



**ANALISIS SENTIMEN PADA ULASAN PRODUK
DI *E-COMMERCE* DENGAN METODE *NAIVE BAYES***

Tugas Akhir
diajukan untuk melengkapi
persyaratan mencapai
gelar sarjana

NAMA : ASSAMI MUZAKI
NPM : 201943502262

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS INDRAPRASTA PGRI
2023**

LEMBAR PERSETUJUAN UJIAN TUGAS AKHIR

Nama : Assami Muzaki
NPM : 201943502262
Program Studi : Teknik Informatika
Fakultas : Teknik dan Ilmu Komputer
Judul : Analisis Sentimen pada Ulasan Produk di *E-Commerce*
dengan Metode *Naïve Bayes*

Telah diperiksa dan disetujui untuk diujikan

Pembimbing Materi

Pembimbing Teknik

Vici Febriana, M.Kom.

Wahyu Nur Cholifah, M.Kom.

LEMBAR PENGESAHAN

Nama : Assami Muzaki
NPM : 201943502262
Program Studi : Teknik Informatika
Fakultas : Teknik dan Ilmu Komputer
Judul : Analisis Sentimen pada Ulasan Produk di E-Commerce
dengan Metode Naïve Bayes

Panitia Ujian

Ketua : Prof. Dr. H. Sumaryoto
Sekretaris : Ir. H. Soepardi Harris, M.T.
Anggota :

No	Nama Penguji	Tanda Tangan

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini

Nama : Assami Muzaki

NPM : 201943502262

Program Studi : Teknik Informatika

Dengan ini menyatakan bahwa skripsi/tugas akhir dengan judul “Analisis Sentimen pada Ulasan Produk di *E-Commerce* dengan Metode *Naïve Bayes*” beserta seluruh isinya adalah karya saya sendiri. Atas pernyataan ini, saya siap menanggung risiko/sanksi apabila di kemudian hari ditemukan seluruh atau sebagian isi skripsi/tugas akhir ini bukan hasil tulisan saya sendiri, saya bersedia menerima sanksi sesuai dengan Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional Bab VI Pasal 25 ayat 2 dan Bab XX Pasal 70.

Demikian Pernyataan ini saya buat untuk dimanfaatkan sesuai dengan keperluan.

Jakarta, Agustus 2023

Yang Menyatakan

Assami Muzaki

ABSTRAK

- A. Assami Muzkai, 201943502262:
- B. Analisis Sentimen pada Ulasan Produk di E-Commerce dengan Metode Naive Bayes. Skripsi/Tugas Akhir : Jakarta : Fakultas : Teknik dan Ilmu Komputer. Program Studi Teknik Informatika, Universitas Indraprasta PGRI, Agustus, 2023
- C. xi, 5 BAB, 100 halaman
- D. Kata Kunci: Analisis Sentimen, Naïve Bayes, Ulasan Produk, E-commerce
- E. Penelitian ini bertujuan untuk mengembangkan sistem analisis sentimen pada ulasan produk di platform e-commerce menggunakan metode Naive bayes. Sistem ini dirancang untuk mengklasifikasikan sentimen ulasan menjadi positif atau negatif. Tujuan utama penelitian ini adalah menciptakan sebuah website interaktif yang mampu melakukan analisis sentimen secara real-time dan juga dapat menerima file berbentuk CSV untuk melakukan analisis sentimen sekaligus pada banyak ulasan. Adapun metode penelitian yang digunakan adalah penelitian eksploratif dengan pendekatan pengembangan sistem. Data ulasan produk diperoleh dari platform e-commerce dan diproses melalui tahap preprocessing untuk mempersiapkan data teks. Hasil pembahasan meliputi pengembangan sistem analisis sentimen berbasis metode Naïve Bayes serta implementasinya dalam bentuk website interaktif. Berdasarkan hasil penelitian, disimpulkan bahwa pengembangan sistem analisis sentimen menggunakan metode Naïve Bayes berhasil menciptakan sebuah website interaktif yang efektif dalam melakukan analisis sentimen ulasan produk di platform e-commerce.
- F. Daftar Pustaka:
 - 1. Buku 1 Buah
 - 2. Jurnal 1 Buah
 - 3. Skripsi 1 Buah
- G. Nama Pembimbing: Vici Febriana, M.Kom.
Wahyu Nur Cholifah, M.Kom

LEMBAR PERSEMBAHAN

“may god shower us with blessings today and always”

KATA PENGANTAR

Puji syukur penulis ucapkan atas ke hadirat Allah SWT yang telah melimpahkan rahmat taufik serta hidayah-nya. Shalawat serta salam penulis sanjungkan kepada junjungan kita, Nabi yang pemberi rahmat dan Nabi penyempurna segala kema'rifan Nabi Muhammad SAW, sehingga akhirnya dapat menyelesaikan skripsi ini tepat pada waktunya. Skripsi yang berjudul “Analisis Sentimen pada Ulasan Produk di E-Commerce dengan Metode Naïve Bayes” ini ditulis untuk memenuhi salah satu syarat guna memperoleh gelar sarjana pada Universitas Indraprasta PGRI. Pada kesempatan yang baik ini, izinkanlah penulis menyampaikan rasa hormat dan terima kasih kepada semua pihak yang dengan tulus dan ikhlas memberikan bantuan dan dorongan kepada penulis dalam menyelesaikan skripsi ini.

1. Ibu Vici Febriana, M.Kom. selaku Dosen Pembimbing Materi Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Indraprasta
2. Ibu Wahyu Nur Cholifah, M.Kom. selaku Dosen Pembimbing Teknik Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Indraprasta
3. Bapak Prof. Dr. H. Sumaryoto, selaku Rektor Universitas Indraprasta PGRI
4. Bapak Ir. H. Soepardi Harris, M.T. selaku Dekan Fakultas Teknik dan Ilmu Komputer, Universitas Indraprasta PGRI
5. Ibu Mei Lestari, M.Kom. selaku Ketua Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Indraprasta PGRI








6. Kedua orang tua penulis yang sudah mendidik dan menyayangi penulis dari lahir hingga sekarang.
7. Teman-temanku seperjuangan Angkatan 2019, khususnya mahasiswa Teknik Informatika kelas R3 yang tidak dapat disebutkan satu per satu








Penulis menyadari bahwa masih terdapat kekurangan dalam penulisan dan penyusunan skripsi ini. Namun, hal tersebut telah diusahakan semaksimal mungkin kesempurnaanya sesuai dengan batas kemampuan yang ada. Oleh karena itu, dengan kerendahan hati penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa yang akan datang.

Jakarta, Agustus 2023

Penulis

DAFTAR SIMBOL

SIMBOL	NAMA	KETERANGAN
	<i>Actor</i>	Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i> .
	<i>Initial State</i>	Menyatakan awal dimulainya suatu aktivitas.
	<i>Final State</i>	Menyatakan berakhirnya suatu aktivitas.
	<i>Use Case</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor.
	<i>Activity</i>	Memperlihatkan bagaimana masing-masing kelas antar muka saling berinteraksi satu sama lain.
	<i>Decision</i>	Simbol yang digunakan untuk memilih proses atau keputusan berdasarkan kondisi yang ada.
	<i>Fork/Join</i>	Suatu aliran yang pada tahap tertentu berubah menjadi beberapa aliran.

	<i>Association</i>	Simbol yang menghubungkan antara objek satu dengan objek lainnya.
	<i>Extend</i>	Menspesifikasikan bahwa <i>use case</i> target memperluas perilaku dari <i>use case</i> sumber pada suatu titik yang diberikan.
	<i>A message</i>	Menggambarkan pengiriman pesan.
	<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung pada elemen yang tidak mandiri.
	<i>Message</i>	Spesifikasi dari komunikasi antar objek yang memuat informasi-informasi tentang aktifitas yang terjadi.
	<i>A Focus of Control & A Life Line</i>	Menggambarkan tempat mulai dan berakhirnya message.
	<i>System</i>	Menspesifikasikan paket yang menampilkan sistem secara terbatas.

DAFTAR ISI

LEMBAR PERSETUJUAN UJIAN TUGAS AKHIR	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN	iii
ABSTRAK	iv
LEMBAR PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR SIMBOL	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah.....	1
B. Identifikasi Masalah	2
C. Batasan Masalah.....	3
D. Rumusan Masalah	4
E. Tujuan Penelitian	5
F. Manfaat Penelitian	5
G. Sistematika Penulisan.....	7
BAB II LANDASAN TEORI DAN PENELITIAN YANG RELEVAN	9
A. Landasan Teori.....	9
B. Penelitian Yang Relevan	23
BAB III METODOLOGI PENELITIAN	27
A. Waktu dan Tempat Penelitian	27
B. Tahapan Penelitian	28
C. Algoritma.....	31
BAB IV HASIL DAN PEMBAHASAN	34
A. Definisi Masalah	34
B. Pembahasan Algoritma.....	35
C. Pemodelan Perangkat Lunak	39
1. <i>Unified Modeling Language</i>	39
2. Rancangan Layar	45
3. Tampilan Layar.....	48
D. Kelebihan dan Kelemahan Penelitian	53

BAB V SIMPULAN DAN SARAN.....	56
DAFTAR PUSTAKA	58
RIWAYAT PENULIS.....	61
LAMPIRAN.....	62

DAFTAR TABEL

Table 3.1 Jadwal Kegiatan Penelitian	11
--	----

DAFTAR GAMBAR

Gambar 2. 1 Confussion Matrix.....	17
Gambar 3. 1 Tahapan Penelitian	29
Gambar 3. 2 Diagram Kerja Algoritma	32
Gambar 4. 1 Diagram Use Case.....	39
Gambar 4. 2 Diagram Activity Melakukan Analisis Sentimen	40
Gambar 4. 3 Diagram Activity Melakukan Analisis Sentimen Melalui File CSV ...	41
Gambar 4. 4 Diagram Activity Menggunakan API Analisis Sentimen	42
Gambar 4. 5 Sequence Diagram Index	43
Gambar 4. 6 Sequence Diagram Analisis Sentimen	43
Gambar 4. 7 Sequence Diagram Analisis Sentimen dengan File CSV	44
Gambar 4. 8 Class Diagram	44
Gambar 4. 9 Rancangan Layar Halaman Utama.....	45
Gambar 4. 10 Rancangan Layar Cara Pakai API dan Akurasi Model	45
Gambar 4. 11 Rancangan Layar Upload File CSV Multi Ulasan	46
Gambar 4. 12 Rancangan Layar History Analisis.....	46
Gambar 4. 13 Rancangan Layar Keluaran Analisis Sentimen	47
Gambar 4. 14 Rancangan Layar Keluaran Analisis File CSV Multi Ulasan	47
Gambar 4. 15 Rancangan Layar Cetak Tabel Laporan	48
Gambar 4. 16 Tampilan Layar Halaman Utama	48
Gambar 4. 17 Tampilan Layar Cara Pakai API dan Akurasi Model	49
Gambar 4. 18 Tampilan Layar Upload File CSV Multi Ulasan	50
Gambar 4. 19 Tampilan Layar History Analisis	50
Gambar 4. 20 Tampilan Layar Keluaran Analisis Sentimen	51
Gambar 4. 21 Tampilan Layar Keluaran Analisis File CSV Multi Ulasan	52

Gambar 4. 22 Tampilan Layar Cetak Tabel Laporan.....	53
--	----

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Dalam era digital dan kemajuan teknologi, *e-commerce* telah menjadi salah satu platform utama untuk berbelanja produk secara online. Seiring dengan pertumbuhan *e-commerce*, ulasan produk oleh pengguna atau konsumen telah menjadi faktor penting dalam pengambilan keputusan pembelian. Ulasan-ulasan ini mencerminkan pendapat dan pengalaman pelanggan terkait kualitas, performa, layanan, dan kepuasan terhadap produk yang mereka beli.

Dalam konteks penelitian ini, peneliti akan berfokus pada analisis sentimen pada ulasan produk di platform *e-commerce* tertentu. Platform tersebut dipilih karena memiliki basis pengguna yang besar dan beragam. Data ulasan yang dapat diakses dari platform tersebut melimpah, sehingga memberikan sampel yang tepat untuk dilakukan analisis sentimen. Dalam penelitian ini, data ulasan produk diambil dari platform Tokopedia melalui sumber data Kaggle, serta dengan pengumpulan data secara manual.

Pemilihan Tokopedia sebagai lokasi penelitian memiliki beberapa alasan yang kuat. Pertama, Tokopedia adalah salah satu platform *e-commerce* terkemuka dengan pasar yang besar di Indonesia. Hal ini memberikan kesempatan untuk mengumpulkan volume data yang mencukupi untuk analisis sentimen. Kedua, keberagaman produk yang tersedia di Tokopedia memungkinkan penelitian ini

melibatkan berbagai kategori produk, sehingga memperluas cakupan data untuk analisis sentimen.

Dalam melakukan analisis sentimen pada ulasan produk di *e-commerce*, tantangan utama adalah jumlah ulasan yang sangat besar dan kerumitan dalam melakukan analisis secara manual. Proses manual membutuhkan waktu dan sumber daya yang besar, sehingga tidak efisien. Oleh karena itu, diperlukan metode analisis sentimen yang otomatis dan efisien untuk mengklasifikasikan ulasan menjadi kategori sentimen yang berbeda.

Dalam penelitian ini, metode yang digunakan untuk analisis sentimen adalah metode *Naive Bayes*. Metode *Naive Bayes* telah terbukti efektif dalam mengklasifikasikan teks dan dapat mengotomatisasi proses analisis sentimen. Adapun tujuan dari penelitian ini adalah membangun sebuah website yang menyediakan API untuk pengguna yang ingin menggunakan layanan analisis sentimen pada produk atau layanan mereka. Website yang dibangun juga dapat melakukan analisis sentimen secara langsung serta dapat mengklasifikasikan ulasan dalam format file CSV. Dengan adanya website ini, pengguna dapat dengan mudah mengakses layanan analisis sentimen untuk mendapatkan wawasan tentang sentimen ulasan produk mereka.

B. Identifikasi Masalah

Berdasarkan latar belakang di atas, dapat diidentifikasi beberapa masalah, yaitu:

1. Analisis sentimen pada ulasan-ulasan produk dapat menjadi tugas yang rumit dan memakan waktu jika dilakukan secara manual.
2. Volume data ulasan produk di platform *e-commerce* bisa menjadi sangat besar dan bervariasi, sehingga diperlukan metode yang efisien dan akurat untuk mengklasifikasikan ulasan menjadi kategori sentimen yang tepat, seperti positif, atau negatif.
3. Ketepatan dan keakuratan analisis sentimen pada ulasan produk di *e-commerce* menjadi faktor penting dalam memberikan wawasan bagi pelaku *e-commerce* dalam memahami pendapat dan evaluasi pelanggan serta meningkatkan kualitas produk dan layanan.

C. Batasan Masalah

Berdasarkan rumusan masalah dan batasan kemampuan peneliti, perlu diadakan batasan-batasan masalah agar penelitian lebih terarah. Batasan-batasan tersebut diantaranya:

1. Data ulasan penelitian ini hanya difokuskan kepada satu platform yaitu Tokopedia
2. Penelitian ini hanya menggunakan data ulasan produk yang tersedia dari Tokopedia, bukan bertujuan membuat sistem yang akan digunakan oleh Tokopedia
3. Data yang digunakan dalam penelitian hanya ulasan produk yang tersedia dan bersifat public

4. Penelitian ini hanya akan menggunakan ulasan produk yang berupa teks, tidak memperhitungkan aspek lain seperti gambar dan video
5. Data latih yang digunakan dalam pengembangan model klasifikasi *naive bayes* akan dibatasi hanya ulasan produk yang telah dilabeli dari Kaggle serta pengumpulan secara manual yang dilakukan penulis
6. Analisis sentimen hanya akan berfokus pada dua kategori utama yaitu: positif, atau negatif.
7. Penelitian ini tidak akan mempertimbangkan faktor selain ulasan produk, seperti harga produk, deskripsi, ataupun rating produk
8. Penelitian ini tidak akan melibatkan ulasan langsung dari pengguna, melainkan hanya menggunakan ulasan produk yang telah ada.
9. Penelitian ini hanya membahas metode klasifikasi *Naive Bayes*, tidak akan membahas metode lainnya
10. Evaluasi performa model *Naive Bayes* hanya akan menggunakan metrik akurasi.
11. Penelitian ini tidak akan menyelidiki dampak system yang dibuat terhadap perbaikan produk atau layanan di *e-commerce*, melainkan lebih berfokus pada pengembangan metode analisis sentimen itu sendiri

D. Rumusan Masalah

Berdasarkan latar belakang dan batasan masalah di atas, maka rumusan masalah untuk topik ini adalah:

1. Bagaimana mengatasi kompleksitas dan kebutuhan waktu yang tinggi dalam melakukan analisis sentimen secara manual pada volume data ulasan produk di platform *e-commerce*?
2. Bagaimana mengembangkan metode otomatis untuk menganalisis sentimen ulasan produk pada data yang dimiliki sebuah *e-commerce*?
3. Bagaimana cara memprediksi sentimen ulasan produk baru secara otomatis dengan menggunakan metode *Naive Bayes*?

E. Tujuan Penelitian

Adapun tujuan dari penelitian skripsi ini adalah:

1. Mengembangkan metode analisis sentimen otomatis untuk mengatasi kompleksitas dan kebutuhan waktu yang tinggi terhadap data ulasan produk di platform *e-commerce*
2. Merancang dan mengimplementasikan metode analisis sentimen ulasan produk pada data yang terdapat di sebuah *e-commerce* menggunakan metode *Naive Bayes* sebagai solusi untuk menganalisis sentimen ulasan secara efisien.
3. Menguji dan memvalidasi performa model klasifikasi *Naive Bayes* dalam memprediksi sentimen ulasan produk baru secara otomatis, dengan model yang telah dilatih sebelumnya.

F. Manfaat Penelitian

Adapun manfaat penelitian terdiri atas dua manfaat, manfaat teoritis dan praktis.

1. Manfaat Teoritis

- a. Penelitian ini dapat memberikan kontribusi terhadap penelitian dalam bidang analisis sentimen pada ulasan produk di *e-commerce* dengan metode *Naive Bayes*. Hasil penelitian ini dapat menjadi acuan dan sumber referensi untuk penelitian lanjutan dalam pengembangan metode analisis sentimen yang lebih baik.
- b. Penelitian ini juga dapat membantu memperluas pemahaman tentang penerapan metode *Naive Bayes* dalam analisis sentimen teks pada skala besar, khususnya pada ulasan produk di platform *e-commerce*.

2. Manfaat Praktis

- a. Penelitian ini akan memberikan kontribusi dalam mengembangkan metode yang efisien dan akurat untuk analisis sentimen pada ulasan produk di *e-commerce*. Hal ini akan mempermudah pelaku *e-commerce* dalam memahami pendapat dan evaluasi pelanggan dengan cepat dan efektif..
- b. Hasil analisis sentimen yang akurat dan relevan dapat memberikan wawasan yang berharga bagi pelaku *e-commerce* dalam pengambilan keputusan terkait perbaikan kualitas produk, peningkatan layanan, serta pengembangan strategi pemasaran dan branding yang lebih tepat sasaran.

- c. Dengan pemahaman yang lebih baik tentang pendapat dan evaluasi pelanggan melalui analisis sentimen, pelaku *e-commerce* dapat lebih responsif terhadap kebutuhan dan keinginan pelanggan. Hal ini dapat berkontribusi pada peningkatan kepuasan pelanggan, memperkuat loyalitas, dan meningkatkan reputasi merek.
- d. Metode analisis sentimen yang dikembangkan dapat mengatasi kompleksitas dan kebutuhan waktu yang tinggi dalam melakukan analisis secara manual. Hal ini akan membantu pelaku *e-commerce* untuk mengoptimalkan penggunaan waktu dan sumber daya yang ada dalam melakukan analisis ulasan produk dengan cepat dan efisien.

G. Sistematika Penulisan

Untuk memberikan mempermudah melihat dan mengetahui pembahasan yang ada pada skripsi ini, maka diperlukan sebuah gambaran berupa sistematika yang akan menjadi sebuah pedoman dan kerangka penulisan skripsi. Adapun sistematika penulisannya sebagai berikut:

BAB I PENDAHULUAN

Terdiri dari latar belakang masalah, identifikasi masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan.

BAB II LANDASAN TEORI DAN PENELITIAN YANG RELEVAN

Bab ini berisi mengenai uraian landasan teori yang membahas berbagai teori-teori yang berkaitan dengan penulisan skripsi serta memberikan beberapa penelitian yang relevan dengan skripsi yang penulis ambil

BAB III METODOLOGI PENELITIAN

Bab ini berisi mengenai waktu dan tempat penelitian, tahapan penelitian, serta algoritma atau langkah langkah pengembangan system yang akan dibuat

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi mengenai definisi masalah terkait program dan penyelesaiannya, pembahasan terkait algoritma, pemodelan perangkat lunak yang dibuat, mencakup *Unified Modeling Language* (UML), rancangan layer, dan tampilan layer. Bab ini juga membahas terkait kelebihan serta kelemahan dalam penelitian yang penulis buat

BAB V SIMPULAN DAN SARAN

Bab ini berisi mengenai simpulan dan saran dari penulis yang diperoleh dari penulisan, dan penyusunan skripsi ini agar pembaca dapat memperbaiki dan mengembangkan skripsi ini lebih baik lagi

BAB II

LANDASAN TEORI DAN PENELITIAN YANG RELEVAN

A. Landasan Teori

Bab ini berisi teori-teori yang berkaitan dengan analisis sentimen terhadap ulasan produk di *e-commerce* dengan metode *Naive Bayes*.

1. Text Mining

Dalam kalimatnya (Allahyari et al., 2017) menyebutkan

Text mining atau *knowledge discovery from text* (KDT) - yang pertama kali diperkenalkan oleh Fledman dkk. - merujuk pada proses ekstraksi informasi berkualitas tinggi dari teks (baik yang terstruktur seperti data RDBMS, semi-terstruktur seperti XML dan JSON, maupun tidak terstruktur seperti dokumen kata, video, dan gambar).

Dikutip juga dari Feldman & Sanger dalam (Sunoto & Wasito, 2014) *Text mining* adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen dimana *text mining* merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar.

Dari kedua pernyataan diatas, penulis menyimpulkan bahwa *text mining* atau *knowledge discovery from text* (KDT) merujuk pada proses ekstraksi informasi berkualitas tinggi dari berbagai jenis teks, baik yang terstruktur maupun tidak terstruktur. Teknik ini digunakan untuk melakukan klasifikasi dokumen dan merupakan variasi dari data mining yang bertujuan untuk menemukan pola menarik dari data tekstual yang besar.

2. Machine Learning

Dalam klausanya (Zhou, 2021) menyatakan bahwa “*Machine learning* adalah cabang dari kecerdasan buatan yang berfokus pada pengembangan teknik dan algoritma yang memungkinkan komputer untuk belajar dan meningkatkan kinerjanya secara otomatis melalui pengalaman”.

Disebutkan juga oleh Putra dalam (Diantika & Firmanto, 2019) “Machine learning adalah teknik untuk melakukan inferensi terhadap data dengan pendekatan matematis”.

Dengan dua pernyataan tersebut penulis menyimpulkan *Machine Learning* yaitu sebuah teknik atau algoritma dengan pendekatan matematis yang dilakukan untuk inferensi terhadap data, dan dengan kemampuannya itu memungkinkan komputer untuk belajar secara otomatis melalui pengalaman.

3. Analisis Sentimen

“*Sentimen analysis* atau analisis sentimen adalah jenis klasifikasi teks yang mengkategorikan teks berdasarkan orientasi sentimen dari opini yang terkandung dalam teks tersebut” (Devika et al., 2016). Disebutkan juga “Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi” (Sari & Wibowo, 2019).

Disebutkan dalam kutipan Sari & Wibowo, Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Serta oleh Devika disebutkan bahwa analisis sentiment merupakan klasifikasi teks yang dikategorikan berdasarkan orientasi sentiment dari opini yang terkandung dalam text.

Maka dari kedua hal tersebut penulis menyimpulkan bahwa analisis sentiment merupakan teknik klasifikasi text yang mengkategorikan sebuah opini menjadi sentiment positif, netral, dan negatif, berdasarkan orientasi sentiment yang terkandung dalam teks itu sendiri.

4. **Text Pre-processing**

“*Preprocessing* merupakan proses awal yang akan mentransformasikan data masukan menjadi data dengan format yang sesuai dan siap untuk diproses” (Setyohadi et al., 2017).

Disebutkan juga dalam kutipan lain “Metode *Preprocessing* memainkan peran yang sangat penting dalam teknik dan aplikasi *text mining*. Metode ini merupakan langkah pertama dalam proses *text mining*.” (Vijayarani et al., 2015).

Dalam kedua kutipan sebelumnya disebutkan bahwa *preprocessing* merupakan langkah awal yang dilakukan dalam prosesi pengaplikasian *text mining*. Adapun *preprocessing* merupakan sebuah teknik

mentransformasikan sebuah data masukan yang masih mentah menjadi data yang matang atau siap di proses untuk tahap tahap berikutnya.

5. Naive Bayes

“*Naive Bayes* merupakan salah satu metode *machine learning* yang menggunakan perhitungan probabilitas”(Heliyanti Susana, 2022).

“Klasifikasi dengan metode *Naive Bayes* secara umum dilakukan dengan pendekatan peluang atau probabilitas. Algoritma ini memprediksi probabilitas di masa depan berdasarkan pengalaman yang sudah ada di masa lalu.” (Sipayung et al., 2016).

Adapun penulis menarik kesimpulan dari kedua pernyataan diatas yakni, metode *Naive Bayes* adalah sebuah algoritma dalam *machine learning* yang digunakan untuk melakukan klasifikasi data ke dalam kategori atau kelas tertentu dengan menggunakan perhitungan probabilitas.

Asumsi utama dari metode *Naive Bayes* adalah bahwa setiap fitur atau atribut dalam data adalah independen satu sama lain, artinya tidak ada ketergantungan atau interaksi antara fitur-fitur tersebut. Hal ini disebutkan oleh (Chatrina, Siregar et al., 2020) “Kelemahan dari metode ini sendiri yaitu adanya asumsi atau dengan kata lain kondisi kelas saling bebas atau independen, sehingga kurang akurat”. Meskipun dalam kenyataannya hal ini mungkin tidak selalu benar, namun asumsi ini digunakan untuk

mempermudah perhitungan dan membuat algoritma menjadi sederhana dan cepat.

Metode naïve bayes berdasarkan teorema bayes memiliki bentuk umum serbagai berikut:

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)}$$

Keterangan:

$P(H|X)$ = Probabilitas hipotesis H pada kondisi X

$P(H)$ = Probabilitas hipotesis H

$P(X|H)$ = Probabilitas X pada kondisi H

$P(X)$ = Probabilitas X

Dalam konteks probabilitas, terutama saat menggunakan probabilitas yang sangat kecil, seperti dalam perhitungan *Naive Bayes* dengan kata-kata yang jarang muncul, perkalian sejumlah probabilitas kecil dapat menghasilkan probabilitas yang lebih kecil lagi. Hal ini dapat menyebabkan *underflow numerik* ketika nilai probabilitas menjadi sangat dekat dengan nol dan kehilangan presisi. Untuk menghindarinya, maka akan digunakan perhitungan logaritma probabilitas untuk menghindari *underflow numerik*. “Jika suatu bilangan tidak mampu direpresentasikan oleh komputer karena $e < L$ atau $e > U$, maka akan terjadi underflow atau overflow” (Luknanto, 2001).

$$\log(P(H|X)) = \log(P(X|H)) + \log(P(H)) - \log(P(X))$$

Keterangan:

$\log(P(H|X))$ = Logaritma dari probabilitas posterior, yaitu probabilitas hipotesis H (kelas sentimen) terjadi setelah mengamati data X (ulasan). Ini adalah nilai yang ingin kita hitung atau prediksi.

$\log(P(X|H))$ = Logaritma dari probabilitas likelihood, yaitu probabilitas data X (ulasan) terjadi mengingat hipotesis H (kelas sentimen) yang ada. Ini menggambarkan seberapa mungkin ulasan dengan kelas sentimen H akan menghasilkan data X (kata-kata dalam ulasan). Pada implementasi Naive Bayes untuk klasifikasi sentimen, ini diperoleh dengan menghitung logaritma dari perkalian probabilitas kata-kata dalam ulasan yang dikaitkan dengan kelas sentimen H.

$\log(P(H))$ = Logaritma dari probabilitas prior, yaitu probabilitas hipotesis H (kelas sentimen) sebelum mengamati data X (ulasan). Ini adalah pengetahuan awal atau asumsi tentang distribusi kelas sentimen sebelum melihat ulasan. Dalam klasifikasi sentimen, ini dapat

diperoleh dengan menghitung logaritma dari probabilitas kelas sentimen H berdasarkan data pelatihan sebelumnya.

$\log(P(X))$ = Logaritma dari probabilitas data X (ulasan) secara keseluruhan. Ini adalah faktor normalisasi untuk memastikan bahwa probabilitas posterior yang dihasilkan berada dalam rentang yang valid (antara 0 dan 1). Dalam klasifikasi sentimen, ini bukan faktor yang mempengaruhi pemilihan sentimen tertentu, tetapi digunakan untuk memperhitungkan distribusi probabilitas dari semua kelas sentimen.

6. Evaluasi Kinerja Metode

Terdapat banyak Teknik klasifikasi yang telah di terapkan dalam banyak aplikasi dalam berbagai bidang ilmu, termasuk *Naive Bayes* yang umum digunakan dalam analisis sentimen.

Disebutkan oleh (Tharwat, 2021)

Metode klasifikasi telah diterapkan dalam banyak aplikasi di berbagai bidang ilmu. Ada beberapa cara untuk mengevaluasi algoritma klasifikasi. Analisis metrik-metrik tersebut dan signifikansinya harus diinterpretasikan dengan benar untuk mengevaluasi algoritma pembelajaran yang berbeda. Sebagian besar metrik ini adalah metrik skalar dan beberapa di antaranya adalah metode grafis.

Setelah mengimplementasikan metode *Naive Bayes* untuk melakukan klasifikasi sentimen, Langkah selanjutnya yakni melakukan evaluasi kinerja dari model yang telah dibangun

“Hasil olahan kasifikasi dapat dievaluasi untuk mengetahui tingkat akurasi dan efisiensi, beberapa metrik yang perlu diketahui dalam evaluasi model yaitu akurasi, *confussion matrix*, f1-score, presisi, *recall*” (Tharwat, 2021).

Metrik akurasi adalah salah satu metrik evaluasi yang umum digunakan dalam pembelajaran mesin untuk mengukur seberapa akurat suatu model dalam melakukan prediksi kelas atau kategori dari data yang diberikan. Metrik akurasi dihitung dengan membandingkan jumlah prediksi yang benar (*True Positive dan True Negative*) dengan jumlah total data yang dievaluasi.

Disebutkan oleh (Normawati & Prayogi, 2021) Confusion matrix adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah. Metrik *confusion* atau *confusion matrix*, bekerja dengan cara membandingkan prediksi model dengan nilai yang sebenarnya dari data yang dievaluasi. Matriks confusion biasanya digunakan untuk masalah klasifikasi dua kelas (*binary*), di mana model klasifikasi memprediksi data menjadi dua kategori yang memungkinkan, yaitu positif dan negatif (atau benar dan salah). F1-score dan *confusion matrix* saling berhubungan dalam evaluasi performa suatu model klasifikasi. *Confusion matrix* umumnya

terdiri dari empat sel, yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN).

		Actual Condition		
		FALSE	TRUE	
Predicted Condition	FALSE	TN	FN	Predicted Negative
	TRUE	FP	TP	Predicted Positive
		Actual Negative	Actual Positive	

Gambar 2.1 Confussion Matrix

Sumber: <https://benhay.es/posts/demystifying-confusion-matrix/>

Akurasi (Accuracy) adalah metrik yang mengukur sejauh mana model klasifikasi dapat memprediksi dengan benar. Akurasi dihitung sebagai persentase prediksi yang benar (*True Positive* dan *True Negative*) dibandingkan dengan total jumlah contoh.

$$\frac{(TP + TN)}{(TP + TN + FP + FN)}$$

Presisi (*precision*) merupakan perbandingan antara jumlah prediksi yang benar positif (*true positive*) dengan jumlah prediksi positif (*true positive* dan *false positive*). Rumus presisi adalah:

$$Presisi = TP \frac{TP}{TP + FP}$$

Recall, juga dikenal sebagai sensitivitas atau *true positive rate* (TPR), merupakan perbandingan antara jumlah prediksi yang benar positif (*true positive*) dengan jumlah data positif sebenarnya (*true positive* dan *false negative*). Rumus *recall* adalah:

$$Recall = \frac{TP}{TP + FN}$$

F1-score dihitung sebagai harmonic mean antara presisi dan recall. Harmonic mean merupakan suatu jenis rata-rata yang memberikan bobot yang seimbang antara dua nilai yang dihitung. Rumus *F1-score* adalah:

$$F1 - score = 2 * \frac{(Presisi * Recall)}{(Presisi + Recall)}$$

F1-score memiliki rentang nilai antara 0 dan 1, dimana nilai 1 menunjukkan performa yang sempurna (presisi dan *recall* yang optimal) dan nilai 0 menunjukkan performa yang buruk (salah satu atau keduanya dari presisi dan *recall* sangat rendah). *F1-score* digunakan untuk menggabungkan presisi dan *recall* menjadi suatu metrik yang tunggal, sehingga dapat memberikan gambaran yang lebih komprehensif tentang performa keseluruhan suatu model klasifikasi.

7. Application Programming Interface

“*Application Programming Interface* atau API merupakan integrasi dari dua bagian dari sistem aplikasi” (Yanti & Rihyanti, 2021).

“API juga biasa dianggap sebagai suatu kumpulan teknik yang jelas untuk menciptakan suatu komunikasi antara perangkat lunak yang berbeda beda komponen” (Khoirul Akmal & Dasaprawira, 2022).

Dari dua kutipan diatas dapat disimpulkan bahwa, *Application Programming Interface* atau API merupakan teknik integrasi antar perangkat lunak yang berbeda komponen untuk dapat menciptakan komunikasi.

8. **Node.JS**

Dijelaskan dalam (Fajrin, 2017) bahwa “Node.JS adalah sistem perangkat lunak yang didesain untuk pengembangan aplikasi web”

Menurut (Kurniawan et al., 2020) dalam jurnalnya

Bila selama ini kita mengenal JavaScript sebagai bahasa pemrograman yang berjalan di sisi client / browser saja, maka Node.js ada untuk melengkapi peran JavaScript sehingga bisa juga berlaku sebagai bahasa pemrograman yang berjalan di sisi server, seperti halnya PHP, Ruby, Perl, dan sebagainya.

Dengan dua pernyataan diatas, maka penulis menyimpulkan bahwa Node.JS merupakan sebuah sistem perangkat lunak yang di design untuk pengembangan website, yang ada untuk melengkapi peran JavaScript sehingga mampu berjalan di sisi server

9. **E-commerce**

Menurut Shely Cashman dalam (Aco & Endang, 2017)

E-commerce atau kependekan dari *electronic commerce* (perdagangan secara elektronik), merupakan transaksi bisnis yang terjadi dalam

jaringan elektronik, seperti internet. Siapapun yang dapat mengakses komputer, memiliki sambungan ke internet, dan memiliki cara untuk membayar barang-barang atau jasa yang mereka beli, dapat berpartisipasi dalam *e-commerce*.

Dengan pernyataan diatas penulis menarik kesimpulan, *e-commerce* memberikan kemudahan bagi konsumen untuk membeli produk dari berbagai negara atau wilayah tanpa harus keluar rumah selama mereka dapat mengakses computer dan memiliki sambungan ke internet. Selain itu, para pelaku bisnis juga mendapatkan manfaat dari *e-commerce* karena dapat menjangkau pasar yang lebih luas dan memperluas jaringan distribusi produk mereka. Seiring dengan pertumbuhan teknologi dan internet yang semakin pesat, *e-commerce* diprediksi akan terus berkembang dan menjadi salah satu bentuk perdagangan yang dominan di masa depan.

a. Ulasan Produk

Ulasan produk merupakan pendapat atau komentar yang diberikan oleh pelanggan atau pengguna produk tertentu setelah mereka menggunakannya. Ulasan produk dapat berisi pengalaman pribadi, penilaian, kritik, dan saran untuk produk tersebut.

E-commerce juga memiliki fitur yang memungkinkan konsumen memberikan ulasan atau penilaian produk yang telah mereka beli. Ulasan produk di *e-commerce* menjadi sangat penting karena dapat memengaruhi keputusan pembelian calon konsumen. Konsumen seringkali mempercayai ulasan dari pengguna sebelumnya untuk

menilai kualitas, keandalan, dan kepuasan produk yang ditawarkan. Oleh karena itu, analisis sentimen terhadap ulasan produk di *e-commerce* menjadi penting bagi penjual atau pedagang dalam memahami bagaimana konsumen merespons produk mereka.

10. UML (Unified Modeling Language)

“Unified Modelling Language merupakan alat perancangan sistem yang berorientasi pada objek” (Haviluddin, 2011).

“UML sendiri juga memberikan standar penulisan sebuah sistem blueprint, yang meliputi konsep proses bisnis, penulisan kelas-kelas dalam bahasa program yang spesifik, skema database, dan komponen yang diperlukan dalam sistem software” (Sonata, 2019).

Dari penjelasan dua kutipan diatas dapat di Tarik kesimpulan bahwa UML (Unified Modeling Language) merupakan sebuah alat perancangan sistem berorientasi object yang memberikan standar penulisan untuk sebuah sistem blueprint

a. Diagram Use Case

Disebutkan oleh Fowler & Scott dalam (Setiyani, 2021) “Use case adalah seperangkat skenario yang diikat bersama oleh user untuk mencapai tujuan”

“Use case merupakan diagram yang bekerja dengan cara mendeskripsikan tipikal interaksi antara user (pengguna) sebuah

sistem dengan suatu sistem tersendiri melalui sebuah cerita bagaimana sebuah sistem dipakai.” (Ismail, 2020)

Dari dua kutipan diatas jelas bahwa diagram use case merupakan sebuah diagram atau scenario yang bekerja secara mendeskripsikan user dan sistem untuk menjelaskan bagaimana sistem dipakai

b. Diagram Activity

“Activity diagram menggambarkan berbagai aliran aktivitas dalam sistem yang sedang di rancang, bagaimana masing-masing aliran berawal, decision yang mungkin terjadi dan bagaimana mereka berakhir.” (Ismail, 2020)

Dapat ditarik kesimpulan dari kutipan berikut bahwa diagram activity menggambarkan alur aktivitas dari sebuah sistem, bagaimana sistem berawal hingga berakhir

c. Diagram Sequence

“Sequence diagram digunakan untuk menggambarkan interaksi antar objek di dalam dan di sekitar sistem yang berupa message yang digambarkan terhadap waktu” (Ismail, 2020)

Dari kutipan berikut penulis mengambil kesimpulan bahwa diagram sequence merupakan sebuah diagram yang menggambarkan interaksi antar objek yang digambarkan terhadap waktu

d. Diagram Class

“Class diagram digunakan untuk melakukan visualisasi struktur kelas-kelas dari suatu sistem dan merupakan tipe diagram yang paling banyak 36 digunakan” (Ismail, 2020)

Dalam kutipan diatas menjelaskan bahwa diagram class digunakan untuk memvisualisasikan kelas kelas yang ada dari sebuah sistem yang digunakan

B. Penelitian Yang Relevan

1. Skripsi

- | | | |
|----|-------------------|--|
| a. | Nama | : Tata Ilham Rais |
| | Universitas | : Universitas Islam Negeri Syarif Hidayatullah Jakarta |
| | Tahun | : 2023 |
| | Judul | : Analisis sentimen terhadap komentar video youtube Raiden Shogun - Judgment of Euthymia menggunakan metode majority voting |
| | Masalah | : Bagaimana analisis sentimen terhadap komentar dalam video youtube dapat dilakukan menggunakan metode majority voting? |
| | Tujuan Penelitian | : Melakukan suatu analisis sentimen yang dapat menghasilkan gambaran umum terkait hal-hal yang disukai penonton dari video “Raiden Shogun – Judgement of Euthymia” yang dapat menjadi masukan untuk pembuatan video atau |

desain karakter selanjutnya, serta menguji akurasi dari model analisis sentimen majority voting

- b. Nama : Gerald Halim Al Rasyid Noer
 Universitas : Universitas Islam Negeri Syarif Hidayatullah Jakarta
 Tahun : 2023
 Judul : Implementasi Algoritma Naïve Bayes dan TF-IDF Dalam Analisis Sentimen Data Ulasan (Studi Kasus: Ulasan Review Aplikasi *E-commerce* Shopee di Situs Google Playstore
 Masalah : Bagaimana Performa Naive Bayes dan TF-IDF dalam melakukan analisis sentimen terhadap data ulasan Shopee di situs Google Playstore?
 Tujuan Penelitian : Mengklasifikasikan apakah algoritma Naive Bayes dan TF-IDF dapat digunakan dalam klasifikasi ulasan ulasan pengguna aplikasi Shopee melalui situs Google Playstore

2. Jurnal

- a. Nama Jurnal : Jurnal Sistem Informasi (JSI)
 Nama Penulis : Evasaria M. Sipayung, Herastia Maharani, Ivan Zefanya
 Judul : Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier

- E-ISSN : 2355-4614
 Volume : Vol. 8, No. 1, April 2016
 Hasil : Perhitungan kehandalan sistem analisis sentimen dengan metode naïve bayes classifier ini dari 175 data latih dapat dibagi menjadi dua bagian, perhitungan klasifikasi kategori dan perhitungan klasifikasi sentimen, untuk hasil dari akurasi sistem memiliki nilai accuration kategori 77.14% dan untuk precision sentimen 99.12%, recall sentimen 72.9%, dan accuration sentimen 75.42%, sehingga sistem mampu mengembalikan dokumen dan kecocokan data yang tinggi.
- b. Nama Jurnal : Jurnal Edukasi dan Penelitian Informatika (JEPIN)
 Nama Penulis : Billy Gunawan, Helen Sastypratiwi, Enda Esyudha Pratama
 Judul : Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes
 E-ISSN : 2548-9364
 Volume : Vol. 4, No. 2, (2018)
 Hasi : Metode Naive Bayes dapat memprediksi kelas sentimen pada ulasan produk online sesuai dengan sistem yang disiapkan. Sistem analisis sentimen pada ulasan produk online menggunakan metode Naive Bayes menghasilkan nilai akurasi terendah pada

pengujian 5 kelas menggunakan dataset 80% latih dan 20% data uji sebesar 52.66%, sedangkan pada pengujian 3 kelas menggunakan dataset 90% data latih dan 10% data uji memiliki akurasi tertinggi sebesar 77.78%.

BAB III

METODOLOGI PENELITIAN

A. Waktu dan Tempat Penelitian

1. Waktu Penelitian

- a. Waktu yang digunakan peneliti untuk penelitian ini dilaksanakan sejak tanggal dikeluarkannya ijin penelitian. Berikut dibawah ini adalah tabel perincian tahapan penelitian berdasarkan waktu dan kegiatan

Table 3.1 Jadwal Kegiatan Penelitian

No	Aktifitas	Maret 2023				April 2023				Mei 2023				Juni 2023				Juli 2023			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1.	Tahap Perancangan																				
2.	Pengumpulan Data																				
3.	Tahap Pengolahan Data																				
4.	Analisis dan Interpretasi Data																				
5.	Penulisan Laporan																				

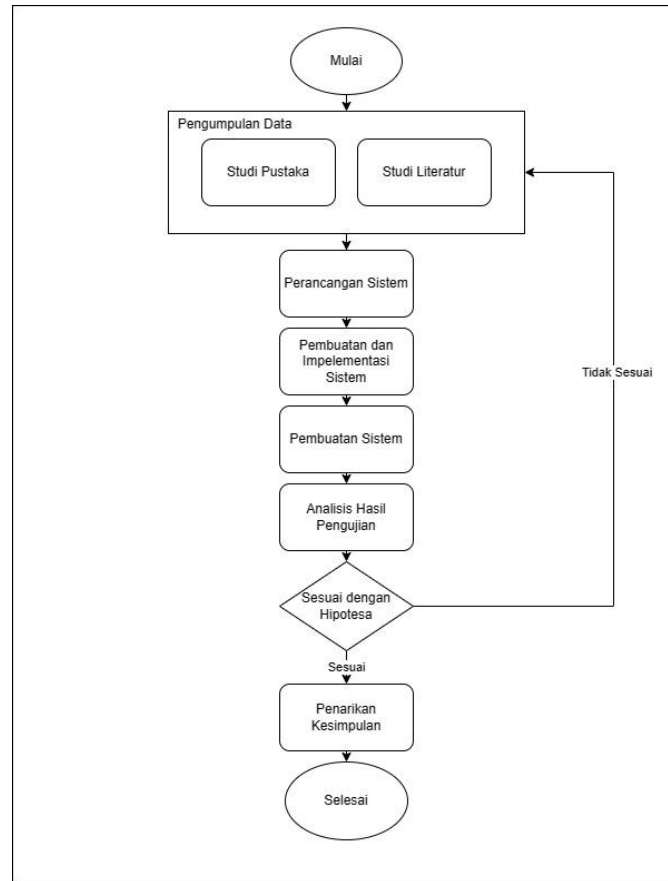
2. Tempat Penelitian

Penelitian ini akan dilakukan dengan menggunakan data ulasan produk dari platform *e-commerce* Tokopedia yang diambil dari website penyedia data Kaggle serta pengumpulan secara manual oleh penulis.

Kaggle adalah platform yang menyediakan berbagai kumpulan data yang dapat digunakan untuk penelitian dan pengembangan di berbagai bidang, termasuk dalam konteks analisis sentimen pada ulasan produk di *e-commerce*. Data ulasan produk Tokopedia yang disediakan oleh Kaggle akan menjadi sumber data utama yang akan digunakan dalam penelitian ini.

B. Tahapan Penelitian

Tahapan penelitian merupakan tahapan yang dilakukan untuk melakukan penelitian mulai dari awal hingga selesai. Penelitian yang dilakukan sesuai dengan tahapan sebagai berikut:



Gambar 3.1 Tahapan Penelitian

Berikut adalah penjelasan untuk tahapan penelitian dari tahapan-tahapan yang diberikan:

1. Tahapan pengumpulan data melalui studi pustaka dan studi literatur

Tahap awal penelitian adalah pengumpulan data melalui studi pustaka dan literatur. Dalam hal ini, penulis akan mengumpulkan informasi tentang teori dan konsep terkait analisis sentimen dan metode *naive bayes*. Dalam hal ini penulis juga akan mengumpulkan data terkait ulasan produk

2. Perancangan sistem

Tahapan selanjutnya adalah perancangan sistem yang akan digunakan untuk melakukan analisis sentimen terhadap ulasan produk. Penulis akan merancang sistem yang terdiri dari beberapa komponen seperti proses *Pre-processing* data, proses pelatihan dan pengujian model *Naive Bayes*, dan proses klasifikasi sentimen.

3. Pembuatan dan implementasi sistem

Setelah perancangan sistem selesai dilakukan, penulis kemudian akan memulai tahap pembuatan dan implementasi sistem. Dalam tahap ini, penulis akan membuat program komputer atau aplikasi yang dapat melakukan analisis sentimen terhadap ulasan produk menggunakan metode *Naive Bayes* yang telah dirancang sebelumnya.

4. Pelatihan model *Naive Bayes*

Tahapan berikutnya adalah pelatihan model *Naive Bayes* menggunakan data pelatihan yang telah dikumpulkan sebelumnya. Data pelatihan akan digunakan untuk melatih model *Naive Bayes* agar dapat mengklasifikasikan sentimen dari ulasan produk dengan akurasi yang tinggi.

5. Analisis hasil pengujian

Setelah model *Naive Bayes* terlatih, tahap selanjutnya adalah melakukan pengujian pada data pengujian. Data pengujian akan digunakan untuk menguji akurasi model *Naive Bayes* dalam mengklasifikasikan sentimen dari ulasan produk. Hasil pengujian akan dianalisis untuk mengetahui

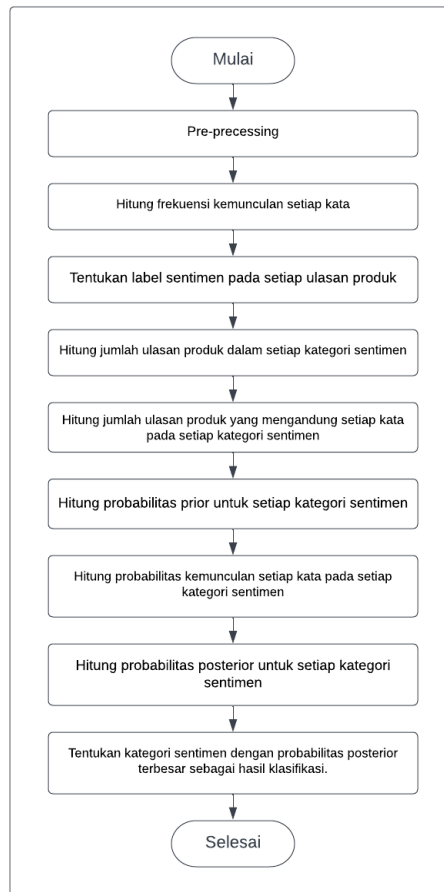
seberapa baik model *Naive Bayes* dalam melakukan analisis sentimen terhadap ulasan produk.

6. Penarikan kesimpulan

Tahap terakhir adalah penarikan kesimpulan dari hasil penelitian yang telah dilakukan. Penulis akan melakukan analisis terhadap hasil pengujian dan menarik kesimpulan mengenai efektivitas dan kegunaan model *Naive Bayes* dalam melakukan analisis sentimen terhadap ulasan produk. Selain itu, penulis juga dapat memberikan rekomendasi untuk pengembangan penelitian selanjutnya.

C. Algoritma

Adapun berikut adalah langkah-langkah algoritma *Naive Bayes* dalam analisis sentimen terhadap ulasan produk di *e-commerce*:



Gambar 3.2 Diagram Kerja Algoritma

Diagram kerja di atas menunjukkan bahwa algoritma *Naive Bayes* untuk analisis sentimen dimulai dengan:

1. membersihkan setiap ulasan produk dari tanda baca, kata-kata yang tidak relevan, dan merubah kata-kata menjadi bentuk dasarnya.
2. Kemudian, algoritma menghitung frekuensi kemunculan setiap kata pada setiap ulasan produk dan menentukan label sentimen (positif, negatif, atau netral) pada setiap ulasan produk.

3. Selanjutnya, algoritma menghitung jumlah ulasan produk pada setiap kategori sentimen dan jumlah ulasan produk yang mengandung setiap kata pada setiap kategori sentimen. Dari sini, algoritma menghitung probabilitas prior dan probabilitas kemunculan setiap kata pada setiap kategori sentimen.
4. Setelah itu, algoritma menghitung probabilitas posterior untuk setiap kategori sentimen dan menentukan kategori sentimen dengan probabilitas posterior terbesar sebagai hasil klasifikasi.
5. Terakhir, algoritma menghasilkan hasil klasifikasi sentimen pada setiap ulasan produk.

BAB IV

HASIL DAN PEMBAHASAN

A. Definisi Masalah

Berdasarkan latar belakang yang telah dijelaskan dalam bab bab sebelumnya, dapat diidentifikasi masalah sebagai berikut

1. Volume ulasan produk yang sangat besar di platform *e-commerce*, menyebabkan sulitnya pelaku *e-commerce* dalam menganalisis sentimen ulasan produk secara manual.
2. Ketidakmampuan pelaku *e-commerce* untuk dengan cepat memahami dan mengevaluasi pendapat serta persepsi pelanggan terhadap produk yang mereka jual.
3. Keterbatasan metode analisis sentimen yang saat ini digunakan oleh penjual atau pemilik toko di sebuah *e-commerce* dalam mengolah ulasan produk secara efisien.

Dengan identifikasi masalah diatas penulis membuat solusi untuk penyelesaian permasalahan sebagai berikut

1. Penggunaan metode *Naive Bayes* dalam analisis sentimen pada ulasan produk di *e-commerce* dapat menjadi solusi yang efektif. Metode ini akan memungkinkan penjual atau pemilik toko untuk dengan cepat dan otomatis menganalisis sentimen ulasan produk, sehingga mereka dapat memahami pendapat dan persepsi pelanggan dengan lebih baik.

2. Dengan menerapkan metode *Naive Bayes*, penjual atau pemilik toko di sebuah *e-commerce* dapat memprediksi sentimen ulasan produk baru secara otomatis. Hal ini akan membantu mereka untuk dengan cepat mengetahui tanggapan pelanggan terhadap produk yang baru mereka jual di platform.
3. Dengan memiliki pemahaman yang lebih baik tentang sentimen pelanggan, penjual atau pemilik toko di suatu *e-commerce* dapat melakukan perbaikan produk yang diperlukan dan meningkatkan layanan yang ditawarkan. Dengan demikian, mereka dapat meningkatkan kepuasan pelanggan dan meningkatkan kualitas bisnis mereka.

B. Pembahasan Algoritma

Dalam pembahasan algoritma ini penulis akan menyajikan perhitungan manual menggunakan data berupa sentiment yang bernilai positive dalam pandangan penulis, berikut merupakan alur pengimplementasian algoritma terhadap studi kasus penelitian.

1. Masukan Text

“Bagus banget barang ini aku suka”
2. Preprocessing Masukan Text
 - a. Tokenisasi

Dalam preprocessing tokenisasi, masukan text akan dirubah menjadi token token berbentuk array berupa

['bagus', 'banget', 'barang', 'ini', 'aku', 'suka']

b. Penghapusan *Stopword*

Setelah pembuatan tokenisasi terlaksana, Langkah selanjutnya dalam *preprocessing* berupa penghapusan *stopword* atau kata-kata umum yang sering muncul dalam teks namun tidak memiliki makna khusus dalam proses analisis teks, menjadi

['bagus', 'banget', 'barang', 'suka']

3. Perhitungan Naive Bayes

a. Menghitung Probabilitas dalam sentiment positive

1) Menghitung $\log(P(H1))$

$(P(H1)) = \text{jumlah sentiment positive} / \text{total sentiment}$

$$(P(H1)) = 270 / 539$$

$$(P(H1)) = 0.5009276437847866$$

$$\log(P(H1)) = -0.6912936119106224$$

2) Menghitung $\log(P(X|H))$

$$\log(P(\text{Bagus}|H1)) = -2.950 \quad \log(P(\text{Banget}|H1)) = -3.886$$

$$\log(P(\text{Barang}|H1)) = -2.784 \quad \log(P(\text{Suka}|H1)) = -5.166$$

3) Menghitung $\log(P(H1|X))$

$$\log(P(H1|X)) = \log(P(X|H1)) + \log(P(H1)) - \log(P(X))$$

$$\log(P(H1|X)) = -0.691 + (-2.950) + (-3.886) + (-2.784) + (-5.166)$$

$$\log(P(H1|X)) = -15.479$$

b. Menghitung Probabilitas dalam sentiment negative

1) Menghitung $\log(P(H2))$

$(P(H2)) = \text{jumlah sentiment negative} / \text{total sentiment}$

$$(P(H2)) = 269 / 539$$

$$(P(H2)) = 0.49907235621521334$$

$$\log(P(H2)) = -0.6950041913071581$$

2) Menghitung $\log(P(X|H2))$

$$\log(P(\text{Bagus}|H2)) = -5.219 \quad \log(P(\text{Banget}|H2)) = -5.367$$

$$\log(P(\text{Barang}|H2)) = -4.142 \quad \log(P(\text{Suka}|H2)) = -6.958$$

3) Menghitung $\log(P(H2|X))$

$$\log(P(H2|X)) = \log(P(X|H2)) + \log(P(H2)) - \log(P(X))$$

$$\log(P(H2|X)) = -0.695 + (-5.219) + (-5.367) + (-4.142) + (-6.958)$$

$$\log(P(H2|X)) = -22.383$$

c. Menentukan kelas sentimen prediksi

Karena $\log(P(H1|X))$ memiliki nilai tertinggi, maka kelas sentimen prediksi untuk ulasan tersebut adalah "Positif".

4. Menghitung Akurasi Model

Untuk menghitung akurasi menggunakan *Confusion Matrix* atau menggunakan matrik akurasi, penulis akan menggunakan data test yang sebelumnya telah sediakan untuk menguji hasil klasifikasi dari model yang telah dilatih.

a. Informasi Data

Data test yang akan digunakan memiliki total sebanyak 89 ulasan beserta label, dimana label positive sebanyak 63 dan label negative sebanyak 25

b. Perhitungan *True Positive* (TP) dan *False Positive* (FP)

Perhitungan confusion matrix dilakukan dengan melakukan klasifikasi dari data test menggunakan model yang telah dilatih, dimana apabila hasil prediksi label positive sesuai dengan label sebenarnya maka jumlah *True Positive* akan ditambahkan dengan 1 (satu), sebaliknya apabila hasil prediksi tidak sesuai dengan label sebenarnya maka *False Positive* akan ditambahkan dengan 1 (satu).

Dari hasil perhitungan yang dilakukan penulis didapati *True Positive* sebanyak 56, dan *False Positive* sebanyak 7

c. Perhitungan *True Negative* (TN) dan *False Negative* (FN)

Perhitungan ini dilakukan dengan membandingkan data test berlabel sentimen negative, dimana apabila hasil prediksi label negative sesuai dengan label sebenarnya maka jumlah *True Negative* akan ditambahkan dengan 1 (satu), sebaliknya apabila hasil prediksi tidak sesuai dengan label sebenarnya maka *False Negative* akan ditambahkan dengan 1 (satu).

Dari hasil perhitungan yang dilakukan penulis didapati *True Negative* sebanyak 18, dan *False Negative* sebanyak 7

d. Menghitung Metrik Akurasi

$$\text{Akurasi} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Akurasi} = (56 + 18) / (56 + 18 + 7 + 7)$$

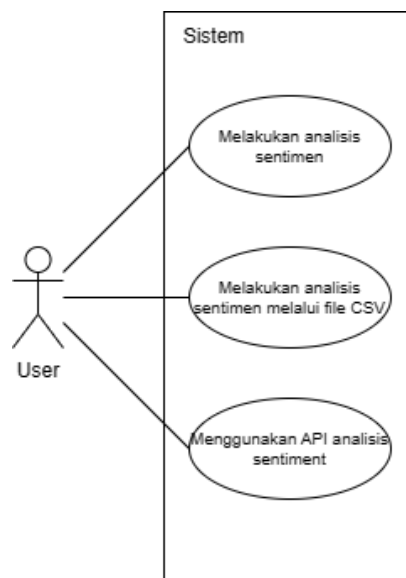
$$\text{Akurasi} = 0.8409 \text{ atau } 84.09\%$$

C. Pemodelan Perangkat Lunak

1. *Unified Modeling Language*

a. *Use Case Diagram*

Diagram berikut menggambarkan interaksi aktor di dalam sistem analisis sentimen pada ulasan produk di *e-commerce*.

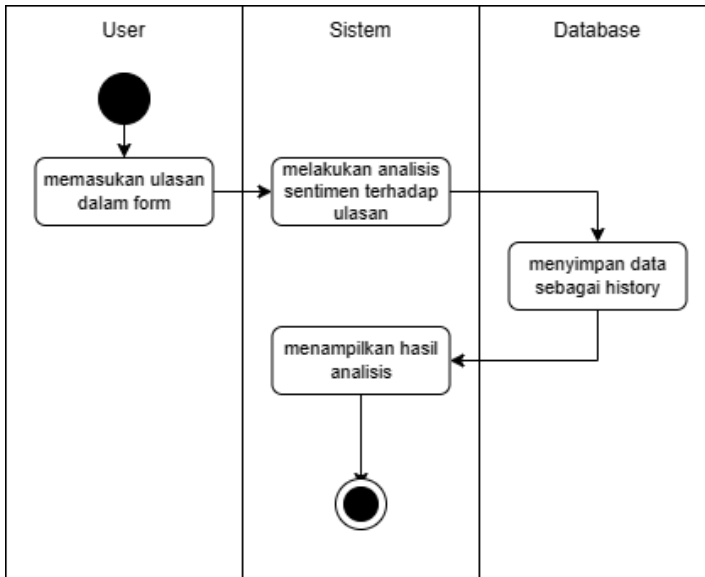


Gambar 4.1 Diagram Use Case

b. *Activity Diagram*

Berikut adalah beberapa diagram aktivitas yang terbentuk dari kegiatan bisnis dan use case diagram yang telah dibahas sebelumnya

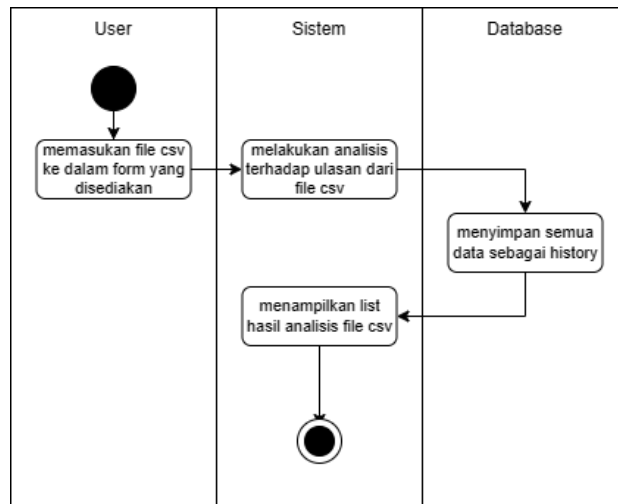
1.) *Activity diagram* dari *Use Case* Melakukan Analisis Sentimen



Gambar 4.2 Diagram Activity Melakukan Analisis Sentimen

Aktivitas ini menggambarkan aktor atau user yang ingin melakukan analisis sentimen terhadap suatu ulasan. Dimulai dari aktor memasukkan ulasan yang ingin di analisis kedalam sebuah form yang telah disediakan. Kemudian sistem akan melakukan analisis terhadap input yang dimasukan oleh user. Setelah analisis berhasil dilakukan sistem akan menyimpan data ulasan dan hasil sentimen ke dalam database history. Terakhir sistem akan menampilkan hasil analisis dari input ulasan yang dimasukan user

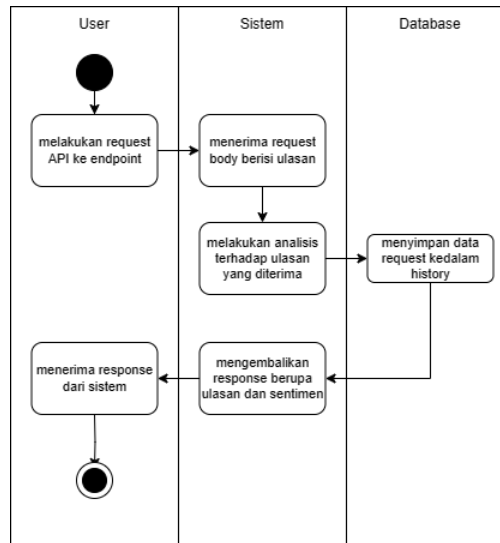
2.) *Activity diagram* dari *Use Case* Melakukan Analisis Sentimen Melalui file CSV



Gambar 4.3 Diagram Activity Melakukan Analisis Sentimen Melalui file CSV

Aktivitas ini menggambarkan aktor atau user yang ingin melakukan analisis terhadap banyak ulasan sekaligus menggunakan file dengan format CSV. Dimulai dari aktor melakukan input file CSV kedalam form input yang telah disediakan. Kemudian sistem akan melakukan analisis terhadap semua ulasan yang terdapat di dalam file CSV. Apabila proses analisis berhasil maka sistem akan menyimpan semua input ulasan serta hasil analisis ke dalam database history. Terakhir sistem akan menampilkan hasil dari analisis berupa tabel list dengan atribut ulasan dan hasil sentimen.

3.) Activity diagram dari Use Case Menggunakan API Analisis Sentimen

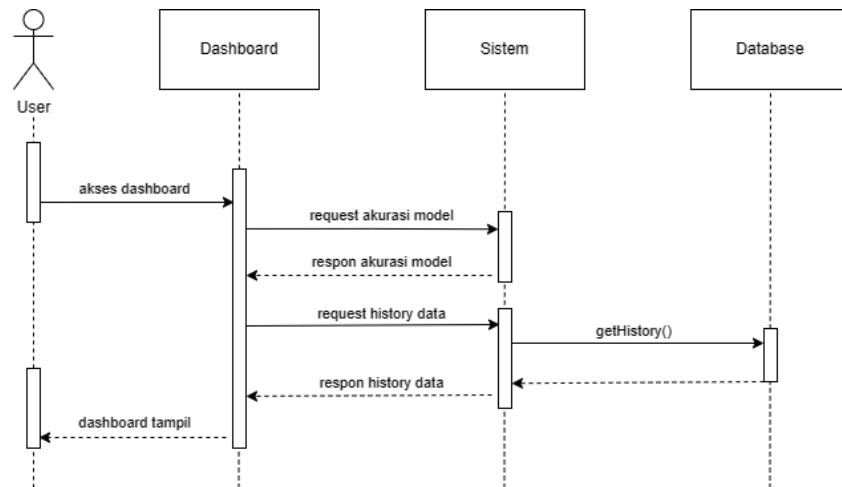


Gambar 4.4 Diagram Activity Menggunakan API Analisis Sentimen

Aktivitas ini menggambarkan aktor atau user yang ingin mengimplementasi API analisis sentimen. Dimulai dari user melakukan request terhadap endpoint yang telah disediakan sistem. Kemudian sistem akan menerima request dari user berupa request body berupa ulasan yang akan di analisis. Sistem menjalankan proses analisis terhadap request body yang dikirim oleh user. Setelah proses analisis berhasil sistem akan menyimpan data ulasan beserta hasil sentiment kedalam database history. Terakhir sistem mengembalikan response berupa ulasan dan hasil sentiment kepada user.

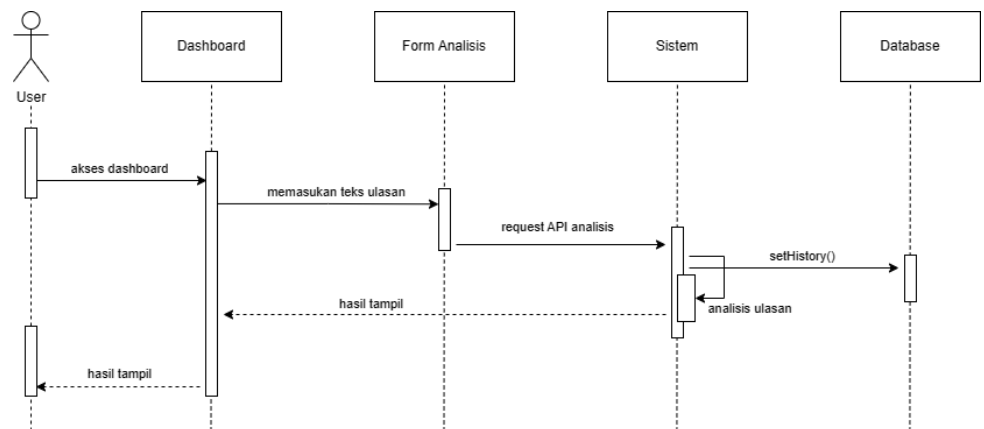
c. Sequence Diagram

1.) Sequence Diagram Indeks



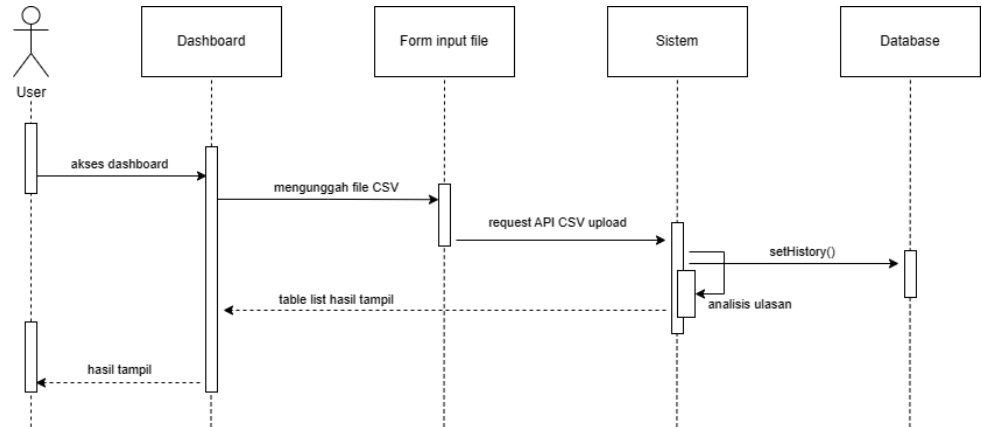
Gambar 4.5 Sequence Diagram Indeks

2.) Sequence Diagram Analisis Sentimen



Gambar 4.6 Sequence Diagram Analisis Sentimen

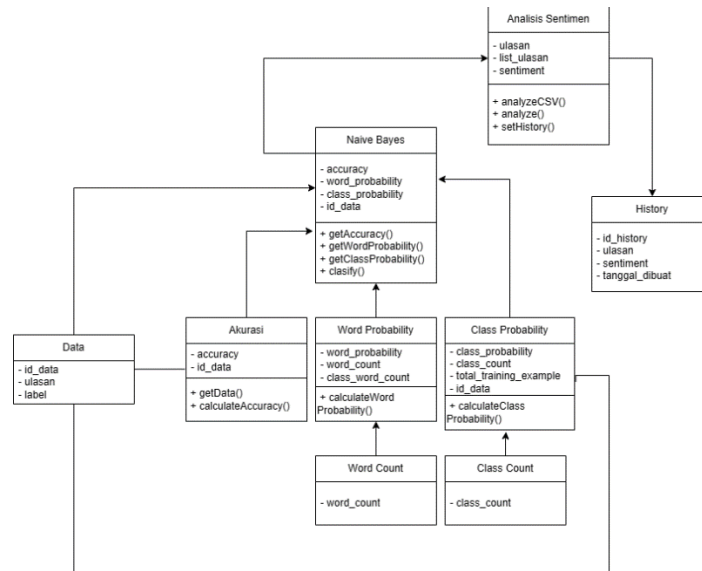
3.) Sequence Diagram Analisis Sentimen dengan file CSV



Gambar 4.7 Sequence Diagram Analisis Sentimen dengan file CSV

d. Class Diagram

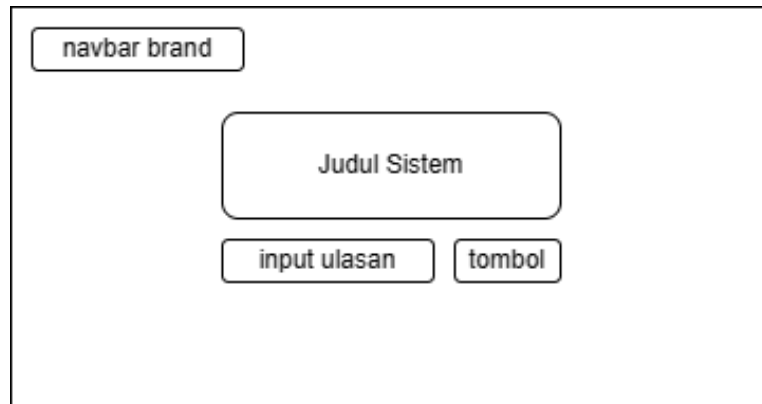
Adapun untuk class diagram untuk Analisis Sentimen pada Ulasan Produk di *E-commerce* sebagai berikut



Gambar 4.8 Class Diagram

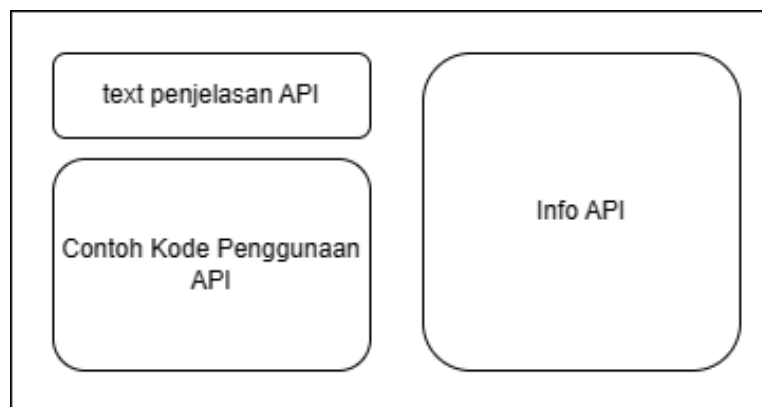
2. Rancangan Layar

a. Halaman Utama



Gambar 4.9 Rancangan Layar Halaman Utama

b. *API Info and Usage*



Gambar 4.10 Rancangan Layar Cara pakai API dan Akurasi

Model

c. Upload file CSV multi ulasan

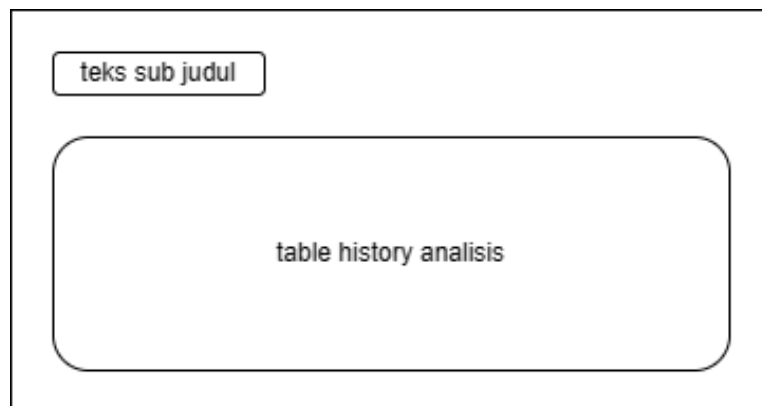


teks penjelasan

input file tombol upload

Gambar 4.11 Rancangan Layar Upload File CSV Multi Ulasan

d. History Analisis

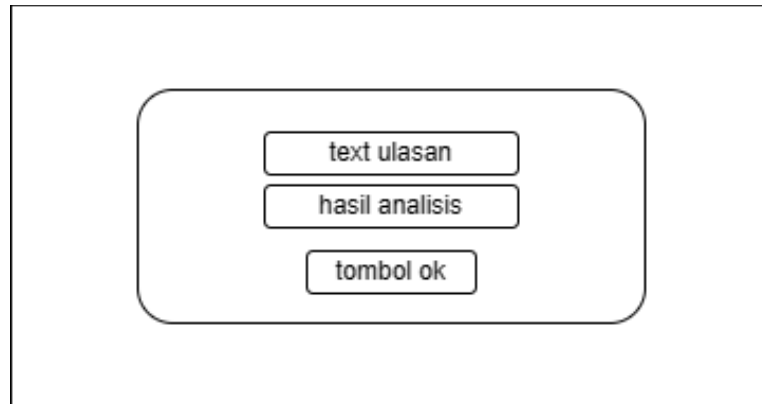


teks sub judul

table history analisis

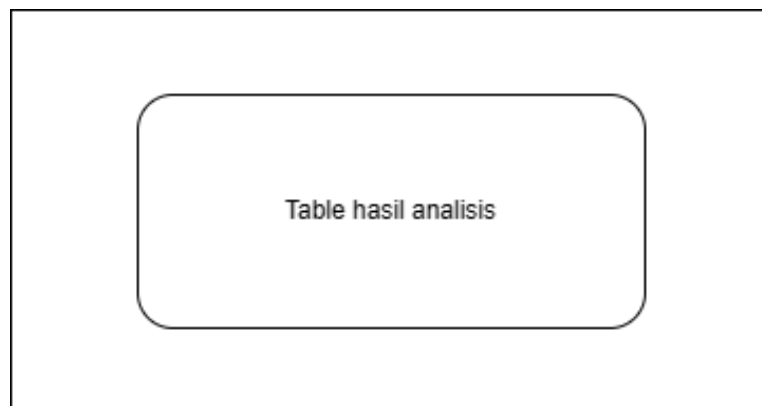
Gambar 4.12 Rancangan Layar History Analisis

e. Keluaran Analisis Sentimen



Gambar 4.13 Rancangan Layar Keluaran Analaisis Sentimen

f. Keluaran Analisis file CSV multi ulasan



**Gambar 4.14 Rancangan Layar Keluaran Analisis file CSV multi
ulasan**

g. Cetak Tabel Laporan



Gambar 4.15 Rancangan Layar Cetak Tabel Laporan

3. Tampilan Layar

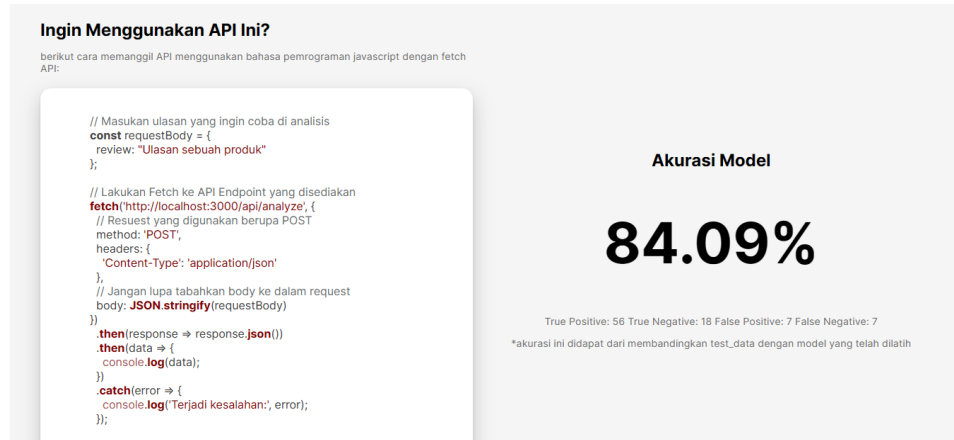
a. Halaman Utama



Gambar 4.16 Tampilan Layar Halaman Utama

Gambar 4.16 merupakan tampilan halaman utama dimana menampilkan judul sebuah program dan penjelasan singkat. Halaman Utama juga memiliki kolom input yang dapat melakukan analisis sentimen terhadap satu kalimat yang pengguna masukan.

b. Cara Pakai API dan Akurasi Model



Gambar 4.17 Tampilan Layar Cara Pakai API dan Akurasi Model

Gambar 4.17 menampilkan dua hal, yaitu cara pakai API apabila pengguna ingin melakukan request API untuk menggunakan servis analisis sentimen ini di dalam program yang mereka buat. Serta tampilan layer ini juga menampilkan informasi terkait hasil akurasi dari model analisis sentimen yang telah penulis buat.

c. Upload file CSV untuk multi ulasan



Gambar 4.18 Tampilan Layar Upload File CSV Multi Ulasan

Gambar 4.18 menampilkan halaman dimana pengguna dapat melakukan upload file dengan format CSV yang berisi banyak ulasan produk sekaligus, namun dengan syarat file CSV harus memiliki header berupa “ulasan”

d. History Analisis

History Analisis			
Show <input type="text" value="10"/> entries		Search: <input type="text"/>	
No	Ulasan	Sentiment	Tanggal dibuat
1	hati hati terkena tipu	negative	2023-08-01
2	seller fast respon	positive	2023-08-01
3	seller ramah, saya suka	positive	2023-08-01
4	jangan dibeli	negative	2023-08-01
5	barang yang sangat buruk	negative	2023-08-01
6	barang ini sangat bagus	positive	2023-08-01
7	Produk ini tidak memberikan kemudahan dalam proses penginstalan atau penggunaan awal. Membutuhkan konfigurasi yang rumit.	negative	2023-08-01
8	Kualitas tali atau pegangan pada produk ini buruk. Mudah putus atau tidak nyaman digunakan.	negative	2023-08-01
9	jangan dibeli produk jelek	negative	2023-08-01
10	tidak rekomended untuk dibeli	positive	2023-08-01

Showing 1 to 10 of 77 entries

Previous **1** 2 3 4 5 ... 8 Next

Gambar 4.19 Tampilan Layar History Analisis

Gambar 4.19 menampilkan history dari semua analisis yang pernah pengguna lakukan, baik itu melalui halaman utama dengan cara satu persatu, maupun melalui upload file CSV, semua history analisis akan tersimpan di tabel ini.

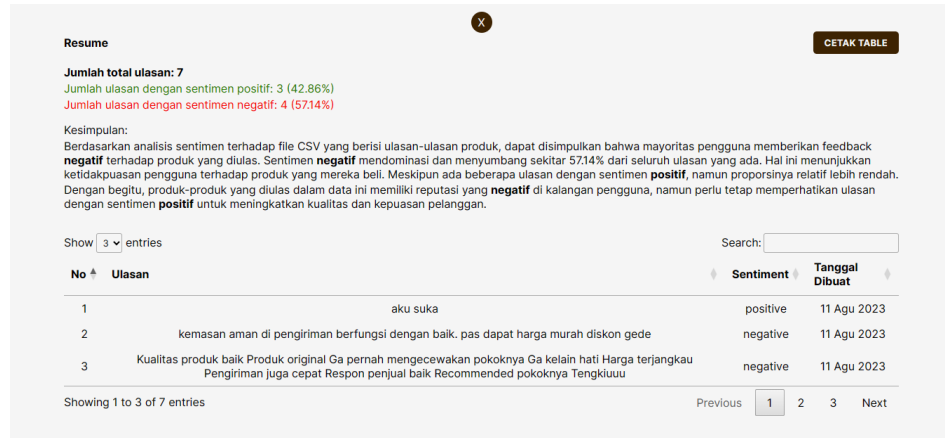
e. Keluaran Analisis Sentimen



Gambar 4.20 Tampilan Layar Keluaran Analisis Sentimen

Gambar 4.20 menampilkan hasil dari analisis sentiment satu kalimat yang pengguna masukan melalui kolom input halaman utama, menu *pop up* akan menampilkan *preview* ulasan yang dimasukan oleh pengguna serta hasil analisis yang berupa sentiment dari kalimat *input*.

f. Keluaran Analisis file CSV multi ulasan



Resume

Jumlah total ulasan: 7
 Jumlah ulasan dengan sentimen positif: 3 (42,86%)
 Jumlah ulasan dengan sentimen negatif: 4 (57,14%)

Kesimpulan:
 Berdasarkan analisis sentimen terhadap file CSV yang berisi ulasan-ulasan produk, dapat disimpulkan bahwa mayoritas pengguna memberikan feedback **negatif** terhadap produk yang diulas. Sentimen **negatif** mendominasi dan menyumbang sekitar 57,14% dari seluruh ulasan yang ada. Hal ini menunjukkan ketidakpuasan pengguna terhadap produk yang mereka beli. Meskipun ada beberapa ulasan dengan sentimen **positif**, namun proporsinya relatif lebih rendah. Dengan begitu, produk-produk yang diulas dalam data ini memiliki reputasi yang **negatif** di kalangan pengguna, namun perlu tetap memperhatikan ulasan dengan sentimen **positif** untuk meningkatkan kualitas dan kepuasan pelanggan.

Show entries

Search:

No	Ulasan	Sentiment	Tanggal Dibuat
1	aku suka	positive	11 Agu 2023
2	kemasan aman di pengiriman berfungsi dengan baik. pas dapat harga murah diskon gede	negative	11 Agu 2023
3	Kualitas produk baik Produk original Ga pernah mengecewakan pokoknya Ga kelain hati Harga terjangkau Pengiriman juga cepat Respon penjual baik Recommended pokoknya Tengkiuuu	negative	11 Agu 2023

Showing 1 to 3 of 7 entries

Previous 2 3 Next

CETAK TABLE

Gambar 4.21 Tampilan Layer Keluaran Analisis file CSV Multi Ulasan

Gambar 4.21 menampilkan hasil dari analisis sentimen menggunakan file CSV, dimana sistem melakukan analisis terhadap banyak ulasan sekaligus. Setelah sistem selesai melakukan ulasan, tampilan layer akan menampilkan *resume* dari hasil ulasan yang dilakukan pengguna, berupa jumlah total ulasan yang dianalisis, jumlah ulasan yang bernilai positif, jumlah ulasan yang bernilai negatif, serta kesimpulan dari analisis tersebut. Tampilan layer ini juga menampilkan tabel yang berisi ulasan dalam file, periode waktu saat ulasan dibuat, serta hasil analisis dari ulasan tersebut berupa sentiment.

Tampilan layer ini juga memiliki tombol “Cetak Tabel” yang berfungsi untuk menampilkan laporan berisi tabel dari analisis sentimen yang telah dilakukan.

g. Cetak Table Laporan

LAPORAN TABEL ANALISIS

No	Ulasan	Sentimen	Tanggal Dibuat
1	aku suka	positive	Jumat, 11 Agustus 2023
2	kemasan aman di pengiriman berfungsi dengan baik. pas dapat harga murah diskon gede	negative	Jumat, 11 Agustus 2023
3	Kualitas produk baik Produk original Ga pernah mengecewakan pokoknya Ga kelain hati Harga terjangkau Pengiriman juga cepat Respon penjual baik Recommended pokoknya Tengkuau	negative	Jumat, 11 Agustus 2023
4	Kurang remote 1 sudah request untuk dikirim remote yang kurangnya tapi malah maunya diretur padahal saya sudah perlu pake. jadi ya sudahlah semoga awet.	negative	Jumat, 11 Agustus 2023
5	barangnya sangat jelek tidak sesuai foto	negative	Jumat, 11 Agustus 2023
6	barang lumayan bagus dengan harga dibawah 100 ribu	positive	Jumat, 11 Agustus 2023
7	WAJIB BELI	positive	Jumat, 11 Agustus 2023

Jakarta, Jumat, 11 Agustus 2023

Admin

Assami Muzaki

Gambar 4.22 Tampilan Layar Cetak Table Laporan

Gambar 4.22 menampilkan hasil cetak tabel yang berupa laporan tabel analisis yang dilakukan, dimana pengguna dapat melihat seluruh hasil analisis terhadap semua ulasan yang mereka masukan menggunakan file CSV.

D. Kelebihan dan Kelemahan Penelitian

1. Kelebihan Penelitian

a. Pengembangan API dan website

Membangun API dan website sebagai hasil dari penelitian ini memiliki nilai praktis yang tinggi. Dengan penelitian ini, aplikasi pihak ketiga

atau user dapat mengintegrasikan fitur analisis sentimen ke dalam produk atau layanan mereka.

b. Penerapan Node.js

Menggunakan Node.js untuk membangun sistem ini memiliki keuntungan dalam hal skalabilitas dan kinerja. Node.js memiliki arsitektur yang non-blok atau *asynchronous*, dan dapat menangani banyak permintaan secara bersamaan, hal ini sangat cocok untuk membangun API serta sistem berbasis web yang responsif dan efisien.

c. Penggunaan Metode *Naive Bayes*

Penggunaan metode *Naive Bayes* dalam analisis sentimen telah terbukti efektif dalam banyak kasus. Metode ini memiliki kinerja yang baik dalam mengklasifikasikan teks berdasarkan sentimen positif, atau negatif. Kelebihan metode ini adalah kemampuannya dalam menangani data teks yang kompleks dengan efisien.

2. Kelemahan Penelitian

a. Keterbatasan data

Penelitian ini mungkin menghadapi keterbatasan dalam hal ketersediaan data ulasan produk. Data yang tidak memadai atau tidak representatif dapat mempengaruhi akurasi dan generalisasi hasil analisis sentimen.

b. Keterbatasan metode *Naive Bayes*

Walaupun metode *Naive Bayes* efektif dalam banyak kasus, ia juga memiliki beberapa kelemahan. *Naive Bayes* cenderung mengabaikan urutan kata dalam teks, sehingga mungkin tidak mampu menangkap konteks yang lebih kompleks. Serta berbagai kekurangan lainnya seperti asumsi independensi yang kuat

BAB V

SIMPULAN DAN SARAN

A. Simpulan

Melalui penelitian ini, penulis berhasil mengatasi kompleksitas dan kebutuhan waktu yang tinggi dengan mengimplementasikan metode analisis sentimen otomatis. Dalam pengembangan metode ini, penulis memanfaatkan pendekatan algoritma klasifikasi Naive Bayes. Pendekatan ini memungkinkan penulis untuk mengolah volume besar data ulasan produk secara efisien dan akurat, sehingga pengambilan keputusan dapat dilakukan dengan lebih cepat dan tepat.

Metode otomatis yang dikembangkan pada penelitian ini membuktikan keandalannya dalam menganalisis sentimen ulasan produk pada data e-commerce. Dengan menggunakan pendekatan Naive Bayes, penulis berhasil memprediksi sentimen ulasan produk baik yang sudah ada maupun secara otomatis. Berdasarkan hasil pengujian, metode ini memiliki tingkat akurasi yang cukup memuaskan dalam mengklasifikasikan sentimen ulasan menjadi positif atau negatif.

B. Saran

Sebagai penutup dari skripsi ini, kami hendaknya menyarankan agar hasil penelitian ini dapat memberikan manfaat yang berkelanjutan. Sebaiknya, peneliti selanjutnya dapat melakukan penelitian lebih lanjut mengenai perbandingan antara metode Naive Bayes dengan metode analisis sentimen lainnya, seperti model berbasis Machine Learning atau Deep Learning. Oleh karena itu, diharapkan

penelitian mendatang dapat memberikan wawasan yang lebih mendalam tentang efektivitas dan efisiensi metode analisis sentimen dalam konteks ulasan produk di platform e-commerce.

Harapan kami, penelitian lebih lanjut juga dapat mempertimbangkan jumlah data yang digunakan dalam penelitian, karena banyaknya jumlah data sangatlah berpengaruh dalam penentu akurasi model. Selain itu, peneliti berikutnya diharapkan dapat mengukur kinerja metode dengan lebih detail melalui evaluasi yang komprehensif, termasuk analisis akurasi, presisi, recall, serta area under curve (AUC) pada kurva ROC.

DAFTAR PUSTAKA

- Aco, A., & Endang, H. (2017). Analisis Bisnis E-Commerce pada Mahasiswa Universitas Islam Negeri Alauddin Makassar. *Jurnal INSYPRO (Information System and Processing)*, 2(1).
- Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *ArXiv Preprint ArXiv:1707.02919*.
- Chatrina, Siregar, N., Ruli, A, Siregar, R., & Yoga, Distra, Sudirman, M. (2020). Implementasi Metode Naive Bayes Classifier (NBC) Pada Komentar Warga Sekolah Mengenai Pelaksanaan Pembelajaran Jarak Jauh (PJJ). *Jurnal Teknologia*, 34(1), 102–110. <https://aperti.e-journal.id/teknologia/article/view/67>
- Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment analysis: a comparative study on different approaches. *Procedia Computer Science*, 87, 44–49.
- Diantika, A. S., & Firmanto, Y. (2019). Implementasi machine learning pada aplikasi penjualan produk digital (studi pada GrabKios). *Jurnal Ilmiah Mahasiswa FEB*, 53(9), 1689–1699. <https://jimfeb.ub.ac.id/index.php/jimfeb/article/view/6928>
- Fajrin, R. (2017). Pengembangan Sistem Informasi Geografis Berbasis Node.JS untuk Pemetaan Mesin dan Tracking Engineer dengan Pemanfaatan Geolocation pada PT IBM Indonesia. *Jurnal Informatika*, 11(2), 40–47. <https://doi.org/10.26555/jifo.v11i2.a6090>
- Haviluddin. (2011). Memahami Penggunaan UML (Unified Modelling Language). *Memahami Penggunaan UML (Unified Modelling Language)*, 6(1), 1–15. <https://informatikamulawarman.files.wordpress.com/2011/10/01-jurnal-informatika-mulawarman-feb-2011.pdf>
- Heliyanti Susana. (2022). Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet. *Jurnal Riset Sistem Informasi Dan Teknologi Informasi (JURSISTEKNI)*, 4(1), 1–8. <https://doi.org/10.52005/jursistekni.v4i1.96>
- Ismail. (2020). Perancangan Sistem Aplikasi Pemesanan Makanan dan Minuman Pada Cafeteria NO Caffe di Tanjung Balai Karimun Menggunakan Bahasa Pemrograman PHP dan MySQL. *Jurnal Tikar*, 1(2), 192–206. https://ejurnal.universitaskarimun.ac.id/index.php/teknik_informatika/article/do

wnload/153/121

- Khoirul Akmal, N., & Dasaprawira, M. N. (2022). RANCANG BANGUN APPLICATION PROGRAMMING INTERFACE (API) MENGGUNAKAN GAYA ARSITEKTUR GraphQL UNTUK PEMBUATAN SISTEM INFORMASI PENDATAAN ANGGOTA UNIT KEGIATAN MAHASISWA (UKM) STUDI KASUS UKM STARLABS. *JURNAL SISTEM INFORMASI DAN TENOLOGI*, 5, 3–10.
- Kurniawan, I., Humaira, & Rozi, F. (2020). REST API Menggunakan NodeJS pada Aplikasi Transaksi Jasa Elektronik Berbasis Android. *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, 1(4), 127–132. <https://doi.org/10.30630/jitsi.1.4.18>
- Luknanto, D. (2001). *Metoda Numerik*. November.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711. <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- Sari, F. V., & Wibowo, A. (2019). ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI. *Jurnal Teknik Industri, Mesin, Elektro Dan Ilmu Komputer SIMETRIS*, Vol 10, No. <https://doi.org/https://doi.org/10.24176/simet.v10i2.3487>
- Setiyani, L. (2021). Desain Sistem : Use Case Diagram. *Prosiding Seminar Nasional : Inovasi & Adopsi Teknologi 2021, September*, 246–260. <https://journal.uui.ac.id/AUTOMATA/article/view/19517>
- Setyohadi, D. B., Kristiawan, F. A., & Ernawati, E. (2017). Perbaikan Performansi Klasifikasi Dengan Preprocessing Iterative Partitioning Filter Algorithm. *Telematika*, 14(01), 12–20. <https://doi.org/10.31315/telematika.v14i01.1960>
- Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *JSI: Jurnal Sistem Informasi (E-Journal)*, 8(1).
- Sonata, F.-. (2019). Pemanfaatan UML (Unified Modeling Language) Dalam Perancangan Sistem Informasi E-Commerce Jenis Customer-To-Customer. *Jurnal Komunika : Jurnal Komunikasi, Media Dan Informatika*, 8(1), 22. <https://doi.org/10.31504/komunika.v8i1.1832>

- Sunoto, Y., & Wasito, B. (2014). Analisis Testimonial Wisatawan Menggunakan Text Mining Dengan Metode Naive Bayes Dan Decision Tree , Studi Kasus Pada Hotel – Hotel Di Jakarta. *Jurnal Informatika Dan Bisnis ANALISIS*, 3(2), 39–49. <https://jurnal.kwikkiangie.ac.id/index.php/JIB/issue/view/26/56>
- Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>
- Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7–16.
- Yanti, S. N., & Rihyanti, E. (2021). Penerapan Rest API untuk Sistem Informasi Film Secara Daring. *Jurnal Informatika Universitas Pamulang*, 6(1), 195. <https://doi.org/10.32493/informatika.v6i1.10033>
- Zhou, Z.-H. (2021). *Machine learning*. Springer Nature.

RIWAYAT PENULIS



Assami Muzaki adalah nama penulis skripsi ini. Lahir pada 18 Oktober 2001, di Kabupaten Bogor, Provinsi Jawa Barat. Penulis merupakan anak bungsu dari pasangan Pak Aji Sukrisno, dan Bu Umi Solihah.

Penulis pertama kali masuk pendidikan di SD Pelita Atsiri Permai pada tahun 2006 dan tamat 2012, pada tahun yang sama penulis melanjutkan pendidikan SMP serta SMA ke MBS Zam-Zam Purwokerto, hingga lulus tahun 2019 dan melanjutkan sebagai Mahasiswa di Universitas Indraprasta PGRI, Fakultas Teknik dan Ilmu Komputer, Jurusan Teknik Informatika.

Alasan utama penulis kuliah di jurusan ini adalah, penulis melihat peluang masa depan yang sangat luas di bidang teknologi informasi.

Akhir kata penulis mengucapkan rasa syukur yang sebesar-besarnya kepada Allah SWT atas terselesaikannya skripsi ini. Terimakasih kepada orang tuadan teman teman yang membantu menyelesaikan proses ini.

LAMPIRAN