



Original Article

Analisis Sentimen Jasa Pengiriman E-Commerce Menggunakan Naïve Bayes

Siti Nor Halizah^{1✉}, Selviana Yunita²

^{1,2}Universitas Darwan Ali, Indonesia

Korespondensi Email: stnurhalizah2003@gmail.com ✉

Abstrak:

Perkembangan e-commerce yang pesat di Indonesia mendorong meningkatnya peran jasa pengiriman sebagai faktor penentu kepuasan konsumen. Banyaknya opini dan keluhan pengguna yang disampaikan melalui media digital menjadikan data komentar sebagai sumber informasi penting untuk dianalisis. Penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna terhadap jasa pengiriman e-commerce, yaitu SPX Express, SiCepat, dan Ninja Xpress. Kajian pustaka yang digunakan meliputi teori analisis sentimen, text mining, kepuasan konsumen, serta algoritma Naïve Bayes. Metode penelitian yang diterapkan adalah pendekatan kuantitatif dengan teknik analisis sentimen. Data diperoleh melalui web scraping komentar pengguna dari website Lacako.com, kemudian dilakukan pelabelan manual, pra-pemrosesan teks, dan pembobotan TF-IDF. Proses klasifikasi dilakukan menggunakan algoritma Naïve Bayes dengan validasi 10-fold cross validation. Hasil penelitian menunjukkan bahwa metode Naïve Bayes mampu mengklasifikasikan sentimen positif dan negatif dengan kinerja yang baik. Kesimpulan penelitian ini menegaskan bahwa analisis sentimen dapat memberikan gambaran objektif mengenai persepsi masyarakat terhadap kualitas jasa pengiriman e-commerce.

Kata kunci: Analisis Sentimen, Jasa Pengiriman, E-Commerce, Naïve Bayes, Kepuasan Konsumen

Submitted	: 15 March 2026
Revised	: 18 March 2026
Acceptance	: 24 March 2026
Publish Online	: 25 March 2026

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.

Copyright © 2026 by Author. Published by Universitas Syiah Kuala



Pendahuluan

Perkembangan teknologi informasi dan komunikasi pada era digital telah membawa perubahan besar dalam berbagai aspek kehidupan, termasuk dalam aktivitas ekonomi dan bisnis. Internet tidak lagi hanya berfungsi sebagai sarana komunikasi, tetapi juga menjadi infrastruktur utama yang mendukung transaksi dan interaksi secara cepat serta tanpa batas geografis (Kurniawan, 2020; Nugroho & Syarief, 2020; Prasetyo, 2021). Dalam konteks Indonesia, perkembangan teknologi digital turut mendorong pertumbuhan e-commerce yang semakin pesat, seiring

dengan meningkatnya penggunaan internet dan perubahan perilaku konsumen yang lebih memilih transaksi secara daring ([Setiawan, 2021; Hidayat et al., 2022](#)). Kemudahan dalam membandingkan harga, membaca ulasan, dan memperoleh berbagai promosi menjadikan e-commerce sebagai salah satu pilihan utama masyarakat dalam berbelanja ([Rahman, 2020; Wibowo, 2021; Pratama, 2021](#)).

Keberhasilan transaksi e-commerce tidak hanya ditentukan oleh kualitas produk dan harga, tetapi juga oleh kualitas layanan logistik atau jasa pengiriman yang berperan sebagai penghubung antara penjual dan pembeli ([Tjiptono, 2020](#)). Permasalahan seperti keterlambatan pengiriman, kerusakan barang, atau pelayanan kurir yang kurang memuaskan seringkali menimbulkan keluhan konsumen dan dapat memengaruhi tingkat kepuasan serta reputasi perusahaan ([Putri, Handayani, & Azzahro, 2022; Rahmawati, Suryanto, & Lestari, 2022; Siregar, 2022](#)). Di era digital, konsumen juga semakin aktif menyampaikan pengalaman dan opini mereka melalui ulasan daring yang dapat memengaruhi persepsi calon pelanggan lainnya ([Rahmawati, 2021; Wijaya, 2020](#)). Oleh karena itu, analisis opini publik melalui pendekatan **sentiment analysis** dalam bidang *Natural Language Processing* menjadi penting untuk mengidentifikasi dan mengklasifikasikan kecenderungan sentimen masyarakat terhadap suatu layanan ([Liu, 2012; Feldman, 2013](#)).

Berbagai penelitian menunjukkan bahwa analisis sentimen berbasis *machine learning* mampu mengklasifikasikan opini publik dengan tingkat akurasi yang baik, terutama dengan dukungan tahapan *pre-processing* teks dan pembobotan fitur seperti TF-IDF ([Saputra & Kurniawan, 2021; Hidayat, Rahman, & Putri, 2022; Pratama & Widya, 2023](#)). Salah satu algoritma yang banyak digunakan adalah Naïve Bayes, yaitu metode klasifikasi berbasis probabilitas yang sederhana namun efektif dalam mengolah data teks berukuran besar ([Utami, 2021; Fauzi & Hidayat, 2021](#)). Evaluasi performa model klasifikasi umumnya dilakukan menggunakan metrik accuracy, precision, dan recall untuk menilai tingkat ketepatan dan keandalan *model* ([Hidayat & Prasetyo, 2020; Wulandari, Lestari, & Putri, 2021](#)). Meskipun demikian, sebagian besar penelitian sebelumnya masih berfokus pada ulasan di marketplace atau media sosial seperti Twitter, sehingga kajian yang secara khusus menganalisis sentimen komentar pengguna jasa pengiriman dari situs ulasan independen masih relatif terbatas ([Saraswati, 2023; Sari, 2022](#)).

Berdasarkan kondisi tersebut, penelitian ini memanfaatkan komentar pengguna dari platform ulasan independen untuk menganalisis sentimen terhadap layanan jasa pengiriman e-commerce di Indonesia. Penelitian ini bertujuan untuk mengklasifikasikan sentimen komentar pengguna terhadap layanan pengiriman menggunakan metode Naïve Bayes serta mengevaluasi performa model berdasarkan metrik accuracy, precision, dan recall. Hasil penelitian diharapkan dapat memberikan gambaran objektif mengenai persepsi masyarakat terhadap kualitas layanan jasa pengiriman serta menjadi dasar evaluasi bagi perusahaan dalam meningkatkan kualitas layanan dalam ekosistem e-commerce.

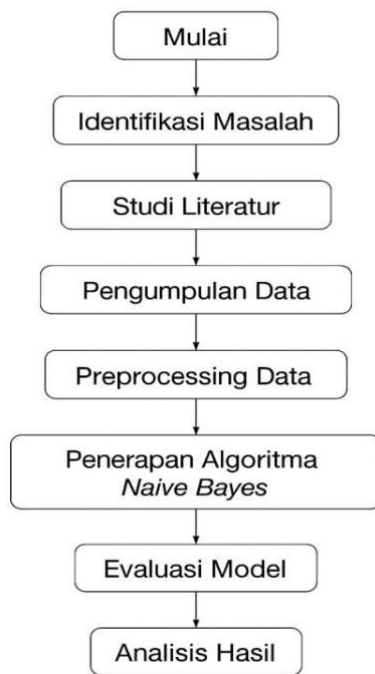
Metode

Tahapan penelitian disusun untuk menggambarkan proses analisis sentimen komentar pengguna terhadap jasa pengiriman e-commerce menggunakan metode Naïve Bayes secara sistematis dan terstruktur. Proses penelitian dilakukan melalui beberapa tahapan yang saling berkaitan, mulai dari identifikasi permasalahan, pengumpulan dan pengolahan data, penerapan metode klasifikasi, hingga analisis

dan interpretasi hasil. Penyusunan tahapan penelitian secara rinci bertujuan untuk memastikan bahwa proses penelitian dapat dipahami dengan jelas, bersifat objektif, serta dapat direplikasi pada penelitian sejenis.

Tahapan Penelitian

Tahapan Penelitian



Gambar 1. Flowchart Tahapan Penelitian

Tahapan penelitian merupakan rangkaian langkah sistematis untuk menganalisis sentimen komentar pengguna terhadap jasa pengiriman e-commerce. Penelitian diawali dengan identifikasi masalah dan studi literatur terkait analisis sentimen, *text mining*, dan algoritma Naïve Bayes. Selanjutnya dilakukan pengumpulan data komentar pengguna SPX Express, SiCepat, dan Ninja Xpress dari website Lacako.com menggunakan teknik *web scraping* melalui Web Scraper Extension pada Google Chrome. Data yang diperoleh kemudian dibersihkan dan dilakukan pelabelan sentimen secara manual ke dalam dua kategori, yaitu positif dan negatif.

Tahap berikutnya adalah preprocessing data yang meliputi *case folding*, *tokenizing*, *stopword removal*, dan *filter token by length* menggunakan RapidMiner. Data teks kemudian diubah menjadi representasi numerik menggunakan TF-IDF agar dapat diproses oleh algoritma klasifikasi. Selanjutnya dilakukan klasifikasi sentimen menggunakan algoritma Naïve Bayes. Evaluasi model dilakukan dengan 10-fold cross validation menggunakan metrik *accuracy*, *precision*, dan *recall*. Tahap akhir adalah analisis dan interpretasi hasil untuk mengetahui distribusi sentimen dan kinerja model.

Lokasi Penelitian

Penelitian dilakukan secara daring dengan memanfaatkan website Lacako.com sebagai sumber data ulasan pengguna jasa pengiriman. Penelitian difokuskan pada komentar pengguna terhadap SPX Express, SiCepat, dan Ninja Xpress yang memiliki

jumlah ulasan cukup banyak. Penelitian dilaksanakan pada periode Oktober 2025 hingga Januari 2026.

Subjek Penelitian

Subjek penelitian adalah komentar pengguna jasa pengiriman e-commerce yang tersedia pada Lacako.com. Populasi penelitian mencakup seluruh komentar terkait SPX Express, SiCepat, dan Ninja Xpress. Sampel dipilih menggunakan random sampling dengan kriteria komentar berbahasa Indonesia yang mengandung sentimen positif atau negatif. Komentar yang berupa spam, promosi, atau tidak relevan dikeluarkan dari dataset.

Definisi Operasional

Variabel dalam penelitian ini meliputi:

1. Sentimen Positif : komentar yang menunjukkan kepuasan terhadap layanan pengiriman.
2. Sentimen Negatif : komentar yang berisi keluhan atau pengalaman buruk pengguna.
3. Jenis Jasa Pengiriman : SPX Express, SiCepat, dan Ninja Xpress.
4. Aspek Layanan : kecepatan pengiriman, keamanan paket, ketepatan waktu, pelayanan pelanggan, dan biaya pengiriman.

Teknik Analisis Data

Analisis data dilakukan melalui tahapan pelabelan sentimen, preprocessing teks, pembobotan TF-IDF, dan klasifikasi menggunakan Naïve Bayes. Evaluasi model dilakukan dengan 10-fold cross validation menggunakan metrik *accuracy*, *precision*, dan *recall* untuk menilai kemampuan model dalam mengklasifikasikan sentimen serta membandingkan kecenderungan sentimen pada masing-masing jasa pengiriman.

Hasil Penelitian

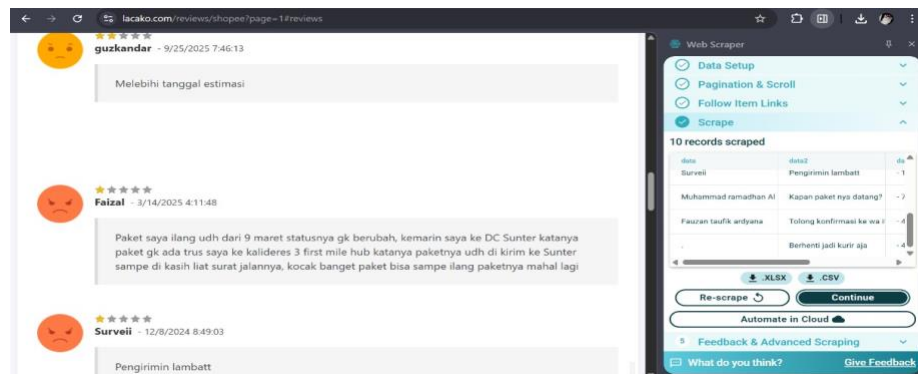
Hasil penelitian ini menyajikan temuan dari proses analisis sentimen komentar pengguna jasa pengiriman e-commerce menggunakan metode Naïve Bayes. Penyajian hasil dilakukan secara sistematis untuk menggambarkan kecenderungan sentimen masyarakat terhadap layanan SPX Express, Ninja Xpress, dan SiCepat berdasarkan komentar yang diperoleh dari website Lacako.com. Hasil penelitian meliputi deskripsi data, proses pengumpulan data, pelabelan sentimen, serta hasil klasifikasi dan evaluasi model. Seluruh proses analisis dilakukan melalui tahapan preprocessing, pembobotan TF-IDF, dan penerapan algoritma Naïve Bayes dengan pendekatan 10-fold cross validation. Hasil yang diperoleh digunakan untuk menjawab rumusan masalah penelitian, yaitu mengetahui kecenderungan pendapat masyarakat terhadap jasa pengiriman e-commerce serta menilai kinerja metode Naïve Bayes dalam mengklasifikasikan sentimen komentar pengguna.

Proses Pengambilan Data

Pengambilan data merupakan tahap awal yang menentukan kualitas penelitian. Data yang digunakan berupa komentar pengguna terhadap jasa pengiriman e-commerce yang diperoleh dari website Lacako.com karena menyediakan ulasan publik yang mencerminkan pengalaman nyata konsumen. Pengambilan data

dilakukan menggunakan teknik web scraping dengan bantuan Web Scraper Extension pada browser Google Chrome untuk mengekstraksi komentar secara otomatis dan efisien.

Proses pengambilan data dilakukan pada halaman ulasan jasa pengiriman dengan memanfaatkan fitur pagination agar seluruh komentar pada beberapa halaman dapat dikumpulkan. Data yang diambil dibatasi pada komentar yang berkaitan dengan pengalaman pengguna terhadap layanan pengiriman. Seluruh data yang diperoleh bersifat publik dan tidak memuat informasi pribadi pengguna, sehingga proses pengambilan data tetap memperhatikan aspek etika penelitian.



Gambar 2. Proses Pengambilan Data Ulasan Pengguna Shopee Menggunakan Web Scraper pada Website Lacako

Hasil dari proses web scraping ini masih berupa data mentah yang terdiri dari berbagai atribut teknis, seperti tautan halaman, informasi pagination, waktu unggahan, serta isi komentar pengguna. Oleh karena itu, data hasil web scraping selanjutnya disimpan dalam format Microsoft Excel untuk dilakukan proses seleksi dan pengolahan data pada tahap berikutnya, sebelum digunakan dalam analisis sentiment.

Gambar 2 menunjukkan proses pengambilan data komentar pengguna dari website Lacako.com menggunakan teknik web scraping dengan bantuan Web Scraper Extension. Pada gambar tersebut terlihat halaman ulasan pengguna yang memuat komentar terkait jasa pengiriman e-commerce, serta panel Web Scraper yang menampilkan tahapan proses scraping yang telah dijalankan.

Melalui tahapan Data Setup, Pagination & Scroll, dan Scrape, sistem berhasil mengekstraksi komentar pengguna secara otomatis dari beberapa halaman ulasan. Proses ini memungkinkan pengambilan data dilakukan secara berkelanjutan tanpa harus berpindah halaman secara manual. Hasil sementara dari proses web scraping ditampilkan dalam bentuk pratinjau data, yang menunjukkan bahwa komentar pengguna berhasil diambil dan tersusun dalam bentuk tabel.

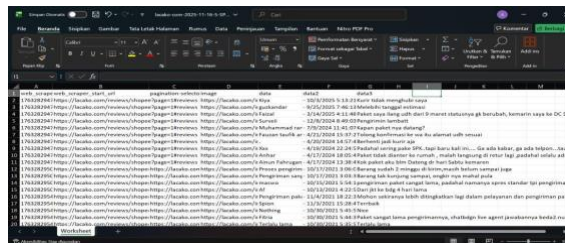
Data hasil web scraping yang ditampilkan pada pratinjau tersebut selanjutnya diekspor ke dalam format file Microsoft Excel. Data inilah yang kemudian digunakan sebagai dataset penelitian dan diproses lebih lanjut melalui tahapan seleksi data, pelabelan sentimen, pra-pemrosesan teks, pembobotan kata, hingga proses klasifikasi sentimen menggunakan algoritma Naïve Bayes. Dengan adanya bukti proses web scraping ini, dapat dipastikan bahwa data yang digunakan dalam penelitian diperoleh secara sistematis dan sesuai dengan metode yang telah dirancang.

a. Hasil Pengumpulan Data

Setelah proses pengambilan data menggunakan teknik web scraping, tahap selanjutnya adalah penyajian hasil pengumpulan data komentar pengguna jasa pengiriman e-commerce. Data yang diperoleh merupakan data mentah hasil scraping yang disimpan dalam format Microsoft Excel dan kemudian diseleksi agar sesuai dengan tujuan penelitian. Komentar yang dikumpulkan mencerminkan berbagai pengalaman pengguna terhadap layanan pengiriman, baik yang bersifat positif maupun negatif. Pada tahap ini, data belum melalui proses pelabelan sentimen maupun pengolahan teks. Pengumpulan data difokuskan pada tiga jasa pengiriman, yaitu SPX Express, SiCepat, dan Ninja Xpress, sehingga memudahkan proses analisis dan perbandingan pada tahap selanjutnya.

1) Hasil Pengumpulan Data SPX Express

Data komentar pengguna SPX Express diperoleh dari halaman ulasan pada website Lacako.com yang berisi pengalaman pengguna terkait layanan pengiriman, seperti kecepatan pengiriman, kondisi barang, dan pelayanan kurir. Berdasarkan proses pengumpulan data, diperoleh sebanyak 1000 komentar pengguna yang membahas berbagai pengalaman terhadap layanan SPX Express. Komentar negatif umumnya berkaitan dengan keterlambatan pengiriman, paket tidak sampai, serta kurangnya informasi dari pihak pengiriman, sedangkan komentar positif menunjukkan kepuasan terhadap layanan yang cepat dan sesuai harapan. Data yang diperoleh masih berupa data mentah hasil web scraping dan pada penelitian ini hanya atribut isi komentar yang digunakan untuk analisis sentimen, sementara atribut lain yang tidak relevan tidak digunakan. Data tersebut selanjutnya diproses pada tahap pelabelan sentimen dan pengolahan teks.



Gambar 3. Data Komentar Pengguna Jasa Pengiriman SPX Express

2) Hasil Pengumpulan Data SiCepat

Data komentar Data komentar pengguna SiCepat diperoleh dari website Lacako.com melalui proses *web scraping* pada halaman ulasan SiCepat. Dari proses tersebut diperoleh 1000 komentar pengguna yang mencerminkan berbagai pengalaman terhadap layanan pengiriman SiCepat. Komentar yang dikumpulkan menunjukkan sentimen yang beragam, baik kepuasan terhadap kecepatan dan kemudahan layanan maupun keluhan seperti keterlambatan pengiriman, kesalahan alamat, dan kesulitan menghubungi layanan pelanggan. Data yang diperoleh masih berupa data mentah dan belum melalui tahap pelabelan sentimen maupun *preprocessing*. Dalam penelitian ini, hanya atribut isi komentar yang digunakan untuk analisis sentimen, sedangkan atribut lain yang tidak relevan diabaikan. Data tersebut selanjutnya diproses pada tahap pelabelan sentimen dan pengolahan teks.

web_scraper_order	web_scraper_start_url	web_scraper_pagination_image	data1	data2
1763230093-1	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-5/4/2021Lambat sekali pengirimannya dan saat ini belum datang, tolong batas waktu sampai besok agar segera sampai.
1763230093-2	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-5/4/2021Paket Regular udh 5 hari nggak sampai dari tanggal 30 paket berangkat dari jakarta sampai sekarang ga datang
1763230093-3	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-5/2/2021Di persepot dong pengirimannya jangen di persepot
1763230093-4	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-4/26/2021Ci tidak kooperatif barang lambat diantar, pihak ekspedisi tidak prosaktif, minta alamat gerai tidak dikasih.
1763230093-5	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	4/28/2021Sangat lama, padahal pesanan sdh tiba
1763230093-6	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-4/28/2021Nawala lama banget
1763230093-7	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	4/27/2021Paket ku blom diterima sampai selingkuh padahal sampai tgl 14 skrg uda tgl 27
1763230093-8	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-4/27/2021Paket ku sampai selingkuh blom diantarakan padahal sampai tgl 14 skrg uda tgl 27
1763230093-9	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	4/24/2021Paket ku kurr telat dikasih paketku, udh lama udh antre, yg diantar ke rumah punya orang
1763230093-10	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-4/24/2021Tidak ada konfirmasi paket tiba2 gagap pengiriman
1763230093-11	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	9/19/2021Tanggapan saya gak di respon sama sekali dan paket sy sampai harti blom sampe di tangan saya. Pelayanan cust...
1763230093-12	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-9/12/2021Lama pul
1763230093-13	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-9/10/2021Puas
1763230093-14	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-9/10/2021Puan
1763230093-15	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-9/10/2021Buat kurr yg namanya fajar fauzi rehatin kalo nganter paket tuh ke rumah yg bersangkutan jangan kerumah org lain
1763230093-16	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-9/10/2021Paket saya ga ke ditetep :)
1763230093-17	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	9/10/2021Paket awan
1763230093-18	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-10/2/2021Udah 3 hari paket gaa sampai untuk pengiriman di lokasi kerja saya yg diantar rumah kurr tapi selulu jauh lebih cep...
1763230093-19	https://lacako.com/reviews/sicepat/https://lacA...		https://lacako.com/reviews/sicepat/https://lacA...	-10/3/2021Cepat sampe

Gambar 4. Data Komentar Pengguna Jasa Pengiriman SiCepat

3) Hasil Pengumpulan Data Ninja Xpress

Pengumpulan data komentar pengguna Ninja Xpress dilakukan pada halaman ulasan Ninja Xpress yang tersedia di website Lacako.com. Komentar-komentar yang diperoleh berisi pengalaman pengguna terhadap layanan pengiriman Ninja Xpress dalam mendukung transaksi e-commerce. Berdasarkan hasil pengumpulan data, komentar pengguna Ninja Xpress menunjukkan adanya opini positif dan negatif terhadap layanan pengiriman. Komentar positif umumnya berkaitan dengan ketepatan waktu pengiriman dan kondisi barang yang diterima dengan baik. Sementara itu, komentar negatif lebih banyak menyoroti permasalahan keterlambatan dan kurangnya kejelasan informasi pengiriman. Data komentar Ninja Xpress yang diperoleh pada tahap ini juga masih berupa data mentah hasil web scraping. Sama seperti pada dua jasa pengiriman sebelumnya, penelitian ini hanya menggunakan isi komentar pengguna sebagai data utama dalam analisis sentimen. Data Ninja Xpress selanjutnya diproses pada tahap pelabelan sentimen, preprocessing data, dan klasifikasi sentimen menggunakan metode Naïve Bayes.

web_scraper_order	web_scraper_start_url	web_scraper_pagination_image	data1	data2
1763230033-1	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-11/6/2021Tidak amanah!!!!!!
1763230033-2	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-11/6/2021Kurir nya tidak bertanggung jawab karena saya tidak puas dengan pelayanan ninja express
1763230033-3	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/29/2021Tt
1763230033-4	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/18/2021Kurir memanipulasi bukti pengiriman gagal, seolah olah wa padahal tidak ada menghubungi penerima
1763230033-5	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/12/2021Bisa kirim no kurir cabang tanente raja
1763230033-6	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/9/2021a
1763230033-7	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/28/2021Paket saya nengstuck, tidak sesuai estimasi
1763230033-8	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/20/2021Waktu kirimn jgn bg alasan kesalahan pd pengguna kalau ianya disebabkan kesalahan sendiri
1763230033-9	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/19/2021Pengiriman lambat...Sudah 22 hari pakainya tidak sampe2
1763230033-10	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/17/2021Kurir ngg ninja van rumpah paket saya stuck dibrosapa gudang skrg uda di bdg stuck di arcamank ganti aja namanya
1763230033-11	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/20/2021Bagus
1763230033-12	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/28/2021Packing yang sangat baik
1763230033-13	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-9/18/2021Tidak ditemukan
1763230041-14	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-11/9/2021Malah paket
1763230041-15	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-11/3/2021Saya ditipu atas pengiriman resi ini. Karena saya merasa tdk ada transaksi apa2...tp disuruh bayar COD dan yg term...
1763230041-16	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/28/2021Puas
1763230041-17	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/23/2021Good
1763230041-18	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/23/2021Paket saya jngan lama lama
1763230041-19	https://lacako.com/reviews/ninja/https://lacA...		https://lacako.com/reviews/ninja/https://lacA...	-10/21/2021Kurirnya cepet, tapi kurang memperhatikan kepedulian terhadap konsumen.

Gambar 5. Data Komentar Pengguna Jasa Pengiriman Ninja Xpress

Berdasarkan proses pengumpulan data yang telah dilakukan, diperoleh sejumlah komentar pengguna yang relevan dengan objek penelitian, yaitu jasa pengiriman SPX Express, Ninja Xpress, dan SiCepat. Data komentar yang dikumpulkan berasal dari website Lacako.com dan mencerminkan pengalaman nyata pengguna dalam menggunakan layanan pengiriman e-commerce, baik pengalaman yang bersifat positif maupun negatif. Seluruh data yang terkumpul telah melalui tahap seleksi dan pembersihan awal untuk memastikan bahwa hanya komentar yang relevan, tidak duplikat, dan memiliki makna sentimen yang jelas yang digunakan

dalam penelitian. Dengan demikian, data yang dihasilkan pada tahap pengumpulan data ini telah memenuhi kriteria sebagai dataset penelitian dan siap untuk diproses pada tahap pelabelan sentimen serta analisis lebih lanjut menggunakan metode klasifikasi yang telah ditentukan.

Pelabelan Data Sentimen

Tahap pelabelan data sentimen merupakan proses penting yang berfungsi sebagai dasar dalam pelatihan dan pengujian model klasifikasi. Pada penelitian ini, pelabelan dilakukan secara manual untuk menentukan kelas sentimen dari setiap komentar pengguna jasa pengiriman e-commerce. Sentimen dibagi menjadi dua kategori, yaitu positif dan negatif. Sentimen positif diberikan pada komentar yang menunjukkan kepuasan terhadap layanan, seperti penggunaan kata “cepat”, “puas”, atau “bagus”, sedangkan sentimen negatif diberikan pada komentar yang berisi keluhan atau pengalaman kurang baik, seperti keterlambatan pengiriman, paket tidak sampai, atau pelayanan yang kurang memuaskan. Proses pelabelan dilakukan dengan memperhatikan konteks kalimat secara keseluruhan untuk menghindari kesalahan interpretasi akibat ambiguitas bahasa. Pelabelan ini diterapkan pada komentar pengguna SPX Express, SiCepat, dan Ninja Xpress sehingga setiap komentar memiliki satu label sentimen. Hasil dari tahap ini berupa dataset komentar yang telah memiliki label sentimen (*ground truth*) yang selanjutnya digunakan pada tahap preprocessing, pembobotan TF-IDF, dan klasifikasi menggunakan algoritma Naïve Bayes.

Tabel 1. Contoh Pelabelan Data Sentimen Komentar Pengguna

Jasa Pengiriman	Komentar Pelanggan	Label Sentimen
SPX Xpress	“Bagus banget pengirimannya cepet dari peku sampai ke rengat cuman 1 hari udah nyampek lanjutkan ya bang kurir semangat dan sehat selalu.”	Positif
SPX Xpress	“Tidak biasanya pengiriman pakatnya lama sekali, biasanya paling lama cuman 5 hari. lah ini sudah lewat waktu estimasi lebih dari 3 hari... Mohon segera diperbaiki”	Negatif
SiCepat	“Kualitas layanan bagus, Juga responsif”	Positif
SiCepat	“Mohon dipercepat lagi ketika mengirim barang, di shope barang masih di Karawang, tapi pas liat di lacako udah di Subang tanggal 19, skrang tanggal 21 masih belum menerima paket”	Negatif
Ninja Xpress	“Saya sangat puas karena kurir ninja express sangat ramah dan baik pokok nya puas deh”	Positif

Ninja Xpress	“Paket saya belum sampai hingga saat ini,sudah melebihi batas akhir estimasi pengiriman, dan tidak ada kejelasan apapun dari ninja express .sangat sangat terlambat dan sangat mengecewakan”	Negatif
--------------	--	---------

Tahap pelabelan data sentimen menghasilkan dataset komentar yang telah diklasifikasikan secara manual ke dalam dua kategori, yaitu sentimen positif dan sentimen negatif. Pelabelan dilakukan dengan mempertimbangkan konteks keseluruhan isi komentar sehingga label yang diberikan mencerminkan persepsi pengguna terhadap kualitas layanan jasa pengiriman. Dataset yang telah diberi label ini menjadi dasar dalam proses pelatihan dan pengujian model klasifikasi Naïve Bayes, sehingga analisis sentimen dapat dilakukan secara terstruktur dan menghasilkan gambaran yang objektif mengenai kecenderungan sentimen pengguna.

Preprocessing Data

Preprocessing data dilakukan untuk membersihkan dan menormalkan data komentar sebelum dilakukan pembobotan fitur dan proses klasifikasi sentimen. Pada penelitian ini, preprocessing dilakukan menggunakan perangkat lunak RapidMiner melalui operator Process Documents from Data, sehingga seluruh tahapan pengolahan teks dapat dilakukan secara terintegrasi. Proses ini mengubah data teks mentah menjadi data yang lebih terstruktur dan siap dianalisis.

a. Transform Cases

Transform cases merupakan tahap awal dalam preprocessing yang bertujuan mengubah seluruh teks komentar menjadi huruf kecil (*lowercase*). Langkah ini dilakukan untuk menyeragamkan bentuk penulisan kata sehingga perbedaan huruf besar dan kecil tidak dianggap sebagai fitur yang berbeda dalam analisis. Sebagai contoh, kata “Paket”, “PAKET”, dan “paket” setelah melalui proses ini akan direpresentasikan dalam bentuk yang sama, yaitu “paket”. Dengan demikian, data teks menjadi lebih konsisten dan siap untuk diproses pada tahap selanjutnya.

Tabel 2. Hasil Proses Transform Cases pada Data Komentar Pengguna

Sebelum Transform Cases	Sesudah Transform Cases
Tidak biasanya pengiriman pakatnya lama sekali,biasanya paling lama cuma 5 hari.lah ini sudah lewat waktu estimasi lebih dari 3 hari. Mohon segera diperbaiki	tidak biasanya pengiriman pakatnya lama sekali,biasanya paling lama cuma 5 hari lah ini sudah lewat waktu estimasi lebih dari 3 hari mohon segera diperbaiki
Tolong tingkatkan kecepatan jasa kirimnya, padahal sudah Reguler tapi berasa Ekonomi, pakatnya 2 minggu baru di antar ke alamat tujuan, coba di cek kurir tiap-tiap kota, ada kurir yang curang pemalas mengantarkan paket.	tolong tingkatkan kecepatan jasa kirimnya padahal sudah reguler tapi berasa ekonomi pakatnya 2 minggu baru di antar ke alamat tujuan coba di cek kurir tiap tiap kota ada kurir yang curang pemalas mengantarkan paket.

Tabel 2. menunjukkan hasil data komentar pengguna setelah melalui tahap

transform cases. Pada tahap ini, seluruh teks komentar yang sebelumnya memiliki variasi penggunaan huruf besar dan huruf kecil telah diubah menjadi huruf kecil (lowercase) secara konsisten. Proses ini bertujuan untuk menyeragamkan bentuk kata sehingga perbedaan penulisan huruf tidak dianggap sebagai fitur yang berbeda dalam analisis sentimen. Setelah tahap transform cases, data teks menjadi lebih konsisten dan siap untuk diproses pada tahap tokenisasi.

b. Tokenize

Tahap tokenize merupakan lanjutan dari proses transform cases dalam preprocessing data teks. Pada tahap ini, data komentar pengguna yang telah diseragamkan dalam bentuk huruf kecil dipecah menjadi unit-unit kata tunggal yang disebut token. Proses tokenisasi bertujuan untuk memisahkan setiap kata dalam komentar agar dapat dianalisis secara individual sebagai fitur dalam proses analisis sentimen. Melalui proses tokenisasi, komentar yang semula berbentuk kalimat utuh diubah menjadi kumpulan kata-kata penyusun kalimat tersebut. Sebagai contoh, komentar “pengiriman sangat lambat” setelah melalui tahap tokenize akan dipecah menjadi token “pengiriman”, “sangat”, dan “lambat”. Token-token ini merepresentasikan kata-kata yang digunakan pengguna dalam menyampaikan opini terhadap layanan jasa pengiriman. Hasil dari tahap tokenize menunjukkan bahwa struktur data teks menjadi lebih terorganisasi, karena setiap token dapat diperlakukan sebagai elemen fitur yang berdiri sendiri. Dengan demikian, proses tokenisasi mempermudah tahapan selanjutnya, seperti penghapusan kata umum (stopword removal) dan penyaringan token, sehingga hanya kata-kata yang relevan yang dipertahankan dalam analisis sentimen.

Tabel 3. Hasil Tokenisasi Data Komentar Pengguna

Sebelum Tokenize	Sesudah Tokenize
tidak biasanya pengiriman pakatnya lama sekali, biasanya paling lama cuman 5 hari lah ini sudah lewat waktu estimasi lebih dari 3 hari mohon segera diperbaiki	“tidak”, “biasanya”, “pengiriman”, “pakatnya”, “lama”, “sekali”, “biasanya”, “paling”, “lama”, “cuman”, “5”, “hari”, “lah”, “ini”, “sudah”, “lewat”, “waktu”, “estimasi”, “lebih”, “dari”, “3”, “hari”, “mohon”, “segera”, “diperbaiki”
tolong tingkatkan kecepatan jasa kirimnya padahal sudah reguler tapi berasa ekonomi pakatnya 2 minggu baru di antar ke alamat tujuan coba di cek kurir tiap tiap kota ada kurir yang curang pemalas mengantarkan paket.	“tolong”, “tingkatkan”, “kecepatan”, “jasa”, “kirimnya”, “padahal”, “sudah”, “reguler”, “tapi”, “berasa”, “ekonomi”, “pakatnya”, “2”, “minggu”, “baru”, “di”, “antar”, “ke”, “alamat”, “tujuan”, “coba”, “di”, “cek”, “kurir”, “tiap”, “tiap”, “kota”, “ada”, “kurir”, “yang”, “curang”, “pemalas”, “mengantarkan”, “paket”

Tabel 3. menunjukkan hasil tokenisasi data komentar pengguna, di mana setiap komentar yang sebelumnya berupa kalimat utuh telah dipecah menjadi token- token

kata. Token-token tersebut menjadi dasar dalam proses pengolahan data teks selanjutnya sebelum dilakukan penghapusan stopwords dan pembobotan fitur. Stopword Removal

Tahap *stopword removal* merupakan proses lanjutan dalam preprocessing data teks yang bertujuan untuk menghapus kata-kata umum yang sering muncul dalam komentar pengguna tetapi tidak memiliki kontribusi yang signifikan terhadap makna sentimen. Kata-kata tersebut umumnya berfungsi sebagai penghubung dalam kalimat dan tidak mencerminkan opini pengguna terhadap layanan jasa pengiriman. Pada tahap ini, token-token hasil proses tokenisasi diseleksi dengan menghilangkan kata-kata umum seperti “tidak”, “nya”, “ini”, “dari”, dan kata sejenis lainnya. Penghapusan stopwords dilakukan untuk mengurangi jumlah fitur yang tidak relevan serta memfokuskan analisis pada kata-kata yang mengandung informasi sentimen, seperti kata yang menunjukkan kepuasan atau keluhan pengguna. Hasil dari proses stopwords removal menunjukkan bahwa token-token yang tersisa merupakan kata-kata yang lebih representatif terhadap isi komentar pengguna. Dengan berkurangnya kata-kata umum yang tidak bermakna sentimen, data teks menjadi lebih ringkas dan informatif, sehingga dapat meningkatkan efektivitas proses pembobotan fitur dan klasifikasi sentimen pada tahap selanjutnya.

Tabel 4. Hasil Stopword Removal pada Data Komentar Pengguna

Sebelum Stopword Removal	Sesudah Stopword Removal
“tidak”, “biasanya”, “pengiriman”, “paketnya”, “lama”, “sekali”, “biasanya”, “paling”, “lama”, “cuman”, “5”, “hari”, “lah”, “ini”, “sudah”, “lewat”, “waktu”, “estimasi”, “lebih”, “dari”, “3”, “hari”, “mohon”, “segera”, “diperbaiki”	pengiriman, paketnya, lama, 5, hari, lewat, waktu, estimasi, 3, hari, mohon, segera, diperbaiki
“tolong”, “tingkatkan”, “kecepatan”, “jasa”, “kirimnya”, “padahal”, “sudah”, “reguler”, “tapi”, “berasa”, “ekonomi”, “paketnya”, “2”, “minggu”, “baru”, “di”, “antar”, “ke”, “alamat”, “tujuan”, “coba”, “di”, “cek”, “kurir”, “tiap”, “tiap”, “kota”, “ada”, “kurir”, “yang”, “curang”, “pemalas”, “mengantarkan”, “paket”	tingkatkan, kecepatan, jasa, kirim, reguler, berasa, ekonomi, paketnya, 2, minggu, antar, alamat, tujuan, coba, cek, kurir, kurir, kota, curang, pemalas, mengantarkan, paket

Tabel 4. menunjukkan hasil penghapusan stopwords pada data komentar pengguna, di mana kata-kata umum yang tidak memiliki makna sentimen telah dihilangkan. Proses ini bertujuan untuk memfokuskan analisis pada kata-kata yang relevan dan berkontribusi terhadap penentuan sentimen komentar pengguna.

c. Filter Token by Length

Tahap filter token by length merupakan proses lanjutan dalam preprocessing data teks yang bertujuan untuk menyaring token berdasarkan panjang karakter tertentu. Proses ini dilakukan untuk menghapus token-token yang terlalu pendek dan tidak bersifat informatif, seperti token berupa angka atau singkatan pendek, yang

sering kali tidak memiliki kontribusi yang signifikan terhadap analisis sentimen. Pada penelitian ini, penerapan filter token by length dilakukan setelah proses tokenisasi dan penghapusan stopwords. Token-token yang memiliki panjang karakter di bawah batas minimum yang telah ditentukan tidak digunakan dalam proses analisis selanjutnya. Dengan demikian, hanya token-token yang memiliki panjang karakter memadai dan mengandung makna yang relevan terhadap opini pengguna yang dipertahankan. Hasil dari proses filter token by length menunjukkan bahwa data teks menjadi lebih ringkas dan terfokus, karena token-token yang tidak informatif telah dihilangkan. Token-token yang tersisa umumnya merupakan kata-kata kunci yang mencerminkan pengalaman dan penilaian pengguna terhadap layanan jasa pengiriman, sehingga dapat meningkatkan kualitas fitur teks yang digunakan pada tahap pembobotan dan klasifikasi sentimen.

Tabel 5. Hasil Filter Token by Length pada Data Komentar Pengguna

Sebelum Filter Token by Length	Sesudah Filter Token by Length
pengiriman, paketnya, lama, 5, hari, lewat, waktu, estimasi, 3, hari, mohon, segera, diperbaiki	pengiriman, paketnya, lama, hari, lewat, waktu, estimasi, hari, mohon, segera, diperbaiki
tingkatkan, kecepatan, jasa, kirim, reguler, berasa, ekonomi, paketnya, 2, minggu, antar, alamat, tujuan, coba, cek, kurir, kurir, kota, curang, pemalas, mengantarkan, paket	tingkatkan, kecepatan, jasa, kirim, reguler, berasa, ekonomi, paketnya, minggu, antar, alamat, tujuan, coba, cek, kurir, kurir, kota, curang, pemalas, mengantarkan, paket

Tabel 5. menunjukkan hasil penerapan filter token by length pada data komentar pengguna, di mana token-token dengan panjang karakter terlalu pendek telah dihapus. Proses ini bertujuan untuk mempertahankan token-token yang lebih informatif dan relevan sehingga data teks siap digunakan pada tahap pembobotan fitur dan klasifikasi sentimen.

d. Pembobotan TF - IDF

Setelah melalui tahap preprocessing, data komentar pengguna selanjutnya dilakukan pembobotan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Metode ini memberikan bobot lebih tinggi pada kata yang sering muncul dalam satu komentar namun jarang muncul pada keseluruhan dataset, sehingga dapat menunjukkan tingkat kepentingan kata terhadap suatu dokumen.

Pada penelitian ini, pembobotan TF-IDF dilakukan menggunakan operator *Process Documents from Data* pada RapidMiner setelah proses tokenisasi, stopwords removal, dan filter token by length. Hasilnya berupa word vector yang merepresentasikan setiap token sebagai atribut numerik. Kata-kata yang berkaitan dengan pengalaman pengguna terhadap layanan pengiriman cenderung memiliki bobot lebih tinggi karena lebih merepresentasikan sentimen komentar. Dataset numerik hasil pembobotan ini kemudian digunakan sebagai input pada tahap klasifikasi sentimen menggunakan algoritma Naïve Bayes.

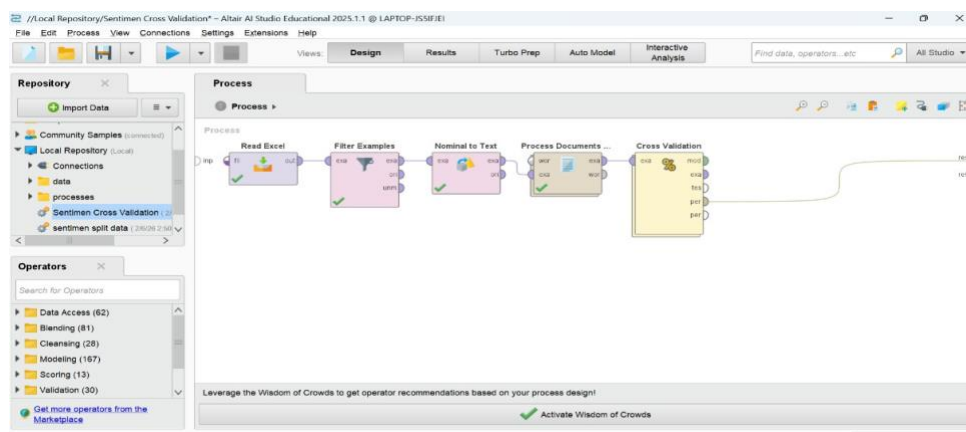
Row No.	sentimen	aaya	abang	abdulrasyid...	acara	adankomfr...	admin	adminnya	agen	agustus	ahk
187	positif	0	0	0	0	0	0	0	0	0	0
188	positif	0	0	0	0	0	0	0	0	0	0
189	negatif	0	0	0	0	0	0	0	0	0	0
190	positif	0	0	0	0	0	0	0	0	0	0
191	positif	0	0	0	0	0	0	0	0	0	0
192	positif	0	0	0	0	0	0	0	0	0	0
193	positif	0	0	0	0	0	0	0	0	0	0
194	positif	0	0	0	0	0	0	0	0	0	0
195	positif	0	0	0	0	0	0	0	0	0	0
196	positif	0	0	0	0	0	0	0	0	0	0
197	negatif	0	0	0	0	0	0	0	0	0	0
198	positif	0	0	0	0	0	0	0	0	0	0
199	negatif	0	0	0	0	0	0	0	0	0	0

Gambar 6. Hasil Pembobotan Kata Menggunakan Metode TF-IDF

Gambar 6 menunjukkan hasil pembobotan TF-IDF pada data komentar pengguna. Setiap baris merepresentasikan satu komentar, sedangkan setiap kolom merepresentasikan kata hasil preprocessing. Nilai numerik pada setiap sel menunjukkan bobot TF-IDF dari kata tersebut dalam suatu komentar. Nilai nol menunjukkan bahwa kata tersebut tidak muncul pada komentar yang bersangkutan, sedangkan nilai lebih besar dari nol menunjukkan tingkat kepentingan kata dalam komentar tersebut. Representasi numerik ini selanjutnya digunakan sebagai input pada proses klasifikasi sentimen menggunakan algoritma Naïve Bayes.

Klasifikasi Sentimen Menggunakan Naïve Bayes

Setelah data komentar pengguna melalui tahap preprocessing dan pembobotan fitur menggunakan metode TF-IDF, tahap selanjutnya dalam penelitian ini adalah melakukan klasifikasi sentimen menggunakan algoritma Naïve Bayes. Tahap klasifikasi ini bertujuan untuk mengelompokkan komentar pengguna jasa pengiriman e-commerce ke dalam dua kelas sentimen, yaitu sentimen positif dan sentimen negatif, berdasarkan representasi numerik dari data teks. Algoritma Naïve Bayes digunakan untuk memproses data hasil pembobotan TF-IDF dengan cara memberikan label sentimen pada setiap komentar pengguna. Melalui proses ini, setiap komentar dianalisis berdasarkan pola kemunculan kata-kata yang merepresentasikan sentimen positif maupun negatif, sehingga diperoleh hasil klasifikasi sentimen secara otomatis.

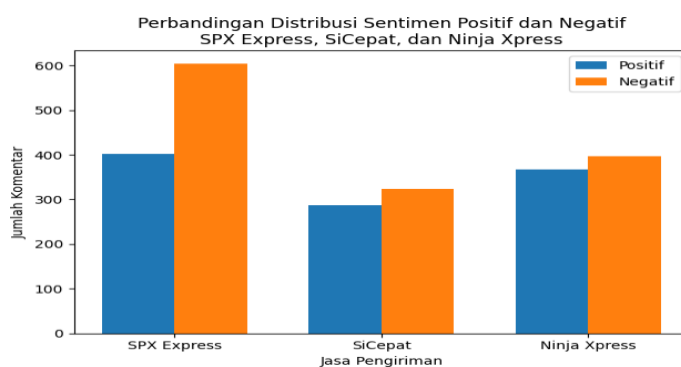


Gambar 7. Proses Klasifikasi Sentimen Menggunakan Algoritma Naïve Bayes

Gambar 7 menunjukkan alur proses klasifikasi sentimen menggunakan algoritma Naïve Bayes pada RapidMiner. Proses dimulai dari pembacaan data komentar pengguna yang telah melalui tahap pra-pemrosesan dan pembobotan fitur. Selanjutnya, data tersebut digunakan sebagai input dalam proses klasifikasi sentimen untuk menghasilkan label sentimen positif dan negatif pada setiap komentar pengguna. Hasil klasifikasi ini kemudian digunakan sebagai dasar untuk mengetahui kecenderungan sentimen pengguna terhadap jasa pengiriman e-commerce. Hasil klasifikasi sentimen ini selanjutnya digunakan sebagai dasar untuk melakukan evaluasi kinerja model klasifikasi.

Hasil Klasifikasi Sentimen

Hasil klasifikasi sentimen pada ketiga jasa pengiriman menunjukkan adanya perbedaan distribusi sentimen positif dan negatif yang mencerminkan persepsi pengguna terhadap kualitas layanan masing-masing jasa pengiriman.



Gambar 8. Distribusi Sentimen SPX Express, SiCepat, dan Ninja Xpress

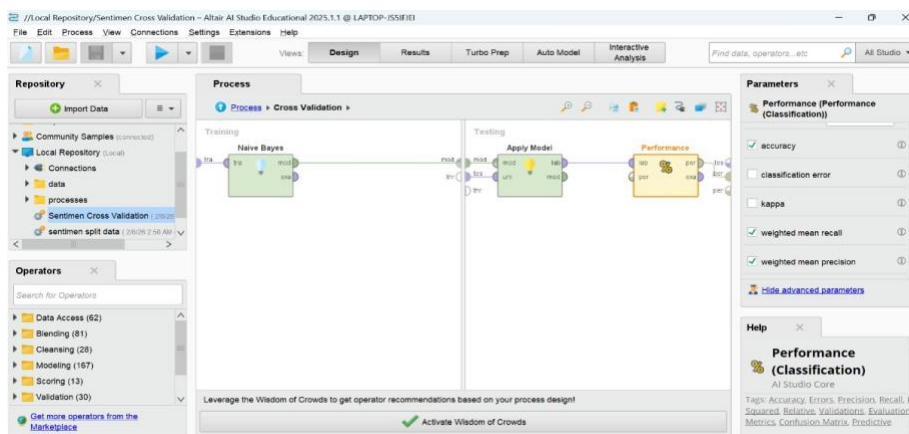
Berdasarkan hasil klasifikasi sentimen menggunakan metode Naïve Bayes terhadap masing-masing 1.000 komentar pengguna pada website Lacako.com, SPX Express menunjukkan dominasi sentimen negatif dengan 603 komentar negatif dan 402 komentar positif, yang mengindikasikan tingkat ketidakpuasan pengguna relatif lebih tinggi. SiCepat juga didominasi sentimen negatif, yaitu 324 komentar negatif dan 286 komentar positif, namun dengan selisih yang tidak terlalu besar. Sementara itu, Ninja Xpress memiliki distribusi sentimen yang lebih seimbang, dengan 396 komentar negatif dan 367 komentar positif, yang mencerminkan persepsi pengguna yang lebih moderat.

Hasil ini menunjukkan bahwa terdapat perbedaan persepsi masyarakat terhadap masing-masing jasa pengiriman dan sekaligus membuktikan bahwa metode Naïve Bayes mampu mengklasifikasikan sentimen komentar pengguna secara efektif, sehingga tujuan penelitian untuk mengetahui persepsi pengguna serta mengevaluasi kinerja metode klasifikasi sentimen dapat tercapai. Diketahui bahwa jumlah komentar positif dan negatif yang ditampilkan pada hasil klasifikasi sentimen tidak selalu sama dengan jumlah data awal yang dikumpulkan. Hal ini disebabkan oleh adanya beberapa komentar yang tereliminasi pada tahap preprocessing, seperti komentar yang menjadi kosong atau tidak memiliki token yang relevan setelah proses stopword removal dan filter token by length, sehingga komentar tersebut tidak dapat diproses lebih lanjut pada tahap klasifikasi sentimen.

Evaluasi Model Klasifikasi

Setelah proses klasifikasi sentimen dilakukan, tahap selanjutnya adalah melakukan evaluasi model klasifikasi untuk menilai kinerja algoritma Naïve Bayes

yang digunakan. Evaluasi kinerja model klasifikasi sentimen dilakukan menggunakan metode Naïve Bayes dengan cross validation pada dataset SPX Express, SiCepat, dan Ninja Xpress. Kinerja model diukur menggunakan metrik accuracy, precision, dan recall yang dihasilkan oleh operator Performance (Classification) pada RapidMiner, serta disajikan dalam bentuk PerformanceVector dan confusion matrix.



Gambar 9. Proses Evaluasi Model Klasifikasi Naïve Bayes Menggunakan 10-Fold Cross Validation

Gambar 9 menunjukkan proses pelatihan (training) dan pengujian (testing) model klasifikasi sentimen menggunakan algoritma Naïve Bayes dengan metode cross validation. Pada tahap training, model Naïve Bayes dibangun menggunakan data latih. Model yang terbentuk kemudian diterapkan pada data uji melalui operator Apply Model, selanjutnya kinerja model dievaluasi menggunakan operator Performance untuk mengukur kemampuan klasifikasi sentimen.



Gambar 10. Hasil Evaluasi Model Pada SPX Express

Gambar 10 menyajikan hasil pengujian model klasifikasi sentimen menggunakan algoritma Naïve Bayes yang dilakukan dengan bantuan RapidMiner. Berdasarkan hasil pengujian tersebut, diperoleh nilai akurasi sebesar 71,39% pada dataset komentar pengguna. Confusion matrix menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen, khususnya pada kelas sentimen negatif, yang ditandai dengan jumlah prediksi benar yang lebih dominan dibandingkan kesalahan klasifikasi. Selain itu, nilai precision sebesar 64,99% dan recall sebesar 75,02% mengindikasikan bahwa model mampu memberikan prediksi sentimen yang cukup tepat serta memiliki kemampuan yang baik dalam mengenali sentimen positif dan negatif pada komentar pengguna.

Tabel 6. Hasil Evaluasi Model Naïve Bayes pada SPX Express

Metrik	Nilai (%)
Accuracy	56,76
Precision	59,54
Recall	59,54

Tabel 6 menunjukkan hasil evaluasi kinerja model klasifikasi Naïve Bayes berdasarkan nilai precision dan recall yang diperoleh dari confusion matrix. Nilai precision sebesar 64,99% menunjukkan bahwa sebagian besar komentar yang diprediksi oleh model sebagai sentimen tertentu telah sesuai dengan label sentimen yang sebenarnya. Sementara itu, nilai recall sebesar 75,02% mengindikasikan bahwa model mampu mengenali sebagian besar komentar yang benar-benar termasuk ke dalam kelas sentimen yang bersangkutan. Hasil evaluasi ini menunjukkan bahwa model Naïve Bayes memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar pengguna, sehingga dapat digunakan untuk merepresentasikan kecenderungan sentimen positif dan negatif pada dataset yang dianalisis.



Gambar 11. Hasil Evaluasi Model Naïve Bayes pada SiCepat

Gambar 11 menyajikan hasil evaluasi kinerja model klasifikasi sentimen menggunakan algoritma Naïve Bayes pada dataset komentar pengguna. Nilai akurasi yang diperoleh sebesar 56,76% menunjukkan bahwa kemampuan model dalam mengklasifikasikan sentimen masih berada pada tingkat sedang. Confusion matrix memperlihatkan bahwa distribusi prediksi sentimen positif dan negatif cenderung seimbang, namun jumlah kesalahan klasifikasi yang dihasilkan masih relatif tinggi. Selain itu, nilai precision dan recall yang masing-masing sebesar 59,54% mengindikasikan bahwa model memiliki tingkat ketepatan dan kemampuan pengenalan sentimen yang cukup, tetapi belum optimal dalam membedakan sentimen positif dan negatif pada komentar pengguna.

Tabel 7. Hasil Evaluasi Model Naïve Bayes pada SiCepat

Metrik	Nilai (%)
Accuracy	56,76
Precision	59,54
Recall	59,54

Tabel 7 menampilkan hasil evaluasi kinerja model klasifikasi Naïve Bayes berdasarkan nilai precision dan recall yang dihasilkan dari confusion matrix. Nilai precision dan recall yang diperoleh masing-masing sebesar 59,54% menunjukkan bahwa model memiliki tingkat ketepatan dan kemampuan pengenalan sentimen yang relatif seimbang. Hasil ini mengindikasikan bahwa model Naïve Bayes mampu mengklasifikasikan sentimen positif dan negatif pada komentar pengguna secara cukup konsisten, meskipun tingkat kinerjanya masih berada pada kategori sedang dibandingkan dengan dataset lainnya.



Gambar 12. Hasil Evaluasi Model Naïve Bayes pada Ninja Xpress

Gambar 12 menyajikan hasil evaluasi kinerja model klasifikasi sentimen menggunakan algoritma Naïve Bayes pada dataset Ninja Xpress. Nilai akurasi yang diperoleh sebesar 54,45% menunjukkan bahwa kemampuan model dalam mengklasifikasikan sentimen masih berada pada tingkat yang relatif rendah dibandingkan dengan dataset lainnya. Confusion matrix memperlihatkan bahwa model belum mampu membedakan sentimen positif dan negatif secara konsisten, yang ditandai dengan masih tingginya jumlah kesalahan klasifikasi. Selain itu, nilai precision sebesar 58,57% dan recall sebesar 61,57% mengindikasikan bahwa meskipun model dapat mengenali sebagian besar sentimen komentar, tingkat ketepatan prediksi yang dihasilkan masih perlu ditingkatkan.

Tabel 8. Hasil Evaluasi Model Naïve Bayes pada Ninja Xpress

Metrik	Nilai (%)
Accuracy	54,45
Precision	58,57
Recall	61,57

Tabel 8 menyajikan hasil evaluasi kinerja model klasifikasi sentimen menggunakan algoritma Naïve Bayes berdasarkan nilai precision dan recall yang diperoleh dari confusion matrix. Nilai precision sebesar 58,57% menunjukkan bahwa sebagian besar prediksi sentimen yang dihasilkan oleh model telah sesuai dengan label sentimen aktual. Sementara itu, nilai recall sebesar 61,57% mengindikasikan bahwa model mampu mengenali sebagian besar komentar yang termasuk ke dalam kelas sentimen yang bersangkutan. Meskipun demikian, hasil ini menunjukkan bahwa tingkat kesalahan klasifikasi yang dihasilkan masih relatif lebih tinggi.

Pembahasan

Pembahasan penelitian ini bertujuan untuk menafsirkan temuan dari hasil analisis sentimen komentar pengguna terhadap jasa pengiriman e-commerce SPX Express, SiCepat, dan Ninja Xpress menggunakan metode Naïve Bayes. Berdasarkan hasil klasifikasi sentimen, ketiga jasa pengiriman menunjukkan dominasi sentimen negatif dengan tingkat yang berbeda pada masing-masing layanan. SPX Express memiliki jumlah sentimen negatif tertinggi dengan 603 komentar negatif dan 402 komentar positif, diikuti oleh SiCepat dengan 324 komentar negatif dan 286 komentar positif, serta Ninja Xpress dengan 396 komentar negatif dan 367 komentar positif. Dominasi sentimen negatif ini menunjukkan bahwa secara umum persepsi masyarakat terhadap layanan jasa pengiriman e-commerce masih didominasi oleh ketidakpuasan pengguna.

Tingginya jumlah sentimen negatif pada SPX Express menunjukkan bahwa pengguna lebih banyak menyampaikan keluhan dibandingkan apresiasi terhadap layanan yang diterima, terutama terkait keterlambatan pengiriman, ketidaksesuaian estimasi waktu, dan kurangnya kejelasan informasi pelacakan paket. Pada SiCepat, perbedaan antara sentimen positif dan negatif relatif lebih kecil, sehingga menunjukkan persepsi pengguna yang lebih seimbang. Sementara itu, pada Ninja Xpress jumlah sentimen positif dan negatif hampir seimbang, yang menunjukkan bahwa pengalaman pengguna terhadap layanan ini cukup beragam, meskipun masih terdapat keluhan terkait keterlambatan pengiriman dan informasi status paket.

Kemampuan metode Naïve Bayes dalam mengklasifikasikan sentimen diuji menggunakan pendekatan 10-fold cross validation. Hasil evaluasi menunjukkan bahwa dataset SPX Express memiliki performa terbaik dengan nilai accuracy 71,39%, precision 64,99%, dan recall 75,02%. Pada dataset SiCepat diperoleh accuracy 56,76% serta precision dan recall sebesar 59,54%, sedangkan pada Ninja Xpress diperoleh accuracy 54,45%, precision 58,57%, dan recall 61,57%. Hasil ini menunjukkan bahwa model cukup mampu mengenali sebagian besar sentimen aktual, meskipun masih terdapat kesalahan klasifikasi terutama pada komentar yang mengandung opini campuran atau menggunakan bahasa informal.

Perbedaan performa klasifikasi pada masing-masing dataset menunjukkan bahwa kinerja model tidak hanya dipengaruhi oleh algoritma, tetapi juga oleh karakteristik data teks yang digunakan. Dataset dengan pola bahasa yang lebih konsisten cenderung menghasilkan performa klasifikasi yang lebih baik. Secara keseluruhan, penelitian ini menunjukkan bahwa analisis sentimen berbasis teks dapat digunakan sebagai pendekatan yang efektif untuk mengetahui persepsi masyarakat terhadap jasa pengiriman e-commerce. Selain itu, metode Naïve Bayes terbukti memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar pengguna, meskipun masih memiliki keterbatasan dalam memahami konteks bahasa yang kompleks.

Kesimpulan

Berdasarkan hasil penelitian mengenai analisis sentimen komentar pengguna terhadap jasa pengiriman e-commerce SPX Express, SiCepat, dan Ninja Xpress menggunakan metode Naïve Bayes, dapat disimpulkan bahwa sentimen masyarakat terhadap ketiga layanan tersebut cenderung didominasi oleh sentimen negatif, meskipun dengan tingkat yang berbeda pada masing-masing jasa pengiriman, di mana SPX Express menunjukkan dominasi sentimen negatif paling tinggi, diikuti

oleh Ninja Xpress dan SiCepat. Hal ini menunjukkan bahwa masih terdapat ketidakpuasan pengguna terhadap layanan pengiriman, terutama terkait keterlambatan pengiriman, ketidaksesuaian estimasi waktu, serta kurangnya kejelasan informasi pelacakan paket, meskipun sentimen positif juga tetap ditemukan yang mencerminkan kepuasan pengguna ketika layanan berjalan dengan baik, seperti pengiriman tepat waktu dan kondisi paket yang diterima dengan baik. Selain itu, hasil evaluasi model menggunakan 10-fold cross validation menunjukkan bahwa metode Naïve Bayes memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen komentar pengguna dengan nilai accuracy, precision, dan recall yang bervariasi pada setiap dataset, serta performa terbaik pada dataset SPX Express, sehingga dapat disimpulkan bahwa metode ini efektif digunakan untuk menggambarkan persepsi masyarakat terhadap layanan jasa pengiriman e-commerce melalui tahapan preprocessing, pembobotan TF-IDF, dan proses klasifikasi sentimen.

Saran

Berdasarkan hasil penelitian yang telah diperoleh, disarankan agar perusahaan jasa pengiriman dan platform e-commerce menjadikan temuan ini sebagai bahan evaluasi untuk meningkatkan kualitas layanan, terutama pada aspek yang sering dikeluhkan pengguna seperti ketepatan waktu pengiriman, kejelasan informasi pelacakan, dan kualitas pelayanan pelanggan guna meningkatkan kepuasan serta kepercayaan konsumen. Bagi peneliti selanjutnya, penelitian ini dapat dikembangkan dengan menambahkan kategori sentimen netral, membandingkan metode klasifikasi lain seperti Support Vector Machine (SVM), K-Nearest Neighbor (KNN), atau metode berbasis deep learning, serta memperluas sumber dan periode pengambilan data dari berbagai platform agar analisis sentimen menjadi lebih komprehensif, representatif, dan memiliki tingkat akurasi yang lebih baik melalui pengembangan teknik preprocessing serta penanganan komentar dengan opini campuran.

Daftar Pustaka

- Ahmad, R., & Sari, N. (2022). Efektivitas fitur penataan dokumen pada Microsoft Word untuk proses revisi akademik. *Jurnal Manajemen Pendidikan dan Teknologi*, 15(3), 200–209.
- Amalia, L. (2020). Pengaruh ulasan konsumen terhadap reputasi jasa pengiriman di e-commerce. *Jurnal Manajemen dan Bisnis*.
- Ardiansyah, R., & Lestari, S. (2020). Pengaruh ulasan online terhadap keputusan pembelian. *Jurnal Sistem Informasi*, 12(2), 145–156.
- Fauzan, M. (2023). Persaingan jasa pengiriman di era e-commerce. *Jurnal Transportasi dan Logistik*, 6(1), 40–50.
- Fauzi, A. (2021). Analisis sentimen ulasan pengguna marketplace menggunakan Naive Bayes. *Jurnal Informatika*.
- Fauzi, A., & Hidayat, F. (2021). Klasifikasi sentimen teks menggunakan algoritma Naive Bayes. *Jurnal Informatika*, 8(2), 101–109.
- Fauzi, A., & Pratama, Y. (2023). Klasifikasi sentimen menggunakan algoritma Naive Bayes berbasis RapidMiner. *Jurnal Data Mining dan Sistem Informasi*, 4(1), 45–54.
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82–89.
- Fitri, N. (2020). Implementasi analisis sentimen pada ulasan jasa pengiriman di sosial media. *Jurnal Informatika dan Multimedia*.
- Hanif, M., et al. (2022). Pengolahan data teks dalam skala besar. *Jurnal Data Science Indonesia*, 4(1), 25–34.

- Hidayat, F. (2021). Analisis sentimen pelanggan e-commerce menggunakan Naive Bayes dan K-Nearest Neighbor. *Jurnal Ilmu Komputer*.
- Hidayat, F., Rahman, A., & Putri, N. (2022). Tahapan pra-pemrosesan teks pada analisis sentimen menggunakan metode Naive Bayes. *Jurnal Informatika*, 9(2), 123–132.
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59–68.
- Kurniawan, A. (2020). Perkembangan teknologi informasi dan komunikasi di era digital. *Jurnal Teknologi Informasi*, 8(1), 1–10.
- Laksono, A. (2023). Analisis opini konsumen jasa pengiriman berbasis web review. *Jurnal Informatika*, 10(2), 120–130.
- Lestari, D. (2022). Ekonomi digital dan pertumbuhan e-commerce. *Jurnal Ekonomi Digital*, 4(1), 21–30.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Nugroho, D., & Lestari, M. (2021). Pemanfaatan Microsoft Word dalam penyusunan laporan penelitian: Pengelolaan tabel, gambar, dan lampiran. *Jurnal Sistem Informasi dan Aplikasi Pendidikan*, 14(1), 55–64.
- Nugroho, Y., & Syarief, S. S. (2020). Transformasi digital dan dampaknya terhadap masyarakat Indonesia. *Jurnal Ilmu Komunikasi*, 18(2), 101–115.
- Prasetyo, B. (2021). Ekosistem digital dan perkembangan sistem informasi. *Jurnal Sistem Informasi*, 17(2), 85–95.
- Pratama, A. (2021). Perilaku konsumen dalam belanja online. *Jurnal Perilaku Konsumen*, 3(2), 45–55.
- Putra, A. R., & Santoso, P. (2021). Evaluasi kinerja algoritma klasifikasi pada analisis sentimen teks. *Jurnal Teknologi Informasi dan Komputer*, 9(1), 55–64.
- Putri, A. R. (2022). Peran platform digital dalam bisnis e-commerce. *Jurnal Bisnis Digital*, 4(1), 30–40.
- Putri, A. R., & Hidayat, R. (2022). Analisis kepuasan konsumen terhadap kualitas layanan jasa pengiriman. *Jurnal Manajemen dan Bisnis*, 9(1), 33–44.
- Rahman, D. (2020). Perkembangan e-commerce di Indonesia. *Jurnal E-Commerce Indonesia*, 1(1), 1–9.
- Rahmawati, D. (2021). Analisis sentimen berbasis natural language processing. *Jurnal Teknologi Informasi*, 14(3), 210–220.
- Ratnasari, C. I. (2023). Perbandingan metode web scraping dalam pengambilan data ulasan daring. *Jurnal AUTOMATA*.
- Sari, M. (2022). Analisis sentimen pada layanan jasa pengiriman. *Jurnal Analisis Sentimen*, 2(1), 15–25.
- Santoso, P. (2022). Implementasi klasifikasi sentimen review pengiriman pada e-commerce menggunakan Naive Bayes. *Jurnal Ilmiah Informatika*.
- Saputra, R., & Kurniawan, D. (2021). Analisis sentimen komentar pengguna media sosial menggunakan RapidMiner. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(3), 489–498.
- Setiawan, R. (2021). Transformasi digital dalam kehidupan masyarakat. *Jurnal Transformasi Digital*, 2(1), 11–20.
- Tjiptono, F. (2020). Manajemen kualitas layanan dan kepuasan pelanggan. *Jurnal Manajemen*, 11(1), 15–25.
- Tjiptono, F., & Diana, A. (2020). Kepuasan pelanggan dalam perspektif manajemen jasa. *Jurnal Manajemen dan Kewirausahaan*, 22(1), 1–10.
- Tsytsarau, M., & Palpanas, T. (2012). Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 24(3), 478–514.
- Utami, E. (2021). Penerapan metode Naive Bayes dalam klasifikasi teks. *Jurnal Informatika Indonesia*, 7(2), 95–104.
- Wibowo, T. (2021). Manfaat e-commerce dalam aktivitas jual beli. *Jurnal Manajemen Informatika*, 9(2), 50–59.

- Widya, R. (2022). Efektivitas Naive Bayes dalam analisis sentimen review pelanggan marketplace Shopee. *Jurnal Informatika dan Data Mining*.
- Widya, R., & Pratama, Y. (2023). Analisis ulasan pelanggan terhadap kinerja jasa pengiriman e-commerce menggunakan pendekatan sentimen. *Jurnal Informatika dan Data Science*, 4(1), 55–66.
- Wulandari, R., & Sari, M. (2020). Penerapan text preprocessing pada analisis sentimen ulasan pelanggan e-commerce. *Jurnal Sistem Informasi*, 16(1), 45–55.