Analisis Sentimen Pengguna *YouTube* Terhadap Program Makan Bergizi Gratis Menggunakan Algoritma *Support Vector Machine*

1st Dewa Brata
Sistem Informasi, Fakultas Rekayasa Industri
Universitas Telkom Purwokerto
Purwokerto, Indonesia
dewabrata@student.telkomuniversity.ac.id

2nd Sena Wijayanto
Sistem Informasi, Fakultas Rekayasa Industri
Universitas Telkom Purwokerto
Purwokerto, Indonesia
senawijayanto@telkomuniversity.ac.id

Abstrak — Program Makan Bergizi Gratis merupakan inisiatif pemerintah Indonesia di bawah Kabinet Merah Putih yang diluncurkan pada Januari 2025 untuk meningkatkan kesejahteraan masyarakat melalui penyediaan makanan bergizi, khususnya bagi kelompok rentan. Program ini memicu beragam opini di media sosial, khususnya di YouTube karena pelaksanaannya menyangkut banyak pihak dan anggaran negara yang sangat besar. Sehingga, penting untuk menganalisis sentimen agar mengetahui penerimaan publik terhadap PMBG. Penelitian ini menggunakan data komentar sebanyak 2760 data komentar lima video dari YouTube berita nasional hasil crawling, kemudian data tersebut diproses melalui tahap preprocessing. Selanjutnya, data komentar diberi kategori polaritas, yaitu sentimen positif dan negatif menggunakan bantuan kamus SenticNet. Setelah itu, dibobot menggunakan TF-IDF, data dipisahkan ke dalam dua skenario proporsi pelatihan dan pengujian, yaitu 80:20 dan 90:10. SMOTE digunakan untuk menyeimbangkan distribusi kelas, lalu data diklasifikasikan menggunakan Support Vector Machine yang menguji empat tipe kernel berbeda. Evaluasi model dilakukan melalui pendekatan confusion matrix untuk memperoleh nilai akurasi, precision, recall, dan f1-score. Hasil skenario 80:20 menggunakan SMOTE, didapatkan kernel linear sebesar 86%, dan akurasi tertinggi tanpa SMOTE didapatkan oleh kernel sigmoid sebesar 86%. Hasil skenario 90:10 menggunakan SMOTE kernel linear menunjukkan akurasi terbaik sebesar 89%, dan akurasi tertinggi 91% tanpa SMOTE pada kernel sigmoid.

Kata kunci— analisis sentimen, youtube, makan bergizi gratis, support vector machine, SVM

I. PENDAHULUAN

Program Makan Bergizi Gratis (PMBG) adalah salah satu inisiatif yang dijalankan oleh pemerintah Indonesia Kabinet Merah Putih, diharapkan program ini mampu meningkatkan taraf hidup masyarakat dengan pemberian makan bergizi gratis [1]. PMBG menjadi topik pembicaraan publik, dikarenakan adanya kontroversi yang menyertai program ini. Seperti dalam pelaksanaan program membutuhkan banyak pihak yang terlibat dan anggaran yang sangat besar dengan estimasi biaya mencapai 450 triliun rupiah [2].

PMBG banyak di bahas di media sosial, khusunya platform *YouTube*. Banyak video yang membahas program tersebut, sehingga memicu berbagai komentar dari pengguna *YouTube*, baik pro maupun kontra. Oleh karena itu, diperlukan analisis sentimen untuk mendapatkan pemahaman mengenai sentimen publik terhadap PMBG.

Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk melakukan klasifikasi sentimen komentar pengguna *YouTube* terhadap PMBG. SVM dipilih karena kemampuannya yang baik dalam klasifikasi dan pengolahan data analisis sentimen . SVM juga sangat efektif dalam memisahkan data teks yang bermuatan sentimen positif dan negatif [3].

Penelitian sebelumnya menunjukan kelebihan SVM dibandingkan algoritma yang lain [4][5]. Data yang digunakan dalam penelitian ini berasal dari komentar pengguna *YouTube* yang dikumpulkan melalui teknik *crawling*. Data diolah dengan preprocessing, pelabelan otomatis menggunakan bantuan SenticNet, dan balancing menggunakan SMOTE.

II. KAJIAN TEORI

A. *Analisis Sentimen*

Analisis sentimen adalah bidang ilmu yang berkaitan dengan proses mengidentifikasi pikiran, pendapat, dan emosi yang diungkapkan dalam suatu teks. Topik ini berfokus pada pengelompokan sentimen berdasarkan teks opini terkait isu atau permasalahan yang menarik [6].

B. *Support Vector Machine* (SVM)

SVM adalah algoritma yang digunakan untuk penyelesaian masalah dengan memprediki klasifikasi dan regresi [7]. Cara kerja SVM adalah menemukan *hyperplane* terbaik yang berfungsi memisahkan dua kelas dalam ruang input dengan margin maksimal.

SVM juga dapat menyelesaikan masalah data yang tidak linear dengan bantuan kernel di ruang yang berdimensi tinggi [8]. Empat jenis kernel yang umum dipakai, antara lain kernel linear, RBF, polynomial, dan sigmoid Berikut merupakan rumus dari SVM dan semua kernelnya [9]:

1. *Support Vector Machine* (SVM)

$$f(x) = sign(w \cdot x + b) \quad (1)$$

Keterangan
f(x) $= fungsi\ prediksi$
w $= vektor\ bobot$
x $= vektor\ fitur\ input$
b $= bias$

2. Kernel Linear

$$K(x, y) = (x.y) \quad (2)$$

Keterangan
$K(x, y) = nilai\ kernel\ data\ x\ dan\ y$
x $= nilai\ data\ latih$
y $= nilai\ data\ uji$

3. Kernel RBF

$$K(x, x') = \exp(-y(x, y)^2 \quad (3)$$

Keterangan
$K(x, y) = nilai\ kernel\ data\ x\ dan\ y$
$exp = eksponensial\ (basis\ e)$
$-y(x, y) = jarak\ euclidean\ antar\ x\ dan\ y$
x $= nilai\ data\ latih$
y $= nilai\ data\ uji$

4. Kernel Polynomial

$$K(x, y) = (x, y + c)^d \quad (4)$$

Keterangan
$K(x, y) = nilai\ kernel\ data\ x\ dan\ y$
x $= nilai\ data\ latih$
y $= nilai\ data\ uji$
$d = derajat$

5. Kernel Sigmoid

$$K(x, x') = tanh\frac{T}{i}(x + r) \quad (5)$$

Keterangan
$K(x, y) = nilai\ kernel\ data\ x\ dan\ y$
tanh $= aktivitasi\ sigmoid\ hiperbolik$
x $= vektor\ data$
$r = bias$

C. YouTube dan *Shorts*

YouTube adalah situs web terbesar dan paling populer di internet untuk berbagi video secara *online* [10]. *Shorts* merupakan video *YouTube* yang lebih pendek dengan durasi antara 15 detik sampai 1 menit dengan format video yang vertikal [11]

D. SenticNet

SenticNet adalah kamus dari kumpulan kata dengan sentimen positif dan negatif. Lalu menghitung total *polarity score* dengan menjumlahkan nilai dari keseluruhan kata. Sentimen positif diberikan nilai lebih dari nol, dan sentimen negatif diberikan nilai kurang dari nol [12].

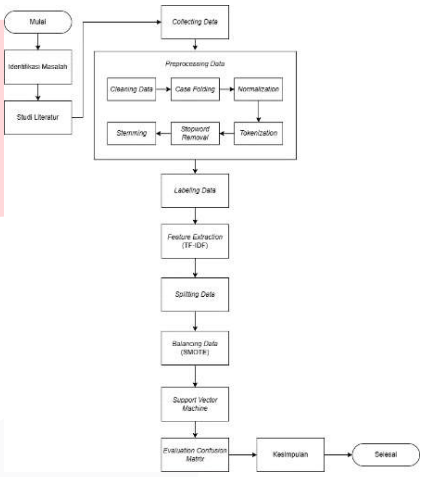
E. Synthetic Minority Over-Sampling Technique (SMOTE)

SMOTE adalah teknik penyeimbangan data untuk menangani permasalahan ketidakseimbangan kelas. Cara kerja dari SMOTE adalah menambahkan data sintetik pada kelas minoritas, tanpa mengurangi jumlah data yang sudah ada [13].

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

TF-IDF merupakan metode untuk pembobotan kata, dengan mengukur seberapa penting suatu kata dalam dokumen. semakin sering suatu kata muncul dalam satu dokumen, semakin besar kontribusinya. Namun, semakin sering suatu kata muncul di keseluruhan dokumen, maka kontribusinya akan menurun [14].

III. METODE



GAMBAR 1
(ALUR PENELITIAN)

Penelitian dimulai dengan identifikasi masalah terhadap tanggapan atau opini masyarakat khususnya pengguna *YouTube* terhadap PMBG. Pengambilan data dilakukan dengan *crawling* menggunakan *YouTube* API dan ID video, yang berhasil mendapatkan 2700 komentar pengguna *YouTube* terhadap PMBG dari 5 video yang diunggah di *channel* berita nasional. Setelah mendapatkan data, dilakukan proses *preprocessing* yang mencakup *cleaning*, *case folding*, *normalization*, *tokenization*, *stopword removal*, dan *stemming* untuk membersihkan data agar siap digunakan dalam analisis sentimen. Kemudian, seluruh data yang telah dibersihkan diberi label sentimen secara otomatis dengan bantuan kamus *SenticNet* untuk memberikan *polarity score* pada setiap komentar. Selanjutnya, tahap TF-IDF dilakukan untuk memberi bobot pada setiap kata dan mentranformasi format data yang awalnya teks menjadi numerik. Data dibagi menjadi data latih dan data uji dengan dua skenario, yaitu skenario rasio 80:20, dan skenario rasio 90:10. Setelah melakukan pembagian data, ditemukan adanya masalah ketidakseimbangan data, sehingga dilakukan penyeimbangan data menggunakan SMOTE. Selanjutnya, pengklasifikasikan data melalui pendekatan SVM dengan menggunakan bantuan empat kernel, yaitu linear, RBF, polynomial, dan sigmoid. Hasil dari skenario pemodelan dievaluasi menggunakan *confusion matrix* agar mendapatkan nilai *accuracy*, *precission*, *recall*, dan *f1-score* untuk memberikan gambaran yang jelas terkait kemampuan setiap kernel dalam mengklasifikasikan data.

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data penelitian ini menggunakan 2700 komentar pengguna *YouTube* terhadap PMBG yang diambil dari 5 video yang diunggah di *channel* berita nasional. Setelah dilakukan *preprocessing*, jumlah data yang benar-benar bersih menjadi 2688 data komentar.

TABEL 1
(HASIL CRAWLING)

Text
Luar biasa 🙂👍
Mulai kemakan rayuan
Program konyall
Penipu ulung

B. Preprocessing

Tahap preprocessing dilakukan dengan tujuan membersihkan 2700 data, karena masih terdapat data yang mengandung kata-kata atau karakter tidak relevan dengan yang dibutuhkan untuk analisis sentimen. Tahap dari *preprocessing* mencakup *cleaning data, case folding*, *normalization*, *tokenization*, *stopword removal*, dan *stemming*.

a) Cleaning Data

Tahap *cleaning data* dilakukan untuk menghapus tanda baca, angka, spasi berlebih, tautan karena dapat menganggu nilai dari sebuah teks [15].

TABEL 2
(HASIL CLEANING DATA)

Sebelum	Sesudah
Saya sangat setuju program ini, utk anak2 di sekolah khususnya di Papua. Sebab makanan bergizi akan meningkatkan IQ anak akan lebih baik, seperti anak2 org barat yg pola makannya diatur. Kalau para org tua tdk perlu makan gratis.	Saya sangat setuju program ini utk anak di sekolah khususnya di Papua Sebab makanan bergizi akan meningkatkan IQ anak akan lebih baik seperti anak org barat yg pola makannya diatur Kalau para org tua tdk perlu makan gratis

b) Case Folding

Tahap *case folding* bertujuan mengubah seluruh huruf dalam komentar menjadi huruf kecil atau seluruh huruf kapital agar format penulisan menjadi konsisten dan seragam [16]. Tanpa proses ini, kata yang sama tetapi ditulis dengan format yang berbeda akan dianggap sebagai kata yang berbeda oleh algoritma.

TABEL 3
(HASIL CASE FOLDING)

Sebelum	Sesudah
Saya sangat setuju program ini utk anak di sekolah khususnya di Papua Sebab makanan bergizi akan meningkatkan IQ anak akan lebih baik seperti anak org barat yg pola makannya diatur Kalau para org tua tdk perlu makan gratis.	saya sangat setuju program ini utk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak org barat yg pola makannya diatur kalau para org tua tdk perlu makan gratis

c) Normalization

Tahap *normalization* merupakan proses mengonversi teks ke dalam format yang konsisten serta lebih mudah diolah, seperti mengoreksi kata-kata yang salah atau disingkat menjadi bentuk yang benar dan baku [17].

TABEL 4
(HASIL NORMALIZATION)

Sebelum	Sesudah
saya sangat setuju program ini utk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak org barat yg pola makannya diatur kalau para org tua tdk perlu makan gratis	saya sangat setuju program ini untuk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak orang barat yang pola makannya diatur kalau para orang tua tidak perlu makan gratis

d) Tokenization

Tahap tokenizing merupakan proses membersihkan data dengan cara memecah rangkaian kata atau istilah individu yang berdiri sendiri [18].

TABEL 5
HASIL TOKENIZATION

Sebelum	Sesudah
saya sangat setuju program ini untuk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak orang barat yang pola makannya diatur kalau para orang tua tidak perlu makan gratis	['saya', 'sangat', 'setuju', 'program', 'ini', 'untuk', 'anak', 'di', 'sekolah', 'khususnya', 'di', 'papua', 'sebab', 'makanan', 'bergizi', 'akan', 'meningkatkan', 'iq', 'anak', 'akan', 'lebih', 'baik', 'seperti', 'anak', 'orang', 'barat', 'yang', 'pola', 'makannya', 'diatur', 'kalau', 'para', 'orang', 'tua', 'tidak', 'perlu', 'makan', 'gratis']

e) Stopword Removal

Tahap *stopword removal* merupakan untuk menghilangkan kata-kata yang sering muncul namun tidak memberikan kontribusi besar terhadap proses analisis data [16].

TABEL 6
(HASIL STOPWORD REMOVAL)

Sebelum	Sesudah
saya sangat setuju program ini untuk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak orang barat yang pola makannya diatur kalau para orang tua tidak perlu makan gratis	sangat setuju program untuk anak sekolah khususnya papua makanan bergizi meningkatkan iq anak lebih baik anak orang barat pola makannya diatur kalau orang tua perlu makan gratis

f) Stemming

Tahap *stemming* adalah tahap terakhir dari *preprocessing*, di mana teks diolah dengan menghapus awalan dan akhiran dari sebuah kata untuk mendapatkan versi dasar dari kata. Tujuannya untuk menyederhanakan keragaman kata. Sehingga, kata dengan arti yang sama namun dengan bentuk berbeda akan dianggap sebagai satu bentuk yang sama [16].

TABEL 7
(HASIL STEMMING)

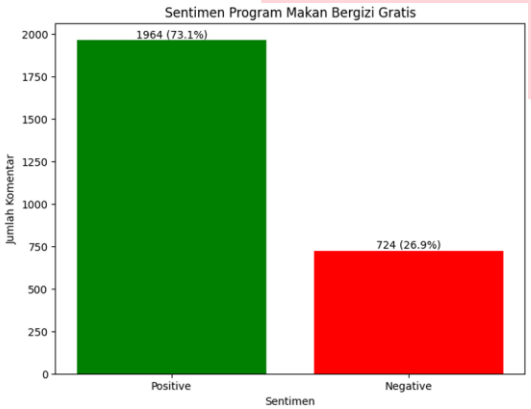
Sebelum	Sesudah
sangat setuju program untuk anak di sekolah khususnya di papua sebab makanan bergizi akan meningkatkan iq anak akan lebih baik seperti anak orang barat yang pola makannya diatur kalau para orang tua tidak perlu makan gratis	sangat tuju program untuk anak sekolah khusus papua makan gizi tingkat iq anak lebih baik anak orang barat pola makan atur kalau orang tua perlu makan gratis

C. Pelabelan Data

Tahap pelabelan data adalah langkah data diberi label tertentu untuk menunjukkan sentimen yang terkandung di dalamnya. Proses pelabelan melibatkan penentuan apakah setiap data seharusnya masuk dalam label sentimen positif atau negatif [16]. Penelitian ini menggunakan kamus *SenticNet* untuk memberikan dan menghitung *polarity score* dengan menjumlahkannya. Opini dengan sentimen positif diberikan nilai 1 atau lebih dari nol, sebaliknya, opini sentimen negatif diberikan nilai -1 atau kurang dari nol [12].

TABEL 8
(HASIL PELABELAN DATA)

Text	*Polarity Score*	Sentimen
dukung program pimpin selagi baik	1,536	Positif
perut kenyang koruptor	1,892	Negatif



GAMBAR 2
(DISTRIBUSI DATA SENTIMEN)

D. Splitting Data

Tahapan dalam membagi data menjadi dua subset yang terpisah untuk keperluan analisis atau pelatihan model dikenal sebagai *splitting data* [19]. Pada penelitian ini, *splitting data* dilakukan menggunakan dua skenario, yaitu 80:20 dan 90:10.

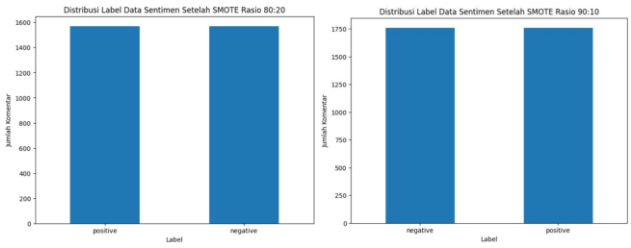
```
Original class distribution:
polarity
positive    1964
negative     724
Name: count, dtype: int64

Training set class distribution:
polarity
positive    1570
negative     580
Name: count, dtype: int64

Testing set class distribution:
polarity
positive    394
negative    144
```

```
Original class distribution:
polarity
positive    1964
negative     724
Name: count, dtype: int64

Training set class distribution:
polarity
positive    1759
negative     660
Name: count, dtype: int64

Testing set class distribution:
polarity
positive    205
negative     64
```

GAMBAR 3
SPLITTING DATA SKENARIO 80:20 DAN 90:10

E. Balancing Data

SMOTE digunakan pada penelitian ini untuk mengatasi ketidakseimbangan data pada hasil proses *splitting data*. Data yang tidak seimbang akan diterapkan SMOTE untuk menghasilkan data sintetik dengan membuat sampel baru dari kelas minoritas agar seimbang dengan kelas mayoritas.



GAMBAR 4
(HASIL DISTRIBUSI SMOTE)

F. Pemodelan SVM

Algoritma SVM digunakan untuk pemodelan klasifikasi data komentar pengguna *YouTube* terhadap PMBG dengan memakai empat kernel, yaitu linear, RBF, polynomial, dan sigmoid. Terdapat empat skenario pengujian algoritma SVM pada penelitian ini seperti di bawah:

TABEL 9
(SKENARIO PENGUJIAN SVM)

Skenario	Kernel Terbaik	*Accuracy*
Rasio 80:20 SMOTE	Linear	86%
Rasio 80:20 Tanpa SMOTE	Sigmoid	86%
Rasio 90:10 SMOTE	Linear	89%
Rasio 90:10 Tanpa SMOTE	Sigmoid	91%

G. Evaluasi Model SVM

Skenario rasio 80:20 SMOTE, kernel linear memperoleh *accuracy* sebesar 86% dari total 538 data uji. Pada kelas negatif, model *precision* negatif 0.71, *recall* 0.78, dan *f1score* 0.74. Sementara, performa pada kelas positif lebih unggul dengan *precision* 0,92, *recall* 0,88, dan *f1-score* 0,90 yang menunjukkan kemampuan model dalam mengenali kelas positif secara akurat.

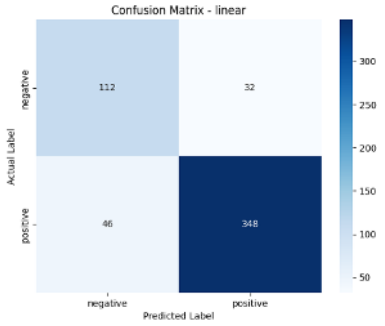
Skenario rasio 80:20 tanpa SMOTE, kernel sigmoid memperoleh *accuracy* sebesar 86% dari total 538 data uji. Pada kelas negatif, model *precision* negatif 0.79, *recall* 0.62, dan *f1-score* 0.70. Sementara, performa pada kelas positif lebih unggul dengan *precision* 0.87, *recall* 0,94, dan *f1-score* 0.90.

Skenario rasio 90:10 SMOTE, kernel linear memperoleh *accuracy* sebesar 89% dari total 269 data uji. Pada kelas negatif, model *precision* negatif 0.72, *recall* 0.88, dan *f1score* 0.79. Pada kelas positif menghasilkan *precision* 0,96, *recall* 0,89, dan *f1-score* 0,92 yang menunjukkan kemampuan model sangat baik dalam mengenali kelas positif dan cukup baik dalam mengenali kelas negatif.

Pada skenario rasio 90:10 tanpa SMOTE, kernel sigmoid memperoleh *accuracy* mencapai 91%. Performa model pada kelas kelas negatif mendapatkan nilai *precision* 0,82, *recall* 0,72, dan *f1-score* 0,79. Sementara, pada kelas positif nilai *precision* 0,92, *recall* 0,97, dan *f1-score* 0,94. Hasil tersebut menunjukkan bahwa model sangat baik dalam mengenali kedua kelas.

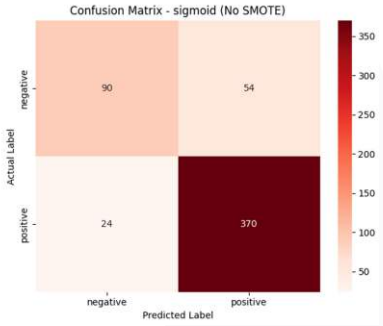
Hasil *confusion matrix* dari kernel linear pada skenario rasio 80:20 SMOTE menunjukkan dari hasil pengujian terhadap 538 data uji, 112 data negatif berhasil diklasifikasikan dengan benar, sementara 32 data negatif salah diklasifikasikan sebagai data positif. Untuk kelas positif, 348 data diklasifikasikan dengan benar, dan 46 data

salah diklasifikasikan sebagai negatif. Hal tersebut menunjukkan bahwa kernel linear pada skenario rasio 80:20 SMOTE mempunyai performa yang cukup baik.



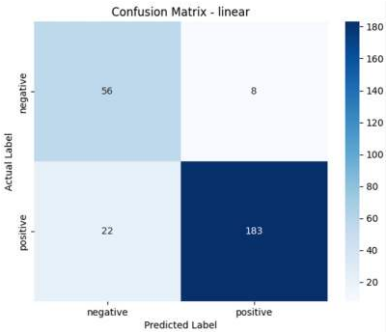
GAMBAR 5 :
(CONFUSION MATRIX KERNEL LINEAR 80:20)

Hasil *confusion matrix* dari kernel sigmoid pada skenario rasio 80:20 tanpa SMOTE menunjukkan dari hasil pengujian terhadap 538 data uji, 90 data negatif berhasil diklasifikasikan dengan benar, 54 data negatif salah diklasifikasikan sebagai data positif. Untuk kelas positif, 370 data diklasifikasikan dengan benar, dan 24 data salah diklasifikasikan sebagai negatif. Hal tersebut menunjukkan bahwa kernel sigmoid pada skenario rasio 80:20 tanpa SMOTE cukup baik dan simbang dalam mengklasifikasikan kedua kelas.

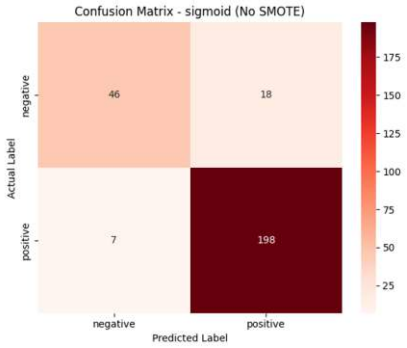


GAMBAR 6 :
(CONFUSION MATRIX KERNEL SIGMOID 80:20)

Hasil *confusion matrix* dari kernel linear pada skenario rasio 90:10 SMOTE menunjukkan dari hasil pengujian terhadap 269 data uji, 56 data negatif berhasil diklasifikasikan dengan benar, 8 data negatif salah diklasifikasikan sebagai data positif. Untuk kelas positif, 183 data diklasifikasikan dengan benar, dan 22 data salah diklasifikasikan sebagai negatif. Hal tersebut menunjukkan bahwa kernel linear pada skenario rasio 90:10 SMOTE mempunyai performa yang seimbang pada kedua kelas.



Hasil *confusion matrix* dari kernel sigmoid pada skenario rasio 90:10 tanpa SMOTE menunjukkan dari hasil pengujian terhadap 269 data uji, 46 data negatif berhasil diklasifikasikan dengan benar, 18 data negatif salah diklasifikasikan sebagai data positif. Untuk kelas positif, 198 data diklasifikasikan dengan benar, dan 7 data salah diklasifikasikan sebagai negatif. Hal tersebut menunjukkan bahwa kernel linear pada skenario rasio 90:10 SMOTE mempunyai performa yang seimbang pada kedua kelas secara akurat.



GAMBAR 8
(CONFUSION MATRIX KERNEL SIGMOID 90:10)

V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan telah dilakukan, berikut disajikan kesimpulan dari penelitian ini:

1. Hasil analisis sentimen pengguna *YouTube* terhadap PMBG menggunakan SVM berhasil mengklasifikasikan komentar dengan sentimen positif sebesar 73,1% dan sentimen negatif sebesar 26,9%. Hasil tersebut menunjukkan bahwa masyarakat merespon positif program tersebut.
2. Berdasarkan hasil evaluasi algoritma SVM, skenario rasio 80:20 menggunakan SMOTE menunjukkan kernel linear memiliki performa terbaik dengan tingkat akurasi 86%. Skenario rasio 80:20 tanpa SMOTE, kernel sigmoid menghasilkan nilai akurasi tertinggi dengan nilai 86%. Pada skenario rasio 90:10 menggunakan SMOTE, kernel linear kembali menghasilkan nilai akurasi tertinggi sebesar 89%. Dan pada skenario rasio 90:10 tanpa SMOTE, kernel sigmoid memiliki akurasi terbaik dengan nilai 91%.

REFERENSI

[1] A. Rachman, “Makan Bergizi Gratis Dimulai 2 Januari 2025, Siapa Saja Penerimanya?,” CNBC Indonesia.

[2] P. A. Maharani, A. R. Namira, and T. V. Chairunnisa, “Peran Makan Siang Gratis Dalam Janji Kampaye Prabowo Gibran Dan Realisasinya,” *Jolasos J. Soc. Soc.*, vol. 1, no. 1, pp. 1–10, 2024.

[3] H. S. W. Hovi, A. Id Hadiana, and F. Rakhmat Umbara, “Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM),”

Informatics Digit. Expert, vol. 4, no. 1, pp. 40–45, 2022.

[4] A. Puji Astuti, S. Alam, and I. Jaelani, "Komparasi Algoritma Support Vector Machine dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo," *J. Bangkit Indones.*, vol. 11, no. 2, pp. 1–6, 2022.

[5] A. M. Ndapamuri, D. Manongga, and A. Iriani, "Analisis Sentimen Ulasan Aplikasi Tripadvisor Dengan Metode Support Vector Machine, K-Nearest Neighbor, Dan Naive Bayes," *INOVTEK Polbeng - Seri Inform.*, vol. 8, no. 1, pp. 127–140, 2023.

[6] C. F. Hasri and D. Alita, "Penerapan Metode Naive Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022.

[7] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, "Film Review Sentiment Analysis Using TF-IDF and Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.

[8] M. Muttaqin and K. Iqbal, "Analisis Sentimen Aplikasi Gojek Menggunakan Support Vector Machine Dan K Nearest Neighbor," *Ujme*, vol. 3, no. 3, pp. 22–27, 2021.

[9] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Comparative Evaluation of SVM Kernels for Sentiment Classification in Fuel Price Increase Analysis," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023.

[10] M. Ardiansyah and M. L. Nugraha, "Analisis Pemanfaatan Media Pembelajaran Youtube," *Semin. Nas. Ris. dan Inov. Teknol. (SEMNAS RISTEK)*, pp. 912–918, 2022.

[11] A. T. Tantika *et al.*, "Data Pengguna Youtube di Indonesia dari Tahun 2019 - 2023," *J. Multidisiplin Ilmu*, vol. 3, no. 1, pp. 65–71, 2024.

[12] O. Manullang, C. Prianto, and N. H. Harani, "Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based Dan Random Forest," *J. Ilm. Inform.*, vol. 11, no. 02, pp. 159–169, 2023.

[13] L. Qadrini, H. Hikmah, and M. Megasari, "Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejawa Timur Tahun 2017," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 386–391, 2022.

[14] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 2, pp. 422–430, 2021.

[15] A. A. Kurniawan and M. Mustikasari, "Implementasi Deep Learning Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, pp. 544–552, 2021.

[16] N. M. Farhan and B. Setiaji, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 3925–3940, 2023.

[17] Y. Veren, S. Berutu, and H. Budiati, "Penerapan Metode Decision Tree Pada Sentimen Media Sosial Terkait Komisi Pemilihan Umum (KPU) Sebelum Dan Sesudah 2024," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 4, pp. 2174–2184, 2024.

[18] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. Tekno Kompak*, vol. 15, no. 1, pp. 131–145, 2021.

[19] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 273–281, 2024.