



Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Shopee Dengan Menggunakan Algoritma Naïve Bayes

Rahel Lina Simanjuntak

Universitas Negeri Medan

Theresia Romauli Siagian

Universitas Negeri Medan

Vina Anggriani

Universitas Negeri Medan

Arnita

Universitas Negeri Medan

Universitas Negeri Medan, Jl Willem Iskander / Psr V 20221 Medan North Sumatra
rahelsimanjuntak12@gmail.com, theresiiasiagian11@gmail.com, vinaanggriani81@gmail.com

Abstract. Every year, the e-commerce industry in Indonesia grows and develops at a rapid pace. In Indonesia, many online marketplaces have sprung up, including Tokopedia, Lazada, and Shopee. People are very interested in all e-commerce companies because they are hassle-free and instant. Among the most famous is Shopee, which offers a wide range of services and also presents a ratings and reviews column. This feature allows users to express their feelings about Shopee. Based on the information gathered from previous user comments, consumers can use these ratings to identify and trust both excellent and negative recommendations of the app they want to use. Sentiment analysis results include both favorable and negative user reviews by scoring, classifying, and filtering viewpoints to help businesses or users. The author uses the naïve bayes algorithm in this research. The Naïve Bayes Classifier approach will be used in this research to perform sentiment classification. The author then uses associations between frequently discussed word terms or themes that are related to each other for the extraction and exploration process, as well as descriptive statistics. Naïve Bayes Classifier is a binary classification technique that applies Bayesian principles with a strong assumption of independence, utilizing simple statistical probabilities.

Keywords: Sentimen Analysis, Naive Bayes, Shopee

Abstrak. Setiap tahun, industri e-commerce di Indonesia tumbuh dan berkembang dengan sangat cepat. Di Indonesia, banyak sekali pasar online yang bermunculan, termasuk Tokopedia, Lazada, dan Shopee. Masyarakat sangat tertarik dengan semua perusahaan e-commerce karena tidak sulit dan instan. Di antara yang paling terkenal adalah Shopee, yang menawarkan berbagai layanan dan juga menyajikan kolom

Received November 1, 2023; Revised November 17, 2023; Accepted November 25, 2023

* Rahel Lina Simanjuntak, rahelsimanjuntak12@gmail.com

peringkat dan ulasan. Fitur ini memungkinkan pengguna untuk mengekspresikan perasaan mereka tentang Shopee. Berdasarkan informasi yang dikumpulkan dari komentar pengguna sebelumnya, konsumen dapat menggunakan penilaian ini untuk mengidentifikasi dan mempercayai rekomendasi yang sangat baik maupun negatif dari aplikasi yang ingin digunakan. Hasil analisis sentimen mencakup ulasan pengguna baik yang menguntungkan maupun negatif dengan menilai, mengklasifikasikan, dan menyaring sudut pandang untuk membantu bisnis atau pengguna. Penulis menggunakan algoritma naïve bayes dalam penelitian ini. Pendekatan Naïve Bayes Classifier akan digunakan dalam penelitian ini untuk melakukan klasifikasi sentimen. Penulis kemudian menggunakan asosiasi antar istilah kata atau tema yang sering dibicarakan yang saling berkaitan satu sama lain untuk proses ekstraksi dan eksplorasi, serta statistik deskriptif. Naïve Bayes Classifier adalah teknik klasifikasi biner yang menerapkan prinsip Bayesian dengan asumsi independen yang kuat, dengan memanfaatkan probabilitas statistik sederhana.

Kata kunci: Sentimen Analysis, Naive Bayes, Shopee

LATAR BELAKANG

Perkembangan e-commerce di Indonesia mengalami pertumbuhan dan perkembangan yang sangat pesat setiap tahunnya. Pesatnya perkembangan e-commerce di Indonesia menyebabkan munculnya banyak marketplace seperti Tokopedia, Shopee, dan Lazada. Pasar adalah model bisnis yang memungkinkan penjual menjual di Internet. Konsep dasarnya sama dengan pasar tradisional, dengan toko-toko berbeda yang menjual barang dalam satu kawasan. Namun, pasar memungkinkan pedagang untuk melakukan transaksi online melalui media internet (Ramadhan dkk, 2022). Hal ini sejalan dengan semakin banyaknya pengguna internet Indonesia yang menginginkan segalanya secara instan dan mudah. Selain itu potensi lainnya yang mendorong berkembangnya bisnis e-commerce di Indonesia disebabkan oleh jumlah masyarakat yang tergolong besar dan letak geografis Indonesia sebagai negara kepulauan sehingga bisnis e-commerce dapat menjembatani antara para pelaku bisnis dan konsumennya (Gumilang, 2018).

Shopee saat ini menjadi marketplace yang paling banyak dikunjungi di Indonesia. Menurut iPrice 2022, rata-rata pengunjung bulanan Shopee mencapai 132,8 juta pada kuartal pertama tahun 2022. Shopee merupakan aplikasi belanja, jual beli yang dapat diunduh dari halaman Google Playstore. Halaman Google Playstore menyediakan fitur rating dan review yang luas. Peringkat dan fitur ini ditujukan bagi pengguna untuk mengevaluasi aplikasi (Limbong dkk, 2022).

Dengan bantuan review dari pengguna lama, informasi tentang produk dan layanan yang baik menjadi semakin valid, dan banyak pengguna internet yang mempercayai rekomendasi dan opini yang digunakan oleh pengguna sebelumnya dalam bentuk produk dan layanan (Cahyaningtyas dkk, 2021). Tidak ada cara yang sistematis dan akurat untuk mengkategorikan ulasan sebagai positif atau negatif. Informasi konsumen berdasarkan jenis lampiran memudahkan pengembang memperoleh data lampiran dari konsumen pada aplikasi e-commerce (Limbong dkk, 2022). Aplikasi Shopee sempat berada pada posisi pertama sebagai aplikasi e-commerce yang menawarkan layanan gratis ongkos kirim keseluruh Indonesia, akan tetapi pada bulan Maret 2018 posisi tersebut tergeser oleh salah satu aplikasi e-commerce lainnya yaitu Lazada. Hal ini disebabkan karena adanya perubahan sistem Shopee yang menyebabkan adanya pembatasan biaya gratis ongkos kirim. Layanan shopee tersebut dapat diakses melalui website, aplikasi android (Play Store), ataupun Google TV. Aspek yang menarik dari aplikasi Google Play adalah kemampuannya untuk menyertakan ulasan dari pengguna (Gumilang, 2018).

Analisis sentimen adalah teknik yang menghilangkan kata-kata dan simbol yang tidak relevan dari data dan mengubah data kualitatif menjadi data kuantitatif. Data ulasan pengguna kemudian diklasifikasikan untuk mendapatkan ulasan positif dan negatif (Manik dkk, 2021). Saat ini analisis sentimen menjadi topik yang banyak digunakan oleh para peneliti dengan tujuan memberikan informasi dari dataset yang tidak terstruktur (Arsi & Waluyo, 2021). Analisis sentimen ini dapat diterapkan pada opini di segala bidang, termasuk ekonomi, politik, masyarakat, dan hukum. Analisis sentimen ini juga memungkinkan Anda mengelompokkan polaritas teks untuk mengetahui apakah opini tentang suatu dokumen bersifat positif atau negative (Ramadhan dkk, 2022).

Dari ulasan tersebut kita dapat memperoleh sentimen positif dan negatif berupa penilaian dari masyarakat. Hal ini memungkinkan kami untuk mempengaruhi dan mengevaluasi apakah aplikasi e-commerce Shopee layak untuk diunduh dan digunakan untuk belanja online. Anda dapat melihat hasil analisis sentimen yaitu komentar positif dan negatif dari pengguna, rating aplikasi, pemberian rating, penyaringan opini untuk membantu perusahaan dan penulis. Oleh karena itu, penulis ingin menganalisis ulasan tersebut dengan menggunakan aplikasi algoritma Naive Bayes (Noer, 2023).

Algoritma Naive Bayes adalah algoritma klasifikasi sederhana yang menghitung sekumpulan probabilitas dengan menambahkan dan menggabungkan nilai dari kumpulan

data tertentu (Hayuningtyas, 2019). Dalam penelitian ini, teknik algoritma Naive Bayes digunakan untuk mengklasifikasikan opini masyarakat mengenai aplikasi Tokopedia, Shopee, dan Lazada. Dari berbagai sumber, algoritma klasifikasi Naïve Bayes lebih disukai karena kesederhanaan dan kecepatannya (Goel dkk, 2016). Klasifikasi algoritma Naive Bayes ini memberikan akurasi yang tinggi dan dapat memproses data dalam jumlah yang sangat besar dengan cepat (Ramadhan dkk, 2022).

Dalam penelitian ini proses klasifikasi sentimen akan dilakukan dengan menggunakan metode Naïve Bayes Classifier. Kemudian untuk proses ekstraksi dan eksplorasi penulis menggunakan statistik deskriptif dan asosiasi antar terms (kata atau topik yang sering dibicarakan) yang saling berkaitan. Metode Naïve Bayes Classifier ini adalah metode klasifikasi biner yang memanfaatkan probabilitas statistika sederhana dengan menerapkan aturan Bayesian menggunakan asumsi independen yang kuat. Metode Bayesian merupakan metode analisis berdasar informasi sampel dan informasi prior. Gabungan dari informasi sampel dengan informasi prior tersebut dinamakan peluang posterior. Penerapan metode Naive Bayes Classifier adalah dengan menggunakan data latih untuk menguji data uji (Gumilang, 2018).

Dalam penelitian ini, kami menggunakan metode scraping untuk mengumpulkan data review dari pengguna aplikasi Shopee. Kemudian, dalam proses analisisnya, penulis mencoba mengklasifikasikan teks review pengguna untuk mengidentifikasi mana review yang positif dan mana yang negatif. Setelah melakukan klasifikasi, penulis mencoba mengekstrak dan menggali sekomprehensif mungkin informasi apa saja yang terkandung dalam ulasan tersebut dianggap penting untuk digunakan untuk berbagai keperluan (Gumilang, 2018). Penelitian ini dilakukan dengan menggunakan algoritma multinomial Naive Bayes dan tools Google Colaboratory untuk analisis sentimen pada salah satu aplikasi e-commerce Shopee. Penelitian ini diharapkan dapat membantu pemangku kepentingan Shopee yang tertarik untuk melihat informasi yang terdapat dalam kumpulan ulasan dan komentar mereka serta memberikan wawasan mengenai sentimen yang terkandung dalam ulasan yang disampaikan pelanggan mengenai aplikasi tersebut (Agustina dkk, 2022).

KAJIAN TEORITIS

1. Analisis Sentimen

Analisis sentimen merupakan bagian dari teks mining. Dataset yang nantinya dianalisis bisa berasal dari berbagai sumber, misalnya kolom komentar yang tersedia di aplikasi Shopee atau unggahan orang-orang tentang pendapat dan perasaannya terhadap sesuatu hal. Orang yang berprofesi sebagai data scientist cenderung mendengar analisis sentimen. Analisis sentimen merupakan hasil proses analisis data berupa pandangan dan opini guna menarik kesimpulan dari berbagai opini yang ada. Hasil analisis sentimen dapat berupa persentase sentimen positif, negatif, atau netral (Larasati dkk, 2022).

Analisis sentimen berguna untuk berbagai masalah yang menjadi perhatian para pakar dan peneliti interaksi manusia-komputer, serta bidang-bidang seperti sosiologi, pemasaran dan periklanan, psikologi, ekonomi, dan ilmu politik. Salah satu media sosial yang biasa digunakan masyarakat untuk menyampaikan pendapatnya adalah melalui aplikasi Shopee secara langsung. Anda juga dapat menulis opini Anda di Shopee dengan sangat mudah menggunakan review yang telah disediakan. Shopee banyak digunakan sebagai media sosial oleh para peneliti di bidang pemrosesan bahasa alami (NLP). Selain konsep data teks sederhana yang dapat dicari, Twitter juga menyediakan fungsionalitas API yang memungkinkan peneliti mengambil data tersebut dengan mudah. Sampai saat ini, beberapa penelitian telah dilakukan dengan menggunakan algoritma klasifikasi yang berbeda (Larasati dkk, 2022).

Data yang dikumpulkan dianalisis untuk memberikan wawasan untuk menarik kesimpulan. Hasilnya adalah ulasan yang positif, negatif, atau netral (Bahwari, 2019). Ada empat jenis analisis sentimen:

- a. Analisis sentimen bertingkat digunakan untuk menginterpretasikan peringkat bintang ulasan. 5 bintang sangat positif, 1 bintang sangat negatif .
- b. Pengenalan emosi. Digunakan untuk mendeteksi emosi seperti kebahagiaan, kemarahan, kesedihan, dan frustrasi.
- c. Berbasis aspek. Digunakan untuk mengidentifikasi aspek atau karakteristik yang memiliki emosi positif, negatif, atau netral.
- d. Multilingual: Mengenali bahasa dalam teks sesuai dengan bahasa yang dipilih.

2. E-commerce Shopee

Minat konsumen terhadap bisnis semakin meningkat karena kemajuan teknologi. Meningkatnya teknologi berarti perdagangan semakin terbatas dalam ruang dan waktu. Jika dulu perdagangan hanya bisa dilakukan dengan pertemuan secara langsung, kini perdagangan melalui internet bisa dilakukan tanpa harus bertatap muka. Preferensi pembelian yang diciptakan oleh media online merupakan keinginan seseorang untuk membeli suatu produk yang ditawarkan oleh suatu produsen kepada konsumen melalui media online (Anwar & Adidarma, 2016).

Selain itu, ada faktor penting lain dalam belanja online yang muncul dari dalam diri setiap individu: kepercayaan. Kepercayaan merupakan faktor penting dan krusial dalam kegiatan perekonomian, terutama dalam mencegah perilaku oportunistik yang tidak diharapkan. Khususnya dalam bidang belanja online, menurut David Gefen, Elena Karahanna dan Detmar W Straub dalam jurnalnya yang bertajuk Trust and Tam in Online Shopping (2003), pertanyaan tentang kepercayaan ini lebih penting dibandingkan saat membeli secara langsung, karena jika Anda membeli secara online, pembeli mempunyai lebih sedikit ruang dan waktu untuk mengevaluasi produk yang ingin dibelinya. Biasanya calon pembeli pertama kali akan melihat testimoni produk di toko online yang ingin dibelinya. Shopee merupakan perusahaan e-commerce yang berada di bawah payung SEA Group atau dulunya Garena, sebuah perusahaan Internet di Asia Tenggara. Shopee memimpin dalam hal pengunduhan di Indonesia, namun masih menempati peringkat keempat dalam jumlah pengunjung unik per bulan (Japariato & Adelia, 2020).

3. Algoritma Naïve Bayes

Salah satu fungsi Data Mining adalah klasifikasi data, khususnya memetakan (mengklasifikasikan) data ke dalam satu atau lebih kelas yang telah ditentukan sebelumnya. Salah satu metode klasifikasi data adalah Naïve Bayes, yaitu metode pembelajaran mesin yang menggunakan perhitungan probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Cara kerja Naive Bayes adalah memprediksi probabilitas di masa depan berdasarkan pengalaman sebelumnya (Amolik dkk, 2016). Naïve Bayes atau biasa kita sebut Multinomial Naïve Bayes merupakan model metode Bayesian yang disederhanakan, cocok untuk klasifikasi teks atau dokumen.

Menurut He dan Tan (2012), klasifikasi adalah jenis model data mining di mana pengklasifikasi dirancang untuk menentukan label kategorikal, seperti "aman" atau

"beresiko" untuk data keuangan, "ya" atau "tidak" untuk data penjualan, atau "pengobatan A", "pengobatan B", atau "pengobatan C" untuk data medis. Proses aplikasi dilakukan dengan membuat model yang lebih rinci dari data, yang juga dikenal sebagai data latih. Setelah model terbentuk dari proses pelatihan tersebut, data baru dapat diterapkan dalam proses aplikasi, yang juga dikenal sebagai proses uji.

Menurut Suntoro dkk. (2018), algoritma Naïve Bayes adalah algoritma klasifikasi yang didasarkan pada teori Bayesian dalam statistik. Algoritma Naïve Bayes berguna untuk mengestimasi probabilitas dari kelas yang diberikan. Statistik yang penting dalam data mining adalah Naïve Bayes. Hasil ini diperoleh dengan menggunakan pendekatan trade-off kuantitatif untuk menganalisis berbagai masalah klasifikasi dengan menggunakan probabilitas. Prinsip utama dari Naïve Bayes Classifier adalah asumsi yang sangat kuat (naif) bahwa setiap kondisi atau kejadian akan independen dari yang lain (Anjar dkk, 2020).

4. Penelitian terdahulu

Pada beberapa penelitian yang dilakukan, algoritma Naïve Bayes dapat digunakan untuk memprediksi sentimen ulasan di aplikasi Shopee. Dan akurasi datanya lebih baik dibandingkan metode lain seperti K-Nearest Neighbor. Rata-rata pengguna algoritma ini juga melaporkan bahwa datanya memiliki tingkat akurasi yang tinggi meskipun tidak 100% akurat (Hasugian dkk, 2023) (Yurmalin dkk, 2023).

5. Kesenjangan pengetahuan

Shopee memberi pengguna kemampuan untuk berdagang barang secara online. Namun tidak jarang konsumen merasa ragu ketika membeli produk secara online dengan harga pasar yang cukup tinggi. Review produk memberikan informasi kualitas suatu produk dan mempunyai dampak besar bagi konsumen dan produsen. Ulasan produk mencakup penilaian bintang serta ulasan yang mencakup komentar, pujian dan kritik, serta umpan balik terhadap produk yang dibeli. Kepuasan pelanggan merupakan isu penting yang menciptakan tujuan perusahaan. Disadari atau tidak, ulasan pelanggan yang diposting di media sosial memengaruhi calon pelanggan dengan cara yang berbeda. Saat ini, di situs review produk terdapat banyak opini, sehingga menghasilkan banyak sekali review. Ketersediaan data rating atau review menjadi asal mula penelitian ini sehingga

memungkinkan dilakukannya evaluasi terhadap produk analisis opini (Hasugian dkk, 2023).

METODE PENELITIAN

1. Populasi dan Sampel

Penelitian Populasi dalam penelitian ini adalah website dari aplikasi shopee, yaitu semua data ulasan Shopee dari tahun 2015 – sekarang. Sedangkan untuk sampel yang digunakan adalah ulasan Shopee dari sejak pertengahan upgrade aplikasi Shopee yaitu dari tahun 2015 hingga terakhir upgrade pada akhir pertengahan Oktober.

2. Waktu Penelitian

Data yang diakses dari website shopee diakses pada tanggal 10 oktober 2023 dengan melakukan web scrapping.

3. Jenis dan Sumber Data

Jenis data yang digunakan pada penelitian ini ialah data primer. Data primer itu sendiri menurut Sekarang (2011) merupakan data yang mengacu pada suatu informasi yang diperoleh dari tangan pertama oleh peneliti yang dapat berkaitan dengan variabel minat untuk tujuan spesifik studi. Data pada penelitian ini diperoleh dengan menggunakan teknik scraping dari halaman situs website shopee dengan menggunakan aplikasi Scraper bawaan dari Google Chrome dengan alamat website yaitu <https://play.google.com/store/apps/details?id=com.shopee.id>). Data yang diperoleh berupa ulasan atau review pengguna aplikasi Shopee yang diambil sejak tanggal upgrade mulai dari tahun 2015 sampai terakhir upgrade tanggal 14 Oktober 2023 sebanyak 1000 ulasan.

4. Variabel dan Jenis Penelitian

Variabel penelitian merupakan segala sesuatu yang dapat berbentuk apa saja dan dapat ditetapkan oleh penulis untuk dipelajari sehingga dapat memperoleh informasi yang kemudian dapat ditarik menjadi kesimpulan (Gumilang, 2018). Adapun variabel yang digunakan dalam penelitian ini hanya ada dua macam, yaitu:

- a. Username merupakan nama pengguna yang memberikan ulasan.
- b. Date merupakan tanggal dibuatnya ulasan
- c. Review merupakan ulasan atau isi komentar pengguna.
- d. Score merupakan untuk membuat rating pada ulasan shopee

Jenis penelitian yang digunakan pada penelitian ini yaitu:

- a. Berdasarkan Analisa dan jenis data Variabel yang digunakan pada penelitian ini yaitu variable kualitatif karena data yang digunakan berbentuk kalimat.
- b. Berdasarkan tingkat eksplanasi Variabel yang digunakan pada penelitian ini yaitu penelitian deskriptif karena penelitian yang dilakukan untuk mengetahui nilai variable mandiri, baik satu variable atau lebih tanpa membuat perbandingan.

5. Metode Pengumpulan Data

Pada penelitian ini, menggunakan metode pengumpulan data dengan menggunakan teknik web scraping. Menurut Vargiu (2012) web Scraping merupakan suatu gabungan teknik yang digunakan untuk mendapatkan informasi dari sebuah website secara otomatis tanpa harus menyalinnya secara manual. Adapun alat yang digunakan untuk web scraping ialah software Scraper yang ada di Google Chrome (Gumilang, 2018).

HASIL DAN PEMBAHASAN

1. Web Scrapping

Dengan menggunakan notebook Jupyter dan bahasa pemrograman Python untuk melakukan web scraping pada website Shopee, didapatkan kurang lebih 1000 titik data baru pada tahun 2023. Proses ini dilakukan dengan memasang library google-play-scraper yang digunakan untuk melakukan scraping data dari website Shopee hanya dengan memasukkan ID aplikasi yang perlu dibuka untuk melihat data yang ada pada link aplikasi yang dimaksud. Hasil scraping data harus diekspor dalam format Excel (.xlsx). Contoh dokumen yang dihasilkan melalui web scraping adalah sebagai berikut:

Tabel 1. Data hasil web scrapping

Content	Score	Label
Pesanan saya di batalkan, katanya melakukan aktivitas tdk wajar. Padahal saya cuma Klaim gratis ongkir dan disc. 20 ribu. Wkwkwk Promo elit, klaim vocer sulit hahaha	1	Negatif
Shopee dengki, masa udah gabisa lagi belanja barang dr luar negeri, padahal udah dimasukkin keranjang. Mana sering gak terload lagi gambar produk" yg kita cari, padahal sinyal bagus.	1	Negatif
sangat membantu sya berbelanja tanpa harus berpergian ketempatnya, paalagi harus bawa anak anak agak repot. rmh saya juga jauh dari pusat belanja jadi memnudahkan saya utk mencari sesuatu	5	Positif
Tolong shopee,saya belanja di mnccosmetic Medan tapi mereka ga amanah,saya pesan toner yg dikirim malah facialwash,udah saya ga punya duit,mau beli toner malah di kirim gtu ya Allah sangat' kecewa,admin nya di chat juga ga di bales, astagfirullah harus bagaimana lagi saya,	1	Negatif

ANALISIS SENTIMEN ULASAN PADA APLIKASI E-COMMERCE SHOPEE DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

Apalagi kurir nya gamau antar sampai rumahh,kita disuruh ngambil paket nya ,ga bakal belanja disini lagi sih kalo gini.		
Kita dipaksa update aplikasi, tapi setelah update ttp ga fungsi jug. Minta bantuan CS juga ga guna. Sampe sekarang belum ada penyelesaian. Coba hubungi cs lagi, jawaban yg sama suruh menunggu lagi. Berhari2 ga kelar2 tuh masalah.	1	Negatif
Ni iklan shopee sangat mengganggu, yang nyebar iklannya sangat tolol sekali, klo penyebar iklan cerdas gak mungkin nyebar ke yang sudah punya aplikasi Shopee, dan klo ini dari pihak shopee yang langsung, aku jamin cicilan gak akan aku bayar, sumpah, catat komen saya ini, udah berbagai cara dicoba buat hilangin ni iklan tapi gak berhasil, peduli amat ada yang nagih ke rumah, aku ajak perang, karena iklan ini sudah mengganggu, saat bekerja tiba" datang iklan bener" aku emosi	5	Positif
Sering kecewa si sama aplikasi nya, sering hilang akun tiba2, karna selalu saja tiba2 di suruh login balek, klo masalah belanja fine2 aja	2	Negatif
Jasa pengiriman shopee sangat lama 🤔 Ndk usah belanja di shopee. Mending tempat sebelah aja. Kalau tiktok shop ngk tutup, banyak org belanja disana, karna pengiriman nya cepat. Delete aja APLIKASI SHOPEE teman	1	Negatif
Ini gimana Shopee kok pas saya check out gak ada di list pesanan terus saya check out lagi masih tetap gk ada sampai 4 kali tetep gak ada di list pesanan tapi sekarang malam Minggu pada ada pesanan nya semua kan aneh terus saya harus bayar semua	1	Negatif

	content	score	Label
2	sangat membantu sya berbelanja tanpa harus ber...	5	Positif
5	Ni iklan shopee sangat mengganggu, yang nyebar ...	5	Positif
9	Di fikir malah rugi konsumen bila wajib unboxi...	4	Positif
15	Respon dan layanan bagus dan cepat... Banyak o...	5	Positif
16	Sangat suka banget,dengan shopee udah dapet li...	5	Positif
..	...	...	...
993	Entah salahnya dimana,setiap pengajuan spaylat...	2	Negatif
996	tidak bisa chat dengan agent shopee, untuk pen...	2	Negatif
997	Mengecewakan sekali. Membatalkan orderan sudah...	1	Negatif
998	Ini gimana sih kok di aku apk shopenya kedap k...	1	Negatif
999	Baru saja di perbarui, masih saja sama bikin e...	2	Negatif

[880 rows x 3 columns]

Gambar 1. Pengelompokan ulasan positif dan negative

2. Text Preprocessing Data

Panduan pengguna untuk aplikasi Shopee di Google Play Store adalah informasi yang tidak tersusun dengan rapi; untuk mendapatkan data yang berguna, hasil pencarian harus dilengkapi. Langkah yang dilakukan dengan cara ini yaitu case folding, stopwords, tokenizing, dan steaming (Tania Puspa dkk, 2023). Preprocessing adalah langkah penting yang harus diselesaikan sebelum menerapkan algoritma klasifikasi pada sebuah dokumen. Langkah-langkah yang terlibat dalam preprocessing adalah sebagai berikut.

a) Case Folding

Proses ini dilakukan untuk mengubah setiap huruf menjadi huruf yang lebih kecil untuk menyamakan hasil teks. Selain pengubahan huruf menjadi lebih kecil dilakukan juga penghapusan tanda baca atau karakter-karakter tertentu yang tidak diinginkan (Agustina dkk, 2022).

b) Proses Stopword

Proses Stopword merupakan proses menghilangkan kata penghubung dan kata yang tidak sesuai dengan dokumen (Sanrilla dkk, 2022).

c) Tokenizing

Proses ini dilakukan untuk memecah atau memotong suatu kalimat menjadi term-term yang dipisahkan berdasarkan spasi (Agustina dkk, 2022).

d) Stemming

Stemming adalah proses yang digunakan untuk menghilangkan kata-kata yang tidak sesuai dengan KBBI, seperti kata imbuhan, depan, ganti, dan belakang. Pada penelitian ini, stemming dilakukan dengan menggunakan library Sastrawi, yang dapat digunakan untuk melakukan stemming pada teks berbahasa Indonesia (Agustina dkk, 2022).

Tabel 2. Hasil Text Preprocessing

Content	Score	Label	Text_clean	Text_Stopword	Text_tokens	Teks_Steaming
Pesanan saya di batalkan, katanya melakukan aktivitas tdk wajar. Padahal saya cuma Klaim gratis ongkir dan disc. 20 ribu. Wkwkwk Promo elit, klaim vocer sulit hahaha	1	Negatif	pesanan saya di batalkan katanya melakukan aktivitas tdk wajar padahal saya cuma klaim gratis ongkir dan disc ribu wkwkwk promo elit	pesanan batalkan aktivitas tdk wajar klaim gratis ongkir disc ribu wkwkwk promo elit klaim vocer sulit hahaha	['pesanan', 'batalkan', 'aktivitas', 'tdk', 'wajar', 'klaim', 'gratis', 'ongkir', 'disc', 'ribu', 'wkwkwk', 'promo', 'elit', 'klaim', 'vocer', 'sulit', 'hahaha']	pesan batal aktivitas tdk wajar klaim gratis ongkir disc ribu wkwkwk promo elit klaim vocer sulit hahaha

ANALISIS SENTIMEN ULASAN PADA APLIKASI E-COMMERCE SHOPEE DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

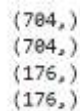
			klaim vocer sulit hahaha			
Shopee dengki, masa udah gabisa lagi belanja barang dr luar negeri, padahal udah dimasukkin keranjang. Mana sering gak terload lagi gambar produk" yg kita cari, padahal sinyal bagus.	1	Negatif	shopee dengki masa udah gabisa lagi belanja barang dr luar negeri padahal udah dimasukkin keranjang mana sering gak terload lagi gambar produk yg kita cari padahal sinyal bagus	shopee dengki udah gabisa belanja barang dr negeri udah dimasukkin keranjang gak terload gambar produk yg cari sinyal bagus	['shopee', 'dengki', 'udah', 'gabisa', 'belanja', 'barang', 'dr', 'negeri', 'udah', 'dimasukkin', 'keranjang', 'gak', 'terload', 'gambar', 'produk', 'yg', 'cari', 'sinyal', 'bagus']	shopee dengki udah gabisa belanja barang dr negeri udah dimasukkin keranjang gak terload gambar produk yg cari sinyal bagus
sangat membantu sya berbelanja tanpa harus berpergian ketempatnya, paalagi harus bawa anak anak agak repot. rmh saya juga jauh dari pusat belanja jadi memnudahkan saya utk mencari sesuatu	5	Positif	sangat membantu sya berbelanja tanpa harus berpergian ketempatnya paalagi harus bawa anak anak agak repot rmh saya juga jauh dari pusat belanja jadi memnudahkan saya utk mencari sesuatu	membantu sya berbelanja berpergian ketempatnya paalagi bawa anak anak repot rmh pusat belanja memnudahkan utk mencari	['membantu', 'sya', 'berbelanja', 'berpergian', 'ketempatnya', 'paalagi', 'bawa', 'anak', 'anak', 'repot', 'rmh', 'pusat', 'belanja', 'memnudahkan', 'utk', 'mencari']	bantu sya belanja pergi tempat paalagi bawa anak anak repot rmh pusat belanja memnudahkan utk cari
Tolong shopee,saya belanja di mnccosmetic Medan tapi mereka ga amanah,saya pesan toner yg dikirim malah facialwash,udah saya ga punya	1	Negatif	tolong shopeesaya belanja di mnccosmetic medan tapi mereka ga amanahsaya pesan toner yg dikirim malah facialwashudah saya ga punya	tolong shopeesaya belanja mnccosmetic medan ga amanahsaya pesan toner yg dikirim facialwashudah ga duitmau beli toner kirim gtu	['tolong', 'shopeesaya', 'belanja', 'mnccosmetic', 'medan', 'ga', 'amanahsaya', 'pesan', 'toner', 'yg', 'dikirim', 'facialwashudah', 'ga', 'duitmau', 'beli', 'toner',	tolong shopeesaya belanja mnccosmetic medan ga amanahsaya pesan toner yg kirim facialwashudah ga duitmau beli toner kirim gtu ya allah kecewaadmin nya chat ga bales

duit,mau beli toner malah di kirim gtu ya Allah sangat' kecewa,admin nya di chat juga ga di bales, astagfirullah harus bagaimana lagi saya, Apalagi kurir nya gamau antar sampai rumahh,kita disuruh ngambil paket nya ,ga bakal belanja disini lagi sih kalo gini.			duitmau beli toner malah di kirim gtu ya allah sangat kecewaadmin nya di chat juga ga di bales astagfirullah harus bagaimana lagi saya apalagi kurir nya gamau antar sampai rumahhkita disuruh ngambil paket nya ga bakal belanja disini lagi sih kalo gini	ya allah kecewaadmin nya chat ga bales astagfirullah kurir nya gamau rumahhkita disuruh ngambil paket nya ga belanja sih kalo gini	'kirim', 'gtu', 'ya', 'allah', 'kecewaadmin', 'nya', 'chat', 'ga', 'bales', 'astagfirullah', 'kurir', 'nya', 'gamau', 'rumahhkita', 'disuruh', 'ngambil', 'paket', 'nya', 'ga', 'belanja', 'sih', 'kalo', 'gini']	astagfirullah kurir nya gamau rumahhkita suruh ngambil paket nya ga belanja sih kalo gin
Kita dipaksa update aplikasi, tapi setelah update ttp ga fungsi jug. Minta bantuan CS juga ga guna. Sampe sekarang belum ada penyelesaian. Coba hubungi cs lagi, jawaban yg sama suruh menunggu lagi. Berhari2 ga kelar2 tuh masalah.	1	Negatif	kita dipaksa update aplikasi tapi setelah update ttp ga fungsi jug minta bantuan cs juga ga guna sampe sekarang belum ada penyelesaian coba hubungi cs lagi jawaban yg sama suruh menunggu lagi berhari ga kelar tuh masalah	dipaksa update aplikasi update ttp ga fungsi jug bantuan cs ga sampe penyelesaian coba hubungi cs yg suruh menunggu berhari ga kelar tuh	['dipaksa', 'update', 'aplikasi', 'update', 'ttp', 'ga', 'fungsi', 'jug', 'bantuan', 'cs', 'ga', 'sampe', 'penyelesaian', 'coba', 'hubungi', 'cs', 'yg', 'suruh', 'menunggu', 'berhari', 'ga', 'kelar', 'tuh']	paksa update aplikasi update ttp ga fungsi jug bantu cs ga sampe selesai coba hubung cs yg suruh tunggu hari ga kelar tuh

3. TF-IDF

Proses machine learning yang digunakan dalam penelitian ini menggunakan metode TF-IDF, yang secara teori dapat menghasilkan sebuah vektor dengan beberapa vektor sehingga setiap kata dapat dimengerti.

- a. Term Frequency (TF): Proses ini dilakukan untuk menghitung jumlah kata yang muncul dalam dataset yang ada.
- b. Inverse Document Frequency (IDF): Proses ini dilakukan untuk mengetahui kuantitas dokumen yang terdapat term yang dicari (Agustina dkk, 2022).



```
(784,)  
(784,)  
(176,)  
(176,)
```

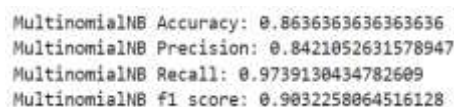
Gambar 2. Hasil TF-IDF

4. Splitting Data dan Implementasi Algoritma Naïve Bayes

Proses pengambilan data pada penelitian ini dibagi menjadi dua bagian untuk melakukan klasifikasi, yaitu menggunakan hold-out dan cross validation. Hold-Out Proses training dan testing data dilakukan dengan menggunakan metode Hold Out yang menghasilkan data testing sebesar 20% dan data training sebesar 80% (Agustina dkk, 2022).

Untuk menghasilkan matriks 2x2, model yang diimplementasikan pada data training dan testing dilakukan pengujian dengan menggunakan confusion matrix. Confusion Matrix digunakan untuk menilai akurasi dan menentukan apakah metode yang digunakan efektif dalam membuat sistem klasifikasi (Bahwari, 2019). Evaluasi data dilakukan dengan membandingkan hasil dari perhitungan manual dan perhitungan sistem dengan menggunakan metode Naive Bayes. Selain itu, membandingkan data mentah, atau data asli, dari data latih dengan data prediksi atau data uji.

Performa algoritma Naïve Bayes dengan menggunakan teknik augmentasi data Hold-Out menghasilkan nilai akurasi sebesar 86% dengan nilai precision, recall, dan f1-score yang ditunjukkan pada Gambar 3.



```
MultinomialNB Accuracy: 0.8636363636363636  
MultinomialNB Precision: 0.8421852631578947  
MultinomialNB Recall: 0.9739138434782609  
MultinomialNB f1_score: 0.9032258064516128
```

Gambar 3. Hasil Akurasi, Precision, Recall, dan f1-score Metode Hold-Out

Kecenderungan sentimen pada ulasan aplikasi Shopee pada website shopee dengan teknik pembagian data Hold-Out dapat dilihat pada gambar 4.

	precision	recall	f1-score	support
Negatif	0.84	0.97	0.90	115
Positif	0.93	0.66	0.77	61
accuracy			0.86	176
macro avg	0.89	0.81	0.84	176
weighted avg	0.87	0.86	0.86	176

Gambar 4. Hasil Sentimen ulasan

Berdasarkan hasil tersebut di atas, dapat diamati bahwa pengalaman pengguna dengan aplikasi Shopee, yang menyediakan bantuan di internet, semakin negatif dengan tingkat precision 84%, tingkat recall 97%, dan persentase f1-score 90%.

KESIMPULAN DAN SARAN

Berdasarkan hasil analisis yang telah dilakukan, langkah-langkah yang sudah dilakukan yaitu mengumpulkan data dengan menggunakan teknik web scraping pada situs Shopee; hal ini menghasilkan kurang lebih 1000 data mentah, yang kemudian diproses lebih lanjut untuk menghasilkan kurang lebih 999 data berkualitas tinggi. Performa Algoritma Naïve Bayes dengan menggunakan teknik pembagian data Hold-Out, yang mendapatkan hasil akurasi sebesar 84%. Berdasarkan hasil penelitian yang telah selesai dilakukan, menunjukkan bahwa hasil dari metode Hold-Out menghasilkan akurasi yang paling baik, maka dapat disimpulkan bahwa user experience dari pengguna Shopee yang memberikan feedback pada website kurang baik dengan tingkat akurasi sebesar 84%, recall sebesar 97%, dan f1-score sebesar 90%.

Saran yang diberikan kepada pengembang penelitian untuk penelitian yang lebih lanjut yang lebih terperinci lagi yaitu dengan menganalisis sentiment ulasan shopee dengan menggunakan metode lain seperti SVM, K-NN, Decision Tree dll dan membuat perbandingan dengan metode metode tersebut untuk mendapatkan hasil terbaik.

DAFTAR REFERENSI

- A. Amolik, N. Jivane, M. Bhandari, and M. Venkatesan, "Twittersentiment analysis of movie reviews using machine learning technique," Int. J. Eng. Technol., vol. 7, no. 6, pp. 2038–2044, 2016.
- Adidarma, R. A. dan W. (2016). Pengaruh Kepercayaan Dan Risiko Pada Minat Beli Belanja Online. Jurnal Manajemen Dan Bisnis Sriwijaya, 14, 1–14.
- Agustina, N., Citra, D. H., Purnama, W., Nisa, C., & Kurnia, A. R. (2022). Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google

- Play Store. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(1), 47–54. <https://doi.org/10.57152/malcom.v2i1.195>
- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, <https://doi.org/10.25126/jtiik.0813944>
- Bahrawi. (2019). Sentiment Analysis Using Random Forest Algorithm Online Social Media Base. *Journal of Information Technology and ITS Utilization*.
- C.Cahyaningtyas, Y. Nataliani, and I. R. Widiarsari, “Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE,” *Aiti*, vol. 18, no. 2, pp. 173–184, 2021, doi: 10.24246/aiti.v18i2.173-184.
- Goel, A., Gautam, J., & Kumar, S. (2017). Real time sentiment analysis of tweets using Naive Bayes. *Proceedings on 2016 2nd International Conference on Next Generation Computing Technologies, NGCT 2016, (October)*, 257–261. <https://doi.org/10.1109/NGCT.2016.7877424>
- Gumilang, Z. A. N. (2018). Implementasi Naive Bayes Classifier dan Asosiasi untuk Analisis Sentimen Data Ulasan Aplikasi E-Commerce Shopee pada Situs Google Play.
- Hasugian, A. H., Fakhriza, M., & Zukhoiriyah, D. (2023). Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naïve Bayes. *J-SISKO TECH (Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD)*, 6(1), 98. <https://doi.org/10.53513/jsk.v6i1.7400>
- Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita. *Jurnal Informatika*, 6(1), <https://doi.org/10.31311/ji.v6i1.4685> [7] 18–22.
- J. A. Josen Limbong, I. Sembiring, K. Dwi Hartomo, U. Kristen Satya Wacana, and P. Korespondensi, “Analisis Klasifikasi Sentimen Ulasan Pada E-Commerce Shopee Berbasis Word Cloud Dengan Metode Naive Bayes Dan K-Nearest Neighbor Analysis of Review Sentiment Classification on E-Commerce Shopee Word Cloud Based With Naïve Bayes and K-Nearest Neighbor Meth,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, pp. 347–356, 2019, doi: 10.25126/jtiik.202294960.
- Japarianto, E., & Adelia, S. (2020). Pengaruh Tampilan Web Dan Harga Terhadap Minat Beli Dengan Kepercayaan Sebagai Intervening Variable Pada E-Commerce Shopee. *Jurnal Manajemen Pemasaran*, 14(1), 35-43.
- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(9), 4305-4313.
- Manik, G., Ernawati, I., & Nurlaili, I. (2021). Analisis Sentimen Pada Review Pengguna E-Commerce Bidang Pangan Menggunakan Metode Support Vector Machine (Studi Kasus : Review Sayurbox dan Tanihub pada Google Play). *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, (September), 64–74

- (Noer, G. (2023). Implementasi Algoritma Naïve Bayes dan TF-IDF Dalam Analisis Sentimen Data Ulasan (Studi Kasus: Ulasan Review Aplikasi E-commerce Shopee di Situs Google Repository.Uinjkt.Ac.Id. [https://repository.uinjkt.ac.id/dspace/handle/123456789/68747%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/68747/1/GERALD HALIM AL RASYID NOER-FST.pdf](https://repository.uinjkt.ac.id/dspace/handle/123456789/68747%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/68747/1/GERALD%20HALIM%20AL%20RASYID%20NOER-FST.pdf)
- Ramadhan, B. Z., Adam, R. I., & Maulana, I. (2022). Analisis Sentimen Ulasan pada Aplikasi E-Commerce dengan Menggunakan Algoritma Naïve Bayes. *Journal of Applied Informatics and Computing*, 6(2), 220–225. <https://doi.org/10.30871/jaic.v6i2.4725>
- Sanrilla, S., Ransi, N., La Surimi, L. S., Andi Tenriawaru, A. T., & La Ode Saidi, L. O. S. (2022). Analisis Sentimen Masyarakat Terhadap Toko Online Aplikasi Shopee Menggunakan Metode Multinomial Naïve Bayes. *Jurnal Matematika Komputasi Dan Statistika*, 2(2), 68–75. <https://doi.org/10.33772/jmks.v2i2.9>
- Tania Puspa Rahayu Sanjaya, Ahmad Fauzi, & Anis Fitri Nur Masruriyah. (2023). Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine. *INFOTECH : Jurnal Informatika & Teknologi*, 4(1), 16–26. <https://doi.org/10.37373/infotech.v4i1.422>
- Wanto, Anjar dkk. 2020. *Data Mining : Algoritma dan Implementasi*. Medan: Yayasan Kita Menulis
- Yumarlin, M. Z., Bororing, J. E., Rahayu, S., & Putra, J. A. (2023). Analisis Sentimen Pengguna Aplikasi Shopee Menggunakan Metode Naive Bayes Classifier dan K-NN. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 12(3), 745-753.