positif dan negatif dari setiap kata [8]. Sebuah teks akan dikategorikan sebagai positif, negatif, atau netral berdasarkan skor total yang diperoleh.

Sementara itu, VADER merupakan metode *rule-based sentiment analysis* yang dirancang untuk teks media sosial dalam bahasa Inggris dan multibahasa. VADER memberikan skor komposit (compound score) antara -1 sampai +1 yang menunjukkan kekuatan emosi dalam teks; skor di atas 0.05 dikategorikan sebagai positif, di bawah -0.05 sebagai negatif, dan sisanya sebagai netral [10]. Studi lain membandingkan VADER dengan beberapa lexicon lain seperti WordNet dan SentiWordNet dalam analisis sentimen ulasan properti di Malaysia, di mana VADER mampu mengidentifikasi kalimat netral dan negatif dengan lebih efektif, meski ada keterbatasan seperti sulit mendeteksi sarkasme dan istilah yang salah eja [11].

InSet Lexicon juga telah diaplikasikan dalam analisis sentimen komentar media sosial, misal pada respons publik terhadap isu nasional, di mana hasilnya merekomendasikan lexicon ini sebagai metode pelabelan yang lebih relevan untuk karakteristik sentimen lokal [12]. Penggunaan kedua metode ini bertujuan untuk membandingkan efektivitas pelabelan berbasis lexicon lokal (InSet) dengan lexicon universal (VADER) terhadap data publik Indonesia.

### D. Split Data

Dataset dipecah menjadi dua bagian menggunakan teknik pemisahan data, dengan proporsi 70% dialokasikan untuk pelatihan dan 30% untuk pengujian dari keseluruhan data. Pembagian ini dipilih karena sudah lazim dipakai dalam studi yang berorientasi pada analisis sentimen dan telah terbukti memberikan keseimbangan yang baik antara jumlah data yang digunakan untuk membangun model dan data yang digunakan untuk mengevaluasi kinerja model secara objektif [13].

### E. Ekstraksi Fitur TF-IDF

Teknik ekstraksi fitur menggunakan TF-IDF (Term Frequency-Inverse Document Frequency) digunakan untuk merubah teks menjadi bentuk angka yang mencerminkan tingkat kepentingan suatu kata dalam konteks seluruh kumpulan dokumen. Representasi TF-IDF sangat berguna dalam studi analisis sentimen karena dapat menyeimbangkan antara kata-kata yang sering muncul dengan kata-kata yang memiliki arti khusus.

### F. Support Vector Machine

Algoritma ini banyak digunakan dalam analisis sentimen karena terbukti efisien dalam mengklasifikasikan data berdimensi tinggi seperti teks, serta memiliki kemampuan generalisasi yang baik [14]. Namun, SVM membutuhkan waktu komputasi yang lebih lama pada dataset. Pada penelitian ini, digunakan kernel linear karena lebih sesuai untuk data yang representasinya sparsity tinggi, seperti TF-IDF.

Penelitian terdahulu mengenai penerapan Support Vector Machine (SVM) dalam analisis sentimen menunjukkan bahwa pemilihan metode pelabelan (labeling) memiliki pengaruh signifikan terhadap hasil klasifikasi. Salah satu studi yang relevan dilakukan oleh [5], yang membandingkan penggunaan InSet Lexicon dan VADER Lexicon dalam analisis sentiment dengan menggunakan SVM menunjukkan bahwa perbedaan karakteristik bahasa antara InSet (yang berbasis leksikon Bahasa Indonesia) dan VADER (yang dikembangkan untuk Bahasa Inggris) berdampak pada variasi tingkat akurasi, di mana InSet lebih unggul dalam menangkap nuansa linguistik lokal.

Penelitian lain oleh [15] juga menerapkan VADER Lexicon dan SVM untuk mengklasifikasi komentar pengguna aplikasi Blu BCA. Studi ini berfokus pada evaluasi kinerja model dalam mengenali opini pengguna terhadap layanan digital perbankan. Hasilnya menunjukkan bahwa kombinasi VADER dan SVM mampu mencapai tingkat akurasi yang kompetitif, terutama dalam mendeteksi sentimen positif dan negatif.

### G. Evaluasi

Penilaian kinerja model SVM dilakukan dengan menggunakan laporan klasifikasi yang mencakup presisi, recall, dan F1-score untuk setiap kategori sentimen serta tingkat akurasi secara keseluruhan. Performa model yang menggunakan pelabelan InSet Lexicon dan VADER dibandingkan secara langsung. Evaluasi ini penting untuk menentukan metode pelabelan yang paling efektif untuk analisis sentimen opini publik program MBG.



kontribusi yang berarti bagi analisis sentimen. Kata-kata ini biasanya muncul sangat sering dalam teks, tetapi tidak membawa makna khusus yang dapat memengaruhi hasil analisis.

TABEL 6  
HASIL STOPWORD REMOVAL

| Token   | Stopword Removal  |
|---|---|
| ['harusnya', 'ganti', 'nama', 'jangan', 'mbg', 'tapi', 'makan', 'bergizi', 'seikhlasnya'] | ['ganti', 'nama', 'mbg', 'tapi', 'makan', 'bergizi', 'seikhlasnya'] |

6) *Stemming*: Stemming merujuk pada metode mengubah kata-kata menjadi bentuk dasar atau bentuk dasarnya dengan menghilangkan sufiks, prefiks, dan bagian-bagian lain yang tidak perlu. Teknik ini memudahkan pengelompokan kata-kata dengan makna yang serupa, yang memungkinkan analisis informasi teks yang lebih efektif. Misalnya, istilah “memasak” dikurangi menjadi “masak,” dan “menulis” disederhanakan menjadi “tuliskan.” [17]. Dalam studi ini, teknik stemming dilakukan menggunakan perpustakaan Sastrawi, yang dirancang khusus untuk bahasa Indonesia.

TABEL 7  
HASIL STEMMING

| Stopword Removal  | Stemming   |
|---|--|
| ['ganti', 'nama', 'mbg', 'tapi', 'makan', 'bergizi', 'seikhlasnya'] | ['ganti', 'nama', 'mbg', 'tapi', 'makan', 'gizi', 'ikhlaskan'] |



Gambar 3. WordCloud Setelah Preprocessing

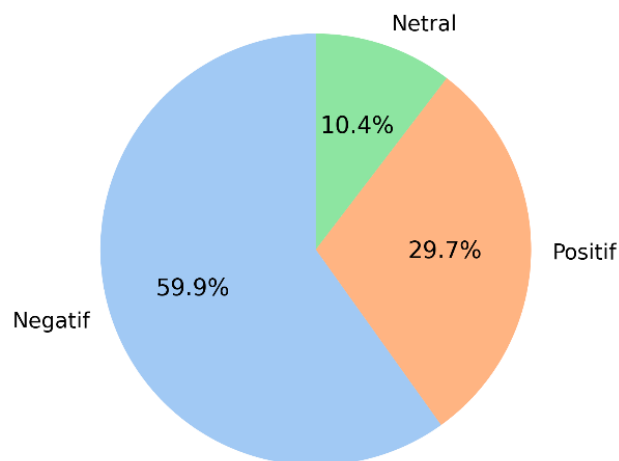
### C. Pelabelan Sentimen

Tahap penandaan dilaksanakan untuk mengelompokkan setiap tweet ke dalam tiga jenis sentimen, yakni positif, negatif, dan netral. Dalam kajian ini, diterapkan dua metode leksikon, yaitu InSet Lexicon (Leksikon Sentimen Indonesia) dan VADER (Valence Aware Dictionary and sEntiment Reasoner) sebagai pembanding dalam proses pelabelan data. Hasil pelabelan dijelaskan dalam Gambar 4 dan Gambar 5.

1) *InSet Lexicon*: Berdasarkan hasil visualisasi pada Gambar 1, distribusi sentimen dengan pendekatan InSet Lexicon menunjukkan bahwa mayoritas tweet termasuk

dalam kategori negatif sebanyak 59.9%, diikuti oleh positif 29.7%, dan netral 10.4%. Hasil ini mengindikasikan bahwa warganet cenderung memberikan tanggapan yang bernada negatif terhadap program MBG. Kecenderungan ini dapat terjadi karena InSet, sebagai kamus sentimen bahasa Indonesia, memiliki cakupan kosakata yang lebih spesifik terhadap ekspresi negatif dalam konteks sosial media berbahasa Indonesia, sehingga mampu menangkap nuansa negatif lebih akurat.

Distribusi Sentimen InSet Lexicon



Gambar 4. Distribusi Sentimen InSet Lexicon

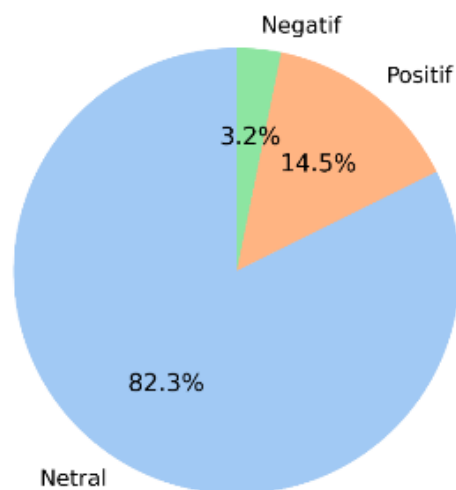
TABEL 8  
HASIL LABELING INSET LEXICON

|   | Tweet   | Score | Sentimen |
|---|---|-------|----------|
| 1 | hebat banget presiden prabowo galang aman pangan bantu program mbg sukses keamananpangan                    | 3     | Positif  |
| 2 | bas balita tasikmalaya duga racun makan mbg orang tua korban susu uht asam                                  | -16   | Negatif  |
| 3 | udah hembus angin yg bilang harga bahan pokok gak turun turun salah satu gara gara mbg                      | -25   | Negatif  |
| 4 | ganti nama mbg tapi makan gizi ikhlaskan  | 0     | Netral   |
| 5 | ratus ajar smp negeri cisarua bandung barat alami racun konsumsi menu mbg racun duga akibat ayam kecap basi | -21   | Negatif  |

2) *VADER Lexicon*: Gambar 5 menunjukkan distribusi yang berbeda signifikan, dengan dominasi sentimen netral sebesar 82.3%, positif 14.5%, dan negatif hanya 3.2%. Hasil ini menandakan bahwa model VADER lebih sering mengklasifikasikan teks berbahasa Indonesia sebagai netral. Hal ini wajar mengingat VADER dikembangkan untuk teks berbahasa Inggris, sehingga kurang mampu mengenali konteks emosional dalam bahasa Indonesia, terutama slang,

singkatan, atau campuran bahasa (*code-mixing*) yang umum digunakan di Twitter Indonesia.

Distribusi Sentimen Lexicon VADER



Gambar 5. Distribusi Sentimen VADER Lexicon

TABEL 9  
HASIL LABELING VADER

|   | <i>Tweet</i>   | <i>Score</i> | <i>Sentiment</i> |
|---|--|--------------|------------------|
| 1 | hebat banget presiden prabowo galang aman pangan bantu program mbg sukses keamananpangan | 0.0          | Netral           |
| 2 | bas balita tasikmalaya duga racun makan mbg orang tua korban susu uht asam               | 0.0          | Netral           |
| 3 | wow dapur mbg daerah bukaaa awas terima lulus giziiii                                    | 0.5          | Positif          |
| 4 | ganti nama mbg tapi makan gizi ikhlas  | 0.0          | Netral           |
| 5 | oke gw buka kritik bangun mbg stop no  | -0.5         | Negatif          |

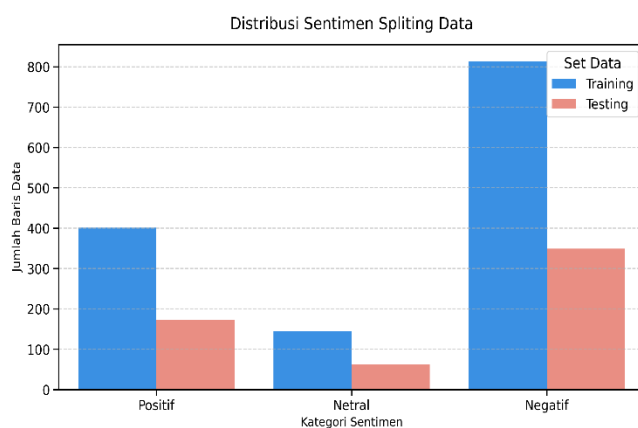
Perbedaan distribusi ini menunjukkan bahwa InSet lebih sensitif terhadap ekspresi negatif dalam bahasa Indonesia, terutama pada kalimat yang mengandung kritik, keluhan, atau sindiran yang secara semantik dikenali sebagai bernuansa negatif. Sementara itu, VADER cenderung mengklasifikasikan banyak tweet sebagai netral, karena leksikon VADER berbasis pada bahasa Inggris dan kurang mampu menangkap nuansa emosional khas bahasa Indonesia seperti kata “lumayan”, “biasa aja”, atau “terlalu” yang dalam konteks lokal dapat menunjukkan kecenderungan negatif.

Temuan ini sejalan dengan hasil penelitian oleh [18] dalam “Performance Comparison of Indonesian Sentiment Lexicon and VADER in Analyzing Local Social Media Data”, yang menyatakan bahwa penggunaan InSet Lexicon efektif untuk menganalisis teks berbahasa Indonesia karena kamus sentimennya telah disesuaikan dengan konteks linguistik lokal. Penelitian tersebut juga menegaskan bahwa

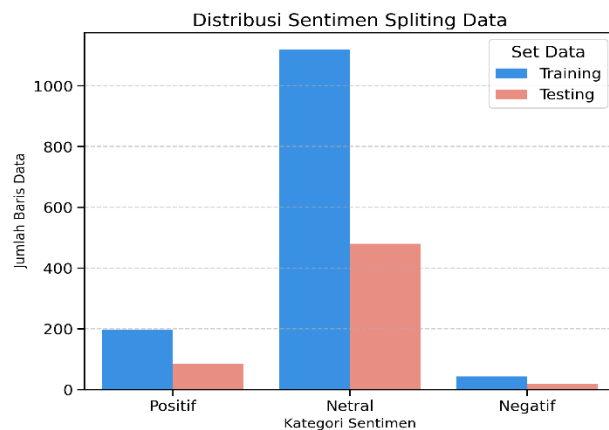
pendekatan lexicon-based menggunakan InSet mampu menghasilkan klasifikasi sentimen yang cukup baik tanpa memerlukan proses pelatihan model, meskipun masih memiliki keterbatasan dalam menangani kalimat ambigu dan sarkastik.

#### D. Split Data (Pembagian Data Latih dan Data Uji)

Untuk tujuan pelatihan dan evaluasi model, dataset berlabel dipecah menjadi dua bagian dengan perbandingan 70% untuk pelatihan dan 30% untuk pengujian (pembagian yang terstratifikasi berdasarkan label sehingga proporsi kelas tetap seimbang pada kedua bagian). Pembagian ini dilakukan secara proporsional terhadap setiap kategori sentimen—positif, netral, dan negatif—baik pada hasil klasifikasi menggunakan InSet Lexicon maupun VADER Lexicon.



Gambar 6. Split Data InSet Lexicon



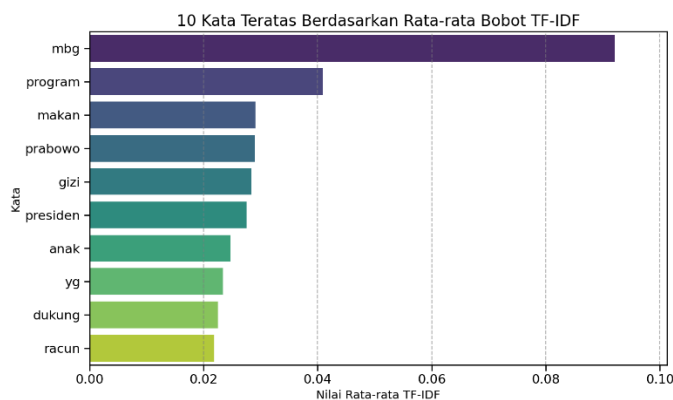
Gambar 7. Split Data VADER Lexicon

#### E. Ekstraksi Fitur TF-IDF

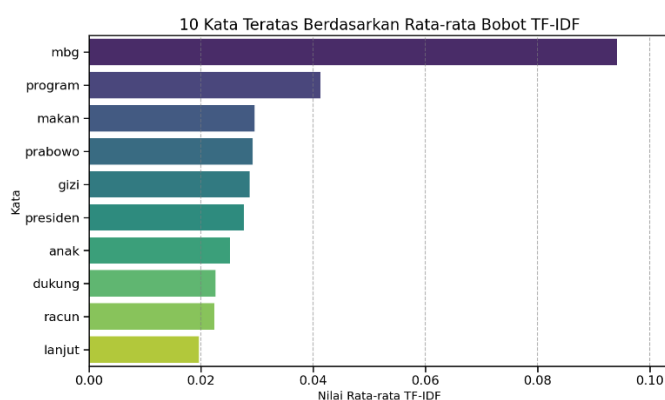
Proses diawali dengan inialisasi TfidfVectorizer dengan parameter `max_features=5000` untuk membatasi jumlah kata yang paling sering muncul agar model tidak kelebihan fitur dan tetap fokus pada kata yang paling representatif. Selanjutnya, data pelatihan ( $X_{train}$ ) di-fit dan di-transform



menggunakan model TF-IDF yang telah diinisialisasi, sedangkan data pengujian ( $X_{test}$ ) hanya melalui proses transform agar menggunakan bobot yang sama dengan data pelatihan.



Gambar 8. Visualisasi TF-IDF InSet Lexicon



Gambar 9. Visualisasi TF-IDF VADER Lexicon

Hasil dari vektorisasi menghasilkan matriks berukuran sesuai dengan jumlah dokumen dan fitur yang terbentuk. Kemudian dilakukan analisis terhadap kata dengan bobot rata-rata TF-IDF tertinggi untuk melihat kata yang paling berpengaruh dalam korpus. Sepuluh kata teratas yang muncul antara lain “mbg”, “program”, “makan”, “prabowo”, “gizi”, “presiden”, “anak”, “dukung”, “racun”, dan “lanjut”, yang menunjukkan konteks dominan dalam pembahasan di Twitter mengenai isu gizi dan politik. Visualisasi hasil ditampilkan dalam bentuk diagram batang horizontal yang menunjukkan besarnya nilai rata-rata TF-IDF setiap kata. Semakin besar nilai TF-IDF, semakin istimewa dan berartinya kata tersebut dalam kumpulan tweet.

#### F. Evaluasi

Tahap penilaian dijalankan untuk menilai kinerja model Support Vector Machine (SVM) dalam mengklasifikasikan sentimen yang telah diberi label melalui dua metode, yaitu InSet Lexicon dan VADER Lexicon. Penilaian dilakukan dengan memanfaatkan metrik laporan klasifikasi, yang mencakup nilai presisi, ingatan, F1-score, dan akurasi. Setiap

metrik memberikan gambaran sejauh mana model mampu mengklasifikasikan data dengan benar serta menghindari kesalahan prediksi pada masing-masing kelas sentimen.

| EVALUASI MODEL SVM |           |        |          |         |
|--------------------|-----------|--------|----------|---------|
| Accuracy: 0.8010   |           |        |          |         |
|                    | precision | recall | f1-score | support |
| Negatif            | 0.80      | 0.93   | 0.86     | 349     |
| Netral             | 0.91      | 0.32   | 0.48     | 62      |
| Positif            | 0.80      | 0.72   | 0.76     | 172     |
| accuracy           |           |        | 0.80     | 583     |
| macro avg          | 0.83      | 0.66   | 0.70     | 583     |
| weighted avg       | 0.81      | 0.80   | 0.79     | 583     |

Gambar 10. Classification Report InSet Lexicon

| EVALUASI MODEL SVM |           |        |          |         |
|--------------------|-----------|--------|----------|---------|
| Accuracy: 0.9383   |           |        |          |         |
|                    | precision | recall | f1-score | support |
| Negatif            | 1.00      | 0.37   | 0.54     | 19      |
| Netral             | 0.93      | 1.00   | 0.96     | 480     |
| Positif            | 0.98      | 0.71   | 0.83     | 84      |
| accuracy           |           |        | 0.94     | 583     |
| macro avg          | 0.97      | 0.69   | 0.78     | 583     |
| weighted avg       | 0.94      | 0.94   | 0.93     | 583     |

Gambar 11. Classification Report VADER Lexicon

TABEL 10  
HASIL PERBANDINGAN EVALUASI SVM

| Metrik Evaluasi          | SVM + InSet Lexicon | SVM + VADER |
|--------------------------|---------------------|-------------|
| Accuracy                 | 0.80                | 0.94        |
| Precision (Weighted Avg) | 0.81                | 0.94        |
| Recall (Weighted Avg)    | 0.80                | 0.94        |
| F1-Score (Weighted Avg)  | 0.79                | 0.93        |
| Recall Kelas Negatif     | 0.93                | 0.37        |
| Recall Kelas Netral      | 0.32                | 1.00        |
| Recall Kelas Positif     | 0.88                | 0.93        |

Berdasarkan Tabel 10, dapat dilihat bahwa model SVM dengan VADER Lexicon unggul pada hampir seluruh metrik utama seperti akurasi, precision, recall, dan F1-score. Hal ini menunjukkan bahwa VADER lebih sensitif terhadap konteks linguistik informal dan ekspresif yang umum digunakan di media sosial. Sebaliknya, InSet Lexicon, yang dikembangkan khusus untuk Bahasa Indonesia formal, menunjukkan

performa baik dalam mendeteksi sentimen negatif namun kurang optimal dalam mengidentifikasi sentimen netral.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian, penerapan Support Vector Machine (SVM) menunjukkan performa yang bervariasi tergantung pada jenis leksikon yang digunakan dalam proses pelabelan dan klasifikasi sentimen. Hasil pelabelan menggunakan VADER Lexicon menghasilkan dominasi sentimen netral yang tinggi, sedangkan InSet Lexicon memberikan distribusi yang lebih seimbang antara sentimen positif, negatif, dan netral. Ketidakseimbangan pada VADER membuat SVM kurang optimal dalam membedakan polaritas sentimen, sementara keseimbangan label pada InSet meningkatkan kemampuan model dalam mengenali variasi emosi secara lebih stabil dan representatif terhadap bahasa Indonesia.

Pada tahap pemodelan, SVM yang dilatih menggunakan label dari InSet Lexicon mencatatkan akurasi sebesar 80,10%, dengan precision 0.81, recall 0.80, dan F1-score 0.79. Sementara itu, SVM yang menggunakan label dari VADER Lexicon memperoleh hasil lebih tinggi, dengan akurasi 93,83%, precision 0.94, recall 0.94, dan F1-score 0.93. Walaupun hasil numerik VADER tampak lebih unggul, perbedaan konteks bahasa menunjukkan bahwa InSet lebih efektif dan relevan untuk teks berbahasa Indonesia, karena mampu menangkap makna dan ekspresi lokal secara lebih akurat. Sebaliknya, VADER lebih sesuai digunakan untuk teks berbahasa Inggris atau campuran dengan struktur informal. Dengan demikian, InSet Lexicon dinilai lebih efisien dan kontekstual dalam pelabelan sentimen bahasa Indonesia, serta mendukung performa SVM yang lebih representatif dalam analisis opini publik di media sosial.

#### DAFTAR PUSTAKA

- [1] G. D'Aniello, M. Gaeta, and I. La Rocca, *KnowMIS-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis*, vol. 55, no. 7. Springer Netherlands, 2022. doi: 10.1007/s10462-021-10134-9.
- [2] F. Aftab *et al.*, "A Comprehensive Survey on Sentiment Analysis Techniques," *Int. J. Technol.*, vol. 14, no. 6, pp. 1288–1298, 2023, doi: 10.14716/ijtech.v14i6.6632.
- [3] H. Firda *et al.*, "Perbandingan Pelabelan Rating - based dan Inset Lexicon - based dalam Analisis Sentimen Menggunakan SVM ( Studi Kasus : Ulasan Aplikasi GoBiz di Google Play Store ) Comparison of Rating - based and Inset Lexicon - based Labeling in Sentiment Analysis usin," vol. 14, pp. 516–528, 2025.
- [4] M. Gultom, J. Marikros, W. Rusli, and V. C. Mawardi, "Penerapan Vader Sentiment untuk Mendeteksi Sentimen Bahasa Inggris berbasis Website," *Semin. Nas. Penelit. (SEMNAS CORISINDO 2024)*, pp. 13–18, 2024, [Online]. Available: <http://corisindo.utb-univ.ac.id/index.php/penelitian/article/view/9>
- [5] Muhammad Fernanda Naufal Fathoni, Eva Yulia Puspaningrum, and Andreas Nugroho Sihananto, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," *Modem J. Inform. dan Sains Teknol.*, vol. 2, no. 3, pp. 62–76, 2024, doi: 10.62951/modem.v2i3.112.
- [6] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
- [7] N. Giovanni, M. M. Olivia Pangaribuan, A. Mulyono, and Z. Muttaqin, "Analisis Sentimen Menggunakan Metode Vader, Sentiar dan Analisis Tematik pada Akun Instagram Pecinta Hewan Peliharaan," *J. Manaj. Pendidik. Dan Ilmu Sos.*, vol. 6, no. 1, pp. 426–443, 2024, doi: 10.38035/jmpis.v6i1.3425.
- [8] A. Rufaida, A. Permanasari, and N. Setiawan, "Lexicon-Based Sentiment Analysis Using Inset Dictionary: A Systematic Literature Review," 2023, doi: 10.4108/eai.5-10-2022.2327474.
- [9] Heti Aprilianti, Khothibul Umam, and Maya Rini Handayani, "Comparative Study of SVM, KNN, and Naïve Bayes for Sentiment Analysis of Religious Application Reviews," *J. Appl. Informatics Comput.*, vol. 9, no. 3, pp. 920–927, 2025, doi: 10.30871/jaic.v9i3.9482.
- [10] J. Wilson and C. Hernández-Hall, "Octava Conferencia Internacional AAAI sobre Weblogs y Redes Sociales," *Eighth Int. AAAI Conf. Weblogs Soc. Media*, p. 18, 2014, [Online]. Available: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109>
- [11] Ahmad Taufik Nursal, "Battle of Sentiment Lexicons: Wordnet, Sentiwordnet, Textblob and Vader in Web Forum Analysis," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 2s, pp. 84–93, 2025, doi: 10.52783/jisem.v10i2s.203.
- [12] A. Okta, K. Adi, A. Sanjaya, A. B. Setiawan, and P. Korespondensi, "Penerapan Inset Lexicon untuk Analisis Sentimen Penonton Video JKT48 di YouTube 1\*," *Inotek*, vol. 9, p. 1276, 2025.
- [13] R. S. Amardita, A. Adiwijaya, and M. D. Purbolaksono, "Analisis Sentimen terhadap Ulasan Paris Van Java Resort Lifestyle Place di Kota Bandung Menggunakan Algoritma KNN," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 1, p. 62, 2022, doi: 10.30865/jurikom.v9i1.3793.
- [14] M. Kusairi and S. Agustioan, "11531-Article Text-43685-1-10-20221130," *J. Teknol. Inf. dan Komun.*, vol. 13, no. 2, pp. 140–150, 2002.
- [15] D. Sabrina, A. D. Sabilla, N. Azizah, and \* Korespondensi, "Kombinasi Vader Lexicon Dan Support Vector Machine Untuk Klasifikasi Sentimen Komentar Aplikasi Blu Bca," *Inser. Inf. Syst. Emerg. Technol. J.*, vol. 6, no. 1, pp. 22–33, 2025.
- [16] M. Apriliyani, "Implementasi analisis sentimen pada ulasan aplikasi Duolingo di Google Playstore menggunakan algoritma Naïve Bayes," *AITI*, vol. 21, no. Analysis Sentiment, p. 14, 2024.
- [17] Rahayu deny danar dan alvi furwanti Alwie, A. B. Prasetyo, R. Andespa, P. N. Lhokseumawe, and K. Pengantar, "Analisis Sentimen Ulasan Aplikasi M-Paspor Menggunakan Naive Bayes Dan Support Vector Machine," *J. Ekon. Vol. 18, Nomor 1 Maret201*, vol. 2, no. 1, pp. 41–49, 2020.
- [18] R. Firdaus, I. Asror, and A. Herdiani, "Lexicon-Based Sentiment Analysis of Indonesian Language Student Feedback Evaluation," *Indones. J. Comput.*, vol. 6, no. 1, pp. 1–12, 2021, doi: 10.34818/indojc.2021.6.1.408.