PENINGKATAN KINERJA ALGORITMA SUPPORT VECTOR MACHINE MELALUI AUGMENTASI SWAP WORD UNTUK KLASIFIKASI SENTIMEN ULASAN PENGGUNA APLIKASI BMKG

Riska Amalia, Heliawati Hamrul, Nurhikma Arifin
Teknik Informatika, Universitas Sulawesi Barat
Jalan Prof. Dr. Baharuddin Lopa, S.H., Talumung, Majene, Sulawesi Barat
ra8646493@gmail.com

ABSTRAK

Ulasan pengguna aplikasi BMKG di *Google Play Store* berpotensi besar digunakan sebagai sumber evaluasi kualitas layanan berbasis persepsi pengguna. Namun, pemanfaatan ulasan tersebut melalui klasifikasi sentimen masih menghadapi permasalahan utama, yaitu ketidakseimbangan distribusi kelas sentimen, di mana jumlah ulasan dengan sentimen tertentu jauh lebih dominan dibandingkan kelas lainnya. Kondisi ini menyebabkan algoritma klasifikasi cenderung bias terhadap kelas mayoritas, menurunkan kemampuan model dalam mengenali sentimen minoritas, serta berdampak langsung pada rendahnya kinerja dan akurasi klasifikasi. Penelitian ini bertujuan untuk meningkatkan kinerja algoritma *Support Vector Machine* (SVM) dalam klasifikasi sentimen ulasan pengguna aplikasi BMKG melalui penerapan teknik *augmentasi* data *Swap Word*. Metode penelitian meliputi pengumpulan data sebanyak 1.691 ulasan pengguna, *preprocessing* teks, ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF), serta pembangunan model klasifikasi berbasis SVM. Untuk mengatasi ketidakseimbangan kelas, diterapkan teknik *augmentasi Swap Word* dengan menukar posisi kata secara acak tanpa mengubah makna kalimat. Hasil penelitian menunjukkan bahwa penerapan *augmentasi Swap Word* meningkatkan akurasi model dari 76% menjadi 87%. Kesimpulannya, *augmentasi Swap Word* efektif dalam mengatasi ketidakseimbangan kelas dan meningkatkan kinerja klasifikasi sentimen ulasan pengguna aplikasi BMKG.

Kata kunci : *Support Vector Machine, Analisis Sentimen, Augmentasi, Textblob, Swap Word, BMKG*

1. PENDAHULUAN

BMKG, sebagai lembaga negara, memiliki tugas yang sangat penting dalam penyediaan data meteorologi, klimatologi, dan geofisika. Tingginya tingkat kerawanan bencana di wilayah Indonesia menempatkan lembaga tersebut pada posisi yang sangat penting. Akurasi data yang dirilis BMKG bukan hanya kebutuhan teknis, melainkan syarat mutlak untuk menunjang mitigasi risiko dan kelancaran aktivitas harian masyarakat [1] .

Untuk memudahkan akses masyarakat terhadap informasi tersebut, BMKG telah mengembangkan aplikasi *mobile,* yang dapat diunduh melalui platform termasuk *Play Store* untuk pengguna perangkat *Android*. Aplikasi ini dirancang untuk memberikan layanan informasi secara *real-time* mengenai prakiraan cuaca, peringatan dini bencana, gempa bumi, dan informasi lainnya berkaitan dengan meteorologi, klimatologi dan geofisika.

Ulasan pengguna tersebut berpotensi besar dimanfaatkan sebagai sumber evaluasi kualitas layanan melalui analisis sentimen. Namun, pemanfaatannya masih menghadapi permasalahan utama, yaitu data ulasan yang tidak terstruktur dan memiliki distribusi kelas sentimen yang tidak seimbang. Ketidakseimbangan jumlah data pada masing-masing kelas sentimen menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas, sehingga menurunkan kemampuan model dalam mengenali sentimen dengan jumlah data yang lebih sedikit. Selain itu, keterbatasan variasi data pelatihan dapat menyebabkan model klasifikasi kurang mampu melakukan generalisasi terhadap data baru.

Dalam analisis sentimen, *Support Vector Machine* (SVM) menjadi salah satu metode *machine learning* yang paling banyak digunakan. Tingginya frekuensi penggunaan algoritma ini disebabkan oleh kapabilitasnya yang unggul dalam klasifikasi teks. SVM terbukti andal dalam menangani data dengan dimensi tinggi serta mampu memberikan generalisasi yang presisi pada data uji. Secara teknis, algoritma ini bekerja dengan mencari *hyperplane* optimal yang memaksimalkan *margin* pemisah antar kelas. Karakteristik inilah yang menjadikan SVM sangat efektif untuk menyelesaikan tugas klasifikasi teks, termasuk dalam memproses ulasan aplikasi [2].

Augmentasi teks merupakan metode dalam *deep learning* yang bertujuan memperbesar jumlah serta memperkaya variasi data pelatihan sehingga dapat meminimalkan terjadinya *overfitting*, terutama pada kondisi data yang sedikit atau tidak seimbang. Pada bidang NLP, teknik ini dilakukan dengan memodifikasi teks asli untuk menghasilkan variasi baru tanpa mengubah makna atau label semantiknya. Salah satu teknik yang digunakan adalah *swap word* (penukaran kata), yang terbukti mampu meningkatkan kinerja model klasifikasi teks, khususnya ketika data pelatihan yang tersedia terbatas [3].

Penelitian ini bertujuan untuk mengetahui bagaimana teknik *Augmentasi* data dapat membantu meningkatkan kinerja model dalam mengklasifikasikan sentimen ulasan pengguna

aplikasi BMKG. Data yang digunakan pada penelitian ini adalah data ulasan pengguna aplikasi BMKG dari *play store* dari rentang waktu September 2024 sampai September 2025, kemudian menerapkan tahapan *preprocessing* teks untuk meningkatkan kualitas data. Selanjutnya, data dipresentasikan dalam bentuk numerik menggunakan metode pembobotan TF-IDF dan diklasifikasikan menggunakan algoritma *Support Vector Machine*. Untuk mengatasi ketidakseimbangna distribusi kelas sentimen, diterapkan teknik *augmentasi* data *swap word* yang bertujuan menambah variasi data tanpa mengubah makna teks. Kinerja model dievaluasi menggunakan metrik akurasi, *precision, racll,* dan *f1-score.*

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Terdapat berbagai penelitian sebelumnya yang telah mengkaji penerapan algoritma Support Vector Machine (SVM) dalam konteks analisis sentimen, di antaranya meliputi:

Penelitian yang mengkaji analisis sentimen masyarakat Indonesia terkait Pemilihan Presiden 2024 dengan menggunakan algoritma *Support Vector Machine* menunjukkan bahwa metode tersebut mampu melakukan klasifikasi terhadap sentimen pada cuitan mengenai Pilpres 2024. Hasil pengujian menunjukkan bahwa algoritma Support Vector Machine mencapai tingkat akurasi sebesar 65%, nilai *recall* sebesar 81%, serta *f1-score* sebesar 74%[4].

Penelitian yang melakukan perbandingan antara Penelitian yang membandingkan kinerja *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) pada analisis sentimen menunjukkan bahwa SVM memiliki performa yang lebih unggul. Hasil eksperimen memperlihatkan SVM mampu mencapai tingkat akurasi 70%, melampaui capaian algoritma KNN yang hanya memperoleh 56%.[5].

Studi yang mengkaji persepsi masyarakat terhadap fenomena pinjaman daring (*online*) di platform Twitter memperlihatkan kinerja positif dari algoritma *Support Vector Machine* (SVM). Dalam pengujian klasifikasi tersebut, diperoleh tingkat akurasi sebesar 62%. Selain itu, analisis distribusi kelas menunjukkan kecenderungan sentimen negatif yang lebih tinggi, yakni 59%, berbanding 41% untuk sentimen positif [6].

Penelitian yang membahas perbandingan kinerja algoritma *Support Vector Machine* dan *Naïve Bayes* dalam klasifikasi sentimen menunjukkan bahwa algoritma *Support Vector Machine* memiliki performa yang lebih baik dibandingkan dengan *Naïve Bayes*. Berdasarkan hasil pengujian, *Support Vector Machine* memperoleh tingkat akurasi sebesar 52%, nilai presisi 58%, *recall* 55%, serta *f1-score* 52%. Sementara itu, algoritma Naïve Bayes hanya mencapai tingkat akurasi sebesar 47%, dengan nilai presisi 58%, *recall* 55%, dan *f1-score* 52%.[7],

Penelitian yang mengkaji perbandingan algoritma *Naïve Bayes* dan *Support Vector Machine* dalam analisis sentimen pengguna metaverse menunjukkan bahwa algoritma *Support Vector Machine* memiliki kinerja yang lebih baik. Hasil pengujian menunjukkan bahwa *Support Vector Machine* mencapai tingkat akurasi sebesar 78% dengan nilai Macro-F1 sebesar 72%, sedangkan algoritma *Naïve Bayes* memperoleh akurasi sebesar 72% dan nilai Macro-F1 sebesar 60% [8].

2.2. Analisis Sentimen

Analisis sentimen merupakan teknik yang digunakan dalam bidang *Natural Language Processing* (NLP) untuk mengidentifikasi dan mengklasifikasikan kecenderungan emosi yang terkandung dalam data berbentuk teks. Tujuan utama penerapan analisis sentimen adalah menentukan polaritas suatu teks, baik positif, negatif, maupun netral, berdasarkan informasi yang terdapat pada kata, kalimat, atau keseluruhan dokumen [4].

2.3. *TF-IDF*

Metode TF-IDF (*Term frequency-inverse Document Frequency*) diterapkan untuk merepresentasikan bobot kepentingan setiap kata dalam suatu dokumen. Pendekatan ini memungkinkan identifikasi kata-kata yang memiliki tingkat relevansi lebih tinggi terhadap keseluruhan kumpulan data, sehingga dapat meningkatkan kualitas representasi fitur dan mendukung ketepatan proses analisis sentimen.

Adapun perhitungan pembobotan TF-IDF ialah sebagai berikut:

$$\text{idf} = log\,\frac{N}{df} \quad (1)$$

$$\text{w, (k, d)} = tf(k,d) * idf \quad (2)$$

$$\text{w, (k, d)} = tf(k,d) * \frac{N}{df} \quad (3)$$

2.4. *Support Vector Machine*

Support Vector Machine adalah *hyperplane* atau garis pemisah, yang mana garis tersebut memiliki peran memisahkan kelas [2]. Persamaan perhitungan *hyperplane* sebagai berikut:

Hyperplane klasifikasi linear SVM dinotasikan dengan:

$$f(x) = w.x + b = 0 \quad (4)$$

$$w.x + b \leq +1 \quad (5)$$

$$w.x + h \geq -1 \quad (6)$$

Vektor w mempresentasikan bidang normal, sedangkan b menunjukkan posisi bidang relatif terhadap pusat koordinat. Dengan melakukan optimasi terhadap jarak antara *hyperplane* dan titik terdekat, dapat diperoleh *margin* maksimum yang dinyatakan sebagai yaitu 1 / $\|w\|$, di mana titik minimum persamaan (4) dengan mengingat syarat dari persamaan (5) tersebut.

$$min\,\frac{1}{2}\|\text{w}\|^2 = \min\frac{1}{2}(\text{w}1^2 + \text{w}2^2) \quad (7)$$

$$y_1\,(\text{w}_{x1} + \text{b}) \geq 1,\ i = 1,2,3\ldots, N \quad (8)$$

2.5. Evaluasi Model

Uji *confusion matrix* dilakukan untuk mendapatkan nilai *presisi*, *recall*, *f1-score* dan akurasi dari hasil performa [9].

a. *Accuracy*

Akurasi merupakan metrik yang digunakan untuk menilai seberapa tepat model melakukan prediksi secara keseluruhan. Persamaan yang digunakan untuk mendapat nilai akurasi sebagai berikut:

$$Accuracy = \frac{(TP+TN)}{(TP=TN+FP+FN)} X\ 100\% \quad (9)$$

b. *Precision*

Presisi (*precision*) didefinisikan sebagai metrik yang mengukur tingkat ketepatan data yang diklasifikasikan sebagai positif. Nilai ini merepresentasikan rasio antara prediksi positif yang benar (*true positive*) terhadap keseluruhan data yang diprediksi positif oleh model. Adapun formulasi matematis untuk menghitung presisi adalah sebagai berikut:

$$Precision = \frac{TP}{(FP+TP)} X\ 100\% \quad (10)$$

c. *Recall*

Recall merupakan ukuran yang menggambarkan kemampuan suatu model dalam mengidentifikasi seluruh data yang benar-benar termasuk ke dalam kelas positif. Metrik ini dihitung dengan membandingkan jumlah prediksi positif yang dilakukan secara tepat terhadap keseluruhan data positif yang sebenarnya.

$$Recall = \frac{TP}{(FN+TP)} X\ 100\% \quad (11)$$

d. F1-Score

F1-Score merupakan ukuran berupa rata-rata harmonis yang digunakan untuk mengevaluasi kemampuan model dalam mempertahankan keseimbangan antara nilai *precision* dan *recall* pada proses klasifikasi.

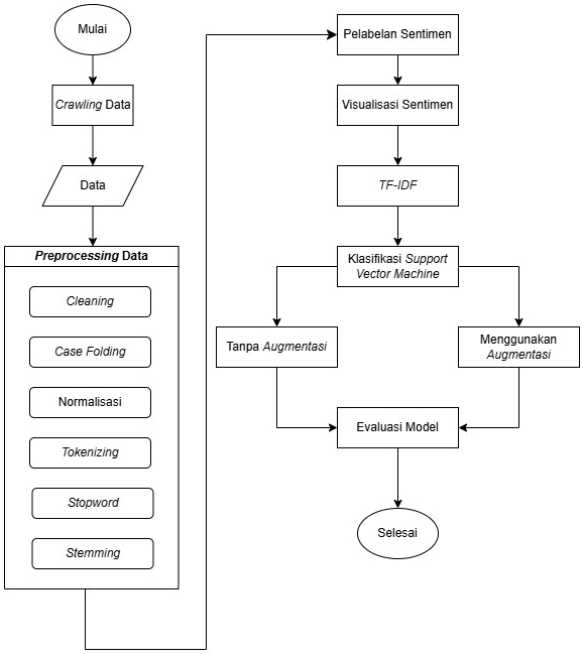
$$F1\text{-}Score = \frac{2(\text{Precision x recall})}{(Precision+recall)} \quad (12)$$

2.6. *Augmentasi*

Tujuan penerapan *augmentasi* dalam analisis sentimen adalah untuk meningkatkan ketahanan serta kemampuan generalisasi model dalam melakukan prediksi. Pada penelitian ini, *augmentasi* data dilakukan untuk mengatasi ketidakseimbangan kelas, yang dapat memengaruhi kinerja model klasifikasi teks berbasis *maching learning*, di mana kelas dengan jumlah data lebih sedikit (kelas minoritas) berisiko mengalami kesalahan klasifikasi. Hal ini pada akhirnya berdampak terhadap menurunnya performa keseluruhan model. *Augmentasi* dilakukan dengan menambah jumlah data pada kategori yang memiliki jumlah baris lebih sedikit [10].

3. METODE PENELITIAN

Rangkaian tahapan atau prosedur sistematis pada penelitian ini:



Gambar 1. Alur penelitian

3.1. Pengumpulan Data

Data utama yang digunakan dalam studi ini bersumber dari ulasan masyarakat terhadap aplikasi BMKG yang diunduh melalui layanan *Google play store* dari rentang waktu September 2024 sampai September 2025. Data dikumpulkan menggunakan *tool scrapping* dengan *library google-play-scraper*.

Data yang dikumpulkan dengan proses *scraping* sebanyak 2000 data ulasan. Data yang telah dikumpulkan akan melalui tahap *preprocessing*. Setelah melaui proses *preprocessing* data sebanyak 1691. Kemudian, dilakukan *Augmentasi* data dengan menggunakan teknik *swap word* data berhasil dikumpulkan sebanyak 2319 data.

3.2. *Preprocessing* Data

Setelah pengumpulan data, selanjutnya *preprocessing* data. *Preprocessing* data dilakukan dengan beberapa tahap yaitu:

3.3. *Cleaning Text*

Cleaning text dilakukan dengan menghapus karakter yang diperlukan seperti tanda baca, angka, tautan, hastag, *mention*, emoji dan lainnya. Tujuan dari tahapan ini adalah meminimalkan gangguan (*noise*) serta menghapus unsur-unsur yang tidak memiliki relevansi dalam data berbentuk teks [15].

Tabel 1. *Cleaning text*

Sebelum *Cleaning*	Sesudah Cleaning
sangat membantu kami sbagai relawan kemanusiaan… Terima kasih	sangat membantu kami sbagai relawan kemanusiaan Terima kasih

3.4. *Case Folding*

Case folding merupakan tahap penyamaan penggunaan huruf dalam dataset dengan cara mengubah seluruh teks menjadi huruf kecil [16].

Tabel 2. *Cace folding*

Sebelum *Case Folding*	Sesudah *Case Folding*
sangat membantu kami sbagai relawan kemanusiaan Terima kasih	sangat membantu kami sbagai relawan kemanusiaan terima kasih

3.5. Normalisasi

Tahap normalisasi berfungsi untuk mentransformasi istilah atau kata yang tidak memenuhi kaidah kebahasaan menjadi bentuk standar. Proses penyeragaman ini dilakukan dengan berpedoman pada aturan ejaan resmi yang termuat dalam Kamus Besar Bahasa Indonesia (KBBI) [17].

Tabel 3. Normalisasi

Sebelum Normalisasi	Sesudah Normalisasi
hbs update mlah sering keluar aplikasi sendiri	habis update malah sering keluar aplikasi sendiri

3.6. *Tokenizing*

Tokenizing merupakan tahap pemrosesan teks yang bertujuan untuk memisahkan setiap kata dalam sebuah kalimat menjadi urutan kata dapat dianalisis secara lebih terstruktur dan sistematis dalam penelitian [18].

Tabel 4. *Tokenizing*

Sebelum *Tokenizing*	Sesudah *Tokenizing*
sangat membantu kami sebagai relawan kemanusiaan terima kasih	[“sangat”, “membantu”, “kami” “sebagai”, “relawan” “kemanusiaan”, “terima”, “ kasih”]

3.7. *Stopword*

Stopwords merupakan tahapan dalam pengolahan teks yang bertujuan untuk menghilangkan kata-kata yang memiliki tingkat kebermaknaan rendah atau tidak memberikan kontribusi yang berarti terhadap proses analisis [16].

Tabel 5. *Stopword*

Sebelum *Stopwords*	Sesudah *Stopwords*
[“sangat”, “membantu”, “kami” “sebagai”, “relawan” “kemanusiaan”, “terima”, “ kasih”]	[“membantu”, “relawan”,”kemanusiaan”, “terima”,”kasih”]

3.8. *Stemming*

Stemming merupakan proses dalam pengolahan teks yang bertujuan untuk mengonversi token yang memiliki imbuhan menjadi bentuk kata dasar dengan menghilangkan seluruh imbuhan yang melekat pada token tersebut [19].

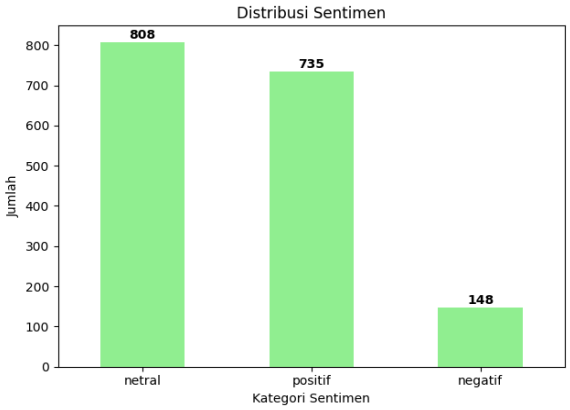
Tabel 6. *Stemming*

Sebelum *Stemming*	Sesudah *Stemming*
[“membantu”]	bantu

3.9. Pelabelan Sentimen

Pelabelan sentimen menggunakan *Textblob* merupakan tahapan pemberian kelas sentimen pada setiap komentar untuk menentukan kecenderungan opini yang terkandung di dalamnya, seperti sentimen positif, negatif dan netral. Salah satu pendekatan untuk melakukan pelabelan secara otomatis adalah memanfaatkan pustaka *Textblob*, yaitu sebuah *library* pemrosesan bahasa alami berbasis *Python* yang mampu mengidentifikasi polaritas sentimen teks secara otomatis [11].

Pada tahap ini teks dilabeli menggunakan pelabelan otomatis *Textblob*, label teks akan dikelompokkan menjadi 3 kategori sentimen, dilabeli dengan netral, positif, dan negatif [11].

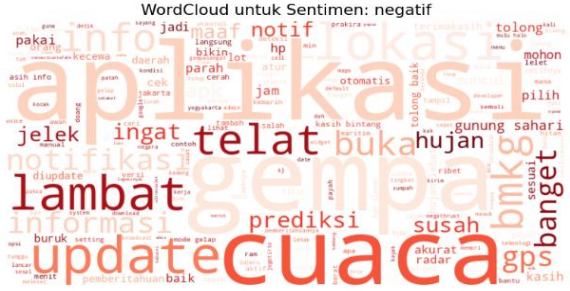


Gambar 2. Distribusi sentimen

Grafik pada Gambar 2 menampilkan distribusi sentimen pada data ulasan menunjukkan jumlah ulasan dalam tiga kategori sentimen. Berdasarkan gambar grafik di atas, sentimen netral lebih dominan dengan jumlah ulasan mencapai 808, sentimen positif dengan jumlah 735 ulasan dan sentimen negatif dengan jumlah 148 ulasan.

3.10. Visualisasi Sentimen

Visualisasi *wordcloud* dalam analisis sentimen merupakan bentuk penyajian grafis yang digunakan untuk menampilkan kata-kata yang memiliki tingkat kemunculan tertinggi dalam seuatu kumpulan teks. Setiap kata dipresentasikan dengan ukuran huruf yang bervariasi, di mana ukuran yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi dibandingkan kata-kata dengan ukuran yang lebih kecil [2].



Gambar 3. *Wordcloud* negatif

Pada gambar 3 menampilkan wordcloud sentimen negatif, kata yang sering muncul adalah "aplikasi", "cuaca", "update", "telat", "lambat", "banget", "prediksi", "jelek", "notif" dan "susah". Ini menunjukkan bahwa dalam sentimen negatif menandakan ketidakpuasan pengguna aplikasi disebabkan adanya bug setelah melakukan pembaruan aplikasi.

Gambar 4. *Wordcloud* positif

Pada gambar 4 menampilkan *wordcloud* sentimen positif kata yang sering muncul adalah "aplikasi", "akurat", "bagus", "gempa", "lokasi", "info", "bantu", "bmkg", dan "manfaat". Ini menunjukkan bahwa dalam sentimen positif mendandakan bahwa aplikasi bmkg membantu dan memudahkan pengguna untuk mengetahui info cuaca, bencana dan lainnya.

Gambar 5. *Wordcloud* netral

Pada gambar 5 menampilkan *wordcloud* sentimen netral, kata yang sering muncul adalah "aplikasi", "gempa", "bantu", "update", "info", "lokasi", "bmkg", "cuaca", "informasi" dan "notifikasi". Ini menunjukkan bahwa sentimen netral sebagian besar pengguna memberikan ulasan yang informatif dan deskriptif aplikasi bmkg.

3.11. TF-IDF

Tahapan yang sangat penting dalam pengolahan data teks adalah pembobotan kata (*term weighting*), yang bertujuan untuk memberikan nilai numerik pada setiap istilah (*term*) untuk menggambarkan bobot kepentingannya. Metode TF-IDF merupakan pendekatan yang paling umum diterapkan dalam tahap ini. Secara spesifik, metode ini membedakan antara *Term Frequency* (TF), yang mengukur frekuensi kehadiran sebuah kata dalam satu dokumen, dengan *Document Frequency* (DF), yang merepresentasikan banyaknya dokumen dalam koleksi (*corpus*) yang memuat kata tersebut [12].

3.12. *Support Vector Machine*

Keunggulan utama *Support Vector Machine* (SVM) terletak pada kapabilitasnya menangani data berdimensi tinggi, khususnya data teks, dengan performa yang konsisten meski pada dataset yang rumit. Secara teknis, SVM beroperasi dengan mengidentifikasi *hyperplane* optimal yang memberikan *margin* pemisah terlebar antar kategori. Mekanisme inilah yang menjadikan SVM metode yang sangat andal untuk klasifikasi teks, termasuk dalam studi analisis sentimen [13].

3.13. *Augmentasi*

Augmentasi pada penelitian ini dilakukan dengan memanfaatkan layanan *Google Translate API* yang diintegrasikan melalui bahasa pemrograman *Python*. *Augmentasi* data diterapkan untuk meningkatkan kemampuan generalisasi model dan mengurangi risiko *overfitting.* [14].

3.14. Evaluasi

Confusion matrix adalah *table* yang dibuat untuk menghubungkan hasil klasifikasi dengan data yang diperoleh untuk pengujian akurasi.

4. HASIL DAN PEMBAHASAN

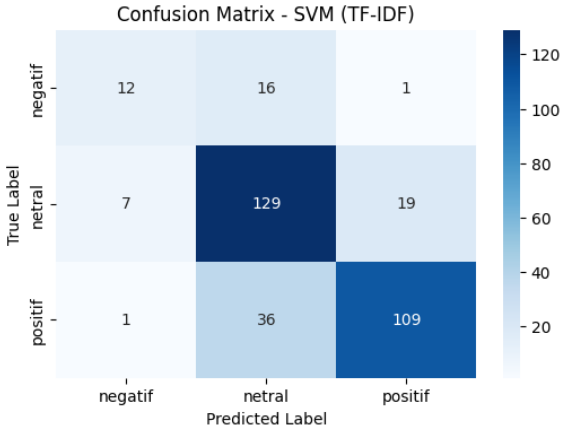
Klasifikasi model *Support Vector Machine* dilakukan pada penelitian ini menggunakan 2 pengujian yaitu menggunakan *Augmentasi* dan tanpa *Augmentasi*.

4.1. Klasifikasi Tanpa *Augmentasi*

Hasil dari klasifikasi dan evaluasi model *Support Vector Machine* tanpa *Augmentasi* dapat dilihat pada Tabel 8 dan Gambar 7.

Tabel 7. Svm *(tf-idf)*

Kategori	*Precision*	*Recall*	*F1-Score*	*Accuracy*
Negatif	60%	41%	49%	76%
Netral	71%	83%	77%	
Positif	84%	75%	79%	



Gambar 6. *Confusion matrix* tanpa *augmentasi*

Tabel 7 dan Gambar 6 menyajikan hasil evaluasi model yang menggunakan metode *Support Vector Machine* dengan pembobotan TF-IDF. Hasil pengujian menunjukkan bahwa model memperoleh tingkat akurasi sebesar 76%, yang mengindikasikan bahwa model memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen ulasan pengguna ke dalam tiga kategori, yaitu positif, negatif, dan netral

Dari hasil pengujian, kelas positif menunjukkan performa terbaik dengan *precision* sebesar 84%, *recall* 75% dan *f1-score* sebesar 79%, yang menandakan bahwa model mampu mengenali ulasan positif dengan tingkat ketepatan yang tinggi serta konsistensi yang baik. Pada kelas netral, model juga menunjukkan performa yang kuat dengan *precision* 71%, *recall* sebesar 83% dan f1-score 77%, yang berarti Sebagian besar ulasan netral berhasil diklasifikasikan dengan benar meskipun *precision*-nya sedikit lebih rendah dibanding kelas positif.

Di sisi lain, kinerja model pada kelas sentimen negatif masih menghadapi kendala, dengan nilai *precision* sebesar 60%, *recall* sebesar 41%, dan *f1-score* sebesar 49%. Kondisi ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan sebagian ulasan negatif dari kategori sentimen lainnya, yang kemungkinan disebabkan oleh jumlah data sentimen negatif yang relatif lebih sedikit atau keragaman penggunaan bahasa yang lebih kompleks.

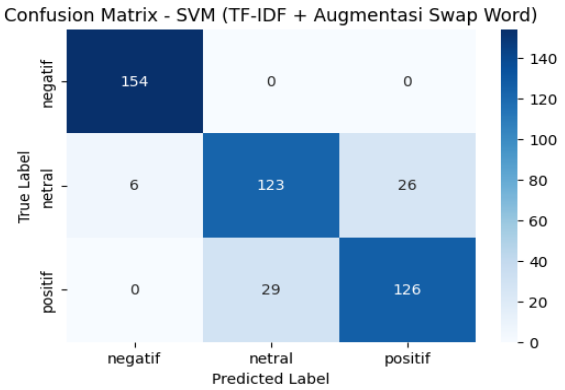
Secara keseluruhan, perolehan nilai *macro average f1-score* sebesar 68% dan *weighted average f1-score* sebesar 0,75 mengindikasikan bahwa model memiliki keseimbangan kinerja yang sangat baik pada seluruh kelas sentimen.

4.2. Klasifikasi Menggunakan *Augmentasi*

Hasil dari klasifikasi dan evaluasi model *Support Vector Machine* menggunakan *Augmentasi* dapat dilihat pada Tabel 8 dan Gambar 7.

Tabel 8. Svm (*tf-idf+augmentasi*)

Kategori	*Precision*	*Recall*	*F1-Score*	*Accuracy*
Negatif	96%	1.00%	98%	87%
Netral	81%	79%	80%	
Positif	83%	81%	82%	



Gambar 7. *Confusion matrix* menggunakan *augmentasi*

Tabel 8 dan Gambar 7 menyajikan hasil evaluasi model yang menerapkan metode *Support Vector Machine* dengan pembobotan TF-IDF pada data hasil augmentasi. Hasil pengujian menunjukkan kinerja yang sangat baik dengan tingkat akurasi sebesar 87%. Penerapan *augmentasi* pada kedua bagian data tersebut bertujuan untuk memperkaya keragaman teks serta meningkatkan kemampuan model dalam mengenali berbagai pola bahasa, sehingga menghasilkan klasifikasi yang lebih stabil dan bersifat general.

Pada kelas sentimen negatif, model memperlihatkan kinerja paling optimal dengan nilai *precision* sebesar 96%, *recall* sebesar 100%, dan *f1-score* sebesar 98%. Temuan ini menunjukkan bahwa hampir seluruh ulasan negatif dapat diidentifikasi secara sangat akurat dengan tingkat kesalahan klasifikasi yang minimal. Untuk kelas netral, model memperoleh *precision* sebesar 81%, *recall* 79%, dan *f1-score* 80%, yang mengindikasikan konsistensi kinerja yang cukup baik dalam mengenali ulasan bersifat netral, meskipun nilainya sedikit lebih rendah dibandingkan kelas negatif. Sementara itu, pada kelas positif, model mencapai *precision* 83%, *recall* 81%, dan *f1-score* 82%, yang menunjukkan kemampuan model dalam mendeteksi sebagian besar ulasan positif dengan tingkat ketepatan yang baik, meskipun masih terdapat variasi bahasa yang belum sepenuhnya teridentifikasi.

Secara keseluruhan, perolehan *nilai macro average f1-score* sebesar 0,87 dan *weighted average f1-score* sebesar 0,87 menunjukkan bahwa model memiliki keseimbangan kinerja yang sangat baik pada seluruh kelas sentimen.

Hasil evaluasi menunjukkan adanya peningkatan kinerja yang signifikan pada model klasifikasi sentimen setelah diterapkan teknik *augmentasi swap word*. Pada tahap awal sebelum optimasi, model memperoleh tingkat akurasi sebesar 76%, dengan kinerja yang relatif baik pada kelas sentimen positif dan netral, yang masing-masing memiliki *f1-score* sebesar 79% dan 77%. Namun demikian, performa model pada kelas sentimen negatif masih tergolong rendah, sehingga mengindikasikan keterbatasan model dalam mengidentifikasi ulasan bernada negatif.

Setelah *augmentasi* data diimplementasikan, kinerja model mengalami peningkatan yang cukup signifikan dengan akurasi mencapai 87%. Peningkatan paling tinggi terjadi pada kelas sentimen negatif, di mana nilai *recall* meningkat secara substansial hingga 100% dan *f1-score* naik menjadi 96%. Penerapan *augmentasi swap word* mampu meningkatkan kemampuan model dalam mengenali pola sentimen negatif secara akurat.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian peningkatan kinerja algoritma *Support Vector Machine*, algoritma *Support Vector Machine* (SVM) dengan pembobotan *TF-IDF* mampu melakukan klasifikasi sentimen ulasan pengguna aplikasi BMKG dengan tingkat akurasi yang

baik. Pada pengujian tanpa *augmentasi* data, model memperoleh akurasi sebesar 76%, sedangkan setelah dilakukan *augmentasi* data menggunakan teknik *swap word* (penukaran kata), akurasi meningkat menjadi 87%. Hal ini menunjukkan bahwa penerapan *augmentasi* data berhasil meningkatkan performa model dengan metode *augmentasi* berbasis *swap word* terbukti efektif dalam meningkatkan kemampuan generalisasi dan akurasi model dalam menganalisis sentimen pengguna aplikasi BMKG.

Untuk pengembangan penelitian di masa depan, direkomendasikan agar menerapkan algoritma klasifikasi alternatif guna mendapatkan perbandingan performa yang lebih komprehensif. Selain itu, upaya optimasi model dapat dilakukan dengan memperbanyak volume data ulasan serta mendiversifikasi teknik augmentasi data agar hasil prediksi semakin akurat.

DAFTAR PUSTAKA

[1] BMKG, "Tentang BMKG."

[2] N. A. Maulana and D. Darwis, "Perbandingan Metode SVM dan Naïve Bayes untuk Analisis Sentimen pada Twitter tentang Obesitas di Kalangan Gen Z," *J. Pendidik. dan Teknol. Indones.*, vol. 5, no. 3, pp. 655–666, 2025, doi: 10.52436/1.jpti.691.

[3] R. Gemino, K. Sagala, T. N. Fatyanosa, and P. P. Adikara, "Klasifikasi Berita Palsu Politik berbasis DistilBERT dengan Peningkatan Data Melalui Teknik Augmentasi Teks," vol. 9, no. 9, pp. 1–10, 2025.

[4] N. Alexander Radja Bria and A. Witanti, "Analisis Sentimen Masyarakat Indonesia Menggunakan Algoritma Support Vector Machine Tentang Pilpres 2024," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3328–3333, 2024, doi: 10.36040/jati.v7i6.8184.

[5] A. Baita, Y. Pristyanto, N. Cahyono, P. Covid-, K. N. N. Akurasi, and K. Kunci, "Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN) Abstraksi Keywords :," vol. 4, no. 2, pp. 42–46, 2021.

[6] D. S. Utami and A. Erfina, "Analisis Sentimen Pinjaman Online di Twitter Menggunakan Algoritma Support Vector Machine (SVM)," pp. 299–305, 2021.

[7] A. F. and H. D. D. Atmajaya, "Metode SCM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," vol. 12, no. 1, pp. 2173–2181, 2023.

[8] S. D. Parameswari, M. Lubis, S. Suakanto, Y. Z. Ramadhan, N. Amanah, and R. A. Dila, "Studi Perbandingan Naï ve Bayes dan Support Vector Machine (SVM) dalam Analisis Sentimen Pengguna Metaverse," vol. 4, no. 3, pp. 1059–1065, 2025.

[9] J. P. Hidayat and I. Nurhaida, "Analisis Sentimen pada Ulasan LMS Pembelajaran Menggunakan Metode Natural Language Processing," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 1, pp. 565–575, 2025, doi: 10.29100/jipi.v10i1.5899.

[10] E. Septiana and C. Caroline, "Optimalisasi Evaluasi Pelaksanaan Pelatihan Melalui Analisis Sentimen Otomatis Dengan Model Text Classification," pp. 141–154, 2024.

[11] S. N. Adhan *et al.*, "Analisis sentimen ulasan aplikasi wattpad di google play store dengan metode random forest," vol. 2, no. 1, pp. 6–15, 2024.

[12] W. E. Nurjanah, R. Setya Perdana, and M. A. Fauzi, "Analisis sentimen terhadap tayangan televisi berdasarkan opini masyarakat pada media sosial twitter menggunakan metode k-kearest neighbor dan pembobotan jumlah retweet," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 12, pp. 1750–1757, 2017.

[13] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Mengunakan Algoritma Support Vector Machine (SVM)," *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023, doi: 10.37905/jjeee.v5i1.16830.

[14] I. A. Rahma, L. H. Suadaa, J. Timur, and P. Korespondensi, "Penerapan Text Augmentation untuk Mengatasi Data yang Tidak Seimbang pada Klasifikasi Teks Berbahasa Indonesia Studi Kasus: Deteksi Judul Clickbait dan Komentar Hate Speech Application Of Text Augmentation To Handling Imbalanced Data Problem In Indonesia," vol. 10, no. 6, pp. 1329–1340, 2023, doi: 10.25126/jtiik.2023107325.

[15] E. Engineering, "Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Mengunakan Algoritma Support Vector Machine (SVM)," vol. 5, pp. 32–35, 2023.

[16] I. P. Angga, P. Widiarta, R. Dwiyansaputra, and A. Aranta, "Analisis Sentimen Masyarakat Terhadap Kebijakan Penerapan PPKM di Media Sosial Twitter dengan Menggunakan Metode XGBOOST (Analysis Of Community Sentiment On The Policy Of Implementation Of PPKM On Twitter Social Media Using Xgboost Method)," vol. 5, no. 2, pp. 154–163, 2023.

[17] N. Mahadiar, "Perbandingan Kinerja Algoritma Naive Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes," 2024, [Online]. Available: http://repository.unpam.ac.id/12843/

[18] A. bil Faqih and D. Avianto, "Jurnal Teknologi Terpadu WATERFALL," *J. Teknol. terpadu*, vol. 6, no. 22, pp. 72–78, 2020.

[19] S. Ernawati and R. Wati, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel," *J. Khatulistiwa Inform.*, vol. 6, no. 1, pp. 64–69, 2018, [Online]. Available: https://www.trustpilot.com/categories/tr