Analisis Sentimen Brand Ambassador Artis Korea Selatan pada Produk Indonesia dengan *Lexicon*

Galuh Etha Pratiwi1, Tiani Wahyu Utami2, Rochdi Wasono3

1, 2, 3Prodi Statistika Universitas Muhammadiyah Semarang

Corresponding author email: [tianiutami@unimus.ac.id](mailto:tianiutami@unimus.ac.id)

***Abstract:*** *The use of South Korean artists as brand ambassadors for Indonesian products is one way that e-commerce and local companies do to attract consumers. However, it reaped positive and negative responses from the public. The social media used to issue opinions regarding this matter is Twitter. This study conducts sentiment analysis from tweets on Twitter regarding this topic. The method used is Lexicon sentiment. The sentiment Lexicon works by first making a dictionary of opinion words and matching them with the sentiment words contained in the Lexicon dictionary. The Lexicon dictionary used is InSet Lexicon. Sentiment classification is divided into 2 namely, positive and negative. Data collection uses the Twitter API with the keyword &quot;BA Artis Korea&quot; as many as 366 data tweets. The research phase consists of data collection, text preprocessing, Lexicon sentiment, visualization, and testing. The results of the classification using the Lexicon method obtained 67.1% positive and 32.9% negative. It was found that the use of South Korean artists to become brand ambassadors for Indonesian products received positive sentiment. Accuracy, precision, and recall values ​​obtained by looking at the confusion matrix table are 78% accuracy, 68% precision, and 61% recall.*

***Keywords:*** *Brand Ambassador, Lexicon, Sentiment, Twitter*

**Abstrak:** Penggunaan artis Korea Selatan sebagai brand ambassador produk Indonesia merupakan salah satu cara yang dilakukan *e-commerce* dan perusahaan lokal untuk menarik konsumen. Namun, hal tersebut menuai tanggapan positif dan negatif dari masyarakat. Media sosial digunakan untuk mengeluarkan pendapat mengenai hal tersebut yaitu Twitter. Penelitian ini melakukan analisis sentimen dari cuitan pada Twitter mengenai topik tersebut. Metode yang digunakan *Lexicon sentiment*. *Lexicon sentiment* bekerja dengan membuat kamus kata opini terlebih dahulu dan mencocokkan dengan kata sentimen terdapat pada kamus *Lexicon*. Kamus *Lexicon* yang digunakan yaitu InSet *Lexicon*. Klasifikasi sentimen dibagi menjadi 2 yaitu, positif dan negatif. Pengumpulan data menggunakan Twitter API dengan kata kunci “BA Artis Korea” sebanyak 366 data tweet. Tahap penelitian terdiri dari pengumpulan data, *text preprocessing, Lexicon sentiment*, visualisasi, dan pengujian. Hasil dari klasifikasi dengan metode *Lexicon* diperoleh 67,1% positif dan 32,9% negatif. Diperoleh bahwa penggunaan artis Korea Selatan menjadi brand ambassador produk Indonesia mendapatkan sentimen positif. Nilai akurasi, presisi, dan *recall* yang didapatkan dengan melihat tabel confusion matrix yaitu akurasi 78%, presisi 68%, dan recall 61%.

**Kata kunci:** Brand Ambassador, *Lexicon*, Sentiment, Twitter

# PENDAHULUAN

Seiring dengan pesatnya perkembangan teknologi dan arus informasi memudahkan masyarakat Indonesia lebih terbuka dalam pengetahuan secara luas. Termasuk dengan kemajuan perkembangan internet dan media sosial yang mudah diakses. Hadirnya kemudahan di internet yang membuat internet tidak lagi menjadi sarana untuk mencari informasi. Internet juga bisa digunakan sebagai lahan bisnis yang lebih dikenal sebagai bisnis online atau *e-commerce*. Laporan Survei Status Literasi Digital di Indonesia berkolaborasi dengan Katadata Insight Center (KIC) dan Kementrian Komunikasi dan Informatika (KOMINFO) menyatakan bahwa 27,6% mayoritas masyarakat Indonesia berbelanja online minimal sebulan sekali [1]. *E-commerce* merupakan kegiatan usaha yang mengimpilkasikan manufaktur, konsumen, perdangangan perantara, dan pelayanan jasa dengan menggunakan jaringan komputer atau dikenal dengan internet [2]. Dengan adanya *e-commerce* memudahkan pelaku bisnis untuk mengalihkan cara penjualan secara online atau media *e-commerce*.

Menurut Asosiasi Penyelenggaraan Jasa Internet Indonesia (APJII) yang dirilis pada bulan November 2022 mencatat sejumlah alasan konsumen memilih untuk belanja kebutuhan secara online [3]. Alasan yang paling banyak yaitu harga yang jauh lebih murah dibandingkan membeli langsung di toko dengan persentase sebanyak 15,2%. Alasan kedua yang paling banyak yaitu dapat dilakukan dimana saja sehingga memudahkan konsumen yang ingin membeli suatu barang namun tidak memiliki waktu untuk datang langsung ke tempat. Semakin banyak *e-commerce* dan perusahaan lokal yang melayani pembelian online maka akan semakin ketat persaingan di antara satu sama lain. Berbagai cara dilakukan para pelaku bisnis online untuk menarik konsumen, dengan cara memberikan gratis ongkir untuk pengiriman, menawarkan promo-promo menarik, dan beberapa cara lainnya. Cara lain yang dilakukan pelaku usaha online adalah menggunakan brand ambassador yang tengah diminati oleh masyarakat Indonesia. Brand ambassador merupakan seseorang yang dipercayai untuk mewakili suatu merk dan menjadi wajah dari merk tersebut [4]. Di zaman pasar modern seperti sekarang, tanpa adanya brand ambassador akan susah untuk menjual produk. Untuk mengingat suatu produk, dibutuhkan pemilihan brand ambassador yang tepat agar sutau produk akan tetap berjalan dengan stabil di antara banyaknya produk dan *e-commerce* serupa. Dewasa ini, beberapa perusahaan lokal Indonesia termasuk *e-commerce* mulai berbondong-bondong menggunakan artis asal Korea Selatan sebagai brand ambassador. Produk industri hiburan asal Korea Selatan mulai banyak digemari di Indonesia, baik dalam bentuk musik, drama, makanan, dan budaya. Menurut laporan *Katadata Insight Center* (KIC) bersama *zigi.id* mengatakan bahwa sekitar 1000 responden yang disurvei, sebanyak 10,9% responden terkena gelombang korea atau disebut dengan korean wave pada tahun 2020. Gelombang tersebut terlihat sudah menguat di awal tahun 2010 hingga puncaknya di awal pandemi Covid-19 tahun 2020. Sebanyak 2,1% perbedaan dari tahun 2019 ke 2020 yang menjadi puncak digemarinya hiburan yang berasal dari Korea Selatan tersebut. Oleh karena itu, banyak perusahaan lokal dan *e-commerce* melihat sebuah kesempatan untuk menggandeng artis Korea Selatan sebagai brand ambassador yang akan menjadi wajah bagi perusahaan mereka.

Penggunaan artis korea sebagai brand ambassador menjadi trend di dalam pemasaran online juga dilakukan oleh beberapa *e-commerce*. Beberapa *e-commerce* yang sudah akrab dengan keseharian seperti Tokopedia, Shopee, Blibli, Lazada, dan lainnya pun mulai memakai artis asal Korea Selatan menjadi brand ambassador. Tokopedia yang menggandeng boygrup BTS, Shopee yang menggandeng girlgrup Blackpink, dan lainnya. Penggunaan artis Korea Selatan menjadi brand ambassador seakan menjadi trend di kalangan bisnis online. Target pemasaran kali ini adalah masyarakat yang menyukai artis-artis yang menjadi brand ambassador produk lokal dan *e-commerce* tersebut. Keterlibatan brand ambassador merupakan salah satu faktor kuat untuk mempengaruhi keputusan minat beli dari konsumen. Pemanfaatan artis Korea Selatan mewakili suatu merk produk atau *e-commerce* memiliki keuntungan dari sisi penjualan dikarenakan target pasar mereka yaitu penggemar artis tersebut yang dimana dikenal loyal dan bersedia untuk membeli [5].

Namun, penggunaan artis Korea Selatan sebagai brand ambassador menuai sentimen positif dan negatif. Twitter merupakan salah satu media sosial yang banyak digunakan oleh konsumen untuk mengeluarkan pendapat mengenai hal tersebut. Twitter memiliki kelebihan dibandingkan media sosial lainnya yaitu tampilan yang simpel, topik terupdate, dan cepat dalam menggambarkan sebuah opini pengguna. Fungsi pada Twitter yang digunakan untuk berbagi bahan menjadi sarana yang digunakan di berbagai aspek salah satunya menjadi media pelatihan berupa sentimen analisis, *text mining,* dan *artificial intelligence* [6]. Untuk mengetahui berbagai sentimen masyarakat Indonesia terhadap penggunaan artis dari Korea Selatan sebagai brand ambassador produk lokal maka diperlukan analisis sentimen terhadap pengguna Twitter.

Analisis sentimen adalah penelitian yang menerapkan teknik untuk mengumpulkan opini, emosi, dan sentimen dalam bentuk dan ciri yang dimiliki yang dituangkan dalam bentuk tekstual [7]. Terdapat 2 pendekatan untuk melakukan analisis sentimen, yang pertama yaitu machine learning dengan melatih data latih pada dataset yang dilabelkan secara manual. Untuk dapat mengklasifikasikan suatu opini (negatif atau positif) perlu dilakukan identifikasi terlebih dahulu apakah suatu kalimat tersebut mengandung opini atau tidak. Pendekatan 2 yaitu berbasis leksikal (*Lexicon*) yang tidak membutuhkan data latih untuk mendapatkan polaritas dari suatu sentimen [8]. Data diproses dengan cara memberikan skor sentimen untuk setiap opini yang ada. Metode *Lexicon* bekerja dengan cara membuat kamus kata opini (*Lexicon*) terlebih dahulu. Metode ini akan mengklasifikasikan teks opini secara otomatis dengan melihat data latih berupa teks opini yang sudah diklasifikasikan secara manual sebelumnya [9]. *Lexicon* bekerja dengan cara memeriksa data yang sudah diproses dan bersih kemudian dibandingkan dengan kamus *Lexicon* yang sudah ada sebelumnya. Kamus *Lexicon* yang digunakan pada penelitian ini adalah InSet (Indonesian *Lexicon Sentiment*) yang dikembangkan oleh [10]. Kamus InSet mengandung kata-kata yang sudah ada bobot atau nilainya per setiap kata. Bobot tersebut nantinya akan digunakan untuk menilai emosi di dalam suatu kalimat. Bobot ini terbagi menjadi positif dan negatif yang akan dijumlahkan untuk menilai suatu sentimen itu mengarah ke positif atau negatif. Kamus InSet terdiri dari 3609 kata positif dan 6609 kata negatif.

Penelitian terdahulu dilakukan oleh [11] yang berjudul “Penerapan *Lexicon-Based* Untuk Analisis Pada Twitter Terhadap Isu Covid-19” didapatkan kesimpulan yaitu opini masyarakat untuk isu Covid-19 memperoleh sentimen positif dengan kesimpulan bahwa pengguna Twitter percaya bahwa Covid-19 nyata adanya masyarakat pengguna Twtter percaya bahwa covid itu nyata dibandingkan dengan masyarakat yang mempercayai isu konspirasi yang megatakan bahwa covid tidak nyata. Penelitian sebelumnya juga pernah dilakukan oleh [12] yang berjudul “Analisis Sentimen Topik Viral Desa Penari Pada Media Sosial Twitter Dengan Metode *Lexicon Based*” dalam penelitian ini hasil akhir dari opini yang diperoleh dari Twitter dibagi menjadi 3 kategori yaitu positif, negatif, dan netral sehingga bisa disimpulkan sentiemn masyarakat terhadap topik tersebut. Penelitian mengenai penggunaan artis Korea Selatan sebagai brand ambassador pernah dilakukan oleh [5] yang berjudul “Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai Brand Ambassador Produk Kecantikan Lokal”, penelitian ini terfokus kepada brand ambassador produk kecantikan lokal dan menggunakan metode *support vector machine* (SVM) dengan hasil analisis menunjukkan kecenderungan masyarakat bersifat positif dengan adanya keputusan tersebut. Berdasarkan uraian tersebut, maka penelitian ini akan meneliti mengenai sentimen brand ambassador pada artis Korea Selatan pada produk Indonesia menggunakan metode *Lexicon sentiment*. Sehingga diperoleh gambaran umum sentimen pengguna Twitter, hasil klasifikasi dan visualisasi, dan tingkat akurasi, presisi, dan recall. Sumber kamus *Lexicon* yang digunakan pada penelitian ini yaitu InSet (Indonesian *Sentiment Lexicon*) yang dikembangkan oleh [10].

# METODE PENELITIAN

* 1. *Sumber Data dan Variabel*

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan data cuitan (tweet) pengguna Twitter. Data dikumpulkan dari hasil *crawling* yang menggunakan kata kunci yang sudah ditentukan yaitu “BA Artis Korea”. Proses *crawling* data menggunakan salah satu fitur dari Twitter yaitu Twitter API (*Aplication Programming Interface*). Twitter API adalah fitur dimana Twitter mengizinkan pihak ketiga untuk mengakses informasi dari Twitter yang terdapat pada website *https://developers.twitter.com/*. Data tweet dikumpulkan pada bulan Januari – Februari 2023 dengan jumlah data tweet sebanyak 366 tweet. Variabel penelitian yang digunakan yaitu tanggal tweet dibuat dan data tweet yang berisi cuitan dari pengguna Twitter.

* 1. *Langkah Penelitian*

Metode penelitian yang digunakan yaitu *Lexicon* sentiment dan menggunakan kamus InSet *Lexicon*. Penelitian ini bertujuan untuk mendapatkan gambaran umum sentimen pengguna Twitter terhadap brand ambassador artis Korea Selatan pada produk Indonesia, mendapatkan hasil klasifikasi dan visualisasi menggunakan metode *Lexicon* sentiment terhadap brand ambassador artis Korea Selatan pada produk Indonesia, dan mendapatkan tingkat akurasi, presisi, dan recall metode *Lexicon* dalam pengklasifikasian sentimen pengguna Twitter terhadap brand ambassador artis Korea Selatan pada produk Indonesia. Adapun langkah-langkah sebagai berikut:

1. Proses *crawling* data tweet menggunakan Twitter API.
2. Text preprocessing yang bertujuan untuk memperbaiki data awal menjadi data yang sudaah bisa digunakan.Tahap *text preprocessing* yang dilakukan, yaitu: *Case folding, Punctuation removal, Tokenizing, Stopword removal, Stemming.*
3. Klasifikasi menggunakan pendekatan *Lexicon* dengan menggunakan kamus InSet *Lexicon*:
   * + *Polarity tagging*, menganalisis polaritas dalam kata dan membagi menjadi positif dan negatif.
     + *Word frequency*, menghitung banyaknya satu kata dalam satu kalimat sentimen.
     + *Word attitude*, mengubah polaritas kata positif dan negatif menjadi angka (nilai).
     + *Overall attitude*, menjumlahkan nilai per kata positif dan negative
     + *Sentiment score*, menghitung total nilai sentimen dalam satu kalimat. Tahap ini merupakan tahap terakhir dalam analisis *Lexicon*.
4. Skenario pengujian, dilakukan pengujian dengan menghitung presisi, *recall*, akurasi, dan *f-measure.*

# HASIL DAN PEMBAHASAN

Data *tweet* di *crawling* menggunakan *software RStudio* dalam rentang bulan Januari – Februari 2023 dengan kata kunci pencarian BA Artis Korea. Jumlah data yang berhasil di *crawling* sebanyak 366 *tweet.* *Text preprocessing* merupakan tahap awal yang dilakukan dalam analisis text mining yang bertujuan untuk memperbaiki data sesuai yang dibutuhkan dalam analisis selanjutnya. Analisis *text preprocessing* menggunakan *Google Colab*.

* *Case Folding*

Tahap ini dilakukan untuk proses penyamaan huruf yang terdapat dalam data dengan mengubah huruf kapital menjadi huruf tidak kapital.

* *Punctuation Removal*

Tahap *punctuation removal* dilakukan pembersihan terhadap data tweet yang didalamnya masih terdapat link, username, hashtag, dan lain-lain yang tidak digunakan dalam proses klasifikasi. Selain itu, dilakukan penghapusan angka dan tanda baca. Sehingga data tweet hanya berisi data yang bersih dan bisa digunakan untuk proses selanjutnya.

* *Tokenizing*

Tahap ini dilakukan untuk membagi kata- kata yang menyususn dokomen atau kalimat pada data tweet. Proses ini memudahkan untuk melakukan pembobotan pada setiap kata dan menghasilkan kata-kata yang sudah terpisah satu sama lain tidak berbentuk satu kalimat.

* *Stopword Removal*

Tahap *stopword removal* dilakukan utuk penghapusan kata yang tidak digunakan dalam analisis. Tabel di atas menunjukkan hasil dari tahapan *stopword removal*. Pada data *tweet* pertama dilakukan penghapusan pada kata “[‘sih’, ‘soalnya’, ‘kan’, ‘ada’, ‘jadi’, ‘gitu’, ‘pasti’]”.

* *Stemming*

Tahap *stemming* bertujuan untuk mengubah kata berimbuhan menjadi kata dasar. Imbuhan pada awal kata, imbuhan akhir kata, dan imbuhan pada tengah kata sehingga diperoleh kata dasar yang akan memudahkan proses analisis sentimen selanjutnya. Tabel di atas menunjukkan hasil dari proses stemming. Data tweet pertama dilakukan perubahan pada kata “[‘untungnya’, ‘meningkatkan’, ‘penjualan’, ‘idolanya’]” menjadi “[‘untung’, ‘tingkat’, ‘jual’, ‘idol’]”. Kata “untungnya” dan “idolanya” dilakukan penghapusan imbuhan pada akhir kata yaitu “- nya”. Kata “penjualan” dan “meningkatkan” dilakukan penghapusan imbuhan pada awal dan akhir kata sehingga didapatkan kata dasar yaitu “jual” dan “tingkat”.

* 1. *Lexicon Sentiment*

Setelah dilakukan proses text preprocessing *yaitu case folding, punctuation removal, tokenizing, stopoword removal,* dan *stemming* maka diperoleh data *tweet* bersih yang akan digunakan untuk proses. Setelah dilakukan proses *text preprocessing* yang dilakukan untuk mempersiapkan data untuk bisa dilanjutkan dengan analisis selanjutnya yaitu *Lexicon* sentiment. *Lexicon* sentiment dibangun sesuai dengan kata-kata sentimen yang terdapat di dalam kamus yang digunakan. Pada analisis ini menggunakan kamus *Lexicon* yaitu Indonesia *Sentiment Lexicon (InSet)* dan *Google Colab*. Kamus pada *Lexicon sentiment* digunakan untuk menentukan polaritas sentimen dari suatu opini. Hasil dari perhitungan polaritas sentimen bisa dilihat pada Tabel di bawah ini:

**Tabel 1.** Hasil Skor Polaritas

|  |  |  |  |
| --- | --- | --- | --- |
| Opini | *Clean Tweet* | Skor Polaritas | Polaritas |
| 1 | orang indonesia perasaan kaga seputih salju skinker lokal pake ba orang korea sehat | 3 | Positif |
| 2 | thats why produk skincare brand lokal gw review recommend always range harga item please skincare brand lokal stop bayar mahal pake ba artis korea harga skincare ikutan mahal marketnya geser | -1 | Negatif |

Pada Tabel 1 terdapat hasil akhir dari skor polaritas pada masing-masing terhadap setiap kalimat sentimen. Proses ini dilakukan dengan membandingkan data *tweet* yang sudah melalui tahap *text preprocessing* terlebih dahulu. Kemudian dibandingkan dengan kata-kata sentimen yang terdapat di dalam kamus InSet *Lexicon*. Kamus ini berisi kata-kata sentimen beserta bobot nilai yang berfungsi untuk menentukan nilai polaritas. Nilai polaritas diperoleh dengan menjumlahkan seluruh bobot kata yang terdapat dalam sebuah kalimat opini. Pada Tabel 1 diketahui skor polaritas dari setiap kata yang membentuk kalimat opini. Kalimat opini yang berbunyi “orang indonesia perasaan kaga seputih salju skinker lokal pake ba orang korea sehat” mengandung 1 kata sentimen yang terdapat dalam kamus InSet *Lexicon* yaitu “sehat”. Tahap-tahap yang dilakukan untuk menghitung skor polaritas diuraikan sebagai berikut:

* *Polarity tagging*: tahap ini menganalisis suatu kata yang terdapat dalam kalimat opini dan membagi dalam kelas positif dan negatif. Kata “sehat” dalam kalimat opini 1 terdapat pada kelas positif dan negatif sehingga kata “sehat” digolongkan sebagai kata positif dan negatif.
* W*ord frequency*: tahap ini menghitung banyaknya satu kata dalam sebuah kalimat sentimen. Dalam opini 1 terdapat 1 kata sentimen yaitu “sehat”. Kata “sehat” pada kalimat opini 1 hanya ada satu. Jika dalam satu kalimat opini terdapat 2 kata sentimen maka akan dihitung menjadi dua dan seterusnya.
* *Word attitude*: tahap ini bertujuan untuk mengubah polaritas kata positif dan negatif menjadi angka atau nilai. Kata “sehat” yang terdapat dalam kelas positif dan negatif diubah menjadi nilai sesuai dalam kamus yang dipakai yaitu InSet *Lexicon.* Kata “sehat” mempunyai 2 nilai, yaitu (4) untuk kelas positif dan (-1) untuk kelas negatif.
* O*verall attitude*: tahap ini menghitung nilai keseluruhan dari suatu kalimat opini sehingga mendapatkan nilai polaritas untuk setiap kata sentimen. Kata “sehat” mempunyai nilai 4 kelas positif dan -1 kelas negatif. Maka total nilai kata “sehat”, yaitu:

Kata “sehat” terdapat 1 kali dalam kalimat opini 1 maka *word frequency* = 1. *Word frequency* tersebut dikalikan dengan nilai sentimen pada kelas positif yaitu 4 dan nilai sentimen pada kelas negatif yaitu -1. Sehingga diperoleh total nilai polaritas untuk kata “sehat” adalah 3.

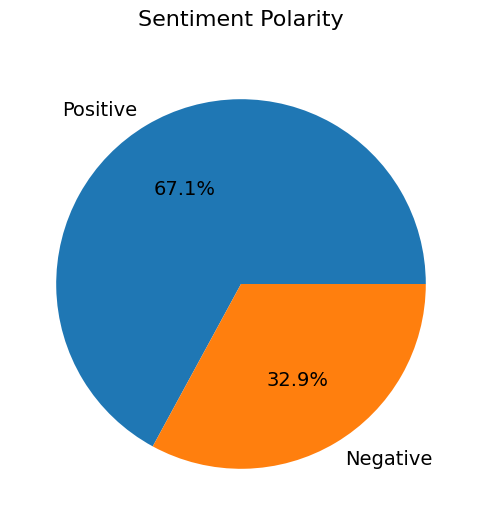
* *Sentiment score*: tahap ini menghitung total nilai sentimen dalam sebuah kalimat.

Karena dalam kalimat opini 1 hanya mengandung 1 kata sentimen yaitu “sehat” maka nilai polaritas total dalam satu kalimat opini adalah 3. Karena *sentiment score* > 0, maka disimpulkan bahwa sentimen kalimat opini 1 adalah positif. Nilai polaritas kemudian diubah menjadi polaritas pada setiap kalimat sentimen yang sudah diproses dengan kamus InSet *Lexicon* untuk mengetahui polaritas sebuah kalimat.

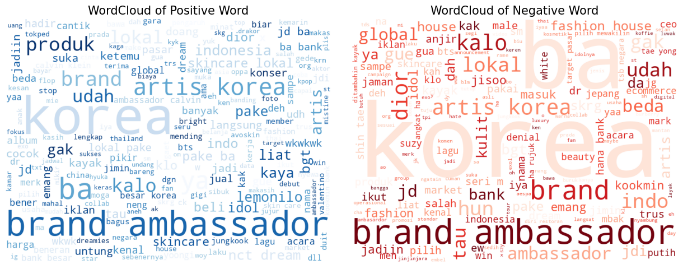
Pada Tabel 1 menampilkan skor polaritas dan polaritas untuk kalimat opini yang telah diproses melalui *text preprocessing* sehingga didapatkan kalimat opini yang sudah bersih. Kalimat opini “orang indonesia perasaan kaga seputih salju skinker lokal pake ba orang korea sehat” mendapatkan skor polaritas 3 dan *sentiment score* > 0 maka polaritasnya positif. Kalimat opini “thats why produk skincare brand lokal gw review recommend always range harga item please skincare brand lokal stop bayar mahal pake ba artis korea harga skincare ikutan mahal marketnya geser” mendapatkan skor polaritas -1 dan *sentiment score* < 0 maka polaritasnya negatif.

* 1. *Visualisasi Sentimen*

Setelah diketahui nilai polaritas dan polaritas dari suatu data, dilanjutkan dengan visualisasi. Visualisasi dilakukan untuk mengetahui pembagian dari polaritas positif dan negatif dalam bentuk *piechart* dan *wordcloud* kata-kata yang sering muncul dalam kelas positif dan negatif. Pada gambar 4.2 di bawah ini diketahui sebanyak 67,1% mempunyai polaritas positif dan 32,9% mempunyai polaritas negatif. Dari 325 data *tweet* yang terkait dengan penggunaan artis Korea Selatan sebagai *brand ambassador* didapatkan hasil bahwa 67,1% memiliki sentimen positif terkait topik tersebut dan 32,9% memiliki sentimen negatif terkait dengan topik tersebut. Hal ini menyatakan bahwa *brand ambassador* yang menggunakan artis Korea Selatan pada produk Indonesia mendapatkan sentimen positif.



**Gambar 1.** Polaritas Sentimen



**Gambar 2.** *Wordcloud* Sentimen

Pada Gambar 2 merupakan visualisasi dari kata positif dan negatif yang paling banyak muncul pada data opini masyarakat mengenai penggunaan artis Korea Selatan sebagai *brand ambassador* produk Indonesia. *Wordcloud* untuk kata positif menunjukkan kata *“brand ambassador”* dan “korea” merupakan kata paling banyak digunakan. Hal ini dikarenakan kata kunci yang digunakan untuk *crawling* data *tweet* adalah BA artis Korea yang merupakan singkatan dari *brand ambassador.* Selain itu terdapat beberapa kata lainnya yaitu “seru”, “suka”, “artis”, dan “lokal”. Kata-kata tersebut dapat diartikan bahwa artis Korea Selatan yang menjadi *brand ambassador* disukai oleh konsumen yakni pengguna Twitter. Konsumen merasakan bahwa artis-artis tersebut menjadi lokal dikarenakan memakai produk Indonesia dan familiar dalam keseharian.

*Wordcloud* untuk kata negatif menampilkan bahwa kata “korea” yang paling banyak digunakan. Kata “korea” merupakan kata yang termasuk ke dalam kata kunci *crawling* data. Hal ini menyatakan bahwa kata Korea juga termasuk ke dalam *tweet* negatif. Kata lainnya yang muncul dalam *wordcloud* negatif adalah “beda”, “kulit”, “lokal”’, dan “*denial”.* Kata-kata tersebut dapat diartikan bahwa konsumen merasa bahwa *brand ambassador* yang digunakan oleh produk Indonesia berbeda dalam warna kulit masyarakat Indonesia. Selain itu, kecantikan Korea dijadikan standar oleh beberapa produk kecantikan. Dengan memakai wajah artis tersebut, produk diiklankan dengan harapan bahwa konsumen akan membeli produk dan bisa menjadi seperti wajah dari *brand ambassador* yang diiklankan. Kata *denial* atau dalam bahasa Indonesia yaitu menyangkal/penyangkalan. Beberapa konsumen menyangkal dengan penggunaan artis Korea sebagai *brand ambassador* produk Indonesia dan menyampaikan opini negatif dalam *tweet* yang dituliskan dalam Twitter.

**Gambar 3.** Histogram kata yang paling banyak muncul

Gambar 3 menunujukkan frekuensi kata yang paling banyak muncul pada data *tweet* yang sudah di proses melalui *text preprocessing.* Kata pada data *tweet* yang paling banyak muncul adalah “korea” yang muncul sebanyak 245 kali. Kata “*brand”* muncul sebanyak 218 kali dan dilanjutkan dengan kata *“ambassador”* yang muncul sebanyak 174 kali. Oleh karena itu, bisa dikatakan bahwa data *tweet* yang sudah di*crawling* memuat kata kunci yang sesuai dengan yang ditentukan yaitu “BA Artis Korea”. Sehingga data *tweet* sudah memenuhi informasi yang dibutuhkan pada penelitian yang menggunakan *brand ambassador* artis Korea Selatan pada produk Indonesia.

Nilai polaritas untuk sentimen positif yang paling tinggi yaitu 23. Nilai polaritas untuk sentimen negatif yang paling tinggi yaitu -14. Pada Tabel berikut ditampilkan sentimen positif teratas dan sentimen negatif teratas.

**Tabel 2.** Top Positif dan Negatif Sentimen

|  |  |  |
| --- | --- | --- |
| Data *Tweet* | Skor Polaritas | Polaritas |
| @arcus\_\_aorta skg skin care lokal aja brand ambassador nya artis korea ya bun..etapi iklan mie sama kopi/susu pun skg artis kepop. Kadang pengen tau bayarannya dibanding artis lokal apa ga beda jauh sampe perusahaannya mutusin pake artis korea/kepop tp ya meh ngopo bar kuwi... | 23 | Positif |
| @waesehoon itu fandom sebelah tuh bilang kalo idolnya 1st ga buat male naa karena dikira sh itu bukan ga tapi ha ada jg ba cmn dr korea grgr brand nya labil ngasih tittle buat sh yaudah dah tuh udah dikasih tau tp tetep denial bilang brandnya boong lah apa lah | -14 | Negatif |

* 1. *Tabel Confusion Matrix*

Tabel *confusion matrix* digunakan untuk mengevaluasi suatu klasifikasi dengan menghitung berdasarkan data penelitian untuk menghindari *false statment.* Tabel 10 di bawah ini merupakan tabel *confusion matrix* dari data *tweet* yang telah diklasifikasikan menggunakan *Lexicon* InSet. Pada tabel *confusion matrix* yang digunakan pada penelitian ini, data prediksi yaitu data dalam format hasil klasifikasi yang dihasilkan oleh alogritma yang digunakan, data aktual dihasilkan dengan penghitungan manual dengan cara membandingkan dengan kamus InSet *Lexicon.*

**Tabel 3.** Hasil Tabel *Confusion Matrix*

|  |  |  |
| --- | --- | --- |
| Aktual | Prediksi | |
| Positif | Negatif |
| Positif | 66 | 30 |
| Negatif | 41 | 188 |

Pada Tabel 3 diperoleh hasil sebanyak 66 data *tweet* termasuk dalam *true positive* yaitu data aktual positif yang diprediksi benar positif, 30 data *tweet* termasuk dalam *false positive* yaitu data aktual positif yang diprediksi salah negatif, 41 data *tweet* termasuk dalam *false negative* yaitu data aktual negatif yang diprediksi salah positif, dan 188 data *tweet* termasuk dalam *true negative* yaitu data aktual negatif yang diprediksi benar negatif. Setelah mendapatkan hasil dari tabel *confusion matrix* maka dilanjutkan dengan menghitung nilai presisi, *recall,* akurasi, dan *f-measure.*

Berdasarkan hasil perhitungan di atas diketahui nilai presisi, *recall,* akurasi, dan *f-measure* dari proses analisis sentimen *Lexicon* menggunakan kamus InSet *Lexicon.* Nilai presisi yang dihasilkan yaitu 68%, *recall* yaitu 61%, akurasi yaitu 78%, dan *f-measure* yaitu 63%.

# KESIMPULAN

Setelah melakukan tahapan analisis dan memperoleh hasil yang telah dibahas sebelumnya, ada beberapa kesimpulan yang diperoleh diantaranya yaitu:

1. Setelah dilakukan klasifikasi sentimen dengan menggunakan kamus InSet *Lexicon* diperoleh total nilai polaritas untuk sentimen positif yang paling tinggi adalah 23. Nilai polaritas untuk sentimen negatif yang paling tinggi adalah -14. *Wordcloud* untuk kata positif menunjukkan kata *“brand ambassador”* dan “korea” merupakan kata paling banyak digunakan. Hal ini dikarenakan kata kunci yang digunakan untuk *crawling* data *tweet* adalah BA artis Korea yang merupakan singkatan dari *brand ambassador.* Selain itu terdapat beberapa kata lainnya yaitu “seru”, “suka”, “artis”, dan “lokal”. Kata-kata tersebut dapat diartikan bahwa artis Korea Selatan yang menjadi *brand ambassador* disukai oleh konsumen yakni pengguna Twitter. Konsumen merasakan bahwa artis-artis tersebut menjadi lokal dikarenakan memakai produk Indonesia dan familiar dalam keseharian. *Wordcloud* untuk kata negatif menampilkan bahwa kata “korea” yang paling banyak digunakan. Kata “korea” merupakan kata yang termasuk ke dalam kata kunci *crawling* data. Hal ini menyatakan bahwa kata Korea juga termasuk ke dalam *tweet* negatif. Kata lainnya yang muncul dalam *wordcloud* negatif adalah “beda”, “kulit”, “lokal”’, dan “*denial”.* Kata-kata tersebut dapat diartikan bahwa konsumen merasa bahwa *brand ambassador* yang digunakan oleh produk Indonesia berbeda dalam warna kulit masyarakat Indonesia. Selain itu, kecantikan Korea dijadikan standar oleh beberapa produk kecantikan. Dengan memakai wajah artis tersebut, produk diiklankan dengan harapan bahwa konsumen akan membeli produk dan bisa menjadi seperti wajah dari *brand ambassador* yang diiklankan. Kata *denial* atau dalam bahasa Indonesia yaitu menyangkal/penyangkalan. Beberapa konsumen menyangkal dengan penggunaan artis Korea sebagai *brand ambassador* produk Indonesia dan menyampaikan opini negatif dalam *tweet* yang dituliskan dalam Twitter.
2. Tingkat akurasi yang diperoleh dari hasil klasifikasi menggunakan kamus InSet *Lexicon* dengan yaitu 78%. Nilai presisi yaitu 68% dan nilai *recall* yaitu 61%.

# UCAPAN TERIMA KASIH

# REFERENSI

1. Yoshio, A. Mayoritas Masyarakat Indonesia Pernah Belanja Online. *Katada.Co.Id*, 2021. https://databoks.katadata.co.id/datapublish/2022/04/27/mayoritas-masyarakat-indonesia-pernah-belanja-online
2. Prasetyo, T., dan Barkatullah, A., *Bisnis E-commerce: Suatu Sistem Keamanan dan Hukum di Indonesia*. Pustaka Pelajar, 2005
3. Annur, C. M. Ragam Alasan Konsumen Pilih Berbelanja Online. *Asosiasi Penyelenggara Jasa Internet Indonesia (APJII)*, *November*, 2020.
4. https://databoks.katadata.co.id/datapublish/2020/11/11/ragam-alasan-konsumen-pilh-berbelanja-online
5. Poghosyan, A. Celebrity Endorsement as one of nowadays Major Ways to Influence Costumer Buying Behaviour. *European Scientific Journal*, 2015; pp 30–38.
6. Slamet, R., Gata, W., Novtariany, A., Hilyati, K., & Jariyah, F. A. Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai Brand Ambassador Produk Kecantikan Lokal. *INTECOMS: Journal of Information Technology and Computer Science*, 2022; *5*(1), 145–153. https://doi.org/10.31539/intecoms.v5i1.393
7. Hartanto. Text Mining dan Sentimen Analisis Twitter. *Jurnal Psikologi Ilmiah*, 2017; *9*(1), 18–25.
8. Liu, B. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 2012; *5*, 1–167.
9. Medhat, W., Hassan, A., & Korashy, H. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*, 2014; *5*(4), 1093–1113.
10. Azhar, Y. Metode *Lexicon*-Learning Based Untuk Identifikasi Tweet Opini Berbahasa Indonesia. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 2018; *6*(3), 237.
11. https://doi.org/10.23887/janapati.v6i3.11739
12. Koto, F., & Rahmaningtyas, G. Y. Inset *Lexicon*: Evaluation of a word list for Indonesian sentiment analysis in microblogs. *Proceedings of the 2017 International Conference on Asian Language Processing, IALP 2017*, *2018*-*Janua*(December), 391–394. https://doi.org/10.1109/IALP.2017.8300625
13. Nooryuda Prasetya, Y., & Winarso, D. Penerapan *Lexicon* Based Untuk Analisis Sentimen Pada Twiter Terhadap Isu Covid-19. *Jurnal Fasilkom*, 2021; *11*(2), 97–103.
14. Arief, R., & Imanuel, K. Analisis Sentimen Topik Viral Desa Penari Pada Media Sosial Twitter Dengan Metode *Lexicon* Based Universitas Gunadarma 1, 2 Jalan Margonda Raya No 100 Depok Jawa Barat 16424 Sur-el : rifiana@staff.gunadarma.ac.id 1 , karel4404@gmail.com 2. *Jurnal Ilmiah MATRIK*, 2019; *21*(3), 242–250