

# Penerapan Algoritma Naïve Bayes dalam Menganalisis Sentimen pada Review Pengguna E-Commerce

Sulindawaty<sup>1,\*</sup>, Erlina Laia<sup>1</sup>, M Yamin<sup>2</sup>

<sup>1</sup> Teknologi Informasi, STMIK Pelita Nusantara, Medan, Indonesia

<sup>2</sup> Teknik Informatika, STMIK Pelita Nusantara, Medan, Indonesia

Email: <sup>1,\*</sup> [sulindawaty@gmail.com](mailto:sulindawaty@gmail.com), <sup>2</sup> [Erlinalaia059@gmail.com](mailto:Erlinalaia059@gmail.com), <sup>3</sup> [muhammadyamin1081@gmail.com](mailto:muhammadyamin1081@gmail.com)

Email Penulis Korespondensi: [sulindawaty@gmail.com](mailto:sulindawaty@gmail.com)

**Abstrak**—Kepuasan pelanggan merupakan hal penting yang menjadi tujuan perusahaan. Disadari atau tidak, opini-opini pelanggan yang dituliskan di media sosial, sedikit ataupun banyak, akan memberikan pengaruh pada calon pelanggan. Dalam perkembangannya sentimen semakin mudah ditemukan dalam berbagai media online, salah satunya seperti shopee. Review produk merupakan sumber informasi yang sangat penting terkait kualitas produk dan sangat berpengaruh, baik bagi konsumen maupun produsen. Dengan jumlah data yang besar untuk setiap produk di toko Shopee, dalam menganalisis dan menyimpulkan informasi ulasan produk pasti akan memakan banyak waktu jika dilakukan secara manual. Untuk mengatasi hal tersebut, diperlukan sistem analisis sentimen yang dapat secara otomatis mengekstraksi informasi penting yang secara objektif dapat menentukan kualitas produk dan menangani informasi tekstual yang besar. Sistem analisis sentimen terdiri dari beberapa tahapan, yaitu crawling, pre-processing, pembobotan kata, dan klasifikasi sentimen. Dengan menerapkan algoritma Naïve Bayes melalui pemilihan fitur range dan frekuensi, maka akan diperoleh hasil akurasi, akurasi dan recall dengan menggunakan uji Confusion Matrix. API digunakan untuk mengumpulkan data dan melakukan simulasi menggunakan ulasan data di situs web Shopee untuk menghasilkan sentimen positif dan negatif. Pengujian akan menampilkan hasil nilai persentil untuk accuracy dan retrieval. Dari hasil pengujian diperoleh hasil accuracy sebesar 99,5%, precision sebesar 99,49%, recall sebesar 100%. Dengan demikian dapat disimpulkan metode klasifikasi algoritma Naïve Bayes cukup relevan meskipun akurasinya belum 100%. Hasil penelitian ini dapat memberikan solusi bagi toko Shopee untuk meningkatkan mutu penjualannya.

**Kata Kunci:** Analisis Sentimen; Review; Shopee; Naïve Bayes; Shopee; Penjualan

**Abstract**—Customer satisfaction is an important thing that is the company's goal. Whether you realize it or not, customer opinions written on social media, a little or a lot, will have an impact on potential customers. In its development, sentiment is increasingly easy to find in various online media, one of which is shopee. Product reviews are a very important source of information related to product quality and are very influential, both for consumers and producers. With a large amount of data for each product in the Shopee store, analyzing and concluding product review information will definitely take a lot of time if done manually. To overcome this, a sentiment analysis system is needed that can automatically extract important information that can objectively determine product quality and handle large textual information. The sentiment analysis system consists of several stages, namely crawling, pre-processing, word weighting, and sentiment classification. By applying the Naïve Bayes algorithm through the selection of range and frequency features, accuracy, accuracy and recall results will be obtained using the Confusion Matrix test. The API is used to collect data and perform simulations using data reviews on the Shopee website to generate positive and negative sentiments. The test will display the results of the percentile values for accuracy and retrieval. From the test results obtained results of accuracy of 99.5%, precision of 99.49%, recall of 100%. Thus it can be concluded that the classification method of the Naïve Bayes algorithm is quite relevant even though the accuracy is not yet 100%. The results of this study can provide solutions for Shopee stores to improve the quality of their sales.

**Keywords:** Sentiment Analysis; Reviews; Shopee; Naïve Bayes; Shopee; Sales

## 1. PENDAHULUAN

Jual beli secara online telah menjadi bagian dari kebutuhan belanja masyarakat saat ini. Salah satu marketplace dengan model e-commerce yang populer di Indonesia adalah Shopee. Berdasarkan data Peta E-Commerce Indonesia yang dikeluarkan oleh iPrice, Shopee adalah salah satu e-commerce regional yang beroperasi di hampir seluruh negara di Asia Tenggara, menduduki peringkat pertama ranking aplikasi e-commerce pada Play Store maupun App Store dengan jumlah pengunjung bulanan sebanyak 134,383,300 pengunjung pada kuartal 3 2021 [9].

Shopee menawarkan pengguna dalam melakukan sebuah transaksi barang secara online. Namun begitu, tidak jarang konsumen merasa ragu untuk membeli produk secara online dengan harga pasar yang cukup tinggi seperti produk smartphone. Hal ini dibuktikan berdasarkan hasil penelitian sebelumnya yang mendapati sebuah kesimpulan bahwa kepercayaan pelanggan mempengaruhi niat pembelian produk smartphone di Shopee Indonesia. Pelanggan dinilai masih merasa kurang percaya terhadap para penjual smartphone dalam memenuhi keinginan para pelanggan [2].

Ulasan produk merupakan salah satu sumber informasi tentang kualitas produk dan sangat berpengaruh pada konsumen dan produsen. Dalam kegiatan pembelian barang di marketplace, pembeli dapat memberikan ulasan setelah menerima barang yang dibeli. Pelanggan atau klien yang merasa tidak puas dengan layanan atau produk yang ditawarkan biasanya akan menuliskan keluhannya di media sosial. Ulasan pembelian produk terdiri dari bintang dan isi komentar ulasan yang berisi tanggapan, apresiasi maupun kritik dan masukan pada produk yang telah dibeli tersebut [7]. Kepuasan pelanggan merupakan hal penting yang menjadi tujuan perusahaan. Disadari atau tidak, opini-opini pelanggan yang dituliskan di media sosial, sedikit atau pun banyak, akan memberikan pengaruh pada calon pelanggan. Dalam perkembangannya sentimen semakin mudah ditemukan dalam berbagai media online seperti shopee.

Menghadapi jumlah data yang besar dan sangat banyak untuk setiap produk pada sebuah toko di shopee, menganalisis dan menyimpulkan informasi secara manual membutuhkan waktu lama dan tidak efektif. Selain jumlah data yang besar dan sangat banyak, masalah lain yang dihadapi adalah informasi yang ada dalam opini. Opini dapat berisi

informasi yang tidak lengkap, informasi yang tidak konsisten. Jadi cara untuk mengekstrak informasi penting dan membangun sistem yang dapat menentukan kualitas produk secara objektif dan secara otomatis untuk menangani informasi teks yang sangat besar dan banyak adalah dengan menggunakan sistem analisis sentimen.

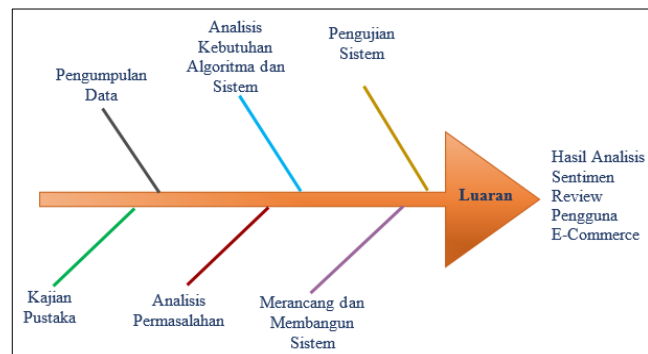
Analisis sentimen adalah cara mengidentifikasi dan mengelompokkan polaritas teks yang diberikan ditingkat dokumen, kalimat dan frasa. Analisis sentimen diaplikasikan untuk mengelompokkan komentar positif dan negatif dari konsumen sehingga mempercepat dan mempermudah tugas perusahaan untuk meninjau kembali kekurangan produk mereka [1]. Apabila ditemukan adanya sentimen negatif, maka perusahaan dapat dengan cepat mengambil tindakan untuk menanggulangnya. Sistem analisis sentimen yang akan dibangun pada penelitian ini menggunakan algoritma klasifikasi Naive Bayes. Ciri utama dari algoritma Naive Bayes adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi atau kejadian. Kelebihan dari Naive Bayes adalah proses klasifikasi data dapat disesuaikan dengan sifat dan kebutuhan masing-masing [3]. Dengan adanya sistem analisis sentimen ini diharapkan dapat masyarakat yang mempromosikan suatu produk pada marketplace mengetahui umpan balik terhadap produknya dan mengetahui opini yang disampaikan oleh pengguna marketplace terhadap produk yang ditawarkan.

Pada penelitian ini produk yang akan di analisis sentimen dari ulasan pelanggan yang terdapat pada produk adalah merupakan produk Iphone secondhand yang dijual oleh Kinophonecell melalui marketplace Shopee. Alasan pemilihan produk tersebut adalah dikarenakan banyaknya penjualan serta banyaknya ulasan yang telah dibuat oleh pelanggan yang telah membeli produk tersebut. Berdasarkan masalah dan pemaparan di atas, maka dilakukan penelitian mengenai feedback atau review dari pengguna e-commerce untuk membantu pelaku bisnis untuk melakukan evaluasi dalam mengetahui kualitas dari suatu produk yang dibeli secara online.

## 2. METODOLOGI PENELITIAN

### 2.1 Kerangka Kerja

Agar penelitian ini memperoleh hasil sesuai dengan tujuan yang diharapkan, maka dilakukan langkah-langkah sebagai berikut:



**Gambar 1.** Tahapan Penelitian

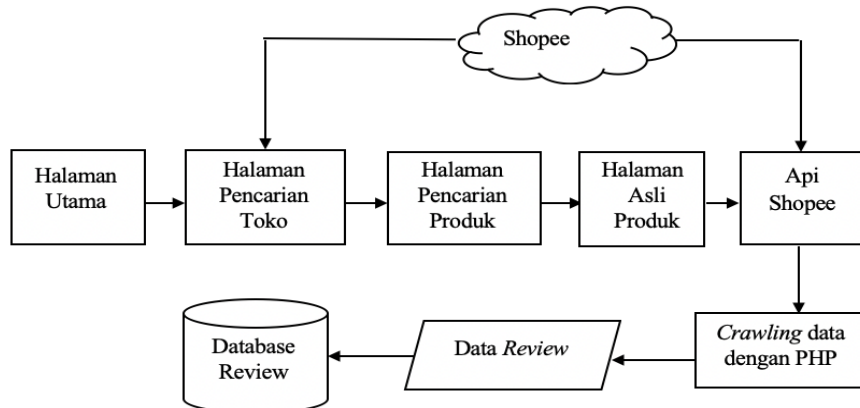
Berdasarkan gambar 1, tahapan penelitian analisis sentimen terhadap riviw pengguna e-commerce diuraikan sebagai berikut:

- Kajian Pustaka**  
Kajian pustaka yang ditelaah terkait dengan review produk, sistem analisis sentimen, algoritma Naïve Bayes dan konsep lainnya yang terkait dengan masalah penelitian ini. Kajian pustaka dilakukan dengan menelaah sumber-sumber yang bersumber dari jurnal-jurnal maupun hasil penelitian lainnya sebagai referensi penelitian terdahulu, buku teks, sumber *online* (internet) dan sumber lainnya yang terkait dengan topik penelitian yang dibahas.
- Pengumpulan Data**  
Pengumpulan data terkait review produk Iphone XR 256GB/128GB/64GB di toko Kinophonecell website Shopee. Data diambil menggunakan API yang telah disediakan oleh Shopee dan di crawling dengan menggunakan bahasa pemrograman PHP, kemudian data akan disimpan ke database.
- Analisis Permasalahan**  
Menganalisis permasalahan terkait review produk Iphone XR 256GB/128GB/64GB oleh pengguna shopee di toko Kinophonecell.
- Analisis Kebutuhan Algoritma dan Sistem**  
Menganalisis kebutuhan penerapan algoritma Naïve Bayes dan pembuatan sistem analisis sentimen yang akan digunakan untuk penyelesaian masalah pada penelitian ini.
- Merancang dan Membangun Sistem Analisis Sentimen**  
Merancang dan Membangun sistem analisis sentimen dengan menerapkan algoritma naïve bayes ke dalam sistem untuk digunakan sebagai perhitungan dalam mengelompokkan sentimen ke dalam kriteria positif, dan negatif.
- Pengujian Sistem**

Dilakukan dengan menggabungkan modul-modul yang sudah dibuat dan dilakukan pengujian ini dilakukan untuk mengetahui apakah aplikasi yang dibuat telah sesuai dengan desain yang telah dirancang sebelumnya dan masih terdapat kesalahan atau tidak. Pengujian yang dilakukan berupa mencoba menguji sistem yang dihasilkan dengan menginputkan data review pelanggan untuk dilihat apakah sistem yang dihasilkan dapat mengelompokkan data review ke dalam kelompok positif, atau negatif.

## 2.2 Analisa Data

Dalam penelitian ini menganalisis data sentiment terhadap review pengguna Shopee pada produk di toko Kinophonecell dengan melakukan pengumpulan data. Proses pengumpulan data dilakukan dengan menggunakan metode *crawling* di website Shopee dengan memanfaatkan fasilitas *Application Interface (API)* yang disediakan oleh Shopee dengan alamat <https://shopee.co.id>. Alur *crawling* data tersebut terlihat seperti gambar 2.



Gambar 2. Alur Sistem Crawling data Review Produk Shopee

Untuk mendapatkan data menggunakan API Shopee, langkah pertama adalah dengan membuat akun Shopee pribadi dan masuk, selanjutnya membuka halaman developer tools, kemudian memilih menu `get_ratings` untuk dapat mengambil data review produk. Untuk menampilkan hasil dan menyimpan hasil crawling data dari API Shopee di halaman website, sistem ini menggunakan bahasa pemrograman PHP pada aplikasi Sublime Text. Proses crawling pada sistem dilakukan dengan mencari toko Kinophonecell, kemudian mencari produk asli yang akan di crawling data reviewnya.

**Text**  
 Alhamdulillah Paketnya sudah sampe seminggu yang lalu yaa kak, cuma emang sengaja kasi riviewnya tunggu seminggu, barangnya bagus ada lecet dikit aja tapi isoke, batrenya dapet yg 94%, sampai skrg masi bertahan, batrenya awet pula alhamdulillah, pokoknya okelah yaa, sukses terus ya kak, tengkyuuu ❤️  
 Pembelian ke 2 dan sumpah ini toko jujur bangeet woi! rekomendasi bgt buat kalian yg mau cari ip semua muluss batree sllu 100 gk ada lecet sedikitpun gk akan kecewa deeh  
 sumpah ini toko recommended! baik bangettt responsible sekaliii makasih banget kakak baikk 😊 barang yg sampe sesuai req ga ada yg lecet samsek huhu pkoknya senang sekali, sudah 6 hari di pakek dan bh nya masih 100% gilasihh ini toko amanah sekali, sukses selalu usahanya kak semoga berkah 😊  
 Biarkan bintang berbicara 😊 Mntep dapat bh 93% ljo semua Lecet dikit gpp org second 😊 Pelayan baik Semoga awet maaf baru kasih penilaian, hpnya bagusssss mulus bgttt. gaada lecet dan battery healthnya masih tinggi, biarpun pas datang 99% dan skrg setelah seminggu pemakaian tinggal 95%. tp gapapa. lama di pengiriman sih kalau pengemasan cepet bgt. thankyous seller.  
 Waktu awal mesen rada ragu yh namanya juga mahal budgetnya, saya pesen white 128 cuma pihak kpc nelfon kalo stock putih lgi kosong jdi yaudah saya bilang warna coral. Daleman bagus minus kamera, akhirnya retur tuh claim garansi alhamdulillah bisa. Untuk retur sy req wrna bebs dan dtg merah mulus bgs  
 Ini sudah pakai seminggu semua normal bh normal, hp sangat seperti baru, tidak rugi beli di sini penjualan ramah banget, batrai ceoat banget ngiai nya hania nya lama lagi,, di jamin puas beli di sini ngga rugi sej makasih min, semoga usaha kita berkah, kamera foto dan video pakai kamera iphone xr, d  
 Mulus banget si ini sdh pemakaian 2 minggu msih aman banget suka banget gak ngecewain pokoknya 🙏 packing ny juga aman dpt case bening juga ,cuman bh berkurang satu asalny 100% jdi 99% tpi gak ppa semoga awet aja min 🙏 seller ny juga ramah 🙏  
 Overall barangnya mulus dan berfungsi normal semuanya,kamernya jernih,layarnya juga masih sangat responsif , jaringannya juga sampe sekarang aman aman aja kok RECOMMEND banget lah pokoknya 😊 tapi kemarin agak kecewa karna gak di double bubble wrap sama di tambah box, tapi gapapa brngnya ttp mantap :)  
 Sepadan dengan Harga:iya Hpnya mulus. Sesuai...tokonya bner2 amanah

Gambar 3. Data Hasil Crawling

Jumlah keseluruhan review yang akan di crawling pada web Shopee sebanyak 600 data atau 600 review yang dijadikan sebagai data set. Kemudian set dibagi dengan data latih dan data uji, dimana data uji berjumlah 200 yang diambil dari data set. Data latih akan dilakukan pemberian label secara otomatis dengan lexicon based ke dalam beberapa kelas diantaranya positif, dan negatif, sesuai dengan sentimen kata yang berada pada review produk. Dari data set yang didapat, kemudian akan diolah pada proses Pre-processing data.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Algoritma Naïve Bayes

##### a. Pembobotan

Setelah tahapan *preprocessing*, kata hasil tokenizing diberikan bobot dengan fitur *tf* (*term frequency*). Pembobotan merupakan proses merubah kata menjadi bentuk angka atau vector. TF merupakan jumlah kemunculan kata dalam dokumen. Kata direpresentasi ke dalam bentuk vector, dimana tiap kata dihitung sebagai satu fitur. Adapun perhitungan bobot yang digunakan merupakan *Term Frequency* (TF).

##### 1. Data Latih Sentimen Positif

**Tabel 1.** Frekuensi Data Latih Positif

| #  | Term       | Frekuensi ( $n_k$ ) |
|----|------------|---------------------|
| 1  | barang     | 2                   |
| 2  | sesuai     | 1                   |
| 3  | deskripsi  | 1                   |
| 4  | sampai     | 1                   |
| 5  | aman       | 1                   |
| 6  | baik       | 1                   |
| 7  | ngemas     | 1                   |
| 8  | cepat      | 1                   |
| 9  | respon     | 1                   |
| 10 | seller     | 1                   |
| 11 | fastrespon | 1                   |
| 12 | semoga     | 1                   |
| 13 | selalu     | 1                   |
| 14 | amanah     | 1                   |
| 15 | good       | 1                   |
| 16 | job        | 1                   |

Pada tabel 1 dapat diketahui bahwa :

- a) Jumlah frekuensi keseluruhan sentiment positif ( $n$ ) = 17
- b) Jumlah kosakata keseluruhan sentiment positif ( $|$ kosakata) = 16

##### 2. Data Latih Sentimen Negatif

**Tabel 2.** Frekuensi data latih negatif

| # | Term    | Frekuensi ( $n_k$ ) |
|---|---------|---------------------|
| 1 | respon  | 1                   |
| 2 | penjual | 1                   |
| 3 | lama    | 1                   |
| 4 | kata    | 1                   |
| 5 | body    | 1                   |
| 6 | lecet   | 3                   |
| 7 | parah   | 1                   |
| 8 | sisi    | 1                   |
| 9 | kecewa  | 1                   |

Pada tabel 2 dapat diketahui bahwa :

- a) Jumlah frekuensi keseluruhan sentimen negatif ( $n$ ) = 11
- b) Jumlah kosakata keseluruhan sentimen negatif ( $|$ kosakata) = 9

##### 3. Data Uji

**Tabel 3.** Frekuensi Data Uji

| # | Term   | Frekuensi ( $n_k$ ) |
|---|--------|---------------------|
| 1 | ngemas | 1                   |
| 2 | cepat  | 1                   |
| 3 | fast   | 1                   |
| 4 | respon | 1                   |
| 5 | admin  | 1                   |
| 6 | ramah  | 1                   |
| 7 | mulus  | 1                   |
| 8 | bh     | 1                   |
| 9 | kasih  | 1                   |

|    |         |   |
|----|---------|---|
| 10 | warna   | 2 |
| 11 | sesuai  | 1 |
| 12 | pesan   | 1 |
| 13 | layar   | 1 |
| 14 | cerah   | 1 |
| 15 | baru    | 1 |
| 16 | pakai   | 1 |
| 17 | iphone  | 2 |
| 18 | overall | 1 |
| 19 | suka    | 1 |
| 20 | mudah   | 1 |
| 21 | awet    | 1 |
| 22 | pertama | 1 |
| 23 | beli    | 1 |
| 24 | kecewa  | 1 |

Pada tabel 3 dapat diketahui bahwa :

- a) Jumlah frekuensi keseluruhan sentiment positif (n) = 26
- b) Jumlah kosakata keseluruhan sentiment positif (|kosakata) = 24

b. Klasifikasi Naïve Bayes

Tahap selanjutnya adalah melakukan proses klasifikasi menggunakan metode Naïve Bayes, Adapun tahapan yang akan dilakukan adalah sebagai berikut :

1. Menghitung nilai priors

Nilai prioritas dapat dicari dengan menggunakan rumus

$$P(V_j) = \frac{|docs_j|}{|contoh_j|}$$

$$P(positif) = \frac{1}{2} = 0,5$$

$$P(negatif) = \frac{1}{2} = 0,5$$

2. Menghitung nilai probabilitas setiap term

a) Data latih

Probabilitas dari setiap term dapat dicari dengan menggunakan rumus :

$$P(x_i|V_j) = \frac{n_k + 1}{n + |kosakata|}$$

1) Probabilitas Data Latih Sentimen Positif

$$P(\text{barang}|\text{positif}) = \frac{2 + 1}{17 + 16} = 0,0909$$

$$P(\text{sesuai}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{deskripsi}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{sampai}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{aman}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{baik}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{ngemas}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{cepat}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{respon}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{seller}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{fastrespon}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{semoga}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{selalu}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{amanah}|\text{positif}) = \frac{1 + 1}{17 + 16} = 0,0606$$

$$P(\text{good}|\text{positif}) = \frac{1+1}{17+16} = 0,0606$$

$$P(\text{job}|\text{positif}) = \frac{1+1}{17+16} = 0,0606$$

2) Probabilitas Data Latih Sentimen Negatif

$$P(\text{respon}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{penjual}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{lama}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{kata}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{body}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{lecet}|\text{negatif}) = \frac{3+1}{11+9} = 0,2$$

$$P(\text{parah}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{sisi}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

$$P(\text{kecewa}|\text{negatif}) = \frac{1+1}{11+9} = 0,1$$

Dengan menghitung nilai probabilitas dari semua kategori sentimen, maka terdapat perubahan nilai probabilitas dari setiap term pada data latih. Perubahan dilakukan dengan menggabungkan total kosakata dari semua sentimen data latih. Total |kosakata| dari semua sentimen adalah 25. Sehingga perubahan probabilitas sebagai berikut:

**Tabel 4.** Perubahan Probabilitas Term Pada Sentimen Positif

| #  | Term       | Frekuensi (n <sub>k</sub> ) | Probabilitas $P(x_i V_j)$ |
|----|------------|-----------------------------|---------------------------|
| 1  | barang     | 2                           | (2+1)/(17+25)=0,071       |
| 2  | sesuai     | 1                           | (1+1)/(17+25)=0,047       |
| 3  | deskripsi  | 1                           | (1+1)/(17+25)=0,047       |
| 4  | sampai     | 1                           | (1+1)/(17+25)=0,047       |
| 5  | aman       | 1                           | (1+1)/(17+25)=0,047       |
| 6  | baik       | 1                           | (1+1)/(17+25)=0,047       |
| 7  | ngemas     | 1                           | (1+1)/(17+25)=0,047       |
| 8  | cepat      | 1                           | (1+1)/(17+25)=0,047       |
| 9  | respon     | 1                           | (1+1)/(17+25)=0,047       |
| 10 | seller     | 1                           | (1+1)/(17+25)=0,047       |
| 11 | fastrespon | 1                           | (1+1)/(17+25)=0,047       |
| 12 | semoga     | 1                           | (1+1)/(17+25)=0,047       |
| 13 | selalu     | 1                           | (1+1)/(17+25)=0,047       |
| 14 | amanah     | 1                           | (1+1)/(17+25)=0,047       |
| 15 | good       | 1                           | (1+1)/(17+25)=0,047       |
| 16 | job        | 1                           | (1+1)/(17+25)=0,047       |

**Tabel 5.** Perubahan Probabilitas Term Pada Sentimen Negatif

| # | Term    | Frekuensi (n <sub>k</sub> ) | Probabilitas $P(x_i V_j)$ |
|---|---------|-----------------------------|---------------------------|
| 1 | respon  | 1                           | (1+1)/(11+25)=0,055       |
| 2 | penjual | 1                           | (1+1)/(11+25)=0,055       |
| 3 | lama    | 1                           | (1+1)/(11+25)=0,055       |
| 4 | kata    | 1                           | (1+1)/(11+25)=0,055       |
| 5 | body    | 1                           | (1+1)/(11+25)=0,055       |
| 6 | lecet   | 3                           | (3+1)/(11+25)=0,111       |
| 7 | parah   | 1                           | (1+1)/(11+25)=0,055       |
| 8 | sisi    | 1                           | (1+1)/(11+25)=0,055       |
| 9 | kecewa  | 1                           | (1+1)/(11+25)=0,055       |

b) Data Uji

Dalam tahap klasifikasi data uji, term yang terdaftar pada tabel 3 frekuensi data uji akan diklasifikasi dicari nilai probabilitasnya dengan membandingkan dan mencocokkan dengan term yang ada pada sentimen positif dan sentimen negatif data latih. Jika terdapat kesamaan term antara data latih dan data uji, maka nilai probabilitas yang sama tersebut dihitung bernilai 1, dan jika tidak sama maka frekuensinya dihitung bernilai 0. Berikut perhitungan manualnya:

“ngemas cepat fast respon admin ramah mulus bh kasih warna sesuai pesan layar warna cerah baru pakai iphone iphone overall suka mudah awet pertama beli kecewa”

1) Kelas Sentimen Positif

$$\begin{aligned}
 P(\text{ngemas}|\text{positif}) &= \frac{1+1}{17+25} = 0,047 \\
 P(\text{cepat}|\text{positif}) &= \frac{1+1}{17+25} = 0,047 \\
 P(\text{fast}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{respon}|\text{positif}) &= \frac{1+1}{17+25} = 0,047 \\
 P(\text{admin}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{ramah}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{mulus}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{bh}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{kasih}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{warna}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{sesuai}|\text{positif}) &= \frac{1+1}{17+25} = 0,047 \\
 P(\text{pesan}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{layar}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{cerah}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{baru}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{pakai}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{iphone}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{overall}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{suka}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{mudah}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{awet}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{pertama}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{beli}|\text{positif}) &= \frac{0+1}{17+25} = 0,023 \\
 P(\text{kecewa}|\text{positif}) &= \frac{0+1}{17+25} = 0,023
 \end{aligned}$$

Berdasarkan nilai probabilitas data uji pada sentimen positif diatas, nilai  $V_{map}$  untuk sentimen positif adalah sebagai berikut:

$$\begin{aligned}
 V_{map} &= \underset{(\text{positif})}{\text{armax}} \prod_{i=1}^{13} P(x_i|V_j)P(V_j) \\
 &= ((0,047)(0,047)(0,023)(0,047)(0,023))(0,023)(0,023)(0,023)(0,023)(0,023) \\
 &\quad (0,023)(0,047)(0,023)(0,023)(0,023)(0,023)(0,023)(0,023)(0,023)(0,023)(0,023) \\
 &\quad (0,023)(0,023)(0,023) * (0,5) \\
 &= 4,1871465e - 39
 \end{aligned}$$

2) Kelas Sentimen Negatif

$$\begin{aligned}
 P(\text{ngemas}|\text{negatif}) &= \frac{0+1}{11+25} = 0,027 \\
 P(\text{cepat}|\text{negatif}) &= \frac{0+1}{11+25} = 0,027 \\
 P(\text{fast}|\text{negatif}) &= \frac{0+1}{11+25} = 0,027 \\
 P(\text{respon}|\text{negatif}) &= \frac{1+1}{11+25} = 0,055
 \end{aligned}$$

$$\begin{aligned}
 P(\text{admin}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{ramah}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{mulus}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{bh}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{kasih}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{warna}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{sesuai}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{layar}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{cerah}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{baru}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{pakai}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{iphone}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{overall}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{suka}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{mudah}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{awet}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{pertama}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{beli}|\text{negatif}) &= \frac{0 + 1}{11 + 25} = 0,027 \\
 P(\text{kecewa}|\text{negatif}) &= \frac{1 + 1}{11 + 25} = 0,055
 \end{aligned}$$

Berdasarkan nilai probabilitas data uji pada sentimen negatif diatas, nilai  $V_{map}$  untuk sentimen negatif adalah sebagai berikut:

$$\begin{aligned}
 V_{map} &= \underset{(\text{negatif})}{\text{armax}} \prod_{i=1}^{13} P(x_i|V_j)P(V_j) \\
 &= ((0,027)(0,027)(0,027)(0,055)(0,027))(0,027)(0,027)(0,027)(0,027)(0,027)(0,027) \\
 &\quad (0,027)(0,027)(0,027)(0,027)(0,027)(0,027)(0,027)(0,027)(0,027)(0,027)(0,027) \\
 &\quad (0,027) * (0,027) * (0,5) \\
 &= 2,29455921e - 37
 \end{aligned}$$

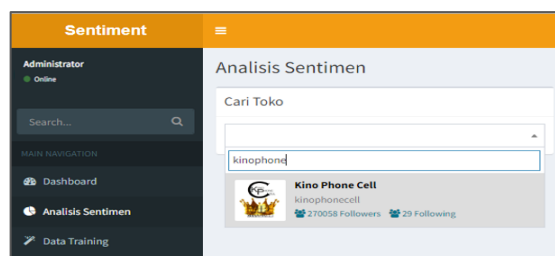
Berdasarkan perbandingan hasil perhitungan  $V_{map}$  kedua sentimen, nilai  $V_{map}$  sentimen positif lebih besar. Sehingga, data uji diklasifikasikan sebagai sentimen kelas **POSITIF**.

### 3.2 Implementasi

Implementasi hasil dari penelitian ini mencakup tampilan interface, pengujian serta penerapan yang dibangun berdasarkan proses tahapan perhitungan manual serta rancangan aplikasi.

#### a. Halaman Pencarian toko

Halaman pencarian toko berfungsi mencari toko untuk melakukan analisis review terhadap produk yang akan dianalisis.

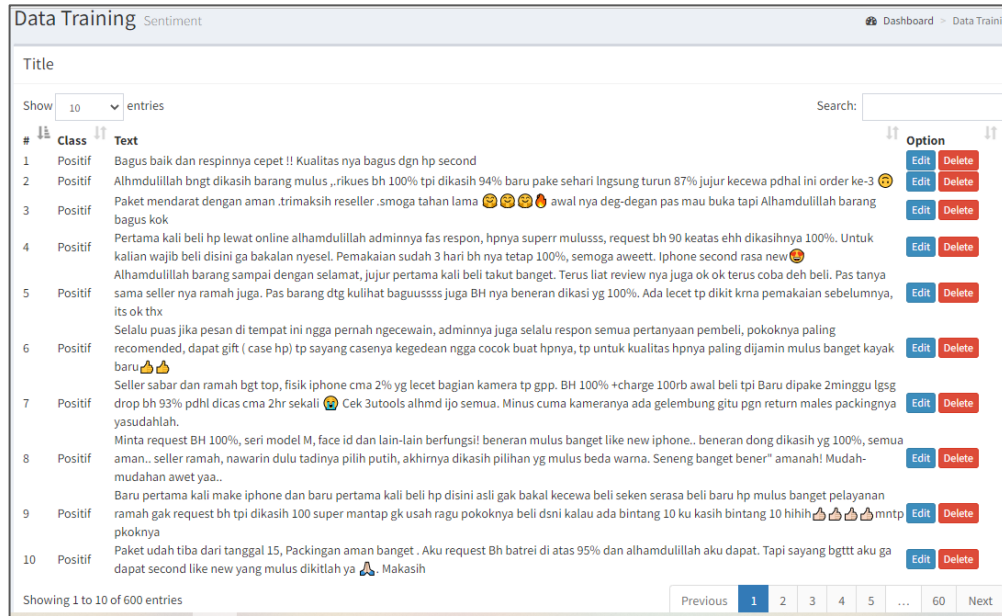


Gambar 4. Halaman Pencarian Toko

Gambar 4 merupakan halaman pencarian toko. Dalam penelitian ini dilakukan pada Kino Phone Cell untuk menjadi sampel review produk yang menjadi objek penelitian.

b. Halaman Data Latih

Halaman ini merupakan implementasi data latih menggunakan algoritma naive bayes. Pada halaman terdapat kumpulan data latih yang dilabeli secara otomatis.

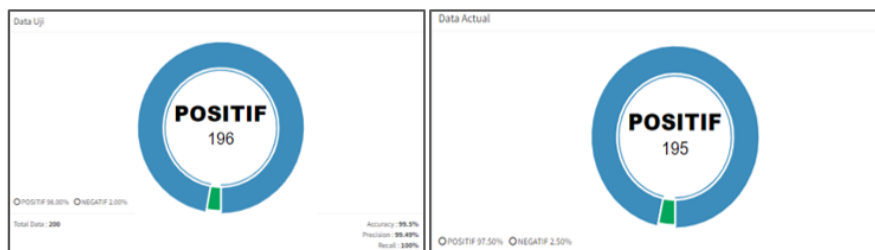


Gambar 5. Halaman Data Latih

Gambar 5 merupakan halaman dari kumpulan data latih. Pada halaman ini terdapat 600 data latih yang sudah dilabeli secara otomatis melalui tahap *lexicon based*.

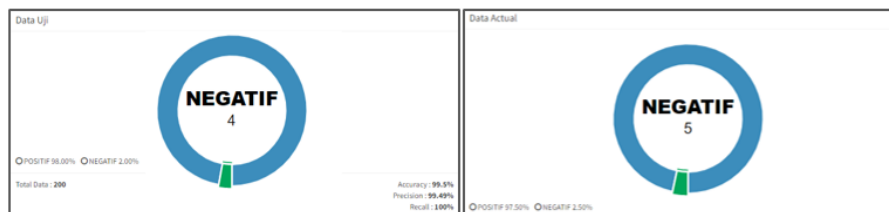
c. Halaman Analisis

Pada penelitian ini, hasil klasifikasi data akan ditampilkan pada halaman hasil analisis pada sistem website yang telah disiapkan.



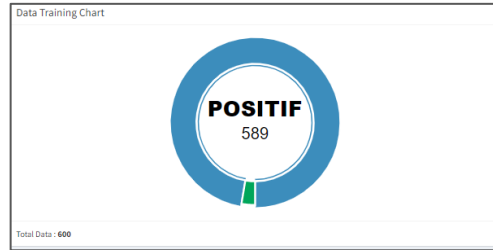
Gambar 6. Visualisasi Sentimen Positif

Berdasarkan gambar 6, perbandingan jumlah data sentiment pada data uji dan data actual divisualisasikan melalui chart, dimana pada data uji terdapat jumlah sentimen positif sebanyak 196. Sedangkan pada data actual terdapat jumlah sentimen positif sebanyak 195.



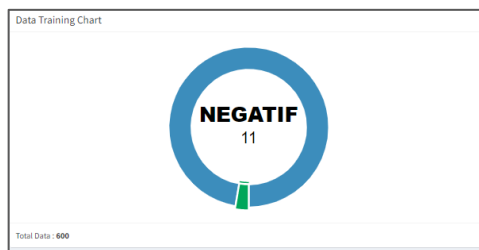
Gambar 7. Visualisasi Sentimen Negatif

Berdasarkan gambar 7, perbandingan jumlah data sentiment pada data uji dan data actual divisualisasikan melalui chart, dimana pada data uji terdapat jumlah sentimen negatif sebanyak 5. Sedangkan pada data actual terdapat jumlah sentimen negatif sebanyak 5.



Gambar 8. Visualisasi Data Latih Positif

Berdasarkan gambar 8, perbandingan jumlah data sentiment pada data training atau data latih divisualisasikan melalui chart, dimana terdapat total jumlah sentimen positif sebanyak 589.



Gambar 9. Tampilan Visualisasi Data Latih Negatif

Berdasarkan gambar 9, perbandingan jumlah data sentiment pada data *training* atau data latih divisualisasikan melalui chart, dimana terdapat total jumlah sentimen negatif sebanyak 11.

| No | Komentar  | Hasil Sentimen |
|----|---|----------------|
| 1  | Bagus baik dan respinnya cepet !! Kualitas nya bagus dgn hp second  | POSITIF        |
| 2  | Alhmdulillah bngt dikasih barang mulus ,rikues bh 100% tpi dikasih 94% baru pake sehari lngsung turun 87% jujur kecewa pdhal ini order ke-3 😞   | POSITIF        |
| 3  | Paket mendarat dengan aman .trimakasih reseller .smoga tahan lama 🤗🤗🤗🤗 awal nya deg-degan pas mau buka tapi Alhamdulillah barang bagus kok  | POSITIF        |
| 4  | Pertama kali beli hp lewat online alhamdulillah adminnya fas respon, hpnya superr mulussss, request bh 90 keatas ehh dikasihnya 100%. Untuk kalian wajib beli disini ga bakal nyesel. Pemakaian sudah 3 hari bh nya tetap 100%, semoga aweett. Iphone second rasa new 🤗                                     | POSITIF        |
| 5  | Alhamdulillah barang sampai dengan selamat, jujur pertama kali beli takut banget. Terus liat review nya juga ok ok terus coba deh beli. Pas tanya sama seller nya ramah juga. Pas barang dtg kulihat baguussss juga BH nya beneran dikasi yg 100%. Ada lecet tp dikit krna pemakaian sebelumnya, its ok thx | POSITIF        |
| 6  | Selalu puas jika pesan di tempat ini ngga pernah ngecewain, adminnya juga selalu respon semua pertanyaan pembeli, pokoknya paling recomended, dapat gift ( case hp) tp sayang casenya kegedean ngga cocok buat hpnya, tp untuk kualitas hpnya paling dijamin mulus banget kayak baru 🤗🤗                     | POSITIF        |
| 7  | Seller sabar dan ramah bgt top, fisik iphone cma 2% yg lecet bagian kamera tp gpp. BH 100% +charge 100rb awal beli tpi Baru dipake 2minggu lgsg drop bh 93% pdhl dicas cma 2hr sekali 🤗 Cek 3utools alhmd ijo semua. Minus cuma kameranya ada gelembung gitu pgn return males packingnya yasadahlah.        | POSITIF        |
| 8  | Minta request BH 100%, seri model M, face id dan lain-lain berfungsi! beneran mulus banget like new iphone.. beneran dong dikasih yg 100%, semua aman.. seller ramah, nawarin dulu tadinya pilih putih, akhirnya dikasih pilihan yg mulus beda warna. Seneng banget bener" amanah! Mudah-mudahan awet yaa.. | POSITIF        |
| 9  | Baru pertama kali make iphone dan baru pertama kali beli hp disini asli gak bakal kecewa beli seken serasa beli baru hp mulus banget pelayanan ramah gak request bh tpi dikasih 100 super mantap gk usah ragu pokoknya beli dsni kalau ada bintang 10 ku kasih bintang 10 hihhih 🤗🤗 🤗🤗mntp pokoknya         | POSITIF        |
| 10 | Paket udah tiba dari tanggal 15, Packingan aman banget . Aku request Bh batrei di atas 95% dan alhamdulillah aku dapat. Tapi sayang bgttt aku ga dapat second like new yang mulus dikitlah ya 🤗. Makasih  | POSITIF        |

Gambar 10. Hasil Data Uji

Gambar 10 merupakan halaman yang menampilkan hasil analisis data review produk yang telah dianalisis sentimennya. Pada penelitian ini, perhitungan pelabelan sentimen telah mendapatkan hasil perbandingan jumlah data kelas sentimen seperti terlihat pada tabel 6.

**Tabel 6.** Perbandingan Jumlah Kelas Sentimen Data Latih

| Kelas Sentimen | Jumlah     |
|----------------|------------|
| Positif        | 589        |
| Negatif        | 11         |
| <b>Total</b>   | <b>600</b> |

Tabel 6 menunjukkan jumlah kelas sentimen pada data latih yang terdiri dari 2 kelas yaitu 589 jumlah sentimen positif dan 11 jumlah sentimen negatif.

**Tabel 7.** Perbandingan Jumlah Kelas Sentimen Data Uji

| Kelas Sentimen | Jumlah     |
|----------------|------------|
| Positif        | 196        |
| Negatif        | 4          |
| <b>Total</b>   | <b>200</b> |

Tabel 7 menunjukkan jumlah kelas sentimen pada data uji yang terdiri dari 2 kelas yaitu 196 jumlah sentimen positif dan 4 jumlah sentimen negatif.

**Tabel 8.** Perbandingan Jumlah Kelas Sentimen Data Actual

| Kelas Sentimen | Jumlah     |
|----------------|------------|
| Positif        | 195        |
| Negatif        | 5          |
| <b>Total</b>   | <b>200</b> |

Tabel 8 menunjukkan jumlah kelas sentimen pada data actual yang terdiri dari 2 kelas yaitu 195 jumlah sentimen positif dan 5 jumlah sentimen negatif.

### 3.2 Pengujian

Pengujian berupa evaluasi hasil didalam penelitian ini menggunakan counfusion matrix. Dari counfusion matrix akan digunakan untuk menghitung nilai *accuracy*, *precision*, dan *recall* berguna untuk membandingkan pengujian mana yang cukup baik untuk membangun model klasifikasi. Pengujian data yang di lakukan dengan cara membandingkan hasil perhitungan manual dan hasil perhitungan secara sistem menggunakan metode Naïve Bayes. Dan juga membandingkan data sentimen asli atau data actual dengan data hasil prediksi atau data uji.

**Tabel 9.** Confusion Matrix Data Uji

| Prediksi | Data Actual |         |
|----------|-------------|---------|
|          | Positif     | Negatif |
| Positif  | 195         | 1       |
| Negatif  | 0           | 4       |

Dari data pada Tabel 9 confusion matrix maka dapat dihitung *accuracy*, *precision*, *recall* sebagai berikut :

$$Accuracy = \frac{195 + 4}{187 + 0 + 1 + 4} \times 100 = 99,5\%$$

$$Precision = \frac{195}{195 + 1} \times 100\% = 99,49\%$$

$$Recall = \frac{195}{195 + 0} \times 100\% = 100\%$$

Dari hasil pengujian diperoleh hasil *accuracy* sebesar 99,5%, *precision* sebesar 99,49%, *recall* sebesar 100%. Dengan demikian dapat disimpulkan metode klasifikasi algoritma Naïve Bayes cukup relevan meskipun akurasi nya belum 100%.

## 4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan maka diperoleh kesimpulan, dengan algoritma Naïve Bayes dapat memprediksi kelas sentimen pada ulasan produk di toko Kinophonecell pada platform e-commerce Shopee. Sistem analisis sentimen review pengguna Shopee menggunakan metode Naïve Bayes ini menghasilkan nilai akurasi dengan satu kali pengujian dengan nilai sebesar *accuracy* sebesar 99,5%, *precision* sebesar 99,49%, *recall* sebesar 100%. Dengan demikian dapat disimpulkan metode klasifikasi algoritma Naïve Bayes cukup relevan meskipun akurasi nya belum 100%. Hasil analisis penelitian ini menghasilkan sentimen positif lebih unggul berjumlah 196 dan jumlah sentimen negatif berjumlah 4. Jumlah data latih dalam sistem memiliki pengaruh terhadap prediksi sistem. Selain jumlah, kualitas data latih juga berperan

karena semakin tinggi kualitas data maka sistem akan mendapatkan kosakata yang semakin besar sehingga akan lebih tepat dalam memprediksi kelas sentimen.

## REFERENCES

- [1] Agustina C.A, N *et al*, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store." MALCOM, 2022.
- [2] Amrullah, A. Z., Sofyan Anas, A., and Hidayat, M. A. J., "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," BITE, 2020.
- [3] Asri Amaliza Fathia Mausea, I. A. S., "RANCANG BANGUN APLIKASI PENDAFTARAN PASIEN ONLINE DAN PEMERIKSAAN DOKTER DI KLINIK PENGOBATAN BERBASIS WEB," Rekayasa Informasi, 2021.
- [4] A.R. Isnain, A. Sihabuddin, and Y. Suyanto, "Bidirectional Long Short Term Memory Method and Word2vec Extraction Approach for Hate Speech Detection," IJCCS, 2020.
- [5] A. Surahman, A. F. O. P. Pasaribu, and D. Darwis, "Ekstraksi Data Produk e-Marketplace Sebagai Strategi Pengolahan Segmentasi Pasar Menggunakan Web Crawler," Sist. J. Sist. Inf., 2020.
- [6] A. Surahman, "PENGEMBANGAN MARKET SEGMENTASI UNTUK MENCAPAI KEUNGGULAN BERSAING PADA E-MARKETPLACE," J. Komput. dan Inform., 2020.
- [7] A. Rahman, E. Utami, and S. Sudarmawan, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," J. Komtika (Komputasi dan Inform., vol. 5, no. 1, pp. 60–71, 2021, doi: 10.31603/komtika.v5i1.5188.
- [8] Azzahra, Sitti Aliyah, & Arief Wibowo. "Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi Dengan Algoritme Naïve Bayes Untuk Ulasan Wisata Kuliner Pada Web Tripadvisor." Jurnal Teknologi Informasi dan Ilmu Komputer 7(4): 737. 2020.
- [9] Darwis, D., Siskawati, N., & Abidin, Z., "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," J. Tekno Kompak, 2021.
- [10] D. Alita, Y. Fernando, and H. Sulistiani, "Implementasi Algoritma Multiclass SVM pada Opini Publik Berbahasa Indonesia di Twitter," J. Tekno Kompak, 2020.
- [11] D. Musfiroh *et al.*, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon," MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 1, no. 1, pp. 24–33, 2021.
- [12] D. Pratmanto, R. Rousyati, F. F. Wati, A. E. Widodo, S. Suleman, and R. Wijianto, "App Review Sentiment Analysis Shopee Application in Google Play Store Using Naive Bayes Algorithm," J. Phys. Conf. Ser., vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012043.
- [13] Giovani, A. P., Ardiansyah, Haryanti, T., Kurniawati, L., & Gata, W. (2020). "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi." Teknoinfo, 2020.
- [14] Gunawan Sudarsono *et al*, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner." Journal of Business and Audit Information Systems, 2021.
- [15] Kosasih, R., & Alberto, A. "Analisis Sentimen Produk Permainan Menggunakan Metode TF-IDF Dan Algoritma K-Nearest Neighbor." InfoTekJar, 2021.
- [16] Masripah, Siti, & Lila Dini Utami. "Algoritma Klasifikasi Naïve Bayes Untuk Analisa Sentimen Aplikasi Shopee." Swabumi 8(2): 114–17, 2020.
- [17] Muktafin, Elik Hari, Kusri Kusri, & Emha Taufiq Luthfi. "Analisis Sentimen Pada Ulasan Pembelian Produk Di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing." Jurnal Eksplora Informatika 10(1): 32–42, 2020.
- [18] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," CESS (Journal Comput. Eng. Syst. Sci., vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186.
- [19] Sari, R., "Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi menggunakan Algoritma K-Nearest Neighbor (K-NN)." Evolusi, 2020.
- [20] S. T. Rizaldi and M. Mustakim, "Perbandingan Teknik Pembagian Data untuk Klasifikasi Sarana Akses Air pada Algoritma K-Nearest Neighbor dan Naïve Bayes Classifier," Semin. Nas. Teknol. Informasi, Komun. dan Ind. 12, pp. 130–137, 2020.