

PENINGKATAN KINERJA SUPPORT VECTOR MACHINE MENGUNAKAN MODEL BAHASA BERT UNTUK KLASIFIKASI SENTIMEN DENGAN DATASET TERBATAS

Marwika Rifattul Iffa¹, Surya Agustian^{2*}, Nazruddin Safaat³, Muhammad Irsyad⁴

^{1,2,3,4}Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri
Sultan Syarif Kasim Riau

Jl.HR.Soebrantas No.155 Km 15, Simpang Baru, Tampan, Pekanbaru, Riau

e-mail: ¹12050122118@students.uin-suska.ac.id, ²surya.agustian@uin-suska.ac.id,

³nazruddin.safaat@uin-suska.ac.id, ⁴irsyadtech@uin-suska.ac.id

Abstrak

Media sosial kini menjadi ruang penting bagi masyarakat untuk mengekspresikan opini secara terbuka terhadap berbagai isu terkini, salah satunya melalui platform X yang populer di kalangan pengguna internet. Platform ini sering digunakan sebagai sumber data klasifikasi sentimen guna mengungkap persepsi masyarakat terhadap peristiwa-peristiwa yang terjadi, khususnya di bidang politik dan pemerintahan. Namun, keterbatasan dataset menjadi tantangan utama dalam proses klasifikasi karena kondisi tersebut dapat mempengaruhi akurasi dan validitas sentimen yang dihasilkan. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan kombinasi algoritma Support Vector Machine (SVM) dengan fitur Bidirectional Encoder Representations from Transformers (BERT) yang terbukti efektif dalam menangkap konteks bahasa secara mendalam. Pendekatan ini bertujuan untuk meningkatkan performa klasifikasi sentimen terkait pengangkatan Kaesang Pangarep sebagai Ketua Umum Partai Solidaritas Indonesia (PSI) pada media sosial X. Metode penelitian meliputi tahap preprocessing text, ekstraksi fitur menggunakan BERT, serta penerapan SVM dalam proses klasifikasi sentimen. Hasil eksperimen menunjukkan bahwa model kombinasi tersebut berhasil meningkatkan F1-Score secara signifikan sebesar 3% pada data uji. Hal ini menandakan model bahasa BERT dapat meningkatkan performa SVM dalam klasifikasi sentimen.

Kata kunci: Klasifikasi Sentimen; SVM; BERT; X; Kaesang Pangarep

Abstract

Social media has now become an important space for the public to openly express opinions on various current issues, one of which is through the X platform, which is popular among internet users. This platform is often used as a source of sentiment classification data to reveal public perceptions of ongoing events, particularly in the fields of politics and governance. However, dataset limitations pose a major challenge in the classification process, as these conditions can affect the accuracy and validity of the resulting sentiments. To address this issue, this study proposes a combination of the Support Vector Machine (SVM) algorithm with features from Bidirectional Encoder Representations from Transformers (BERT), which has proven effective in capturing deep language context. This approach aims to improve sentiment classification performance related to the appointment of Kaesang Pangarep as the Chairman of the Indonesian Solidarity Party (PSI) on the X social media platform. The research method includes text preprocessing, feature extraction using BERT, and the application of SVM in the sentiment classification process. Experimental results show that the combined model successfully improved the F1-Score by a significant 3% on the test data. This indicates that the BERT language model can enhance the performance of SVM in sentiment classification tasks.

Keywords: Sentiment Classification; SVM; BERT; X; Kaesang Pangarep

1. PENDAHULUAN

Kemajuan pesat dalam teknologi informasi dan komunikasi telah menyebabkan perubahan besar dalam cara masyarakat mengakses serta menyebarkan informasi. Salah satu dampak paling nyata dari perkembangan ini adalah meningkatnya peran media sosial sebagai platform utama dalam komunikasi, baik di tingkat individu, organisasi, maupun lembaga pemerintahan[1]. Berdasarkan data terbaru, sekitar 49% populasi Indonesia secara aktif menggunakan media sosial, menjadikannya sebagai salah satu alat yang paling berpengaruh dalam pembentukan serta penyebaran opini publik[2]. Salah satu platform media sosial yang berpengaruh dalam membentuk opini publik yaitu Twitter. Twitter memungkinkan penggunanya untuk mengekspresikan pandangan, berdiskusi, serta menyebarkan informasi secara cepat melalui fitur retweet, hashtag, dan trending topic. Platform ini sering digunakan dalam berbagai isu sosial, politik, hingga ekonomi, sehingga menjadi ruang interaksi publik yang dinamis[3]. Opini mengenai pengangkatan Kaesang Pangarep sebagai Ketua Umum Partai Solidaritas Indonesia (PSI) pada September 2023 menjadi perbincangan luas di media sosial, terutama di Twitter[4]. Reaksi publik terhadap peristiwa ini beragam dan dapat dikategorikan ke dalam tiga jenis sentimen utama: positif, negatif, dan netral[5]. Sentimen positif muncul dari mereka yang melihat langkah ini sebagai upaya untuk menarik generasi muda ke dunia politik serta memperkuat posisi PSI. Sebaliknya, sentimen negatif datang dari pihak yang menganggap pengangkatan Kaesang sebagai bentuk politik dinasti yang mempertahankan dominasi keluarga Presiden Joko Widodo. Sementara itu, kelompok dengan sentimen netral cenderung menunggu perkembangan lebih lanjut sebelum memberikan penilaian[6]. Klasifikasi sentimen ini menunjukkan bagaimana media sosial menjadi tempat di mana opini publik dapat dikumpulkan dan dianalisis untuk memahami trend persepsi masyarakat terhadap isu politik tertentu.

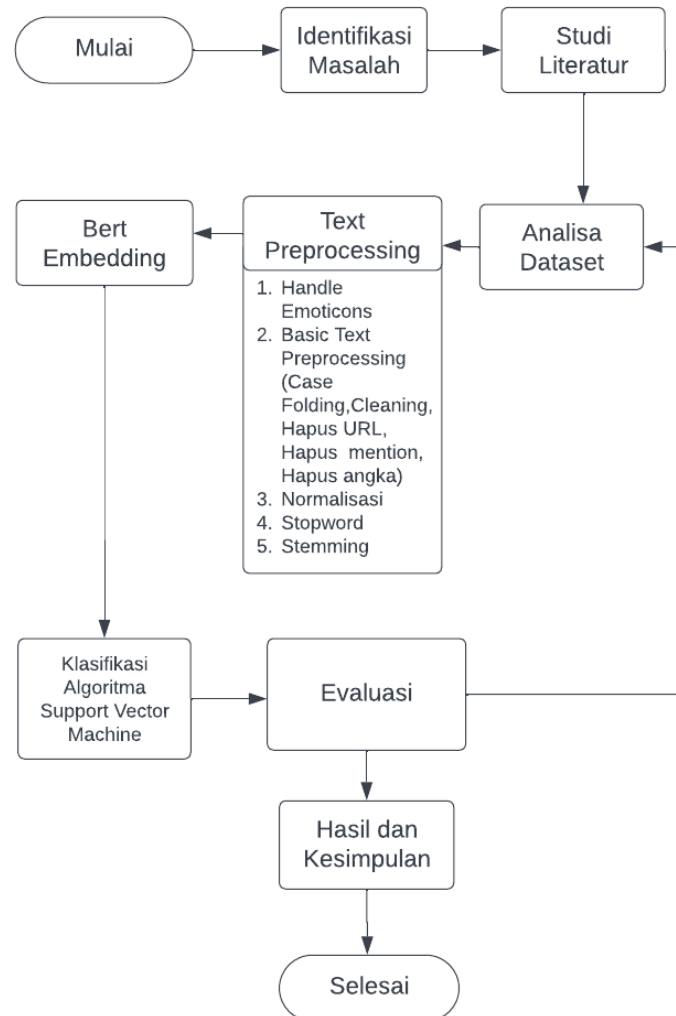
Klasifikasi sentimen adalah teknik dalam pemrosesan bahasa alami (NLP) yang digunakan untuk mengelompokkan sentimen berdasarkan isi dan konteksnya[7]. Dalam penelitian ini, algoritma Support Vector Machine (SVM) dengan fitur BERT digunakan untuk mengklasifikasikan opini publik terkait pengangkatan Kaesang Pangarep sebagai Ketua Umum PSI. Dataset yang digunakan berasal dari sistem shared task yang terdiri dari 300 data opini Kaesang serta sejumlah data open topic. Setelah melalui tahap pra-pemrosesan seperti pembersihan teks dan tokenisasi, fitur diekstraksi menggunakan BERT untuk menghasilkan representasi vektor yang lebih kontekstual. Model SVM kemudian memisahkan opini ke dalam kategori positif, netral, atau negatif. Meskipun dataset terbatas, kombinasi SVM dan BERT memungkinkan analisis sentimen yang lebih akurat[8].

Model ensemble yang menggabungkan BERT dan Support Vector Machine (SVM) untuk mendeteksi emosi dalam tweet telah diusulkan, dengan akurasi 91% dalam mengenali emosi seperti takut, sedih, bahagia, dan marah[9]. Analisis komparatif terhadap model analisis sentimen Multinomial Naive Bayes (MultinomialNB), Support Vector Machine (SVM), dan BERT pada data Twitter Garuda Indonesia memberikan wawasan mengenai efektivitas masing-masing model. BERT menunjukkan kinerja terbaik dengan akurasi 75.6% dan macro-average F1-score 58%, diikuti oleh SVM dengan akurasi 71.6% dan macro-average F1-score 50%, dan MultinomialNB dengan akurasi 64.8% dan macro-average F1-score 54%[10]. Metode Support Vector Machine (SVM) dengan fitur BERT untuk klasifikasi artikel berita mencapai akurasi 83,47%, menunjukkan efektivitas BERT sebagai metode ekstraksi fitur dibandingkan TF-IDF[11].

Berdasarkan penelitian ini, diharapkan penerapan algoritma Support Vector Machine (SVM) yang digabungkan dengan fitur BERT dapat meningkatkan akurasi dalam klasifikasi sentimen pada platform Twitter, meskipun dengan jumlah data yang terbatas. Pendekatan ini, diharapkan model yang dikembangkan dapat memberikan pemahaman yang lebih mendalam mengenai persepsi publik terhadap isu-isu terkini. Selain itu, penelitian ini juga bertujuan untuk memberikan kontribusi dalam pengembangan model klasifikasi sentimen yang lebih fleksibel, yang mampu menangani tantangan seperti data yang tidak lengkap, data noise, atau perubahan topik, tanpa mengurangi kualitas kinerjanya. Dengan demikian, penelitian ini diharapkan dapat memberikan

dampak positif dalam analisis opini publik di media sosial untuk berbagai keperluan aplikasi yang lebih luas.

2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

Metode penelitian yang merujuk pada gambar 1 dimulai dengan melakukan identifikasi masalah sebagai langkah awal. Setelah itu, dilakukan studi literatur untuk mendapatkan pemahaman lebih dalam mengenai topik yang sedang diteliti. Selanjutnya, proses berlanjut dengan analisis dataset, yang meliputi pengumpulan dan pemilahan data yang relevan untuk analisis sentimen. Setelah dataset siap, tahap berikutnya adalah text preprocessing, yang mencakup beberapa langkah penting seperti case folding, pembersihan data, penghapusan URL, mention, angka, serta penanganan emotikon, normalisasi, stopwords, dan stemming. Setelah tahap preprocessing selesai, dilakukan BERT embedding untuk menghasilkan representasi vektor dari teks. Kemudian, dilakukan klasifikasi menggunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan data menjadi kategori yang relevan. Setelah model dikembangkan, tahap selanjutnya adalah evaluasi untuk mengukur kinerja model berdasarkan metrik yang telah ditentukan, seperti F1-score, akurasi, precision, dan recall. Akhirnya, hasil dari evaluasi dianalisis dan disusun dalam hasil dan kesimpulan, sebelum proses penelitian dinyatakan selesai.

2.1. Identifikasi Masalah

Identifikasi masalah ini merupakan langkah awal penting dalam memahami kasus penelitian. Identifikasi masalah memungkinkan untuk merumuskan dan memahami masalah utama yang menjadi fokus penelitian yang dibuat.

Identifikasi masalah dalam penelitian ini adalah bagaimana mengoptimasi klasifikasi sentimen dari dataset yang kecil secara cepat dan handal dengan memanfaatkan data eksternal yaitu data Covid-19 menggunakan algoritma Support Vector Machine (SVM) dengan fitur BERT.

2.2. Studi Literatur

Studi literatur merupakan tahap penting dalam penelitian yang bertujuan untuk mengumpulkan informasi yang relevan dan mendalam dari berbagai sumber, seperti buku, jurnal ilmiah, artikel internasional, dan referensi lainnya. Tujuan utama dari studi literatur ini adalah untuk mendapatkan pemahaman yang lebih baik mengenai topik penelitian yang sedang dilakukan, serta untuk menggali teori, konsep, atau metode yang telah digunakan dalam penelitian sebelumnya yang berkaitan dengan topik tersebut. Dalam konteks penelitian ini, studi literatur akan difokuskan pada pemahaman mengenai klasifikasi sentimen menggunakan algoritma Support Vector Machine (SVM), serta penerapan teknik BERT dalam pengolahan teks. Selain itu, penggunaan data eksternal, seperti data Covid-19, juga akan dibahas untuk memahami bagaimana data tambahan dapat memperkaya dataset dan meningkatkan performa model klasifikasi sentimen pada dataset yang terbatas.

2.3. Analisa Dataset

Penelitian ini berfokus pada tantangan klasifikasi sentimen dengan dataset yang terbatas, yang menjadi salah satu isu utama dalam klasifikasi teks. Proses pengumpulan data dilakukan melalui proses *crawling* pada *platform Twitter* dengan kata kunci “Kaesang PSI” dan pelabelan dataset dilakukan melalui metode *crowdsourcing*, yaitu beberapa anotator memberikan label sentimen (positif, netral, negatif) pada setiap *Tweet*. Pelabelan sentimen dilaksanakan melalui mekanisme *Majority Voting*, dengan ketentuan bahwa *Tweet* yang tidak memperoleh label mayoritas akan dikeluarkan dari kumpulan data. *Dataset* yang digunakan diambil dari penelitian[12]. Berikut adalah tabel *dataset* yang akan digunakan untuk penelitian :

Tabel 1. Dataset

Dataset	Penggunaan	Jumlah Tweet
Data Kaesang V1	<i>Training</i>	300
Data Kaesang V2	<i>Training</i>	300
Data Covid	<i>Training</i>	8.000
Data Kaesang	<i>Testing</i>	924

Berdasarkan Tabel 1, terdapat beberapa data yang akan digunakan, yaitu data *Training* dan data *Testing*. Data *Training* terdiri dari Data Kaesang V1, Kaesang V2, dan Data Covid yang akan digunakan untuk melatih model. Sementara itu, data *Testing* berupa Data Kaesang yang berjumlah 924 tweet yang akan digunakan untuk menguji kinerja model yang telah dilatih sebelumnya. Data *Testing* ini nantinya akan dimasukkan ke dalam sistem *Leaderboard* untuk untuk membandingkan hasil prediksi model dengan model lain yang berpartisipasi dalam penelitian ini, sehingga dapat diperoleh skor evaluasi yang objektif. Sebelumnya, penelitian ini akan menggunakan data train Kaesang V2 yang akan dibagi menjadi data *Training* dan data *Validasi*. Pembagian data train Kaesang V2 yaitu data train sebesar 80% dan data validasi sebesar 20%.

2.4. Text Preprocessing

Text preprocessing adalah tahapan penting dalam pemrosesan bahasa alami (NLP) untuk mempersiapkan data teks agar dapat diolah oleh model *machine learning*[13]. Dalam klasifikasi sentimen, *preprocessing* sangat penting karena data teks dari *platform* seperti *Twitter* sering mengandung *noise*, variasi linguistik, dan bahasa informal seperti singkatan, *emoticon*, atau

mention (@username). Tanpa tahapan *text preprocessing* yang tepat, model kesulitan memahami konteks teks, sehingga hasil analisis kurang akurat. Oleh karena itu, *text preprocessing* diperlukan untuk membersihkan, menstandarisasi, dan mengubah teks mentah menjadi format yang terstruktur dan siap diproses[14]. Berikut beberapa tahapan *text preprocessing*:

- a. *Emoji Handling* yaitu proses mengganti emotikon dalam teks dengan kata-kata yang menggambarkan emosi, seperti ":" menjadi "senyum" atau ":(" menjadi "sedih", untuk memudahkan analisis sentimen. Pada kode di atas, fungsi `handle_emoticons()` melakukan penggantian emotikon dengan kata-kata yang sudah didefinisikan dalam `emoticon_dict`, serta mengonversi emoji menjadi representasi teks menggunakan `emoji.demojize()`.
- b. *Basic Text Processing*:
 - *Case Folding* yaitu proses mengubah semua huruf dalam teks menjadi huruf kecil (lowercase) untuk menghindari perbedaan antara huruf besar dan kecil yang dapat mempengaruhi analisis.
 - *Cleaning* yaitu proses menghilangkan karakter yang tidak diperlukan seperti tanda baca, simbol, atau kata-kata yang tidak relevan.
 - Hapus URL yaitu menghapus alamat situs web (URL) seperti yang dimulai dengan "http" atau "https" dari teks, agar teks hanya berfokus pada kata-kata dan bukan tautan.
 - Hapus *mention* yaitu menghapus referensi akun pengguna (misalnya, "@username") untuk fokus pada konten teks.
 - Hapus angka yaitu menghapus semua angka (baik numerik atau tertulis) agar tidak mengganggu analisis kata.
- c. *Normalisasi* yaitu proses mengubah kata-kata yang tidak baku atau singkatan menjadi bentuk standar.
- d. *Stopword* yaitu proses menghilangkan kata-kata yang sering muncul tetapi tidak memiliki makna signifikan dalam analisis (seperti "dan", "atau", "yang").
- e. *Stemming* yaitu proses menghilangkan imbuhan (*prefix*, *suffix*, atau *infix*) dari suatu kata sehingga kata tersebut dikembalikan ke bentuk dasarnya.

Tabel 2. Tahapan Preprocessing

Langkah	Sebelum	Sesudah
Emoji Handling		Selamat buat Kaesang #ketumpsi yg anak muda di PSI Jakarta 1 senyum mempersiapkan produk baru untuk memperkenalkan ke pasar, colek @grace23 @psi https://t.co/qVeuDrjqV7
Basic Text Processing	Selamat buat Kaesang #ketumpsi yg anak muda di PSI Jakarta 1 :) mempersiapkan produk baru untuk	selamat buat kaesang yg anak muda di psi jakarta senyum mempersiapkan produk baru untuk memperkenalkan ke pasar colek
Normalisasi	memperkenalkan ke pasar, colek @grace23 @psi	selamat buat kaesang yang anak muda di psi jakarta senyum mempersiapkan produk baru untuk memperkenalkan ke pasar colek
Stopword	https://t.co/qVeuDrjqV7	selamat kaesang anak muda psi jakarta senyum produk memperkenalkan pasar colek
Stemming		selamat kaesang anak muda psi jakarta senyum produk kenal pasar colek

2.5. Bidirectional Encoder Representations from Transformers (BERT)

Bidirectional Encoder Representations from Transformers (BERT) adalah model pembelajaran mesin yang dikembangkan oleh Google pada tahun 2018. Model ini dirancang

untuk memahami konteks dalam teks dengan cara menganalisis kata-kata secara dua arah (bidirectional). BERT menggunakan arsitektur Transformer, yang memungkinkan model untuk menangkap hubungan antara kata-kata dalam suatu kalimat dengan lebih baik[15]. Salah satu keunggulan utama BERT adalah kemampuannya untuk melakukan pre-training pada sejumlah besar data teks, sehingga model ini dapat memahami berbagai nuansa bahasa. Setelah pre-training, BERT dapat disesuaikan (fine-tuned) untuk tugas-tugas spesifik seperti klasifikasi teks, analisis sentimen, atau pemahaman pertanyaan. Hal ini membuat BERT sangat fleksibel dan efektif dalam berbagai aplikasi pemrosesan bahasa alami (NLP)[16].

Pada penelitian ini, BERT digunakan untuk melakukan embedding tweet dalam rangka analisis sentimen terhadap data yang diperoleh dari platform Twitter. Data yang telah dibersihkan dan diproses melalui tahap *text preprocessing* kemudian dimasukkan ke dalam model BERT untuk mendapatkan representasi vektor dari teks yang lebih bermakna. Model BERT yang digunakan dalam penelitian ini adalah indolem/indobertweet-base-uncased, sebuah model yang telah dilatih sebelumnya dengan data bahasa Indonesia, khususnya untuk konteks tweet. Langkah pertama dalam proses ini adalah tokenisasi teks, di mana setiap tweet diubah menjadi format yang dapat dipahami oleh model. Tokenisasi ini menghasilkan dua elemen utama yaitu `input_ids` yang berisi urutan token untuk setiap tweet dan `attention_mask` yang menunjukkan token yang relevan untuk diproses oleh model.

Selanjutnya, setelah proses tokenisasi, model BERT digunakan untuk menghasilkan embedding tweet. Guna mendapatkan representasi vektor yang lebih ringkas dari tweet, dilakukan mean pooling, di mana rata-rata dari embedding token-token dalam setiap tweet dihitung. Ini memberikan sebuah vektor representasi untuk setiap tweet yang merefleksikan makna keseluruhan kalimat. Setelah embedding diperoleh, hasilnya diubah menjadi format numpy arrays untuk memudahkan penggunaan dalam algoritma klasifikasi sentimen berikutnya. Dengan menggunakan embedding BERT, model dapat memahami makna dan hubungan antar kata dalam tweet, serta meningkatkan akurasi dalam analisis sentimen.

2.6. Klasifikasi Algoritma Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Prinsip dasar dari SVM adalah mencari *hyperplane* terbaik yang memisahkan kelas-kelas data dengan margin terbesar. Dalam konteks klasifikasi, SVM berusaha untuk menempatkan sebuah *hyperplane* di antara dua kelas data, sehingga jarak dari titik data ke *hyperplane* tersebut menjadi maksimal[17].

Pada penelitian ini, *Support Vector Machine* (SVM) digunakan untuk melakukan klasifikasi sentimen pada tweet yang telah diproses menggunakan model BERT. Setelah mendapatkan embedding dari tweet dengan BERT, hasil embedding tersebut digunakan sebagai fitur untuk melatih model SVM. Data yang telah diproses dalam bentuk vektor numerik ini akan dipisahkan ke dalam tiga kategori sentimen: positif, negatif, dan netral. Algoritma SVM kemudian mencari *hyperplane* yang memisahkan tweet dengan sentimen positif, negatif, dan netral dengan margin terbesar. Kemudian dilanjutkan dengan proses klasifikasi dengan tuning *hyperparameter* menggunakan teknik *GridSearchCV* untuk mencari kombinasi parameter terbaik seperti nilai C, kernel, dan gamma. Hasil terbaik dari pencarian grid tersebut digunakan untuk melatih model SVM yang kemudian diuji menggunakan data validasi.

2.7. Evaluasi

Evaluasi dalam konteks model *machine learning* adalah proses untuk menilai kinerja model setelah dilatih menggunakan data pelatihan. Tujuan evaluasi adalah untuk mengukur seberapa baik dapat menggeneralisasi data yang belum pernah dilihat sebelumnya (data validasi atau data uji)[18]. Evaluasi dilakukan dengan membandingkan hasil prediksi model dengan label atau kelas yang sebenarnya pada dataset yang terpisah, yang tidak digunakan selama pelatihan. Dalam hal ini, evaluasi penting untuk memastikan bahwa model tidak hanya "mengingat" data pelatihan atau dikenal dengan *overfitting*, tetapi juga dapat bekerja dengan baik pada data baru yang belum dikenal[19].

Evaluasi lebih lanjut dilakukan dengan menghasilkan classification report, yang memberikan metrik rincian seperti *precision*, *recall*, dan *f1-score* (persamaan (1)) untuk masing-masing kelas sentimen (positif, netral, negatif), sebagaimana diterangkan di dalam Yohana dkk[20]. Score final, dihitung berdasarkan nilai F1 score rata-rata dari ketiga kelas, sebagaimana persamaan (2).

$$F1_{class} = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall} \right)_{class} \quad (1)$$

$$F1_{ave} = \frac{F1_{pos} + F1_{neg} + F1_{net}}{3} \quad (2)$$

3. HASIL DAN PEMBAHASAN

Pada bagian ini, disajikan hasil eksperimen yang mencakup eksperimen *setup*, pengujian terhadap data uji serta perbandingan pengujian untuk mengevaluasi kinerja model klasifikasi sentimen yang dikembangkan menggunakan kombinasi algoritma *Support Vector Machine* (SVM) dan *Bidirectional Encoder Representations from Transformers* (BERT).

3.1. Eksperimen Setup

Pada penelitian ini, eksperimen dilakukan untuk mengoptimalkan performa klasifikasi sentimen pada *platform* X dengan menggunakan algoritma SVM yang dikombinasikan dengan fitur BERT. Proses eksperimen dimulai dengan tahap eksplorasi fitur, dengan metode ekstraksi fitur IndoBERT, yang memiliki keunggulan dalam mengolah data teks dalam bahasa Indonesia. Eksperimen ini bertujuan untuk mengevaluasi kemampuan model SVM dalam mengklasifikasikan sentimen dengan berbagai representasi fitur, serta melihat seberapa besar kontribusi masing-masing metode terhadap peningkatan akurasi.

Selanjutnya, dilakukam eksperimen penerapan model *baseline* yang digunakan tanpa optimasi tertentu dengan menggunakan dataset Kaesang V1 sebanyak 300 tweet dan Kaesang V2 sebanyak 240 tweet data train dan 60 tweet data validasi serta kombinasi dataset Kaesang V1 dan Kaesang V2 sebanyak 480 tweet dan 120 tweet data validasi. Untuk penambahan data yaitu menggabungkan data training terbatas dengan data eksternal Covid yang ditambah dengan kelipatan 100 tweet untuk setiap label.

Tabel 3. Hasil Penelitian Model Optimal

ID	Komposisi Data		Text Preprocessing					F1-Score (data Val)
	Train	Val	Basic Text Processing	Emoji Handling	Normalisasi	Stopword	Stemming	
A1	Ks v1 : 300 Ks v2 : 240 Cov 19 : 0	60	iya	tidak	tidak	tidak	tidak	0.71
A2	Ks v1 : 300 Ks v2 : 240 Cov 19 : 100	60	iya	iya	iya	iya	iya	0.66
A3	Ks v1 : 300 Ks v2 : 240	60	iya	iya	tidak	iya	tidak	0.70

	Cov 19 : 200							
A4	Ks v1 : 300 Ks v2 : 240 Cov 19 : 300	60	iya	iya	iya	iya	tidak	0.69
A5	Ks v1+v2 : 480 Cov 19 : 0	120	iya	iya	iya	iya	tidak	0.65
A6	Ks v1+v2 : 480 Cov 19 : 200	120	iya	iya	tidak	iya	tidak	0.65

Berdasarkan tabel 3 terdapat 6 eksperimen yang diambil dari beberapa eksperimen terbaik yang dilihat dari hasil F1-Score data validasi. Eksperimen A1 menggunakan 300 tweet dari Kaesang V1 dan 240 tweet dari Kaesang V2, tanpa penambahan data eksternal, dan menghasilkan F1-Score tertinggi yaitu 0.71. Sementara itu, eksperimen A2, A3, dan A4 menambahkan data eksternal *Covid-19* dengan agregasi yang berbeda yaitu menambahkan masing-masing 100, 200, dan 300 tweet dari *Covid-19*. Hasilnya menunjukkan F1-Score masing-masing sebesar 0.66, 0.70, dan 0.69. Eksperimen A5 dan A6 menggunakan gabungan 480 tweet dari Kaesang V1 dan V2, dengan A6 menambahkan 200 tweet dari *Covid-19*. Hasil eksperimen A5 dan A6 menghasilkan F1-Score 0.65 dan 0.65.

3.2. Pengujian Terhadap Data Uji

Setelah eksperimen setup selesai dilakukan, langkah selanjutnya adalah pengujian terhadap data uji yang terdiri dari 924 tweet. Pada tahap ini, model klasifikasi yang telah dikembangkan dengan menggunakan algoritma *Support Vector Machine* (SVM) dan fitur BERT dievaluasi untuk mengukur kemampuannya dalam mengklasifikasikan sentimen secara akurat melalui sistem *leaderboard*. Hasil pengujian disajikan pada tabel berikut:

Tabel 4. Pengujian Data Uji

ID	Run	Data Validasi		Data Test	
		F1	Akurasi	F1	Akurasi
A1	Run 1	0.71	0.72	0.50	0.57
A5	Run 2	0.65	0.65	0.53	0.61
A6	Run 3	0.65	0.65	0.52	0.60

Berdasarkan Tabel 4, terdapat 3 eksperimen yang telah di submit ke sistem *leaderboard* dengan menggunakan data uji yang terdiri dari 924 tweet. Setiap eksperimen memiliki hasil yang berbeda dalam hasil F1-Score dan akurasi, yang mencerminkan seberapa baik model dalam mengklasifikasikan sentimen tweet pada data uji. Eksperimen A1 menggunakan data Kaesang V1 dan V2 tanpa penambahan data eksternal, menghasilkan F1-Score sebesar 0.50 dan akurasi 0.57 pada data uji. Meskipun eksperimen ini menunjukkan hasil terbaik pada data validasi F1-Score 0.71 dan akurasi 0.72, performanya menurun saat diuji pada data uji. Eksperimen A5 merupakan eksperimen terbaik, yang menggabungkan data Kaesang V1 dan V2, memberikan hasil F1-Score 0.53 dan akurasi 0.61 pada data uji, menunjukkan hasil yang lebih baik dibandingkan eksperimen A1. Sementara itu, eksperimen A6, yang menggunakan data Kaesang V1 dan V2 bersama dengan 200 tweet dari data eksternal *Covid-19*, memperoleh F1-Score 0.52 dan akurasi 0.60, yang menunjukkan sedikit peningkatan dibandingkan eksperimen A1 pada data uji.

3.3. Perbandingan Hasil Pengujian

Setelah dilakukan pengujian terhadap data uji melalui sistem *leaderboard*, selanjutnya hasil pengujian dilakukan perbandingan dengan beberapa eksperimen dalam sistem *leaderboard*. Proses ini menghitung nilai F1-Score, *Accuracy*, *Precision* dan *Recall* sebagai skor evaluasi. F1-Score menjadi acuan utama dalam evaluasi performa model karena metrik ini menggabungkan *Precision* dan *Recall*.

Tabel 5. Perbandingan Hasil Leaderboard

Tim	Metode	F1-Score	Accuracy	Precision	Recall
Rank 1	BERT	0.60	0.70	0.59	0.67
Penelitian Ini	SVM + BERT	0.53	0.61	0.54	0.63
Admin	Baseline	0.40	0.45	0.49	0.48

Berdasarkan perbandingan pada Tabel 5, dapat dilihat bahwa model yang dikembangkan dalam penelitian ini, yaitu kombinasi SVM dengan fitur BERT, memiliki performa yang baik meskipun tidak sebaik model BERT yang memperoleh nilai F1-Score 0.60 dan akurasi 0.70. Perbandingan ini menunjukkan bahwa BERT memiliki keunggulan dalam pengolahan konteks bahasa, yang memberikan F1-Score dan akurasi yang lebih tinggi. Meskipun demikian, kombinasi SVM dan BERT masih dapat memberikan hasil yang cukup baik dengan F1-Score 0.53 dan akurasi 0.61. Sementara itu, model *baseline* (admin), yang tidak menggunakan optimasi atau teknik ekstraksi fitur, menunjukkan performa yang lebih rendah, dengan F1-Score 0.40 dan akurasi 0.45. Hal ini menunjukkan bahwa penggunaan metode yang lebih canggih, seperti BERT, memberikan peningkatan yang signifikan pada klasifikasi sentimen di Twitter, meskipun dengan dataset terbatas.

3.4. Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi metode *Support Vector Machine* (SVM) dengan fitur *embedding* dari BERT mampu meningkatkan performa klasifikasi sentimen pada dataset terbatas. Pendekatan ini menghasilkan F1-score sebesar 0.53 dan akurasi 0.61, menunjukkan peningkatan yang signifikan dibandingkan *baseline* [12] yang hanya mencapai F1-score 0.40 dan akurasi 0.45. Selain itu, dibandingkan dengan metode SVM murni yang digunakan oleh penelitian sebelumnya [12] yang menghasilkan F1-score 0.51, pendekatan dalam penelitian ini menunjukkan adanya kontribusi nyata dari pemanfaatan representasi kontekstual berbasis BERT. Bahkan dibandingkan dengan kombinasi SVM dengan Word2Vec [21] dan FastText [8], metode ini menghasilkan nilai presisi dan *recall* yang lebih tinggi, mengindikasikan kemampuan yang lebih baik dalam menangkap nuansa sentimen. Meskipun belum melampaui performa model BERT [22] dan BiLSTM IndoBERT [23], pendekatan ini menawarkan efisiensi komputasi yang lebih baik karena menggunakan model klasifikasi yang lebih ringan, sekaligus mempertahankan performa yang kompetitif. Penelitian ini juga menguji berbagai cara dalam tahap *preprocessing*, termasuk teknik lanjutan seperti penanganan emoji dan *stemming*, yang terbukti mampu meningkatkan akurasi model secara nyata. Pendekatan ini mengisi kekurangan dari penelitian sebelumnya yang belum banyak membahas pentingnya tahap *preprocessing* dalam proses klasifikasi sentimen untuk bahasa Indonesia. Oleh karena itu, penelitian ini tidak hanya membuktikan bahwa *contextual embedding* seperti BERT efektif, tetapi juga memberikan kontribusi tambahan melalui pemanfaatan fitur bahasa lokal dan penerapan *preprocessing* yang lebih menyeluruh.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa kombinasi algoritma *Support Vector Machine* (SVM) dengan fitur BERT dapat meningkatkan performa klasifikasi sentimen pada dataset terbatas, khususnya dalam menganalisis opini publik terkait pengangkatan Kaesang Pangarep sebagai Ketua Umum PSI. Meskipun model BERT yang digunakan secara terpisah menunjukkan performa lebih tinggi dengan F1-Score 0.60 dan akurasi 0.70, kombinasi SVM dan BERT tetap memberikan hasil yang signifikan dengan F1-Score 0.53 dan akurasi 0.61. Hasil ini mengungguli model *baseline* yang hanya mencapai F1-Score 0.40 dan akurasi 0.45, serta model SVM tanpa BERT yang mencapai F1-Score 0.51 dan akurasi 0.61. Distribusi sentimen menunjukkan kecenderungan opini publik terhadap pengangkatan Kaesang Pangarep sebagai Ketua Umum PSI

lebih banyak bersifat netral sebanyak 411 data, diikuti oleh sentimen positif 320 data, dan negatif 193 data. Hal ini menunjukkan bahwa sebagian besar masyarakat menunjukkan sikap yang tidak terlalu berpihak, meskipun opini positif cenderung lebih dominan dibanding opini negatif.

DAFTAR PUSTAKA

- [1] I. N. N. Fauziah, S. A. Saputri, and Y. T. Herlambang, 'Teknologi Informasi: Dampak Media Sosial pada Perubahan Sosial Masyarakat', *Indo-MathEdu Intellectuals Journal*, vol. 5, no. 1, pp. 757–766, Jan. 2024, doi: 10.54373/imeij.v5i1.645.
- [2] M. Mesran and D. P. Indini, 'Analisis Dalam Pendukung Keputusan Seleksi Content Creator Mahasiswa Terbaik Menerapkan Metode EDAS dan ROC', *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 4, pp. 912–921, Aug. 2023, doi: 10.47065/josyc.v4i4.4093.
- [3] O. Fontanelli, D. Hernández, and R. Mansilla, 'Modeling the popularity of twitter hashtags with master equations', *Soc Netw Anal Min*, vol. 12, no. 1, Dec. 2022, doi: 10.1007/s13278-022-00861-4.
- [4] Y. El Saputra, S. Agustian, and S. Ramadhani, 'KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Sentimen SVM Dengan Dataset yang Kecil Pada Kasus Kaesang Sebagai Ketua Umum PSI', *Media Online*, vol. 4, no. 6, pp. 2902–2908, 2024, doi: 10.30865/klik.v4i6.1944.
- [5] T. Y. Pahtoni and H. Jati, 'Analisis sentimen data Twitter terkait ChatGPT menggunakan Orange Data Mining', *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 329–336, Apr. 2024, doi: 10.25126/jtiik.20241127276.
- [6] K. H. Siregar and R. Rahmawati, 'Kaesang's Coat-Tail Effect: Influence and Implications for PSI's Political Branding in the 2024 Election', *Jurnal Ilmu Sosial Indonesia (JISI)*, vol. 5, no. 2, pp. 153–167, Dec. 2024, doi: 10.15408/jisi.v5i2.43589.
- [7] M. Arsyad, 'Analisis sentimen linguistik komputasional: memahami emosi dalam teks', 2023, [Online]. Available: <https://www.researchgate.net/publication/376722063>
- [8] S. Agustian and A. Nazir, 'Klasifikasi Sentimen Terhadap Pengangkatan Kaesang Sebagai Ketua Umum Partai PSI Menggunakan Metode Support Vector Machine', *Technology and Science (BITS)*, vol. 6, no. 1, 2024, doi: 10.47065/bits.v6i1.5340.
- [9] I.-A. Albu and S. Spînu, 'Emotion detection from tweets using a BERT and SVM ensemble model', *U.P.B. Sci. Bull., Series C*, vol. 84, no. 1, 2022, [Online]. Available: <https://pypi.org/project/demoji>
- [10] B. Prasetyo, Ahmad Yusuf Al-Majid, and Suharjito, 'A Comparative Analysis of MultinomialNB, SVM, and BERT on Garuda Indonesia Twitter Sentiment', *PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic*, vol. 12, no. 2, pp. 445–454, Sep. 2024, doi: 10.33558/piksel.v12i2.9966.
- [11] S. Forhad and A. H. Siddique, 'Analyzing TF-IDF and BERT Approach for Bangla Text Classification Using Transformer-Based Embedding for Newspaper Sentiment Classification'. [Online]. Available: <https://www.researchgate.net/publication/384888117>
- [12] S. Agustian *et al.*, 'New Directions in Text Classification Research: Maximizing The Performance of Sentiment Classification from Limited Data Arah Baru Penelitian Klasifikasi Teks: Memaksimalkan Kinerja Klasifikasi Sentimen dari Data Terbatas', 2024. Accessed: Feb. 10, 2025. [Online]. Available: <https://arxiv.org/abs/2407.05627>
- [13] A. Gunawan, E. Altirika, S. Pratama, and Y. B. Pratama, 'Pengembangan Aplikasi Asisten Virtual Menggunakan Machine Learning Berbasis Mobile Untuk Meningkatkan Pelayanan Kampus di Muhammadiyah Bangka Belitung', 2024. [Online]. Available: <https://jits.unmuhbabel.ac.id/>
- [14] A. Addiga and S. Bagui, 'Sentiment Analysis on Twitter Data Using Term Frequency-Inverse Document Frequency', *Journal of Computer and Communications*, vol. 10, no. 08, pp. 117–128, 2022, doi: 10.4236/jcc.2022.108008.

- [15] A. Amelia and R. Yusuf, 'Analisis sentimen masyarakat Indonesia pada platform X terhadap isu fufufafa menggunakan Bidirectional Encoder Representations from Transformers', *Jurnal Informatika Teknologi dan Sains*, vol. 7, no. 1, pp. 72–80, 2025.
- [16] K. Setyo Nugroho, A. Yullian Sukmadewa, H. D. Wuswilahaken, F. Abdurrachman Bachtiar, and N. Yudistira, 'BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews', 2021. Accessed: Feb. 17, 2025. [Online]. Available: <https://arxiv.org/pdf/2107.06802>
- [17] G. Sanhaji, A. Febrianti, and F. Teknik, 'Aplikasi DIATECT Untuk Prediksi Penyakit Diabetes Menggunakan SVM Berbasis Web', *Jurnal Tekno Kompak*, vol. 18, no. 1, pp. 150–163, 2024.
- [18] H. Hendra, 'Prediksi Daya Listrik Pada Pembangkit Listrik Siklus Gabungan Berdasarkan Kondisi Lingkungan Menggunakan Metode Machine Learning', *Elektron : Jurnal Ilmiah*, pp. 72–82, Dec. 2023, doi: 10.30630/eji.0.0.415.
- [19] Y. Prasetyo *et al.*, 'Analisis perbandingan optimizer SGD dan ADAM pada model CNN untuk klasifikasi jamur edible dan poisonous', 2025.
- [20] P. Yohana, S. Agustian, and S. Kurnia Gusti, 'Klasifikasi Sentimen Masyarakat terhadap Kebijakan Vaksin Covid-19 pada Twitter dengan Imbalance Classes Menggunakan Naive Bayes', 2022. [Online]. Available: <https://lp2m.unmul.ac.id/webadmin/public/upload/files/9584b64517cfe308eb6b115847cbe8e7.pdf>
- [21] B. Cahya, 'Implementasi WORD2VEC Sebagai Fitur Pada Algoritma SVM Dalam Klasifikasi Sentimen (Studi Kasus Tweet Kaesang Sebagai KETUM PSI)', Universitas Islam Negeri Sultan Syarif Kasim Riau, 2024.
- [22] J. Pranata, S. Agustian, and E. Haerani, 'Penggunaan Model Bahasa indoBERT pada Metode Random Forest Untuk Klasifikasi Sentimen Dengan Dataset Terbatas', *Technology and Science (BITS)*, vol. 6, no. 3, pp. 1668–1676, 2024, doi: 10.47065/bits.v6i3.6335.
- [23] R. Illahi, S. Agustian, F. Yanto, P. H. Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau Jl Soebrantas, S. Baru, and K. Pekanbaru, 'Klasifikasi sentimen menggunakan bidirectional LSTM dan IndoBERT dengan dataset terbatas.', 2025.



ZONAsi: Jurnal Sistem Informasi

Is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)