

# Analisis Sentimen Terhadap Penggunaan Artis Korea Selatan Sebagai *Brand Ambassador* Pada Produk Makanan Dan Minuman Lokal Menggunakan Algoritma Support Vector Machine (Svm)

1<sup>st</sup> Endar Pariswara  
Fakultas Rekayasa Industri  
Universitas Telkom  
Bandung, Indonesia

endarpariswara@student.telkomunivers  
ity.ac.id

2<sup>nd</sup> Dita Pramesti  
Fakultas Rekayasa Industri  
Universitas Telkom  
Bandung, Indonesia

ditapramesti@telkomuniversity.ac.id

3<sup>rd</sup> Riska Yanu Fa'rifah  
Fakultas Rekayasa Industri  
Universitas Telkom  
Bandung, Indonesia

riskayanu@telkomuniversity.ac.id

**Abstrak** — Penggunaan *brand ambassador* pada suatu produk dinilai dapat meningkatkan daya jual. Biasanya perusahaan menggunakan artis atau *public figure* sebagai *brand ambassador*. Penggunaan artis atau *public figure* bukanlah tanpa suatu alasan. *Public figure* yang memiliki banyak penggemar atau pengikut di berbagai media sosialnya dapat dijadikan sebagai target dalam melakukan pemasaran. Maraknya fenomena *Korean wave* yang sedang terjadi di Indonesia membuat produk lokal saling bersaing dalam menjadikan artis asal Korea Selatan sebagai *brand ambassador*. Penggunaan artis Korea Selatan sebagai *brand ambassador* produk lokal menuai komentar positif maupun negatif dari masyarakat Indonesia. Komentar-komentar tersebut banyak dituangkan oleh masyarakat di berbagai media sosial, contohnya pada media sosial Twitter. Terhadap komentar tersebut akan dilakukan analisis untuk mengetahui sentimen masyarakat mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal. Pada penelitian ini akan digunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasi setiap data ke dalam sentimen positif dan sentimen negatif menggunakan metode *imbalance handling* seperti SMOTE untuk *oversampling*, dan RUS untuk *undersampling*. Setelah melakukan proses pengklasifikasian, selanjutnya diterapkan evaluasi menggunakan *confusion matrix*. Penelitian ini menghasilkan model SVM terbaik dengan penerapan metode SMOTE dengan nilai *accuracy* tertinggi pada fold ke-10 mencapai 83,89%. Selain itu, terdapat nilai *recall* sekitar 80%, *precision* 85%, dan *F1-Score* 82%.

**Kata kunci**— *Brand Ambassador*, Twitter, Analisis Sentimen, Support Vector Machine (SVM), SMOTE, RUS

## I. PENDAHULUAN

*Korean wave* sedang marak-maraknya terjadi di Indonesia. Pada 28 Februari 2022 Statista Research Department merilis sebuah artikel berjudul “Popularity of South Korean pop music (K-Pop) in Indonesia 2019”. Pada

artikel ini dijelaskan data statistik mengenai kepopuleran musik korea di Indonesia, seperti disajikan pada Gambar I.1, didapatkan sebanyak 60% *very popular*, 30% *quite popular*, 9% *popular only for a new people*, dan sebanyak 2% *not really*.

Dengan tingginya fenomena *Korean wave* ini banyak produk makanan dan minuman lokal yang berlomba-lomba dalam menjadikan artis asal Korea Selatan sebagai *brand ambassador* guna mempromosikan produknya. Contohnya seperti NCT Dream yang menjadi *brand ambassador* dari Mie Lemonilo, salah satu anggota grup senior Korea Selatan Siwon Super Junior yang menjadi *brand ambassador* Mie Sedap dan Sasa Santan Omega3, NCT 127 yang menjadi *brand ambassador* dari produk minuman lokal NU Green Tea, Lucas WayV yang menjadi *brand ambassador* salah satu minuman terkenal di Indonesia Neo Coffee, dan yang baru-baru ini grup wanita ITZY ditunjuk sebagai *brand ambassador* dari produk susu kemasan Ultramilk.

Keputusan beberapa *brand* lokal dalam menjadikan artis Korea Selatan sebagai *brand ambassador* untuk mempromosikan produk makanan dan minuman lokal ternyata menuai opini dari masyarakat Indonesia. Dalam kegiatan menyampaikan opininya di berbagai media sosial, salah satu media sosial yang banyak digunakan adalah Twitter. Dari opini-opini pada media sosial Twitter dapat dilakukan analisis sentimen untuk mengetahui pendapat masyarakat mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk lokal.

Slamet (2022) telah melakukan penelitian mengenai Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai *Brand Ambassador* Produk Kecantikan Lokal. Pada penelitian tersebut digunakan algoritma Support Vector Machine (SVM). SVM merupakan algoritma yang memiliki performa terbaik dibandingkan algoritma klasifikasi lainnya. Pada penelitian tersebut SVM menghasilkan sentimen positif sebanyak 266 data dan

sentimen negatif sebanyak 51 data. Penelitian ini menghasilkan nilai *accuracy* sebesar 83,60%, *precision* 83,86%, dan *recall* 99,62%. Kesimpulannya penggunaan artis Korea Selatan sebagai *brand ambassador* dari produk kecantikan lokal banyak menuai sentimen positif dari masyarakat pengguna Twitter.

Hal tersebut menjadi pendorong untuk melakukan penelitian mengenai analisis sentimen penggunaan artis Korea Selatan sebagai *brand ambassador* pada produk makanan dan minuman lokal. Penelitian ini juga akan menggunakan metode klasifikasi. Algoritma klasifikasi yang akan digunakan adalah Support Vector Machine (SVM), yang akan menghasilkan data komentar menjadi opini negatif dan positif.

## II. KAJIAN TEORI

Pada bagian ini dijelaskan dasar-dasar teori yang menjadi bahan acuan untuk melakukan penelitian ini. Teori yang dibahas di antaranya teori mengenai objek dari penelitian, pendekatan yang digunakan, dan teori mengenai metode-metode untuk melakukan analisis serta *tools* yang akan digunakan nantinya untuk penelitian ini.

### A. Brand Ambassador

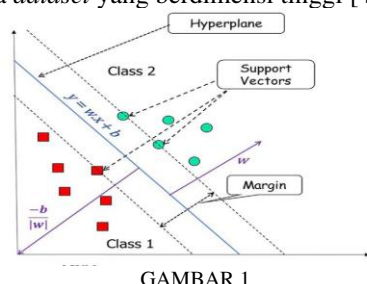
Salah satu usaha yang dilakukan untuk menarik minat pembeli yaitu dengan menggunakan *brand ambassador* untuk membantu memperkenalkan produk kepada calon pembeli. *Brand ambassador* nantinya dapat membantu menciptakan ikatan yang lebih mendalam antara merek dengan calon konsumennya, sehingga secara tidak langsung *brand ambassador* dapat membangun citra produk yang akan berdampak kepada keputusan pembelian produk [2]. Biasanya yang ditunjuk sebagai *brand ambassador* ialah artis untuk menjalankan iklan guna meningkatkan jumlah penjualan dan yang paling utama untuk menarik minat penggemarnya.

### B. Data Preprocessing

*Preprocessing* perlu dilakukan sebelum proses klasifikasi agar dimensi data menjadi lebih kecil. Hal itu akan membuat proses klasifikasi menjadi lebih cepat. Tujuan dari *preprocessing* ini adalah untuk menghilangkan dan membersihkan kata-kata yang tidak sesuai dengan penelitian, menyeragamkan bentuk kata, dan memperkecil volume kata [3].

### C. Support Vector Machine (SVM)

Konsep klasifikasi dari Support Vector Machine adalah mencari *hyperplane* terbaik yang memisahkan dua kelas kata yaitu positif dan negatif. Support Vector Machine mampu bekerja pada *dataset* yang berdimensi tinggi [4].



GAMBAR 1

Ilustrasi Hyperplan dan Support Vector Machine [5]

Pada Gambar 1 di atas, terdapat dua label yaitu label kotak merah dan label lingkaran hijau. Kedua label tersebut dipisahkan oleh garis *hyperplane* yang dapat disebut juga dengan *decision boundary*. *Hyperplane* atau batas keputusan merupakan garis yang berfungsi untuk memisahkan satu kelas dengan kelas lainnya serta memaksimalkan *margin* dari label-label. *Hyperplane* digunakan untuk menentukan kemungkinan terbaik yang dimiliki oleh suatu data. *Support vector* adalah titik-titik dari masing-masing kelas yang berada paling dekat dengan *hyperplane*. *Margin* merupakan jarak antara *hyperplane* dan data *support vector* terdekat.

Persamaan *hyperplane* pada Support Vector Machine dapat diselesaikan menggunakan Persamaan II.1 berikut:

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b \quad (\text{II.1})$$

Keterangan:

- $\mathbf{w}$  : vektor bobot (*weight*).
- $\mathbf{x}$  : data latih berupa vektor yang berisikan nilai-nilai dari atribut.
- $b$  : bias yang berupa skalar.

### D. Imbalance Handling

Penanganan ketidakseimbangan (*imbalance handling*) merujuk pada strategi atau teknik yang digunakan untuk mengatasi ketidakseimbangan dalam data pelatihan. Ketidakseimbangan terjadi saat jumlah sampel di kelas target yang berbeda tidak seimbang atau tidak proporsional. Ketidakseimbangan kelas dapat menyebabkan masalah dalam pembelajaran mesin karena model cenderung memiliki kecenderungan untuk memprioritaskan dan memberikan hasil yang lebih baik pada kelas mayoritas yang lebih dominan, sementara mengabaikan atau kurang akurat dalam memprediksi kelas minoritas yang penting. Oleh karena itu, penanganan ketidakseimbangan menjadi sangat penting agar model pembelajaran mesin dapat secara adil dan efektif mempelajari pola-pola dari semua kelas target. Metode-metode seperti RUS (*Random Undersampling*) dan SMOTE (*Synthetic Minority Over-sampling Technique*) adalah beberapa contoh teknik *imbalance handling* yang umum digunakan.

### E. K-Fold Cross validation

*K-Fold Cross Validation* merupakan metode validasi yang digunakan untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi data independen. Fungsi utama penggunaan metode ini untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model ketika diimplementasikan. Penggunaan *k-fold cross validation* adalah untuk menghilangkan bias pada data [6]. Metode ini membantu dalam mengukur seberapa baik model akan berperforma pada data yang belum pernah dilihat sebelumnya. Cara kerja metode ini adalah melakukan pembagian data menjadi data uji dan data latih.

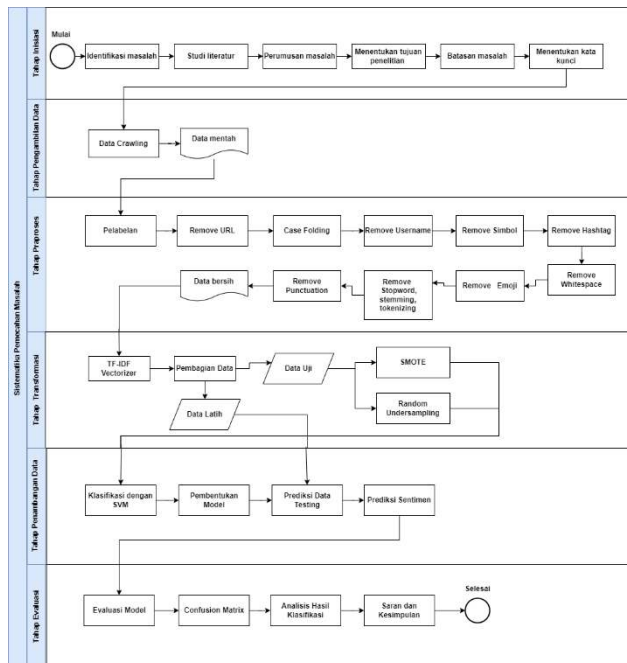
### F. Confusion Matrix

*Confusion matrix* adalah suatu instrumen yang digunakan dalam melakukan evaluasi terhadap performa dari model klasifikasi yang telah dihasilkan sebelumnya. *Confusion matrix* terdiri dari empat kategori, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dalam instrumen ini hasil kelas yang sudah diprediksi

akan dibandingkan dengan kelas data yang sebenarnya [3]. Hasil tersebut akan digunakan untuk melakukan perhitungan nilai *accuracy*, *precision*, *recall*, dan *F1-Score*.

### III. METODE

Pada bagian ini dijelaskan mengenai kerangka pemecahan masalah yang digunakan, membahas sistematika penyelesaian masalah, pengumpulan data, pengolahan data atau proses pengembangan produk/ artifak, serta membahas metode evaluasi yang digunakan.



GAMBAR 2  
Kerangka Pemecahan Masalah

#### 1. Tahap Inisiasi

Pada tahap ini dilakukan studi literatur untuk dapat mengidentifikasi serta merumuskan masalah yang menjadi objek pada penelitian ini yaitu penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal. Setelah melakukan identifikasi masalah, studi literatur, dan merumuskan masalah langkah selanjutnya adalah menentukan tujuan dilakukannya penelitian setelah itu menentukan hal-hal yang menjadi batasan dalam melakukan penelitian.

#### 2. Tahap Pengambilan Data

Pada tahap ini dilakukan pengumpulan data berupa komentar masyarakat mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal pada media sosial Twitter. Pengumpulan data tersebut dilakukan dengan cara *crawling* menggunakan Python dan *library snsrape*. Adapun rentang waktu pengambilan data dimulai pada 01 Agustus 2022 sampai 31 Desember 2022 dan menghasilkan sebanyak 18.484 data. Adapun kata kunci yang digunakan dalam pengambilan data seperti “lemonilo NCT”, “Nu green tea NCT”, “ultra milk ITZY”, “mie sedap siwon”, “lemonilo photocard”, dan “ultra milk photocard”. Kata kunci tersebut dipilih karena relevan dan berkaitan dengan tujuan akhir penelitian.

#### 3. Tahap Praproses

Setelah mendapatkan sebanyak 18.484 data dengan beberapa kata kunci, kemudian dilakukan pelabelan pada setiap datanya. Adapun label yang digunakan adalah label positif dan label negatif. Kemudian dilakukan *remove URL*, *case folding*, *remove username*, *remove simbol*, *remove hashtag*, *remove shitesapce*, *remove emoji*, *remove stop word*, *stemming*, *tokenizing*, dan *remove punctuation*.

#### 4. Tahap Transformasi

Setelah didapatkan data yang sudah benar benar bersih, pada tahap ini dilanjutkan dengan melakukan pembobotan kata menggunakan metode TF-IDF, selanjutnya *dataset* tersebut dilakukan pembagian data menjadi data uji dan data laatih. Setelah itu, diimplementasikan metode *imbalace handling* yaitu RUS dan SMOTE.

#### 5. Tahap Penambahan Data

Pada tahap ini digunakan algoritma SVM dalam melakukan klasifikasi sentimen dengan pembentukan model untuk memprediksi data testing dan memprediksi sentimen.

#### 6. Tahap Evaluasi

Pada tahap ini dilakukan evaluasi performa model berdasarkan *confusion matrix* yang di dalamnya memiliki komponen nilai *accuracy*, *recall*, *precision*, dan *F1-Score*. Dengan menggunakan *confusion matrix* sebagai alat untuk membantu evaluasi penelitian, peneliti dapat menghitung nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dengan melalui proses evaluasi maka kesimpulan akhir akan mudah ditarik terhadap hasil penelitian analisis sentimen masyarakat mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal

### IV. HASIL DAN PEMBAHASAN

#### A. Proses Klasifikasi SVM

Sebelum dilakukan model kasifikasi, dilakukan parameter tuning terlebih dahulu, yaitu menentukan parameter apa saja yang cocok dengan data yang digunakan seperti *kernel*, *C*, dan *gamma*. Proses parameter tuning ini menggunakan metode *gridsearch*. *Grid search* akan melakukan pemilihan nilai-nilai parameter yang cocok untuk data dan model yang digunakan. Adapun hasil dari *Grid search* yaitu *kernel* : *linear*, *C* : 1, dan *gamma* : 0,001.

*Grid search* memprediksi *kernel* terbaik untuk model SVM adalah *linear*. Hal tersebut dapat dikarenakan karakteristik data yang digunakan memiliki pola yang cukup *linear* dan dapat dengan baik dipisahkan oleh garis lurus, maka *kernel linear* dapat memberikan performa yang baik.

Kemudian untuk parameter *C* atau Cost Parameter *Grid search* memprediksi nilai *C* yaitu sebesar 1 (satu). Hal ini dikarenakan semakin tinggi nilai *C*, semakin ketat *margin* dalam pemisahan kelas dan semakin tinggi dalam memberikan terhadap kesalahan.

Kemudian untuk parameter *gamma*, *Grid search* memprediksi nilai *gamma* hanya sebesar 0,001. Hal tersebut dapat terjadi karena parameter *gamma* hanya relevan jika menggunakan *kernel RBF*, sedangkan *kernel* yang diprediksi oleh *Grid search* untuk model pada penelitian ini adalah *linear*. Maka dari itu nilai *gamma* yang dihasilkan *Grid search* terbilang sangat kecil.

Metode SVM untuk memperoleh *hyperplan* optimal yang memisahkan data ke dalam dua kelas. Berdasarkan

pengolahan data *training*, maka diperoleh persamaan *hyperplan* optimal untuk nilai  $k=10$

$$\begin{pmatrix} 0.272016 \\ 0.035536 \\ 0.126352 \\ 0.161121 \\ \vdots \\ 0.076894 \end{pmatrix}_{918 \times 1} \cdot X + (0,978796410518377) = 0$$

Dimana  $X$  adalah variabel independen yang merupakan vektor input yang berasal dari hasil TF-IDF. Berikutnya persamaan diatas direpresentasikan menjadi fungsi  $f(x) = w^T \cdot x + b$  dan digunakan untuk mengklasifikasikan data testing sebagai contoh dokumen ke-1 dan ke-2:

$$d_2 \begin{pmatrix} 0.272016 \\ \vdots \\ -0,081516 \\ \vdots \\ 0.076894 \end{pmatrix}_{918 \times 1} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0,422714 \\ \vdots \\ 0 \end{pmatrix} + (0,978796410518377) \geq 1$$

Karena  $f(x) \geq 1$  maka dokumen ke-1 diatas diklasifikasikan sebagai *tweet* positif.

$$d_2 \begin{pmatrix} 0.272016 \\ 0.035536 \\ 0.126352 \\ 0.161121 \\ \vdots \\ 0.076894 \end{pmatrix}_{918 \times 1} \cdot \begin{pmatrix} 0 \\ 0 \\ 0,255450 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (0,978796410518377) \leq 1$$

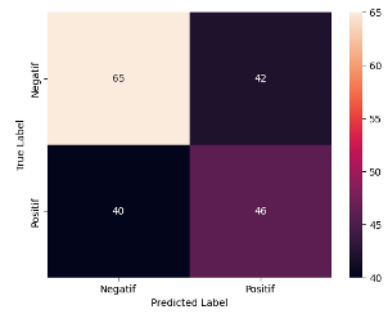
Karena  $f(x) \leq 1$  maka dokumen ke-2 diatas diklasifikasikan sebagai *tweet* negatif.

## B. Evaluasi Performa Model Menggunakan *Confusion Matrix*

*Confusion matrix* merupakan metode evaluasi untuk mengetahui komponen nilai-nilai yang dihasilkan oleh model seperti: True Positive (TP): Jumlah sampel yang benar diprediksi sebagai positif oleh model. True Negative (TN): Jumlah sampel yang benar diprediksi sebagai negatif oleh model. False Positive (FP): Jumlah sampel yang salah diprediksi sebagai positif oleh model. Serta False Negative (FN): Jumlah sampel yang salah diprediksi sebagai negatif oleh model. Berikut merupakan implementasi *confusion matrix* yang diterapkan pada metode SMOTE, dan *Random Undersampling*

### 1. *Random Undersampling* (RUS)

Berikut merupakan visualisasi dari *confusion matrix* dari metode RUS untuk nilai *accuracy* tertinggi yaitu pada *fold* ke-3.



GAMBAR 3  
Confusion Matrix Metode RUS

Pada Gambar 4 di atas, merupakan *confusion matrix* yang dihasilkan dari metode *K-Fold Cross Validation*. Model tersebut menghasilkan nilai TN sebesar 65 data, nilai FP sebesar 42 data, nilai FN sebesar 40 data, dan nilai TP sebesar 46 data.

$$Accuracy = \frac{TN+TP}{TP+FP+FN+TN} = \frac{65+46}{65+42+40+46} = \frac{111}{193} = 57,51\%$$

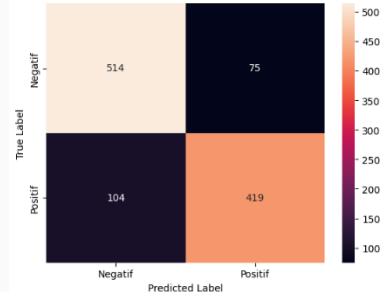
$$Precision = \frac{TP}{TP+FP} = \frac{46}{46+42} = \frac{46}{88} = 52\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{46}{46+40} = \frac{46}{86} = 53\%$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,52 \times 0,53}{0,52 + 0,53} = 53\%$$

### 2. *Synthetic Minority Oversampling Technique* (SMOTE)

Berikut merupakan visualisasi dari *confusion matrix* dari metode SMOTE untuk nilai *accuracy* tertinggi yaitu pada *fold* ke-10.



GAMBAR 4  
Confusion Matrix Metode SMOTE

Pada Gambar 5 di atas, merupakan *confusion matrix* yang dihasilkan dari metode *K-Fold Cross Validation*. Model tersebut menghasilkan nilai TN sebesar 514 data, nilai FP sebesar 75 data, nilai FN sebesar 104 data, dan nilai TP sebesar 419 data.

$$Accuracy = \frac{TN+TP}{TP+FP+FN+TN} = \frac{514+419}{419+75+104+514} = \frac{933}{1112} = 83,9\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{419}{419+75} = \frac{419}{494} = 85\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{419}{419+104} = \frac{419}{523} = 80\%$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,84 \times 0,80}{0,84 + 0,80} = 82\%$$

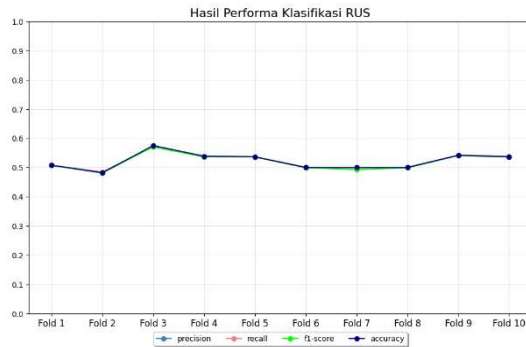


### C. Evaluasi Performa Model Menggunakan Metriks Penilaian Klasifikasi.

Metrik penilaian klasifikasi adalah ukuran yang digunakan untuk mengevaluasi performa model dalam melakukan klasifikasi pada data seperti nilai *accuracy*, *precision*, *recall*, dan *F1-Score*.

#### 1. Random Undersampling (RUS)

Berikut merupakan diagram batang yang menunjukkan hasil *accuracy*, *precision*, *recall*, dan *F1-Score* pada 10 fold dengan metode RUS.

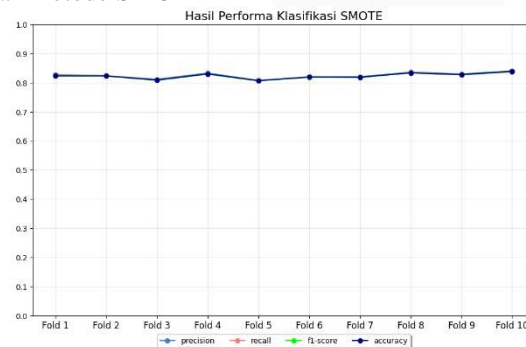


GAMBAR 5  
Hasil Performa Klasifikasi Metode RUS

Pada Gambar 7 di atas menunjukkan hasil performa klasifikasi dari metode RUS. Dapat dilihat bahwa fold ke-3 menghasilkan nilai *accuracy* tertinggi yaitu 57,5%, *recall* sebesar 57%, *precision* 57%, dan *F1-Score* sebesar 57%. Nilai *accuracy* terendah dihasilkan oleh fold ke-2 yaitu sebesar 48,19%, *recall* sebesar 48%, *precision* 48%, dan *F1-Score* sebesar 48%. Adapun rata-rata dari nilai *accuracy* 10 fold ini adalah sebesar 52,18%.

#### 2. Synthetic Minority Oversampling Technique (SMOTE)

Berikut merupakan diagram batang yang menunjukkan hasil *accuracy*, *precision*, *recall*, dan *F1-Score* pada 10 fold dengan metode SMOTE



GAMBAR 6  
Hasil Performa Klasifikasi Metode SMOTE

Pada Gambar 8 di atas menunjukkan hasil performa klasifikasi dari metode SMOTE. Dapat dilihat bahwa fold ke-10 menghasilkan nilai *accuracy* tertinggi yaitu 83,89%, *recall* sebesar 80%, *precision* 85%, dan *F1-Score* sebesar 82%. Nilai *accuracy* terendah dihasilkan oleh fold ke-5 yaitu sebesar 80,67%, *recall* sebesar 78%, *precision* 83%, dan *F1-Score* sebesar 80%. Adapun rata-rata dari nilai *accuracy* 10 fold ini adalah sebesar 82,28%.

TABEL 1  
Hasil Performa Model

	Accuracy	Recall	Precision
RUS	57,5%	57%	57%
SMOTE	83,89%	80%	85%

Berdasarkan Tabel V-2 di atas, maka dapat disimpulkan bahwa performa model dengan menggunakan metode RUS menghasilkan nilai *accuracy* sebesar 57,5%, *recall* sebesar 57%, *precision* sebesar 57%, dan *F1-Score* sebesar 57%. Kemudian untuk metode SMOTE menghasilkan nilai *accuracy* tertinggi sebesar 83,89%, *recall* sebesar 80%, *precision* sebesar 85%, dan *F1-Score* sebesar 82%.

### D. Visualisasi Wordcloud Terhadap Komentar Positif dan Negatif

Setelah dilakukan seluruh tahap untuk melakukan analisis sentimen mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal, berikut didapatkan persebaran kata (*wordcloud*).



GAMBAR 7  
Wordcloud Sentimen Positif

Pada Gambar 9 di atas menunjukkan *wordcloud* dengan sentimen positif. Persebaran kata untuk sentimen bernada positif frekuensi yang paling besar yaitu 'photocard', 'lemonilo', dan 'ultramilk' karena seluruh *tweet* yang dianalisa membicarakan mengenai produk yang mendapatkan photocard pada saat pembeliannya. Kemudian persebaran kata lainnya yaitu artis-artis yang menjadi *brand ambassador* produk seperti 'itzy', 'nct dream', dan 'siwon'. Persebaran kata mengenai produk makan dan minuman seperti 'lemonilo', 'mie sedap' dan 'ultamilk'. Adapun persebaran kata lainnya yang bersifat pujian seperti 'dapat' dan 'suka'.



GAMBAR 8  
Wordcloud Sentimen Negatif

Untuk *wordcloud* sentimen bernada negatif ditunjukkan pada Gambar 10 di atas. Persebaran frekuensi kata terbanyak

yaitu pada kata ‘tidak’, ‘photocard’, dan ‘lemonilo’ karena seluruh *tweet* dengan sentimen negatif berbicara tentang produk mie lemonilo dan juga tidak mendapatkan *photocard* pada saat pembelian produk. Adapun persebaran kata yang menunjukkan kekesalan para konsumen atas pembelian produk menggunakan kata umpatan seperti ‘anjing’.

## V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Dalam mencari kombinasi parameter yang dihasilkan teknik *grid search* untuk model SVM dalam melakukan analisis sentimen penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal, *grid search* menghasilkan kombinasi *kernel* = ‘linear’, *C* = ‘1’, dan *gamma* = ‘0,001’. Kombinasi dari parameter yang dihasilkan *grid search* ini merupakan kombinasi terbaik untuk model dalam penelitian ini.
2. Telah dilakukan klasifikasi menggunakan algoritma SVM dengan metode *imbalance handling* yaitu SMOTE, dan RUS. Hasil *accuracy* tertinggi diperoleh dengan metode SMOTE yaitu dengan rata-rata *accuracy* sebesar 82,28% sedangkan metode RUS menghasilkan rata-rata *accuracy* sebesar 52,18%.
3. Kata kunci yang relevan atau dominan dan yang sering muncul pada *dataset* terkait penggunaan artis Korea Selatan sebagai *brand ambassador* produk makanan dan minuman lokal adalah “photocard”, “itzy”, “Dapat Photocard”, “NCT dream”, “lemonilo”. Dapat diambil kesimpulan bahwa *photocard* menjadi taktik marketing yang digunakan *brand* makanan dan minuman dalam melakukan promosi dan target pemasaran mereka adalah masyarakat Indonesia yang menyukai hal-hal berbau K-Pop.

## REFERENSI

- [1] R. Slamet, W. Gata, A. Novtariany, K. Hilyati, dan F. Ainun jariyah, “ANALISIS SENTIMEN TWITTER TERHADAP PENGGUNAAN ARTIS KOREA SELATAN SEBAGAI BRAND AMBASSADOR PRODUK KECANTIKAN LOKAL,” vol. 5, 2022.
- [2] E. S. Hariandja dan U. P. Harapan, “THE INFLUENCE OF BRAND AMBASSADOR ON BRAND IMAGE AND CONSUMER PURCHASING DECISION: A CASE OF TOUS LES JOURS IN THE INFLUENCE OF BRAND AMBASSADOR ON BRAND IMAGE AND CONSUMER PURCHASING DECISION: local or international brands in the market market in Indon,” no. March, 2016.
- [3] W. Bourequat dan H. Mourad, “Sentiment Analysis Approach for Analyzing iPhone Release using Support Vector Machine,” *Int. J. Adv. Data Inf. Syst.*, vol. 2, no. 1, hal. 36–44, 2021, doi: 10.25008/ijadis.v2i1.1216.
- [4] U. Rofiqoh, R. S. Perdana, dan M. A. Fauzi, “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter dengan Metode Support Vector Machine dan Lexicon Based Features Twitter event detection View project Human Detection and Tracking View project,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1(12), no. October, hal. 1725–1732, 2017, [Daring]. Tersedia pada: <https://www.researchgate.net/publication/320234928>
- [5] A. Rani, N. Kumar, J. Kumar, dan N. K. Sinha, “Machine learning for soil moisture assessment,” *Deep Learn. Sustain. Agric.*, no. January, hal. 143–168, 2022, doi: 10.1016/B978-0-323-85214-2.00001-X.
- [6] F. Tempola, M. Muhammad, dan A. Khairan, “Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, hal. 577, 2018, doi: 10.25126/jtiik.201855983.