

ANALISIS SENTIMEN PUBLIK TERHADAP PROGRAM MAKAN BERGIZI GRATIS (MBG) PADA APLIKASI X MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) DENGAN TEKNIK SMOTE

Khairun Nisak Nabila, Wahyu Fuadi, Fajriana, Nurdin, Lidya Rosnita

Teknik Informatika, Universitas Malikussaleh

Jalan Batam, Blang Pulo, Muara Satu, Lhokseumawe, Aceh, Indonesia

khairun.200170284@mhs.unimal.ac.id

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan salah satu program prioritas pemerintah yang bertujuan meningkatkan kualitas sumber daya manusia melalui pemenuhan kebutuhan gizi peserta didik. Pelaksanaan program ini menuai beragam respons di media sosial X, terutama terkait aspek anggaran dan distribusi, sehingga diperlukan analisis sentimen untuk memahami kecenderungan opini publik. Penelitian ini bertujuan membangun sistem klasifikasi sentimen masyarakat terhadap program MBG melalui metode *Support Vector Machine* (SVM) menggunakan teknik SMOTE guna mengatasi permasalahan ketidakseimbangan kelas yang umum ditemukan pada data media sosial. Dataset penelitian terdiri dari 3.809 *tweet* yang terbagi menjadi tiga kelas sentimen, yakni positif, netral, serta negatif. Tahapan penelitian meliputi *preprocessing* data, pembobotan TF-IDF, serta proses klasifikasi menggunakan metode SVM sebelum dan sesudah penerapan SMOTE. Berdasarkan hasil uji, SMOTE meningkatkan akurasi dari 82,68% menjadi 84,51% dan *F1-score average* dari 81,59% menjadi 83,89%. Hasil tersebut menunjukkan bahwa penerapan SMOTE mampu meningkatkan performa klasifikasi serta mengatasi permasalahan ketidakseimbangan data sehingga menghasilkan distribusi prediksi yang lebih seimbang. Dengan demikian, metode SVM menggunakan teknik SMOTE efektif digunakan untuk analisis sentimen terhadap Program MBG pada aplikasi X.

Kata kunci : Analisis Sentimen, Program MBG, Aplikasi X, SVM, Teknik SMOTE

1. PENDAHULUAN

Program Makan Bergizi Gratis (MBG) termasuk satu dari beberapa program prioritas pemerintahan Presiden Prabowo Subianto, yang memiliki tujuan meningkatkan kualitas Sumber Daya Manusia Indonesia lewat pemenuhan gizi bagi peserta didik. Meskipun dirancang untuk mendukung peningkatan Kesehatan dan kualitas Pendidikan, pelaksanaan program ini menimbulkan beragam respon di Masyarakat. Sebagian masyarakat memberikan dukungan, sementara sebagian lainnya mempertanyakan efektivitas serta kesiapan pelaksanaannya, terutama terkait aspek anggaran dan distribusi [1].

Perbedaan pandangan masyarakat terhadap program MBG menunjukkan perlunya analisis sentimen untuk mengetahui kecenderungan opini publik yang berkembang di media sosial. Informasi tersebut bisa dimanfaatkan sebagai bahan pertimbangan dalam mengevaluasi dan meningkatkan efektivitas pelaksanaan program.

Aplikasi X, yang lebih dulu dikenal dengan *Twitter*, yakni satu dari beberapa media sosial yang banyak dimanfaatkan dalam membagikan pesan singkat serta mengemukakan pendapat, serta berbagi informasi terkini. *Platform* ini juga memungkinkan pengguna mengekspresikan opini mereka secara *real time* atas sejumlah topik kebijakan pemerintah, ekonomi, politik, ataupun sosial [2].

Karakteristiknya yang terbuka dan penyebaran informasi yang berlangsung cepat menjadikan X

sumber data yang sesuai terhadap upaya analisis pendapat publik, terutama terkait Program Makan Bergizi Gratis (MBG).

Analisis sentimen menjadi satu dari beberapa metode untuk mengkategorikan teks dari media sosial menjadi positif, netral, maupun negatif [3].

Seiring meningkatnya perhatian masyarakat terhadap program MBG, penelitian analisis sentimen terkait program tersebut juga mulai banyak dilakukan. Namun, sebagian besar penelitian masih berfokus pada proses klasifikasi sentimen tanpa memperhatikan permasalahan ketidakseimbangan kelas data yang sering ditemukan pada data media sosial. Kondisi tersebut dapat memengaruhi kemampuan model dalam mengklasifikasikan setiap kelas sentimen secara optimal, khususnya pada kelas minoritas.

Metode *Support Vector Machine* (SVM) diterapkan sebab kemampuannya mengatasi data kompleks dengan margin optimal serta sudah terbukti memberikan akurasi yang lebih baik daripada sejumlah algoritma lain termasuk *Naive Bayes* serta *Random Forest* [4].

Akan tetapi, performa SVM dapat dipengaruhi oleh distribusi data yang tidak seimbang. Permasalahan utama yang muncul adalah ketidakseimbangan kelas data sentimen di media sosial, yakni satu kelas cenderung mendominasi sehingga model bias terhadap kelas mayoritas.

Penelitian [5] menunjukkan bahwa dari 1.516 *tweet*, sebanyak 75.8% didominasi oleh sentimen positif.

Dalam menangani permasalahan tersebut, diterapkan teknik *Synthetic Minority Oversampling Technique* (SMOTE) yang dapat memperbanyak jumlah sampel minoritas secara sintesis [6].

Penerapan SMOTE diharapkan mampu meningkatkan representasi kelas minoritas sehingga distribusi data menjadi lebih seimbang dan model SVM dapat menghasilkan performa klasifikasi yang lebih baik. Kombinasi SVM dan SMOTE terbukti meningkatkan akurasi pada berbagai penelitian analisis sentimen [7][8][9].

Selain itu, beberapa penelitian terbaru juga telah menerapkan teknik SMOTE pada analisis sentimen Program Makan Bergizi Gratis untuk mengatasi permasalahan ketidakseimbangan data. Meskipun demikian, penelitian yang mengkaji pengaruh penerapan SMOTE terhadap kinerja model SVM pada analisis sentimen Program Makan Bergizi Gratis masih terbatas sehingga perlu dilakukan pengujian lebih lanjut.

Penelitian ini bertujuan membangun sistem analisis sentimen atas Program MBG pada aplikasi X serta menganalisis pengaruh penerapan teknik SMOTE terhadap kinerja model SVM pada kategori tiga kelas sentimen: netral, negatif, serta positif. Temuan penelitian diharap mampu menyajikan gambaran terkait kecenderungan opini publik atas program tersebut dan menjadi bahan masukan dalam evaluasi kebijakan yang berkaitan dengan pelaksanaannya.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Beberapa penelitian telah dilakukan untuk menganalisis sentimen masyarakat terhadap program MBG maupun kebijakan serupa di media sosial. [10] menerapkan metode *Random Forest* (RF) dan SM dengan teknik SMOTE untuk menangani ketidakseimbangan data sentimen MBG. Hasil penelitian menunjukkan SMOTE mampu memperbaiki distribusi kelas dan mengurangi bias prediksi serta menunjukkan performa SVM lebih unggul dibandingkan RF.

Penelitian lain [11] yang mengkombinasikan metode K-Means dan SVM dengan SMOTE membuktikan bahwa penerapan SMOTE dapat meningkatkan kinerja klasifikasi secara signifikan dengan menyeimbangkan distribusi kelas. Selain itu [12] juga menunjukkan bahwa SMOTE diperlukan untuk meningkatkan kualitas klasifikasi.

Penelitian terkait sentimen program MBG juga dilakukan oleh [13] yang membandingkan algoritma SVM, Naive Bayes dan Random Forest. Hasil penelitian menunjukkan bahwa SVM memperoleh akurasi tertinggi sebesar 85% sehingga dinilai efektif untuk klasifikasi sentimen pada data media sosial.

Sementara itu, penelitian [14] yang menggunakan SVM untuk menganalisis sentimen memperoleh nilai akurasi sebesar 80%. Namun, model menunjukkan performa yang lebih rendah pada kelas

netral dan positif. Hal ini mengindikasikan adanya pengaruh ketidakseimbangan distribusi data terhadap hasil klasifikasi.

2.2. Analisis Sentimen

Analisis sentimen adalah satu bagian dari *text mining* dengan tujuan melakukan identifikasi terhadap opini maupun emosi dalam teks, meliputi komentar maupun cuitan media sosial ke dalam kategori netral, negatif, ataupun positif [15].

Pada pembahasan kebijakan publik, analisis sentimen memudahkan pemerintah untuk mengetahui bagaimana pandangan atau persepsi masyarakat dalam menerima suatu program.

2.3. Text Mining

Text mining dimaknai sebagai proses analisis data yang memiliki bentuk teks di mana sumber data diperoleh melalui dokumen [16].

Tahapan *Preprocessing* mencakup (1) *Case folding* untuk mengganti huruf menjadi huruf kecil; (2) *Cleaning* untuk membersihkan data dari tanda baca, angka, *hashtag*, *mention*, URL, hingga karakter khusus lain; (3) *Normalization* untuk mengganti kata tidak baku menjadi formal; (4) *Tokenizing* untuk memotong struktur kalimat menjadi per kata; (5) *Stopwords Removal* untuk mengeliminasi kata tidak penting; (6) *Stemming* untuk mengganti kata menjadi bentuk kata dasar [17].

2.4. Term Frequency-Inverse Document Frequency (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) termasuk metode dalam menilai masing-masing kata di dokumen. TF diterapkan untuk mengukur frekuensi munculnya *term* pada dokumen, karena panjang dari setiap dokumen berbeda, maka:

$$tf_{t,d} = \frac{n_{t,d}}{\sum_k n_{k,d}} \quad (1)$$

Di mana $n_{t,d}$ merupakan jumlah munculnya kata t dalam dokumen d serta $\sum_k n_{k,d}$ sebagai total semua kata pada dokumen d .

IDF berfungsi untuk mengukur tingkat keunikan suatu *term* berdasarkan kemunculannya pada keseluruhan dokumen. Dihitung dengan persamaan:

$$idf = \log \left(\frac{N}{df_t} \right) \quad (2)$$

Di mana N merupakan total seluruh dokumen serta df_t merupakan total dokumen yang memiliki kata t .

Skor TF-IDF didapatkan melalui hasil perkalian antara *Term Frequency* (TF) serta *Inverse Document Frequency* (IDF). Bobot yang diperoleh mengindikasikan tingkat kepentingan sebuah term menurut frekuensi keberadaannya di dokumen serta seberapa jarang term itu muncul pada keseluruhan koleksi dokumen.

2.5. Support Vector Machine (SVM)

Support Vector Machine (SVM) termasuk satu dari beberapa algoritma *machine learning* yang bekerja melalui menemukan sebuah *hyperplane* yang berfungsi sebagai batas pemisah antar kelas dalam ruang fitur. *Hyperplane* dalam SVM dihitung dengan rumus 3.

$$\vec{w} \cdot \vec{x} + b = 0 \quad (3)$$

Dengan $\vec{w}$ mengindikasikan vektor bobot dan b adalah bias. Data diklasifikasikan berdasarkan tanda nilai $f(x)$; positif apabila $f(x) > 0$, negatif jika $f(x) < 0$, hingga netral jika $f(x) = 0$.

2.6. Synthetic Minority Oversampling Technique (SMOTE)

Synthetic Minority Oversampling Technique (SMOTE) termasuk metode *oversampling* yang diterapkan dalam mengatasi ketidakseimbangan kelas dataset melalui memproduksi data sintesis baru kelas minoritas menurut karakteristik data yang telah ada [18].

Untuk menghasilkan data sintesis, digunakan rumus berikut:

$$x_{syn} = x_i + (x_{knn} - x_i) \times \gamma \quad (4)$$

Di mana x_{knn} adalah tetangga paling dekat yang ditentukan dengan acak serta γ merupakan skor acak pada rentang 0-1. Proses tersebut dilakukan berulang sampai total data minoritas meraih proporsi yang diinginkan terhadap kelas mayoritas.

2.7. Evaluasi Model

Evaluasi kinerja model klasifikasi dilaksanakan melalui *confusion matrix* serta metrik *accuracy*, *precision*, *recall*, serta *F1-score*. *Confusion matrix* memunculkan total prediksi benar serta salah dalam setiap kelas. Metrik evaluasi dihitung menggunakan persamaan (5) hingga (8).

$$Akurasi = \frac{\text{jumlah data diprediksi benar}}{\text{total seluruh data}} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

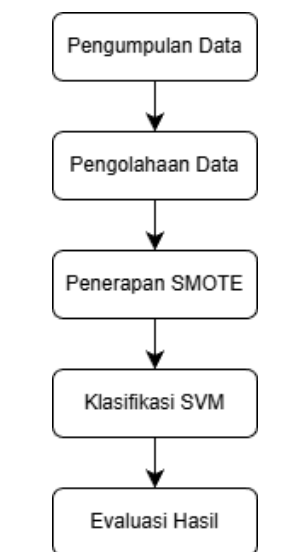
Di mana TP mengindikasikan *true positive*, TN mengindikasikan *true negative*.

3. METODE PENELITIAN

3.1. Waktu dan Sumber

Penelitian ini menerapkan data *tweet* dari aplikasi X yang diambil selama tahun 2025 dengan kata kunci yang berkaitan dengan program Makan Bergizi Gratis. Data hanya mencakup *tweet* dengan bahasa Indonesia dari akun publik. Data awal berjumlah 4.000 *tweet*. Setelah penghapusan data duplikat, diperoleh 3.809 *tweet* bersih yang digunakan dalam penelitian.

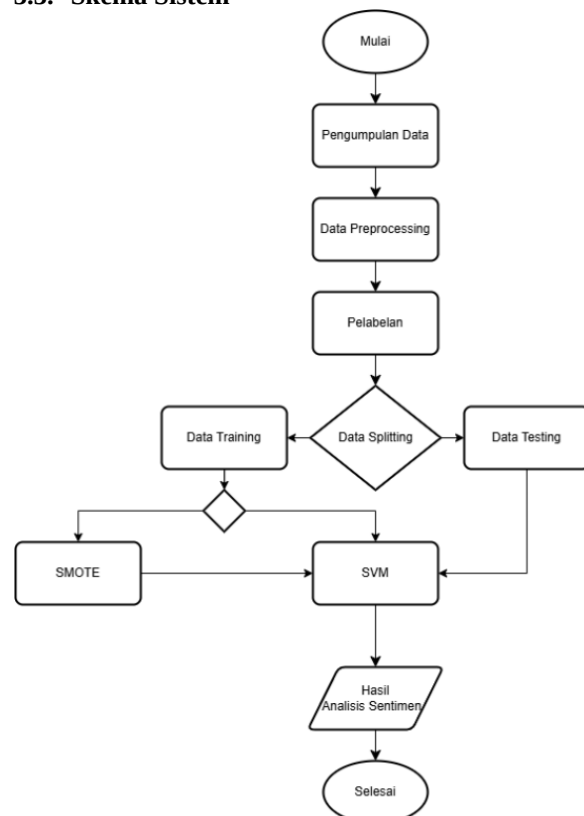
3.2. Alur Penelitian



Gambar 1. Alur Penelitian

Dari Gambar 1. ditemukan bahwa alur penelitian terdiri atas 5 tahapan utama: (1) pengumpulan data *tweet* dari aplikasi X; (2) pengolahan data mencakup *case folding*, *cleaning*, *tokenizing*, *stopword removal*, serta *stemming*; (3) Penerapan SMOTE pada data latih; (4) Pelatihan model SVM menggunakan kernel *linear*; (5) Evaluasi hasil model dengan *confusion matrix*.

3.3. Skema Sistem



Gambar 2. Skema Sistem

Dari Gambar 2. dapat dilihat skema sistem dimulai dengan: (1) mengumpulkan data terkait program MBG dengan menggunakan kata kunci tertentu; (2) Membersihkan teks dari berbagai *noise* untuk menghasilkan data yang siap diproses; (3) pelabelan sentimen ke kategori positif, negatif, dan netral; (4) data berlabel dibagi menjadi *data training*/data latih dan *data testing*/data uji; (5) teknik SMOTE diterapkan pada data latih untuk menghasilkan distribusi data yang seimbang; (6) setelah proses SMOTE, data latih digunakan pada tahap klasifikasi menggunakan SVM; (7) data uji digunakan untuk menguji performa model; (8) hasil klasifikasi digunakan untuk memperoleh hasil analisis sentimen program MBG.

4. HASIL DAN PEMBAHASAN

4.1. Analisis Data

Dataset akhir meliputi 3.809 *tweet* mengenai Program MBG yang diperoleh melalui proses *crawling* dan telah melalui pembersihan data duplikat. Data dibagi menggunakan rasio 80:20, memperoleh 3.047 data latih serta 762 data uji. Distribusi kelas sentimen pada keseluruhan dataset tersaji dalam Tabel 1.

Tabel 1. Distribusi sentimen data *tweet*

Sentimen	Jumlah Data	Persentase
Negatif	1.789	45,96%
Netral	1.122	29,45%
Positif	898	23,57%
Total	3.809	100%

Berdasarkan Tabel 1, kelas negatif mendominasi dataset dengan persentase 45,96%, diikuti kelas netral 29,45%, dan kelas positif sebagai minoritas 23,57%. Ketidakseimbangan ini mengindikasikan perlunya teknik penyeimbangan data seperti SMOTE supaya model tidak bias atas kelas mayoritas.

4.2. Hasil Preprocessing

Proses *preprocessing* meliputi *remove duplicate*, *case folding*, *cleaning*, *tokenizing*, *stopword removal*, serta *stemming*. Tabel 2 merupakan contoh data *tweet* yang telah melalui tahap *preprocessing*.

Tabel 2. Data sebelum serta setelah *preprocessing*

Sebelum <i>Preprocessing</i>	Setelah <i>Preprocessing</i>
Presiden Prabowo menegaskan masa depan Indonesia terletak pada kualitas generasi muda. Karena itu pemerintah memprioritaskan pemenuhan gizi setiap anak Indonesia melalui program Makan Bergizi Gratis (MBG).	presiden prabowo menegaskan masa depan indonesia terletak kualitas generasi muda karena pemerintah memprioritaskan pemenuhan gizi setiap anak

Berdasarkan Tabel 2, *preprocessing* berhasil menghapus karakter tidak relevan dan menyederhanakan kata sehingga teks siap digunakan pada tahap pembobotan TF-IDF.

4.3. Hasil Pembobotan TF-IDF

Pembobotan TF-IDF dilaksanakan terhadap seluruh dataset menghasilkan representasi numerik dari setiap *tweet*. Tabel 3 menampilkan sepuluh *term* dengan nilai TF-IDF tertinggi pada dataset penelitian ini.

Tabel 3. Hasil pembobotan kata

No	Term	TF-IDF Score
1	makan	0.057758
2	gratis	0.057240
3	siang	0.038324
4	mbg	0.037957
5	program	0.033397
6	bergizi	0.027784
7	ada	0.019373
8	yg	0.018959
9	anak	0.015114
10	sekolah	0.013745

Berdasarkan Tabel 3. *term* “makan” dan “gratis” memiliki bobot TF-IDF tertinggi dibandingkan *term* lainnya. Temuan tersebut mengindikasikan apabila kedua kata itu cukup dominan pada *dataset* serta memiliki pengaruh terhadap proses klasifikasi sentimen dengan *Support Vector Machine*.

4.4. Penerapan SMOTE

SMOTE diterapkan khusus pada data latih agar data sintesis tidak mempengaruhi evaluasi model.

Tabel 4. Distribusi data sebelum serta sesudah SMOTE

Sentimen	Sebelum SMOTE	Sesudah SMOTE	Peningkatan
Negatif	1.382	1.382	0
Netral	909	909	0
Positif	756	1.382	+626
Total	3.047	3.673	+626

Berdasarkan Tabel 4, SMOTE berhasil membentuk 626 data sintesis baru terhadap kelas positif sehingga jumlahnya setara terhadap kelas negatif (1.382 data). Distribusi yang lebih seimbang ini memberikan peluang bagi model untuk mempelajari pola setiap kelas sentimen dengan lebih merata.

4.5. Penerapan Metode SVM

Model SVM dilatih dengan data latih hasil TF-IDF dengan kernel *linear*. Parameter SVM yang diterapkan merupakan $C = 1.0$ (parameter regularisasi default) yang memberikan keseimbangan antara margin maksimum serta toleransi terhadap *misclassification*. Selanjutnya, dua model dilatih secara terpisah: model SVM tanpa SMOTE dan model SVM+SMOTE.

4.6. Evaluasi Model

Pengujian dilaksanakan dengan 762 data uji yang tidak terlibat pada proses pelatihan.

Tabel 5. Perbandingan performa

Metode	Accuracy	Precision Avg	Recall Avg	F1-Score Avg
SVM	82,68%	83,71%	80,37%	81,59%
SVM+SMOTE	84,51%	85,41%	82,86%	83,89%
Selisih	+1,83%	+1,70%	+2,49%	+2,30%

Berdasarkan Tabel 5, penerapan SMOTE meningkatkan seluruh metrik evaluasi. Recall mengalami peningkatan tertinggi sebesar 2,49%, menunjukkan model lebih mampu mengenali semua kelas sentimen dengan tepat setelah penyeimbangan data.

4.7. Implementasi Sistem

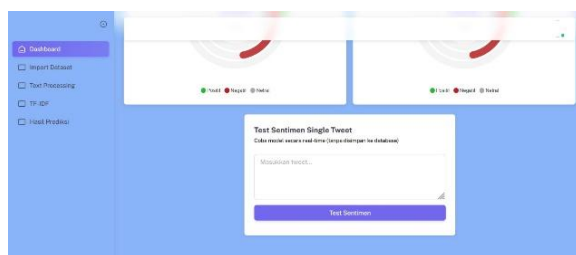
Halaman dashboard: Halaman *dashboard* menampilkan visualisasi distribusi sentimen berupa diagram perbandingan kinerja kedua model, serta fitur pengujian sentimen *single tweet* secara *real-time*.



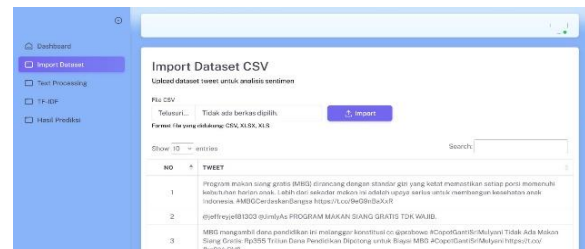
Gambar 3. Menampilkan visualisasi distribusi sentimen



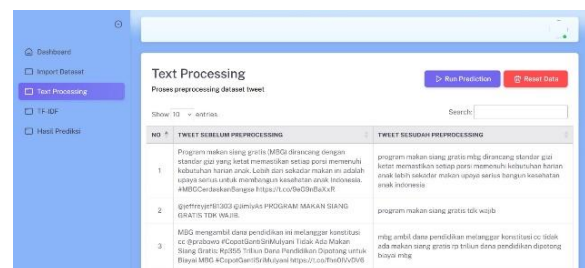
Gambar 4. Menampilkan diagram lingkaran perbandingan kinerja model

Gambar 5. Pengujian *single tweet*

Halaman *Import dataset*: pengguna bisa mengunggah data *tweet* dalam format CSV maupun XLSX untuk dianalisis.

Gambar 6. Halaman *import dataset*

Halaman *text preprocessing*: menampilkan proses pembersihan teks secara transparan beserta tombol "Run Prediction" dan "Reset Data".

Gambar 7. Halaman *text preprocessing*

Halaman ekstraksi fitur TF-IDF: menampilkan nilai bobot setiap *term* sebagai representasi numerik fitur.

NO	TERM	TF-IDF
1	makan	0.057756
2	gratis	0.05724
3	siang	0.036304
4	mba	0.007957
5	program	0.003369
6	bergizi	0.027784
7	ada	0.000075
8	nya	0.000099

Gambar 8. Halaman hasil TF-IDF

Halaman hasil prediksi: menampilkan data asli *tweet*, *clean tweet*, serta hasil prediksi dari kedua model secara berdampingan sehingga pengguna dapat membandingkan langsung output SVM dan SVM+SMOTE

NO	TWEET	CLEAN TWEET	SVM	SVM + SMOTE
1	1. Presiden Prabowo mengatakan masa depan Indonesia terancam oleh ancaman global. Karena itu pemerintah harus meningkatkan program Makan Bergizi Gratis (MBG). #KorupsiTV Apa Perlu diungkap DUKA MBG khusus meningkatkan program makan bergizi gratis. #KorupsiTV. Tengah kerja TP yang belum berakhir.?	presiden Prabowo mengatakan masa depan Indonesia terancam oleh ancaman global. Karena itu pemerintah harus meningkatkan program Makan Bergizi Gratis (MBG). #KorupsiTV Apa Perlu diungkap DUKA MBG khusus meningkatkan program makan bergizi gratis. #KorupsiTV. Tengah kerja TP yang belum berakhir.?	Positif	Positif
2			Negatif	Negatif

Gambar 9. Halaman perbandingan hasil prediksi sentimen

4.8. Pengujian Sistem

Setelah sistem berhasil diimplementasikan, tahap berikutnya adalah melakukan pengujian *black box* untuk mengetahui kesesuaian fungsi-fungsi yang terdapat pada sistem dengan rancangan yang telah dibuat.

Tabel 6. Pengujian Sistem

No	Fitur yang Diuji	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian	Status
1.	Halaman Dashboard	Membuka halaman dashboard utama	Sistem menampilkan halaman dashboard yang berisi jumlah data sentimen, diagram lingkaran, dan pengujian <i>single tweet</i>	Dashboard tampil dengan total data sentimen, diagram lingkaran, dan koom pengujian <i>tweet</i>	Valid
2.	Halaman import dataset	Klik tombol <i>choose file</i>	Sistem menerima file dan menampilkan data pada tabel	Dataset berhasil dimuat dan ditampilkan	Valid
		Klik tombol <i>import</i>	Sistem menampilkan nomor, <i>tweet</i> , dan isi dataset	Data tampil dalam tabel	Valid
3.	Halaman Text Preprocessing	Klik <i>run prediction</i>	Sistem melakukan proses cleaning, case folding, tokenizing, stopword removal, dan stemming	Hasil <i>preprocessing</i> tampil pada tabel	Valid
		Klik <i>reset data</i>	Sistem menghapus data sementara dan mengembalikan kondisi awal	Data berhasil direset	Valid
4.	Halaman TF-IDF	Menampilkan pembobotan kata	Sistem menghitung bobot TF-IDF setiap term	Tabel <i>term</i> dan nilai TF-IDF tampil	Valid
5.	Halaman Hasil Prediksi Sentimen	Melihat hasil klasifikasi dataset	Sistem menampilkan label SVM dan SVM+SMOTE	Label positif, negatif, atau netral tampil pada tabel	Valid
6.	Pengujian Single <i>Tweet</i>	Melakukan prediksi satu teks	Sistem memberikan hasil klasifikasi sentimen	Sistem memberikan hasil klasifikasi sentimen	Valid
7.	Pencarian Data Tabel	Menggunakan fitur <i>search tabke</i>	Sistem memfilter data sesuai kata kunci	Data sesuai pencarian muncul	Valid
8.	<i>Pagination Tabel</i>	Berpindah halaman data	Sistem menampilkan data halaman selanjutnya	<i>Pagination</i> berjalan normal	Valid
9.	Navigasi Sidebar	Berpindah menu halaman	Sistem membuka halaman sesuai menu	Semua menu dapat diakses	Valid

Dari tabel 6. diatas dapat dilihat bahwa hasil pengujian sistem analisis sentimen MBG menggunakan SVM dan SMOTE, seluruh fitur diuji berhasil berjalan sesuai dengan kebutuhan fungsional sistem. Sistem mampu melakukan *import* dataset, *preprocessing* teks, pembobotan TF-IDF, prediksi sentimen, serta menampilkan hasil klasifikasi dengan baik.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun sistem analisis sentimen terhadap Program Makan Bergizi Gratis (MBG) dalam aplikasi X menggunakan metode SVM dengan teknik SMOTE berbasis web. Dataset terdiri dari 3.809 *tweet* berbahasa Indonesia dengan tiga kelas sentimen: negatif (45,96%), netral (29,45%), dan positif (23,57%). Penerapan SMOTE terbukti meningkatkan performa model SVM pada seluruh metrik evaluasi: *accuracy* dari 82,68% menjadi 84,51%, *precision* dari 83,71% menjadi 85,41%, *recall* dari 80,37% menjadi 82,86%, serta *F1-score* dari 81,59% menjadi 83,89%. Temuan tersebut membuktikan apabila SMOTE efektif menangani ketidakseimbangan kelas. Dengan demikian, model memperoleh distribusi prediksi yang lebih seimbang serta akurat. Bagi penelitian berikutnya, disarankan dalam memperbanyak data *tweet* multibahasa dan membandingkan dengan metode lain.

DAFTAR PUSTAKA

- [1] A. Ramadhani, I. Permana, M. Afdal, and M. Fronita, "Analisis Sentimen Tanggapan Publik di
- [2] R. A. Husen, R. Astuti, L. Marlia, R. Rahmaddeni, and L. Efrizoni, "Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 211–218, 2023, doi: 10.57152/malcom.v3i2.901.
- [3] A. D. Cahyani, "Analisa Kinerja Metode Support Vector Machine untuk Analisa Sentimen Ulasan Pengguna Google Maps," *J. Comput. Syst. Informatics*, vol. 4, no. 3, pp. 604–613, 2023, doi: 10.47065/josyc.v4i3.3426.
- [4] I. Irawan, W. Wardianto, M. H. Wathan, and M. B. Prayogi, "Studi Perbandingan: Algoritma Random Forest, Naive Bayes Dan Support Vector Machine Dalam Analisis Sentimen Pada Aplikasi Capcut Di Google Play Store," *J. Pengemb. Sist. Inf. dan Inform.*, vol. 5, no. 4, pp. 152–164, 2024, doi: 10.47747/jpsii.v5i4.1959.
- [5] G. Ainur, N. Latifah, and F. Nugraha, "Analisis Sentimen Publik Terhadap Program Makan Bergizi Gratis Menggunakan Metode Random Forest Pada Platform X," *J. Sist. Inf. Dan Teknol.*, vol. 1, no. 2, pp. 1–6, 2025, doi: 10.24176/sitech.v8i1.15628.
- [6] A. Hadi, B. Bernard, L. N. Hulu, and W. N. Hulu, "Sentimen Ulasan Pengguna E-Commerce Blibli

- Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine Dengan Metode Smote,” *J. Multidisiplin Saintek*, vol. 6, no. 5, pp. 1–9, 2025, doi: 10.8734/Kohesi.v1i2.365.
- [7] N. Fajriyah, N. T. Lapatta, D. W. Nugraha, and R. Laila, “Implementasi Svm Dan Smote Pada Analisis Sentimen Media Sosial X Terhadap Pelantikan Agus Harimurti Yudhoyono,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 2, pp. 1359–1370, 2025, doi: 10.29100/jupi.v10i2.6246.
- [8] T. A. Dewi and E. Mailoa, “Perbandingan Implementasi Metode Smote pada Algoritma Support Vector Machine (SVM) dalam Analisis Sentimen Opini Masyarakat Tentang Mixue,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 3, pp. 849–855, 2023, doi: 10.35870/jimik.v4i3.289.
- [9] R. Gusmansyah, Rahmaddeni, H. Hendrawan, and Rohid, “Peningkatan Kinerja Analisis Sentimen Pada Ulasan Aplikasi Identitas Kependudukan Digital (Ikd) Di Indonesia Menggunakan Algoritma Support Vector Machine (Svm) Dan Smote,” *J. INSTEK (Informatika Sains dan Teknol.*, vol. 10, no. 1, pp. 185–196, 2025, doi: 10.24252/instek.v10i1.55292.
- [10] S. E. Prianto, Berlilana, and R. E. Saputro, “Analisis Sentimen Program Makan Bergizi Gratis Menggunakan Random Forest dan Support Vector Machine dengan Penyeimbangan Data Berbasis SMOTE,” vol. 6, no. 2, pp. 353–365, 2026, doi: 10.52436/1.jpti.1491.
- [11] R. Delfalizanty and Yulhendri, “Analisis Sentimen Persepsi Masyarakat pada Media Sosial X Menggunakan Metode K-Means , SVM , dan Teknik Smote,” vol. 6, no. 1, 2026, doi: 10.55606/juisik.v6i1.2071.
- [12] S. Shafwah and H. Al Azies, “Komparasi SVM dan IndoBERT dalam Klasifikasi Sentimen Program Makanan Bergizi Gratis,” vol. 10, no. 2, pp. 105–112, 2025, doi: 10.31544/jtera.v10.i2.2025.105-112.
- [13] D. W. Brata, W. Purnomo, H. Farisi, and H. Kevin Marcelino Ratu, “Menggali Opini Publik: Sentimen Terhadap Kebijakan Makan Siang Gratis Dengan Supervised Learning,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 3, pp. 533–538, 2025, doi: 10.25126/jtiik.20259565.
- [14] F. Fatkhurrohman, B. I. Nugroho, and N. Fadillah, “Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM,” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 3906–3917, 2025, doi: 10.31004/riggs.v4i3.2533.
- [15] A. Ahmad and W. Gata, “Analisis Sentimen Menggunakan Support Vector Machine Masyarakat Indonesia Di Twitter Terkait Bjorka,” *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 4, pp. 548–555, 2022, doi: 10.37859/coscitech.v3i3.4358.
- [16] Nurhaliza, Risawandi, and L. Rosnita, “Analisis Sentimen Kepuasan Customer Terhadap Ekspedisi Tiki, Sicepat Express dan Ninja Express Menggunakan Algoritma Naive Bayes,” *J. Informatics Comput. Sci.*, vol. 9, no. 1, pp. 1–5, 2023, doi: 10.33143/jics.v9i1.2943.
- [17] Yusril, W. Fuadi, and Y. Afrillia, “Analisis Sentimen Review Aplikasi Stockbit Di Google Play Store Dan X(Twitter) Menggunakan Support Vector Machine,” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 1050–1062, 2025, doi: 10.36341/rabit.v10i2.6446.
- [18] Hermanto, A. Y. Kuntoro, T. Asra, E. B. Pratama, L. Effendi, and R. Ocanitra, “Gojek and Grab User Sentiment Analysis on Google Play Using Naive Bayes Algorithm and Support Vector Machine Based Smote Technique,” *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012102.