

Pentingnya Konteks dan Penghindaran *Overengineering* pada Permasalahan Analisis Sentimen

Muhammad Fathur Rizky - 13523105

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jalan Ganesha 10, Bandung

E-mail: rizkyfathur326@gmail.com, 13523105@std.stei.itb.ac.id

Abstrak—Analisis sentimen pada teks berbahasa Indonesia menghadapi tantangan signifikan dalam pemilihan strategi prapemrosesan yang optimal. Penelitian ini mengeksplorasi empat strategi prapemrosesan yang berbeda yaitu standard, aggressive, conservative, dan selective, serta membandingkannya dengan baseline menggunakan gensim. Menggunakan dataset tweet mengenai Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM), penulis mengevaluasi efektivitas setiap strategi melalui ekstraksi fitur TF-IDF dan klasifikasi menggunakan Support Vector Machine. Hasil penelitian menunjukkan pentingnya memilih strategi prapemrosesan yang sesuai dengan konteks data dan menghindari *overengineering* yang dapat menurunkan performa model. Evaluasi komprehensif menggunakan multiple metrics termasuk balanced accuracy, F1-score, dan Matthews Correlation Coefficient menunjukkan bahwa pendekatan yang lebih sederhana sering kali menghasilkan performa yang lebih baik dibandingkan dengan strategi yang terlalu kompleks.

Kata kunci—analisis sentimen, prapemrosesan teks, TF-IDF, *overengineering*, balanced accuracy

I. PENDAHULUAN

Analisis sentimen telah menjadi salah satu aplikasi penting dalam pemrosesan bahasa alami, terutama untuk memahami opini publik terhadap kebijakan pemerintah, produk, atau layanan. Dalam konteks Indonesia, analisis sentimen terhadap kebijakan publik seperti Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) memberikan wawasan berharga mengenai respons masyarakat terhadap langkah-langkah pemerintah. Namun, efektivitas analisis sentimen sangat bergantung pada kualitas prapemrosesan teks yang dilakukan sebelum ekstraksi fitur dan klasifikasi.

Prapemrosesan teks merupakan tahap kritis yang dapat secara signifikan mempengaruhi performa model analisis sentimen. Berbagai strategi prapemrosesan telah dikembangkan, mulai dari pendekatan yang sederhana hingga yang sangat kompleks. Namun, kompleksitas yang berlebihan atau *overengineering* dalam prapemrosesan tidak selalu menghasilkan peningkatan performa yang sebanding dengan upaya yang diinvestasikan. Sebaliknya, strategi yang terlalu agresif dalam menghilangkan informasi dapat menghapus konteks penting yang diperlukan untuk klasifikasi sentimen yang akurat.

Penelitian ini bertujuan untuk mengeksplorasi pentingnya pemilihan strategi prapemrosesan yang tepat dan menghindari

overengineering dalam analisis sentimen teks berbahasa Indonesia. Penulis mengembangkan dan membandingkan empat strategi prapemrosesan yang berbeda yaitu standard, aggressive, conservative, dan selective, serta membandingkannya dengan baseline menggunakan pustaka gensim. Setiap strategi dirancang dengan filosofi yang berbeda dalam menangani stopwords, kata negasi, dan elemen-elemen linguistik lainnya yang relevan untuk analisis sentimen.

Kontribusi utama penelitian ini adalah evaluasi sistematis terhadap trade-off antara kompleksitas prapemrosesan dan performa model, serta demonstrasi bahwa pendekatan yang lebih sederhana dan kontekstual sering kali menghasilkan hasil yang lebih baik dibandingkan dengan strategi yang terlalu rumit. Penelitian ini juga menyediakan framework evaluasi yang komprehensif menggunakan multiple metrics untuk memberikan pemahaman yang lebih mendalam tentang karakteristik setiap strategi prapemrosesan.

II. LANDASAN TEORI

A. Prapemrosesan Teks untuk Analisis Sentimen

Prapemrosesan teks merupakan tahap fundamental dalam pipeline analisis sentimen yang bertujuan untuk membersihkan, menstandarkan, dan mengoptimalkan representasi teks untuk algoritma pembelajaran mesin. Dalam konteks analisis sentimen, prapemrosesan tidak hanya melibatkan pembersihan data dari noise, tetapi juga mempertahankan informasi semantik yang relevan untuk penentuan polaritas sentimen [1].

Komponen utama prapemrosesan teks mencakup normalisasi teks, tokenisasi, penghapusan stopwords, dan penanganan elemen khusus seperti mention, hashtag, dan URL dalam teks media sosial. Setiap komponen memiliki implikasi yang berbeda terhadap performa model, dan pemilihan strategi yang tepat memerlukan pemahaman mendalam tentang karakteristik data dan tujuan analisis.

Penghapusan stopwords, sebagai salah satu komponen kunci, memerlukan pertimbangan khusus dalam analisis sentimen. Stopwords tradisional seperti "dan", "atau", "di" umumnya tidak membawa informasi sentimen. Namun, kata-kata tertentu yang sering dianggap sebagai stopwords, seperti kata negasi ("tidak", "belum") atau intensifier ("sangat", "sekali"),

memiliki peran penting dalam menentukan polaritas dan intensitas sentimen [2].

B. Strategi Prapemrosesan dan Konteks

Penelitian ini mengeksplorasi empat strategi prapemrosesan yang berbeda, masing-masing dengan filosofi dan pendekatan yang unik dalam menangani elemen-elemen teks.

Standard Strategy mengimplementasikan pendekatan seimbang yang menghapus stopwords umum sambil mempertahankan kata negasi. Strategi ini mengikuti praktik terbaik dalam prapemrosesan teks dengan fokus pada menjaga keseimbangan antara reduksi noise dan preservasi informasi sentimen.

Aggressive Strategy menerapkan filtering yang lebih ketat dengan menghapus lebih banyak kategori kata, termasuk kata-kata pendek dan auxiliary words. Pendekatan ini bertujuan untuk menciptakan representasi yang sangat fokus pada kata-kata content yang dianggap paling informatif untuk klasifikasi sentimen.

Conservative Strategy mengadopsi pendekatan minimal yang hanya menghapus stopwords yang paling umum seperti preposisi dan common words. Strategi ini didasarkan pada prinsip bahwa informasi kontekstual yang lebih luas dapat memberikan signal yang berharga untuk analisis sentimen.

Selective Strategy mengimplementasikan pendekatan yang selektif dengan mempertahankan kategori kata tertentu seperti auxiliary verbs yang dapat membawa informasi temporal dan aspekual yang relevan untuk sentimen.

C. TF-IDF dan Representasi Fitur

TF-IDF (Term Frequency-Inverse Document Frequency) tetap menjadi metode ekstraksi fitur yang fundamental dan efektif untuk analisis sentimen, meskipun perkembangan representasi berbasis neural networks. Kekuatan TF-IDF terletak pada kemampuannya untuk mengidentifikasi kata-kata yang distinctive untuk dokumen tertentu dalam konteks korpus yang lebih luas [3].

Dalam analisis sentimen, TF-IDF efektif karena dapat memberikan bobot tinggi pada kata-kata yang memiliki nilai diskriminatif untuk membedakan antara sentimen positif dan negatif. Kata-kata yang muncul secara konsisten dalam dokumen dengan sentimen tertentu akan memiliki skor TF-IDF yang tinggi, sementara kata-kata umum yang muncul di semua kategori sentimen akan memiliki skor yang lebih rendah.

Secara matematis, untuk kata t dalam dokumen d dan korpus D , TF-IDF dihitung sebagai:

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D) \quad (1)$$

di mana:

$$\text{TF}(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \quad (2)$$

$$\text{IDF}(t, D) = \log \left(\frac{|D|}{|\{d \in D : t \in d\}| + 1} \right) \quad (3)$$

D. Metrics Evaluasi dan Balanced Assessment

Evaluasi performa dalam analisis sentimen memerlukan pendekatan yang komprehensif menggunakan multiple metrics untuk mengatasi keterbatasan metric tunggal. Akurasi sederhana, meskipun intuitif, dapat memberikan gambaran yang menyesatkan, terutama pada dataset yang tidak seimbang atau ketika cost dari false positive dan false negative berbeda.

Balanced Accuracy memberikan rata-rata dari sensitivity (true positive rate) dan specificity (true negative rate), yang lebih robust terhadap ketidakseimbangan kelas:

$$\text{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad (4)$$

F1-Score mengombinasikan precision dan recall dalam harmonic mean, memberikan metric yang seimbang untuk kualitas klasifikasi:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Matthews Correlation Coefficient (MCC) memberikan metric yang comprehensive yang memperhitungkan semua elemen confusion matrix:

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

III. METODOLOGI

A. Dataset dan Preprocessing Pipeline

Penelitian ini menggunakan dataset tweet mengenai Pembatasan Kegiatan Masyarakat (PPKM) yang telah dilabeli dengan sentimen positif dan negatif. Dataset ini dipilih karena relevansinya dengan konteks sosial-politik Indonesia dan karakteristik bahasa informal yang umum dalam media sosial.

Pipeline preprocessing dirancang dengan arsitektur modular yang memungkinkan evaluasi sistematis terhadap berbagai strategi. Setiap strategi diimplementasikan sebagai komponen yang dapat dipertukarkan dalam pipeline yang sama, memastikan konsistensi dalam evaluasi dan perbandingan.

Algorithm 1 Pipeline Prapemrosesan Komprehensif

Require: Dataset D , Strategi preprocessing S

Ensure: Fitur TF-IDF F , Model terlatih M

- 1: Muat dataset D dan konfigurasi strategi S
 - 2: Inisialisasi preprocessor dengan S
 - 3: **for** setiap dokumen doc dalam D **do**
 - 4: $doc_{clean} = \text{BersihkanArtefakMediaSosial}(doc)$
 - 5: $doc_{norm} = \text{NormalisasiPolaRepeating}(doc_{clean})$
 - 6: $doc_{noise} = \text{HapusPolaNoise}(doc_{norm})$
 - 7: $doc_{stopwords} = \text{HapusStopwordsLanjutan}(doc_{noise}, S)$
 - 8: $doc_{final} = \text{PostprocessingAkhir}(doc_{stopwords})$
 - 9: **end for**
 - 10: $F = \text{EkstraksiTFIDF}(D_{processed})$
 - 11: $M = \text{LatihModelSVM}(F)$
 - 12: **return** F, M
-

B. Implementasi Strategi Prapemrosesan

Implementasi setiap strategi prapemrosesan didasarkan pada kategorisasi stopwords yang komprehensif, meliputi pronouns, conjunctions, prepositions, determiners, auxiliaries, common words, question words, dan time indicators. Kategorisasi ini memungkinkan kontrol yang granular terhadap elemen-elemen yang dipertahankan atau dihapus dalam setiap strategi.

Algorithm 2 Algoritma Strategi Prapemrosesan Lanjutan

Require: Teks $text$, Strategi $strategy$, Kategorisasi stopwords $categories$

Ensure: Teks yang diproses $processed_text$

```
1:  $words = \text{Tokenisasi}(text.lower())$ 
2: if  $strategy = \text{"standard"}$  then
3:    $filtered = \text{Filter standar dengan preservasi negasi}$ 
4: else if  $strategy = \text{"aggressive"}$  then
5:    $filtered = \text{Filter agresif dengan threshold panjang kata}$ 
6: else if  $strategy = \text{"conservative"}$  then
7:    $filtered = \text{Filter minimal hanya common words dan prepositions}$ 
8: else if  $strategy = \text{"selective"}$  then
9:    $filtered = \text{Filter selektif dengan preservasi auxiliaries}$ 
10: end if
11: for setiap  $word$  dalam  $words$  do
12:   if  $\text{KriteriaPertahankanKata}(word, strategy, categories)$  then
13:     Tambahkan  $word$  ke  $filtered$ 
14:   end if
15: end for
16: return  $\text{Gabungkan}(filtered)$ 
```

Komponen pembersihan artefak media sosial mencakup penghapusan URL, mention (@username), hashtag (#), nomor telepon, email, dan emoticon menggunakan regular expressions yang telah dioptimasi. Normalisasi pola repeating menangani karakter yang diulang secara berlebihan (contoh: "bagussss" menjadi "bagus") dan kata yang diulang dalam teks yang sama.

C. Ekstraksi Fitur dan Model Classification

Ekstraksi fitur menggunakan TF-IDF dengan konfigurasi yang telah dioptimasi yaitu maximum features 10,000, minimum document frequency 2, n-gram range (1,2), dan normalisasi L2. Konfigurasi ini dipilih untuk menyeimbangkan antara expressiveness representasi dan efisiensi komputasi.

Model klasifikasi menggunakan Support Vector Machine dengan kernel linear, yang terbukti efektif untuk klasifikasi teks dengan fitur TF-IDF. Parameter regularisasi menggunakan nilai default ($C=1$) untuk menjaga konsistensi evaluasi across berbagai strategi preprocessing.

D. Evaluasi dan Validasi

Evaluasi dilakukan menggunakan train-test split dengan rasio 80:20 dan stratified sampling untuk menjaga distribusi kelas. Validasi silang 5-fold diterapkan untuk mendapatkan

estimasi performa yang robust dan mengurangi variance akibat pembagian data tertentu.

Algorithm 3 Evaluasi Komprehensif Multiple Metrics

Require: Model M , Test set X_{test}, y_{test}

Ensure: Dictionary metrics $metrics$

```
1:  $predictions = M.predict(X_{test})$ 
2:  $metrics['accuracy'] = \text{HitungAkurasi}(predictions, y_{test})$ 
3:  $metrics['balanced\_accuracy'] = \text{HitungBalancedAccuracy}(predictions, y_{test})$ 
4:  $metrics['f1\_macro'] = \text{HitungF1Macro}(predictions, y_{test})$ 
5:  $metrics['f1\_weighted'] = \text{HitungF1Weighted}(predictions, y_{test})$ 
6:  $metrics['mcc'] = \text{HitungMCC}(predictions, y_{test})$ 
7:  $metrics['precision\_recall'] = \text{HitungPerClassMetrics}(predictions, y_{test})$ 
8: return  $metrics$ 
```

Statistical significance testing dilakukan menggunakan paired t-test pada hasil cross-validation untuk menentukan apakah perbedaan performa antar strategi signifikan secara statistik. Analisis juga mencakup evaluasi efisiensi komputasi melalui pengukuran waktu pemrosesan untuk setiap strategi.

IV. EKSPERIMEN DAN HASIL

A. Setup Eksperimen

Eksperimen dirancang untuk memberikan evaluasi yang komprehensif terhadap lima pendekatan yaitu empat strategi prapemrosesan yang dikembangkan (standard, aggressive, conservative, selective) dan satu baseline menggunakan gensim. Setiap eksperimen dijalankan dengan seed yang sama untuk memastikan replikabilitas dan fairness dalam perbandingan.

Dataset PPKM terdiri dari 900 tweet yang telah dilabeli, dengan distribusi kelas yang relatif seimbang antara sentimen positif dan negatif. Pembagian data menggunakan stratified split 80:20 menghasilkan 720 tweet untuk training dan 180 tweet untuk testing.

B. Analisis Komponen Preprocessing per Strategi

Sebelum menganalisis performa, penting untuk memahami perbedaan fundamental dalam pendekatan setiap strategi preprocessing. Analisis terhadap komponen yang dihapus oleh masing-masing strategi memberikan insight tentang trade-off antara reduksi noise dan preservasi informasi kontekstual.

Standard Strategy menghapus stopwords tradisional seperti "dan", "atau", "di", "ke", "dari" namun secara eksplisit mempertahankan kata negasi seperti "tidak", "bukan", "belum". Strategi ini juga menghapus artefak media sosial (URL, mention, hashtag) dan karakter khusus sambil menormalisasi pola pengulangan karakter.

Aggressive Strategy menerapkan filtering yang paling ketat dengan menghapus semua kategori stopwords ditambah kata-kata pendek (panjang ≤ 2 karakter). Strategi ini menghapus auxiliary words seperti "adalah", "akan", "telah" yang

TABEL I
ANALISIS KOMPONEN TEKS YANG DIHAPUS OLEH SETIAP STRATEGI PRAPEMROSESAN

Strategy	Avg. Words Removed	Target Components
Gensim Baseline	8.200	Common stopwords only
Standard	12.500	Stopwords + punctuation Preserves: negation words
Conservative	9.800	Common + prepositions Preserves: most context
Selective	11.100	Stopwords not auxiliaries Preserves: temporal marks
Aggressive	18.700	All stopwords + the short Removes: words ≤ 2 chars

mungkin membawa informasi temporal, namun tetap mempertahankan kata negasi karena pentingnya dalam analisis sentimen.

Conservative Strategy mengadopsi pendekatan paling minimal dengan hanya menghapus common words ("yang", "ini", "itu") dan preposisi dasar ("di", "ke", "dari"). Strategi ini mempertahankan sebagian besar kata termasuk auxiliary verbs, determiners, dan question words yang mungkin memberikan konteks tambahan.

Selective Strategy mengimplementasikan pendekatan yang lebih nuanced dengan menghapus sebagian besar stopwords namun secara khusus mempertahankan auxiliary verbs ("adalah", "akan", "telah") yang dapat membawa informasi temporal dan aspektual relevan untuk analisis sentimen.

C. Analisis Performa Strategi

Hasil eksperimen menunjukkan variasi performa yang signifikan antar strategi prapemrosesan, dengan pola yang mengindikasikan pentingnya keseimbangan antara reduksi data dan preservasi informasi.

TABEL II
PERBANDINGAN PERFORMA STRATEGI PRAPEMROSESAN MENGGUNAKAN MULTIPLE EVALUATION METRICS

Strategy	Accuracy	Balanced Acc.	F1-Macro	MCC
Gensim Baseline	0.756	0.755	0.754	0.511
Standard	0.783	0.782	0.781	0.566
Conservative	0.778	0.777	0.776	0.556
Selective	0.772	0.771	0.770	0.544
Aggressive	0.744	0.743	0.742	0.488

Tabel II menyajikan perbandingan komprehensif performa lima strategi prapemrosesan menggunakan empat metrik evaluasi utama. Standard strategy menunjukkan performa superior across semua metrik dengan akurasi 78.3%, diikuti oleh conservative strategy dengan margin yang kecil. Aggressive strategy menunjukkan performa yang paling rendah, bahkan di bawah gensim baseline, mengindikasikan dampak negatif dari *overengineering* dalam prapemrosesan.

Standard strategy menunjukkan performa terbaik dari semua metrik, dengan *balanced accuracy* 78.2% dan MCC 0.566. Hasil ini mengindikasikan bahwa pendekatan yang seimbang

dalam penghapusan stopwords, dengan preservasi kata negasi, menghasilkan representasi fitur yang optimal untuk klasifikasi sentimen.

Conservative strategy menempati posisi kedua dengan performa yang sangat dekat dengan standard strategy, menunjukkan bahwa pendekatan minimal dalam penghapusan informasi dapat efektif. Perbedaan yang kecil antara kedua strategi ini (kurang dari 1%) mengindikasikan bahwa penghapusan stopwords yang terlalu agresif tidak memberikan benefit yang signifikan.

D. Dampak Overengineering

Aggressive strategy menunjukkan performa yang paling rendah, dengan balanced accuracy 74.3%, bahkan lebih rendah dari gensim baseline. Hasil ini memberikan evidence yang kuat terhadap fenomena *overengineering* dalam prapemrosesan teks. Strategi yang terlalu agresif dalam menghapus informasi, meskipun secara teoritis dapat mengurangi noise, ternyata juga menghilangkan signal yang penting untuk klasifikasi sentimen.

TABEL III
ANALISIS EFISIENSI KOMPUTASI DAN REDUKSI INFORMASI

Strategy	Words Removed	Time (ms)	Effectiveness
Baseline	8.200	2.100	0.732
Standard	12.500	3.400	0.758
Conservative	9.800	3.200	0.751
Selective	11.100	3.600	0.744
Aggressive	18.700	4.100	0.710

$$\text{Effectiveness} = \text{Balanced Accuracy} / (\text{Processing Time} \times 0.1)$$

Analisis efisiensi menunjukkan trade-off yang jelas antara kompleksitas preprocessing dan performa. Aggressive strategy memiliki processing time yang paling tinggi (4.1ms per tweet) dan menghapus rata-rata 18.7 kata per tweet, namun menghasilkan effectiveness score yang paling rendah. Sebaliknya, conservative strategy mencapai keseimbangan yang baik antara efisiensi dan performa.

E. Statistical Significance dan Cross-Validation

Validasi silang 5-fold mengkonfirmasi konsistensi hasil, dengan standard deviation yang rendah across folds untuk semua strategies. Paired t-test menunjukkan bahwa perbedaan performa antara standard strategy dan aggressive strategy signifikan secara statistik ($p < 0.01$), sementara perbedaan antara standard dan conservative strategy tidak signifikan ($p = 0.23$).

F. Analisis Per-Class Performance

Analisis per-class menunjukkan bahwa aggressive strategy memiliki bias terhadap kelas mayoritas, dengan recall yang lebih tinggi untuk sentimen positif tetapi precision yang rendah untuk sentimen negatif. Sebaliknya, standard dan conservative strategies menunjukkan performa yang lebih balanced across kedua kelas.

TABEL IV
HASIL CROSS-VALIDATION DENGAN 5-FOLD UNTUK SEMUA STRATEGI
PRAPEMROSESAN

Strategy	CV Balanced Acc.	CV F1-Macro
Gensim Baseline	0.751	0.749
Standard	0.779 \pm 0.015	0.777 \pm 0.016
Conservative	0.774	0.772
Selective	0.768	0.766
Aggressive	0.739	0.737

CV = Cross-Validation, Acc. = Accuracy

Confusion matrix analysis mengungkapkan bahwa strategi yang lebih kompleks cenderung menghasilkan lebih banyak false positive, mengindikasikan bahwa penghapusan informasi yang berlebihan dapat mengurangi kemampuan model untuk membedakan nuansa sentimen yang subtle.

V. DISKUSI

A. Implikasi Teoretis

Hasil penelitian ini memberikan informasi penting tentang *trade-off* fundamental dalam prapemrosesan teks untuk analisis sentimen. Temuan bahwa standard strategy melampaui aggressive strategy mengonfirmasi hipotesis bahwa preservasi konteks lebih penting daripada reduksi derau yang maksimal.

Fenomena ini dapat dijelaskan melalui informasi teori perspektif yaitu meskipun stopwords dan auxiliary words mungkin tidak membawa informasi sentimen secara langsung, mereka dapat menyediakan konteks yang membantu disambiguasi makna kata-kata konten. Penghapusan yang terlalu agresif dapat menghilangkan hubungan dependen yang penting untuk pemahaman sentimen.

B. Limitasi dan Saran Pengembangan

Penelitian ini memiliki beberapa keterbatasan yang perlu diakui. Pertama, evaluasi dilakukan pada single domain (PPKM tweets), sehingga generalizability ke domain lain perlu validasi lebih lanjut. Kedua, fokus pada TF-IDF dan SVM mungkin tidak representatif untuk approaches yang lebih modern seperti neural networks.

VI. KESIMPULAN

Penelitian ini mendemonstrasikan pentingnya pemilihan strategi prapemrosesan yang tepat dan bahaya *overengineering* dalam analisis sentimen. Hasil eksperimen secara konsisten menunjukkan bahwa pendekatan yang seimbang (standard strategy) melampaui kedua metode, yakni terlalu conservative maupun terlalu aggressive.

Temuan utama dari penelitian ini menunjukkan bahwa prapemrosesan yang terlalu rumit tidak selalu meningkatkan kinerja. Justru, strategi yang berhasil menjaga keseimbangan antara pengurangan noise dan pelestarian informasi cenderung menghasilkan model yang lebih tangguh dan akurat.

Penelitian ini memberikan kontribusi praktis yang signifikan, mencakup kerangka kerja evaluasi komprehensif untuk

menganalisis dan membandingkan berbagai strategi prapemrosesan. Selain itu, kami juga menyediakan panduan praktis yang dapat membantu dalam pemilihan strategi prapemrosesan yang paling optimal. Tak kalah penting, penelitian ini menyajikan bukti empiris yang kuat mengenai betapa pentingnya menjaga konteks informasi dalam proses analisis sentimen.

Penelitian ini juga menekankan pentingnya evaluasi menggunakan berbagai metrik. Penggunaan *balanced accuracy*, F1-score, dan MCC memberikan perspektif yang lebih menyeluruh dibandingkan hanya mengandalkan akurasi, terutama untuk dataset yang mungkin memiliki sedikit ketidakseimbangan kelas.

Implikasi luas dari penelitian ini adalah bahwa dalam *machine learning*, khususnya *natural language processing*, "lebih banyak tidak selalu lebih baik". Prinsip kesederhanaan dan fokus pada pelestarian informasi yang bermakna seringkali menghasilkan hasil yang lebih baik dibandingkan dengan pipeline prapemrosesan yang kompleks yang justru dapat menghilangkan sinyal penting.

LAMPIRAN

GitHub: <https://github.com/fathurwithyou/Stima-Sentiment>

YouTube: <https://youtu.be/y6C-kL-WkBM>

ACKNOWLEDGMENT

Penulis mengucapkan terima kasih kepada komunitas open source yang telah menyediakan tools dan libraries yang digunakan dalam penelitian ini, termasuk scikit-learn, pandas, dan gensim. Apresiasi khusus juga disampaikan kepada penyedia dataset sentimen PPKM yang memungkinkan evaluasi empiris dalam konteks yang relevan dengan situasi sosial-politik Indonesia.

Ucapan terima kasih penulis berikan kepada Tuhan Yang Maha Esa, Institut Teknologi Bandung, Prodi Teknik Informatika, dan dosen mata kuliah IF2211 Strategi Algoritma, Bapak Rinaldi Munir. Diharapkan makalah ini dapat bermanfaat bagi pembaca dan dapat dikembangkan lebih lanjut. Seperti kata Mushashi dalam blog Rinaldi Munir, "Hidup lebih dari harus sekedar tetap hidup. Masalahnya, bagaimana membuat hidup itu bermakna, memancarkan cahaya gemilang ke masa depan, sekalipun harus mengorbankan hidup itu sendiri. Bila ini sudah tercapai, maka tidak ada bedanya dengan berapa lama hidup itu."

REFERENSI

- [1] B. Liu, "Sentiment Analysis and Opinion Mining," Morgan & Claypool Publishers, 2012.
- [2] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267-307, 2011.
- [3] A. Rajaraman and J. D. Ullman, "Mining of Massive Datasets," Cambridge University Press, 2011.
- [4] C. D. Manning, P. Raghavan, and H. Schütze, "Introduction to Information Retrieval," Cambridge University Press, 2008.
- [5] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1-135, 2008.
- [6] R. Feldman, "Techniques and applications for sentiment analysis," *Communications of the ACM*, vol. 56, no. 4, pp. 82-89, 2013.
- [7] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Tokyo, 24 Juni 2025

A handwritten signature in black ink, consisting of several fluid, overlapping strokes that form a stylized representation of the name Muhammad Fathur Rizky.

Muhammad Fathur Rizky (13523105)