

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Analisis Sentimen

Analisis sentimen adalah proses yang digunakan untuk mengevaluasi dan mengidentifikasi emosi atau sikap yang terkandung dalam teks, baik itu positif, negatif, maupun netral. Secara umum, tujuan utama dari analisis sentimen adalah untuk memahami persepsi masyarakat terhadap suatu objek atau entitas melalui data yang tersedia, seperti komentar di media sosial atau ulasan daring. Menurut (Mohamed & Bayraktar, 2022), analisis sentimen berperan penting dalam membantu organisasi memantau opini publik secara real-time dan mengambil langkah strategis yang tepat dalam merespons berbagai isu yang berkembang. Dalam konteks penelitian ini, analisis sentimen digunakan untuk mengevaluasi opini masyarakat terhadap Universitas Darma Persada yang tercermin dari komentar-komentar di Instagram serta ulasan di Google Maps

2.1.1.1 Pentingnya Analisis Sentimen di Media Sosial dan Ulasan Daring

Media sosial seperti Instagram dan platform ulasan seperti Google Maps telah berkembang menjadi saluran utama untuk berbagi opini, pengalaman, serta perasaan terkait berbagai entitas, termasuk lembaga pendidikan seperti universitas. Di Instagram, pengguna sering memberikan pendapat mereka terkait pengalaman pribadi, kualitas layanan, atau bahkan fasilitas yang ada di universitas. Sementara

itu, di Google Maps, pengguna memberikan ulasan tentang berbagai aspek, termasuk lokasi dan pelayanan universitas. Menurut (Matulesy & Purnama, 2023), analisis sentimen terhadap ulasan pengguna secara daring berperan penting dalam memahami persepsi publik terhadap suatu layanan atau entitas. Dalam konteks penelitian ini, analisis sentimen digunakan untuk mengevaluasi opini masyarakat terhadap Universitas Darma Persada yang tercermin dari komentar-komentar di Instagram dan ulasan di Google Maps. Dengan memahami opini tersebut, universitas dapat mengidentifikasi kekuatan dan kelemahan yang ada serta merancang strategi untuk meningkatkan citra dan reputasinya di mata publik.

2.1.1.2 Teknik dalam Analisis Sentimen

Ada beberapa teknik yang dapat digunakan untuk melakukan analisis sentimen, yang sebagian besar berfokus pada penerapan algoritma machine learning. Salah satu teknik yang paling populer adalah Naive Bayes, yang merupakan algoritma probabilistik yang mengklasifikasikan teks ke dalam kategori sentimen berdasarkan frekuensi kata-kata yang muncul. Menurut (Fadila & Ikhsan, n.d.), algoritma Naive Bayes efektif dalam mengklasifikasikan sentimen ulasan pengguna terhadap tempat wisata di Google Maps, memungkinkan identifikasi opini positif, negatif, atau netral secara akurat. Dalam konteks penelitian ini, pendekatan serupa digunakan untuk menganalisis opini masyarakat terhadap Universitas Darma Persada berdasarkan komentar-komentar di Instagram dan ulasan Google Maps. Selain Naive Bayes, teknik lain seperti Support Vector Machines (SVM) dan Deep Learning juga dapat digunakan, namun Naive Bayes

tetap menjadi pilihan yang paling efisien dalam aplikasi yang membutuhkan pemrosesan cepat dan hasil yang dapat diandalkan.

2.1.2 Machine Learning dalam Analisis Sentimen

Machine Learning (ML) adalah cabang dari kecerdasan buatan yang memungkinkan sistem untuk belajar dari data dan membuat prediksi atau keputusan berdasarkan pola yang ditemukan dalam data. Dalam konteks analisis sentimen, machine learning digunakan untuk mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya, apakah itu positif, negatif, atau netral. Menurut (Nikhil Sanjay Suryawanshi, 2024), algoritma *machine learning* seperti *Naive Bayes*, *Support Vector Machines* (SVM), dan *Deep Learning* sering digunakan dalam analisis sentimen karena kemampuan mereka dalam mengklasifikasikan teks besar secara efisien. *Naive Bayes* dikenal karena kesederhanaannya dan efektivitasnya dalam menangani data teks besar. Dalam konteks analisis sentimen terhadap Universitas Darma Persada, penggunaan machine learning membantu dalam memproses dan mengklasifikasikan komentar-komentar Instagram dan ulasan di Google Maps secara otomatis. Untuk analisis sentimen terhadap Universitas Darma Persada, penggunaan machine learning membantu dalam memproses dan mengklasifikasikan komentar-komentar Instagram dan ulasan di Google Maps secara otomatis.

2.1.3 Data Mining dalam Analisis Sentimen

Data mining adalah proses menemukan pola atau informasi yang berguna dalam kumpulan data besar. Dalam konteks analisis sentimen, data mining digunakan untuk mengekstraksi sentimen dari data teks yang tidak terstruktur, seperti komentar di media sosial atau ulasan di platform ulasan daring. Menurut (Norlaila et al., 2024), data mining merupakan proses penting dalam mengekstraksi pola dan informasi dari kumpulan data besar yang tidak terstruktur, seperti komentar di media sosial dan ulasan daring. Dalam analisis sentimen, teknik data mining seperti clustering dan klasifikasi efektif dalam mengidentifikasi opini positif, negatif, dan netral, sehingga membantu dalam menganalisis opini publik secara efektif.

2.1.4 Prediksi dalam Analisis Sentimen

Prediksi dalam analisis sentimen merujuk pada kemampuan model untuk memproyeksikan sentimen yang terkandung dalam data baru berdasarkan pelatihan yang telah dilakukan. Dalam hal ini, model machine learning yang dilatih dengan data ulasan dari Instagram dan Google Maps dapat digunakan untuk memprediksi sentimen terhadap Universitas Darma Persada di masa depan. Prediksi ini sangat penting dalam memberikan wawasan bagi pihak universitas untuk memahami perubahan persepsi masyarakat terhadap mereka seiring waktu. Algoritma seperti Naive Bayes digunakan untuk memprediksi apakah ulasan baru yang masuk mengandung sentimen positif, negatif, atau netral terhadap universitas tersebut.

2.1.5 Metode Naive Bayes dalam Analisis Sentimen

Naive Bayes adalah salah satu metode yang banyak digunakan dalam analisis sentimen, terutama pada data teks. Algoritma ini bekerja dengan mengasumsikan bahwa fitur-fitur yang ada bersifat independen, yang memudahkan perhitungan probabilitas untuk menentukan kelas sentimen dalam suatu teks. Proses ini melibatkan pelatihan model dengan data ulasan yang telah diberi label sentimen, dan kemudian model ini dapat digunakan untuk mengklasifikasikan ulasan atau komentar baru ke dalam kategori sentimen yang sesuai. Metode Naive Bayes merupakan algoritma yang sering digunakan dalam analisis sentimen karena kesederhanaannya dan kemampuannya dalam mengklasifikasikan data teks dengan efisien. Algoritma ini mengasumsikan fitur-fitur independen dan menggunakan probabilitas kondisional untuk menentukan kelas sentimen suatu teks. Dengan kecepatan komputasi yang tinggi, Naive Bayes sangat cocok untuk mengolah data komentar dari media sosial dan platform ulasan online (Norlaila et al., 2024).

2.1.6 CRISP-DM (Cross Industry Standard Process for Data Mining)

CRISP-DM adalah metodologi standar yang digunakan untuk mengembangkan solusi data mining. CRISP-DM terdiri dari enam tahap, yakni: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Dalam konteks analisis sentimen, CRISP-DM dapat digunakan untuk mengorganisir proses pengumpulan, pembersihan, analisis, dan pemodelan data ulasan di Instagram dan Google Maps. Tahap pertama, pemahaman bisnis, akan berfokus pada tujuan dari analisis sentimen terhadap Universitas Darma Persada.

Menurut (Laturiuw & Singgalen, 2023), metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining) terdiri dari enam tahap utama: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Dalam konteks analisis sentimen terhadap destinasi wisata Raja Ampat, CRISP-DM digunakan untuk mengorganisir proses pengumpulan, pembersihan, analisis, dan pemodelan data ulasan dari Tripadvisor. Tahap pertama, pemahaman bisnis, berfokus pada tujuan dari analisis sentimen, sementara tahap pemahaman data melibatkan eksplorasi data teks untuk menemukan pola yang relevan, yang kemudian diproses dan dimodelkan menggunakan algoritma machine learning seperti Naive Bayes (NBC), Support Vector Machine (SVM), Decision Tree (DT), dan k-Nearest Neighbor (k-NN).

2.1.7 Pemodelan UML

Unified Modeling Language (UML) adalah bahasa pemodelan standar yang digunakan untuk membantu proses pengembangan perangkat lunak, khususnya dalam tahap analisis dan desain. UML memungkinkan pengembang untuk merepresentasikan sistem secara visual melalui berbagai jenis diagram yang mendeskripsikan struktur dan perilaku sistem.

Menurut (John, 2025), diagram Unified Modeling Language (UML) memainkan peran penting dalam pengembangan perangkat lunak, terutama pada proyek berskala menengah. Diagram seperti use case, class, sequence, dan activity membantu dalam memvisualisasikan struktur dan perilaku sistem, memfasilitasi komunikasi antar pemangku kepentingan, serta meningkatkan pemahaman dan

dokumentasi sistem. Studi ini menekankan pentingnya penerapan UML untuk mengelola kompleksitas, meningkatkan kolaborasi, dan memastikan pengiriman solusi perangkat lunak yang sukses.

Menurut (Deshmukh & Kumar, n.d.), Unified Modeling Language (UML) memainkan peran penting dalam meningkatkan efisiensi pemeliharaan perangkat lunak. Diagram UML, seperti diagram kelas dan urutan, membantu dalam pemahaman kode, mengurangi waktu pemeliharaan, dan memfasilitasi kolaborasi antar pengembang. Studi ini menunjukkan bahwa proyek dengan dokumentasi UML yang terkini mengalami pengurangan waktu pemeliharaan hingga 30% dibandingkan dengan yang tidak memiliki dokumentasi UML yang memadai.

2.1.7.1 Use Case


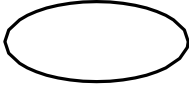
Use Case Diagram digunakan untuk menggambarkan interaksi antara aktor (pengguna atau sistem lain) dengan sistem yang dikembangkan. Diagram ini menampilkan skenario bagaimana pengguna berinteraksi untuk mencapai suatu tujuan tertentu dalam sistem.



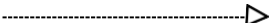
“Use case diagram digunakan untuk menunjukkan fungsi-fungsi utama sistem dan interaksinya dengan aktor, yang sangat membantu dalam mengidentifikasi kebutuhan sistem dari perspektif pengguna” (Refianti & Anggraeni, 2023).

Dalam penelitian ini, use case diagram memodelkan interaksi antara pengguna sistem (admin, mahasiswa, dan publik) dengan fitur-fitur utama seperti

mengambil data komentar dari Instagram, ulasan Google Maps, melakukan preprocessing, klasifikasi sentimen menggunakan Naive Bayes, serta menghasilkan laporan sentimen terhadap Universitas Darma Persada. Berikut adalah komponen yang terdapat pada Use Case Diagram:

Tabel 2. 1 Komponen Usecase Diagram

Nama Komponen	Kegunaan / Fungsi	Notasi / Simbol
<i>Actor</i>	Untuk mewakili setiap pengguna atau sistem eksternal yang berinteraksi langsung dengan sistem utama. <i>Actor</i> berfungsi sebagai titik awal dalam mendefinisikan bagaimana dan kapan suatu fungsi atau proses harus dijalankan.	
<i>Use Case</i>	Berfungsi untuk mendefinisikan dan menggambarkan interaksi antara pengguna (<i>actor</i>) dan sistem yang akan dirancang, serta bagaimana sistem	

	tersebut akan memenuhi kebutuhan pengguna.	
<i>Association</i>	Berfungsi untuk menggambarkan hubungan antara <i>actor</i> dan <i>use case</i> , menunjukkan interaksi yang terjadi antara pengguna atau sistem eksternal dengan sistem yang sedang dikembangkan.	
<i>Generalization atau inheritance (pewarisan)</i>	Berfungsi untuk menggambarkan hubungan hierarkis antara elemen yang lebih umum dan elemen yang lebih spesifik.	
<i>Dependency</i>	Berfungsi untuk menggambarkan hubungan antara dua elemen di mana	

	<p>satu elemen (misalnya, <i>use case</i> atau <i>actor</i>) bergantung atau membutuhkan elemen lain untuk berfungsi dengan baik.</p>	
<i>Include</i>	<p>Berfungsi untuk menunjukkan bahwa satu <i>use case</i> selalu melibatkan atau menyertakan eksekusi <i>use case</i> lainnya sebagai bagian dari alurnya.</p>	<<include>>
<i>Extend</i>	<p>Berfungsi untuk menunjukkan bahwa sebuah <i>use case</i> dapat diperluas atau ditambah dengan fungsionalitas tambahan dalam kondisi tertentu.</p>	<<extend>>
<i>Communicates</i>	<p>Berfungsi untuk menunjukkan interaksi atau pertukaran informasi antara dua elemen, seperti antara dua <i>actor</i> atau antara <i>actor</i> dan sistem.</p>	<<communicates>>


2.1.7.2 Activity



Activity Diagram digunakan untuk memodelkan aliran aktivitas atau proses bisnis dalam sistem, yang menggambarkan langkah-langkah logis dari suatu proses kerja. Diagram ini sangat berguna untuk menunjukkan alur proses dari awal hingga akhir, termasuk pengambilan keputusan dan aktivitas paralel.




“Activity diagram merupakan representasi grafis yang menggambarkan alur kerja aktivitas secara berurutan dalam sebuah proses, termasuk pengambilan keputusan dan aktivitas yang berjalan secara paralel, sehingga memudahkan pemahaman alur sistem secara menyeluruh” (Umesh Awate et al., 2024).

Dalam konteks penelitian ini, activity diagram digunakan untuk memodelkan proses analisis sentimen, mulai dari pengambilan data, pembersihan teks, ekstraksi fitur, klasifikasi dengan metode Naive Bayes, hingga penyimpanan hasil sentimen ke dalam sistem. Berikut komponen yang terdapat dalam Activity Diagram:

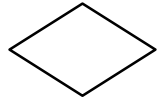

Tabel 2. 2 Komponen Activity Diagram

Nama Komponen	Kegunaan / Fungsi	Notasi / Simbol
<i>Start State</i>	Berfungsi untuk menunjukkan keadaan awal dalam sebuah diagram status (<i>state diagram</i>) yang menggambarkan titik	

	awal dari proses atau alur dalam sistem.	
<i>State</i>	Berfungsi untuk menggambarkan kondisi atau situasi tertentu dalam siklus hidup suatu objek atau entitas dalam sistem, yang dapat berubah berdasarkan interaksi atau peristiwa tertentu.	
<i>Activity</i>	Berfungsi untuk menggambarkan kondisi atau situasi tertentu dalam siklus hidup suatu objek atau entitas dalam sistem, yang dapat berubah berdasarkan interaksi atau peristiwa tertentu.	

<p><i>State Transition</i></p>	<p>Berfungsi untuk menggambarkan perubahan dari satu <i>state</i> (keadaan) ke <i>state</i> lainnya dalam sistem, berdasarkan peristiwa atau kondisi tertentu.</p>	
<p><i>Transition to Self</i></p>	<p>Berfungsi untuk menggambarkan perubahan atau transisi dari satu <i>state</i> ke <i>state</i> yang sama dalam diagram status (<i>state diagram</i>).</p>	
<p><i>Horizontal Synchronization</i></p>	<p>Biasanya merujuk pada koordinasi atau penyelarasan elemen atau proses dalam satu baris atau jalur yang sejajar secara horizontal. Dalam konteks sistem atau diagram, ini bisa</p>	

	<p>digunakan untuk menggambarkan pengaturan waktu atau alur yang terjadi secara bersamaan atau sejajar, namun dengan fokus pada arah horizontal.</p>	
<p><i>Vertical Synchronization</i></p>	<p>Berfungsi untuk menggambarkan sinkronisasi alur aktivitas yang terjadi secara bersamaan dalam urutan vertikal. Hal ini biasanya terjadi ketika ada beberapa aktivitas yang berjalan paralel dan perlu disinkronkan untuk memastikan bahwa semua jalur aktivitas selesai sebelum melanjutkan ke langkah berikutnya dalam alur.</p>	

<i>Decision</i>	Berfungsi untuk menggambarkan titik di mana suatu alur aktivitas bercabang berdasarkan kondisi tertentu.	
<i>End State</i>	Berfungsi untuk menunjukkan titik akhir dari suatu alur aktivitas atau proses dalam sistem.	

2.1.7.3 Sequence Diagram

Sequence diagram adalah jenis diagram dalam UML yang menggambarkan interaksi antar objek atau komponen dalam sistem berdasarkan waktu. Diagram ini menggambarkan bagaimana objek berkomunikasi satu sama lain dan urutan pesan yang dikirimkan untuk menjalankan sebuah fungsi atau proses dalam sistem.

"Sequence diagram merupakan diagram UML kedua yang paling umum digunakan dan berfungsi untuk merepresentasikan bagaimana objek saling berinteraksi serta bertukar pesan seiring waktu". (Al-Fedaghi, 2021)

Pada umumnya, sequence diagram digunakan untuk memodelkan aliran interaksi antara aktor dan sistem, serta antar komponen dalam sistem yang lebih

besar. Diagram ini sangat berguna untuk memvisualisasikan proses yang melibatkan banyak entitas dengan urutan yang kompleks.

2.1.8 Database dan DBMS

Database adalah kumpulan data yang saling berhubungan dan dikelola dengan cara yang terstruktur untuk memungkinkan penyimpanan, pencarian, dan manipulasi data secara efisien. Menurut (Purnaningsih & Yulianto, 2022), "Database merupakan kumpulan data yang saling berhubungan yang disimpan secara sistematis untuk memenuhi kebutuhan informasi suatu organisasi."

Database Management System (DBMS) adalah perangkat lunak yang digunakan untuk mengelola database dan menyediakan antarmuka untuk mengakses data dengan cara yang aman, efisien, dan terstruktur. Menurut (Kalsum Siregar et al., 2024), "Database Management System (DBMS) adalah perangkat lunak yang digunakan untuk mengelola dan mengatur data dalam suatu basis data."

2.1.9 Software Dan Pemrograman Terkait

Dalam pengembangan aplikasi analisis sentimen, penggunaan perangkat lunak seperti Python dengan library Natural Language Toolkit (NLTK) dan algoritma Naïve Bayes terbukti efektif dalam mengolah data teks secara otomatis. Natural Language Toolkit merupakan toolkit untuk pemrosesan bahasa alami yang dibangun dengan Python, sementara Naïve Bayes dikenal sebagai algoritma yang sederhana namun kuat untuk klasifikasi sentimen. Kombinasi alat ini telah berhasil diterapkan dalam analisis ulasan pengguna Google Maps, seperti yang ditunjukkan

dalam studi yang dilakukan oleh Budiman, Lestanti, dan Erwan (2024) yang menggunakan NLTK, TextBlob, dan Naïve Bayes dalam menganalisis sentimen pelanggan terhadap sebuah kafe (Saiful Nur Budiman et al., 2024)

2.1.9.1 Bahasa Pemrograman Python dan Library Terkait

Python adalah bahasa pemrograman yang digunakan dalam proyek ini untuk menganalisis data besar, melakukan pemrosesan bahasa alami, dan membangun model machine learning. Python dikenal karena sintaksisnya yang sederhana dan banyaknya pustaka analisis data yang tersedia.

Menurut (Aisyah et al., n.d.), "Python menjadi bahasa pemrograman utama dalam proses pengolahan data sentimen karena menyediakan berbagai pustaka yang lengkap untuk analisis statistik, pemrosesan bahasa alami, hingga visualisasi data. Dalam penelitian ini digunakan Jupyter Notebook sebagai editor berbasis Python dengan dukungan pustaka Pandas untuk manipulasi data, NLTK dan Sastrawi untuk pemrosesan bahasa alami, serta Matplotlib dan Seaborn untuk visualisasi." Python memungkinkan pengolahan data dilakukan secara efisien dan fleksibel, serta mendukung pembangunan model klasifikasi berbasis machine learning seperti Naïve Bayes untuk analisis sentimen.

2.1.9.2 Library Scikit-learn

Scikit-learn adalah pustaka Python yang menyediakan berbagai algoritma untuk machine learning dan analisis data. Dalam proyek ini, Scikit-learn digunakan untuk implementasi metode Naive Bayes, yang merupakan teknik

klasifikasi yang efektif dalam analisis sentimen. Pustaka ini juga menyediakan alat untuk evaluasi model, seperti confusion matrix, yang berguna untuk menilai akurasi dan performa model.

Dalam studi ini, digunakan library Scikit-learn (sklearn) untuk menerapkan algoritma machine learning. Scikit-learn adalah pustaka machine learning berbasis Python yang menyediakan berbagai algoritma dan tools untuk klasifikasi, regresi, klustering, reduksi dimensi, pemilihan fitur, dan praproses data (Handhayani, 2025). Dalam konteks penelitian ini, sklearn digunakan khususnya untuk membangun dan mengevaluasi model klasifikasi secara efisien dan terstruktur. Penelitian ini menggunakan algoritma Naïve Bayes, lebih spesifiknya Multinomial Naïve Bayes, yang merupakan salah satu algoritma populer untuk klasifikasi teks.

2.2.1 Paper 1

Penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara, Oktariani Nurul Pratiwi, dan Hilman Dwi Anggana (2021) berjudul “*Analisis Sentimen Review Customer Terhadap Layanan Ekspedisi JNE dan J&T Express Menggunakan Metode Naïve Bayes*” yang dipublikasikan dalam *e-Proceedings of Engineering* ini membahas penggunaan algoritma Naïve Bayes dalam mengklasifikasikan sentimen pelanggan terhadap layanan dua ekspedisi besar di Indonesia, yaitu JNE dan J&T Express, berdasarkan ulasan dari platform digital. Latar belakang penelitian ini menunjukkan bahwa meningkatnya aktivitas belanja online juga mendorong masyarakat memberikan review secara digital terhadap jasa pengiriman yang digunakan. Oleh karena itu, penelitian ini bertujuan untuk

mengolah dan mengklasifikasikan sentimen positif, negatif, dan netral dalam review pelanggan menggunakan pendekatan machine learning, khususnya algoritma Naïve Bayes.

Penelitian ini sangat relevan dengan skripsi ini karena menggunakan algoritma Naïve Bayes untuk analisis sentimen berbasis teks, yang merupakan pendekatan yang juga digunakan dalam skripsi ini. Meskipun sumber data yang digunakan berupa ulasan teks dari platform marketplace atau platform digital lainnya (bukan media sosial seperti Instagram), konteks dan tujuannya sama, yaitu untuk mengetahui persepsi publik terhadap suatu layanan menggunakan pemrosesan bahasa alami (NLP).

Dengan demikian, penelitian ini menjadi referensi penting dalam menunjukkan bagaimana algoritma Naïve Bayes dapat diterapkan secara efektif untuk menganalisis opini masyarakat secara otomatis, serta memberikan hasil klasifikasi sentimen yang akurat dalam mendukung evaluasi layanan publik berdasarkan review digital.

2.2.1.1 Tujuan Penelitian

Penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara, Oktariani Nurul Pratiwi, dan Hilman Dwi Anggana (2021) bertujuan untuk menganalisis sentimen ulasan pelanggan terhadap layanan ekspedisi JNE dan J&T Express. Fokus utama dari penelitian ini adalah untuk mengklasifikasikan ulasan tersebut ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, dengan menggunakan metode Naïve Bayes. Tujuan dari klasifikasi ini adalah untuk

mengetahui persepsi masyarakat terhadap kualitas layanan yang diberikan oleh kedua jasa ekspedisi tersebut. Selain itu, hasil dari penelitian ini diharapkan dapat memberikan masukan bagi perusahaan dalam mengevaluasi dan meningkatkan kualitas layanannya berdasarkan opini yang diberikan oleh pelanggan. Penelitian ini juga bertujuan untuk menunjukkan penerapan algoritma Naïve Bayes secara efektif dalam analisis sentimen berbasis teks pada data ulasan digital.

2.2.1.2 Metodologi Yang Digunakan

Metodologi yang digunakan dalam penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara et al. (2021) melibatkan beberapa tahapan utama dalam proses analisis sentimen. Tahapan dimulai dengan pengumpulan data ulasan pelanggan terhadap layanan ekspedisi JNE dan J&T Express yang diperoleh dari platform digital. Data yang terkumpul kemudian melalui proses pra-pemrosesan teks, yang mencakup pembersihan data (cleaning), tokenisasi, penghapusan stopword, dan stemming untuk mereduksi kata ke bentuk dasarnya. Setelah data siap, tahap selanjutnya adalah pelabelan data ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Untuk proses klasifikasi, penelitian ini menggunakan algoritma Naïve Bayes yang dikenal efektif dalam pengolahan data teks dan analisis sentimen. Evaluasi kinerja model dilakukan menggunakan metode confusion matrix untuk mengukur tingkat akurasi klasifikasi yang dihasilkan oleh algoritma. Secara keseluruhan, pendekatan metodologi ini bertujuan untuk menghasilkan klasifikasi sentimen yang akurat dan dapat dijadikan dasar dalam mengevaluasi persepsi pelanggan terhadap layanan ekspedisi.

2.2.1.3 Temuan Utama

Temuan utama dari penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara et al. (2021) menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen ulasan pelanggan terhadap layanan JNE dan J&T Express dengan tingkat akurasi yang cukup baik. Berdasarkan hasil pengujian menggunakan confusion matrix, model klasifikasi yang dibangun berhasil mengidentifikasi sentimen positif, negatif, dan netral secara cukup akurat. Sebagian besar ulasan pelanggan terhadap kedua layanan ekspedisi tersebut cenderung bernada positif, meskipun terdapat pula sentimen negatif yang mengarah pada keluhan terhadap keterlambatan pengiriman, kondisi paket, serta pelayanan kurir. Penelitian ini juga menemukan bahwa metode Naïve Bayes cukup efektif dalam mengolah data ulasan berbahasa Indonesia, dengan hasil yang dapat dijadikan acuan dalam evaluasi kualitas layanan oleh pihak penyedia jasa ekspedisi. Temuan ini menegaskan bahwa analisis sentimen berbasis teks dapat menjadi alat bantu yang penting dalam memahami opini publik secara otomatis dan efisien.

2.2.1.4 Kesimpulan Penelitian

Penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara et al. (2021) menyimpulkan bahwa algoritma Naïve Bayes merupakan metode yang efektif dan akurat dalam melakukan klasifikasi sentimen terhadap ulasan pelanggan mengenai layanan ekspedisi JNE dan J&T Express. Dengan menerapkan tahapan pra-pemrosesan data dan klasifikasi yang sistematis, penelitian ini berhasil mengelompokkan ulasan menjadi tiga kategori sentimen utama, yaitu positif,

negatif, dan netral. Hasil analisis menunjukkan bahwa sebagian besar pelanggan memberikan ulasan positif, namun juga ditemukan sejumlah kritik yang mencerminkan ketidakpuasan terhadap aspek tertentu dalam pelayanan. Kesimpulan utama dari penelitian ini menekankan bahwa analisis sentimen berbasis algoritma machine learning, khususnya Naïve Bayes, dapat dimanfaatkan sebagai alat bantu evaluasi pelayanan yang berbasis data digital. Temuan ini memberikan kontribusi penting bagi perusahaan jasa ekspedisi dalam meningkatkan kualitas layanannya berdasarkan masukan dan opini pelanggan secara real-time.

2.2.2 Paper 2

Penelitian oleh Mahardika Tania Nitami dan Herny Februariyanti (2022) yang berjudul *Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma Naïve Bayes* mengkaji opini pelanggan atas layanan J&T Express dengan menggunakan teknik klasifikasi Naïve Bayes pada sekumpulan ulasan online. Data penelitian terdiri dari 500 ulasan, di mana 400 digunakan untuk pelatihan dan 100 untuk pengujian. Setelah melalui tahapan pra-pemrosesan teks yang mencakup case folding, penghilangan stopword, dan stemming, model Naïve Bayes dikembangkan untuk membedakan sentimen positif dan negatif. Hasil evaluasi menunjukkan model mencapai akurasi sebesar 87 %, dengan presisi 70 % untuk kelas positif dan 95 % untuk kelas negatif, serta error rate sebesar 13 %. Penelitian ini relevan karena menunjukkan bahwa algoritma Naïve Bayes efektif dalam mendeteksi sentimen pengguna terhadap layanan ekspedisi secara otomatis, meski hanya dibedakan menjadi dua kelas, dibandingkan dengan pendekatan tiga kelas pada paper

sebelumnya. Temuan ini dapat dijadikan acuan atau pembandingan mengenai efektivitas model dalam klasifikasi sentimen dua label pada data teks bahasa Indonesia.

2.2.2.1 Tujuan Penelitian

Penelitian yang dilakukan oleh Mahardika Tania Nitami dan Herny Februariyanti (2022) bertujuan untuk menganalisis sentimen pelanggan terhadap layanan ekspedisi J&T Express berdasarkan ulasan yang tersedia secara daring. Penelitian ini bertujuan mengklasifikasikan opini pelanggan ke dalam dua kategori sentimen, yaitu positif dan negatif, dengan menerapkan algoritma Naïve Bayes sebagai metode klasifikasi. Tujuan utama dari penelitian ini adalah untuk mengetahui seberapa besar tingkat kepuasan atau ketidakpuasan pelanggan terhadap layanan J&T Express, serta mengevaluasi efektivitas algoritma Naïve Bayes dalam mengolah data teks berbahasa Indonesia. Melalui hasil klasifikasi yang diperoleh, perusahaan diharapkan dapat memperoleh wawasan mengenai persepsi publik terhadap layanan mereka, sehingga dapat menjadi dasar dalam pengambilan keputusan untuk peningkatan kualitas pelayanan di masa depan.

2.2.2.2 Metodologi yang Digunakan

Penelitian yang dilakukan oleh Mahardika Tania Nitami dan Herny Februariyanti (2022) menggunakan metode kuantitatif dengan pendekatan klasifikasi sentimen berbasis teks menggunakan algoritma Naïve Bayes. Data yang digunakan berupa 500 ulasan pelanggan terhadap layanan J&T Express yang

diambil dari platform daring. Sebanyak 400 data digunakan sebagai data latih dan 100 data sebagai data uji. Sebelum dilakukan proses klasifikasi, data ulasan melalui beberapa tahapan pra-pemrosesan teks, di antaranya adalah case folding (mengubah semua huruf menjadi huruf kecil), tokenisasi (memisahkan kata), stopword removal (menghapus kata-kata umum yang tidak memiliki makna penting), dan stemming (mengubah kata menjadi bentuk dasar). Setelah data dibersihkan, algoritma Naïve Bayes diterapkan untuk membedakan ulasan positif dan negatif. Kinerja model kemudian diuji menggunakan confusion matrix untuk menghitung akurasi, presisi, recall, dan error rate dari hasil klasifikasi.

2.2.2.3 Temuan Utama

Temuan utama dari penelitian yang dilakukan oleh Mahardika Tania Nitami dan Herny Februariyanti (2022) menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan ulasan pelanggan terhadap layanan J&T Express dengan tingkat akurasi yang cukup tinggi. Berdasarkan hasil pengujian terhadap 100 data uji, model menghasilkan akurasi sebesar 87 %, dengan tingkat presisi 70 % untuk sentimen positif dan 95 % untuk sentimen negatif. Selain itu, nilai recall untuk sentimen positif sebesar 95 %, sedangkan untuk sentimen negatif sebesar 73 %, dengan error rate keseluruhan sebesar 13 %. Temuan ini memperlihatkan bahwa meskipun terdapat sedikit ketidakseimbangan dalam klasifikasi sentimen positif dan negatif, algoritma Naïve Bayes tetap efektif dalam memproses dan memahami teks ulasan berbahasa Indonesia. Hasil tersebut memberikan gambaran umum mengenai kecenderungan opini pelanggan terhadap J&T Express dan dapat

dijadikan dasar bagi perusahaan dalam mengevaluasi serta meningkatkan kualitas layanannya.

2.2.2.4 Kesimpulan Penelitian

Kesimpulan dari penelitian yang dilakukan oleh Mahardika Tania Nitami dan Hery Februariyanti (2022) menyatakan bahwa algoritma Naïve Bayes merupakan metode yang efektif dan cukup akurat dalam mengklasifikasikan sentimen ulasan pelanggan terhadap layanan J&T Express. Dengan akurasi sebesar 87% dan tingkat presisi serta recall yang cukup tinggi, penelitian ini membuktikan bahwa Naïve Bayes dapat diandalkan untuk mengolah data teks dalam bahasa Indonesia, khususnya dalam konteks analisis sentimen dua kelas (positif dan negatif). Hasil klasifikasi yang diperoleh tidak hanya menggambarkan persepsi pelanggan secara umum, tetapi juga dapat dimanfaatkan oleh perusahaan sebagai dasar evaluasi dan perbaikan layanan. Penelitian ini juga memperkuat relevansi penggunaan algoritma machine learning dalam memahami opini publik secara otomatis melalui data ulasan daring.

2.2.3 Paper 3

Penelitian yang dilakukan oleh Ansyahry dan Al Amin (2023) berjudul *“Klasifikasi Opini Masyarakat terhadap Jasa Ekspedisi J&T Express pada Media Sosial Twitter dengan Naïve Bayes”* membahas penerapan algoritma Naïve Bayes untuk menganalisis sentimen masyarakat terhadap layanan J&T Express berdasarkan unggahan di media sosial Twitter. Fokus dari penelitian ini adalah

untuk mengidentifikasi kecenderungan opini publik—baik positif maupun negatif—melalui pendekatan klasifikasi teks. Penelitian ini menekankan pentingnya media sosial sebagai sumber data yang bersifat real-time dan sangat dinamis dalam menggambarkan persepsi pelanggan terhadap suatu layanan.

Penelitian ini relevan dengan topik skripsi karena menggunakan algoritma Naïve Bayes untuk mengklasifikasikan opini masyarakat berbasis data teks dari media sosial, sama seperti pendekatan yang digunakan dalam penelitian ini. Selain itu, penggunaan data Twitter sebagai sumber informasi publik juga memperlihatkan kesamaan dengan sumber data utama dalam skripsi ini, yang menggunakan media sosial sebagai bahan analisis. Penelitian ini memberikan gambaran nyata mengenai bagaimana teknik pemrosesan bahasa alami (NLP) dan machine learning dapat dimanfaatkan untuk mengevaluasi layanan publik secara otomatis dan efisien, khususnya dalam sektor logistik dan ekspedisi.

2.2.3.1 Tujuan Penelitian

Penelitian yang dilakukan oleh Ansyahry dan Al Amin (2023) bertujuan untuk mengklasifikasikan opini masyarakat terhadap layanan ekspedisi J&T Express yang disampaikan melalui media sosial Twitter. Tujuan utama dari penelitian ini adalah untuk mengetahui kecenderungan sentimen masyarakat, apakah bersifat positif atau negatif, terhadap layanan yang diberikan oleh J&T Express. Dengan menggunakan algoritma Naïve Bayes, penelitian ini berusaha menghasilkan model klasifikasi yang mampu mengelompokkan tweet pelanggan secara otomatis berdasarkan sentimennya. Hasil dari penelitian ini diharapkan dapat

memberikan informasi yang bermanfaat bagi perusahaan dalam mengevaluasi dan meningkatkan kualitas pelayanan, serta membuktikan efektivitas metode Naïve Bayes dalam pengolahan data teks berbahasa Indonesia di platform media sosial.

2.2.3.2 Metodologi yang Digunakan

Penelitian yang dilakukan oleh Ansyahry dan Al Amin (2023) menggunakan pendekatan kuantitatif dengan metode text mining untuk mengklasifikasikan sentimen opini masyarakat terhadap layanan J&T Express yang diambil dari media sosial Twitter. Data yang digunakan dalam penelitian ini terdiri dari 1.694 tweet yang berkaitan dengan J&T Express. Proses pengolahan data dimulai dengan tahapan pra-pemrosesan teks yang meliputi pembersihan data (cleaning), tokenisasi, penghapusan stopword, dan stemming. Setelah data diproses, algoritma Naïve Bayes digunakan untuk membangun model klasifikasi sentimen. Evaluasi performa model dilakukan dengan menggunakan metrik akurasi, presisi, dan recall untuk mengukur seberapa baik model dapat mengklasifikasikan tweet ke dalam kategori sentimen positif atau negatif. Pendekatan metodologis ini dirancang untuk menangkap persepsi masyarakat secara otomatis dan efisien melalui analisis konten media sosial.

2.2.3.3 Temuan Utama

Temuan utama dari penelitian yang dilakukan oleh Ansyahry dan Al Amin (2023) menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen masyarakat terhadap layanan J&T Express dengan performa yang cukup

baik. Berdasarkan hasil evaluasi model klasifikasi, diperoleh tingkat akurasi sebesar 84%, dengan nilai presisi mencapai 76% dan recall sebesar 87%. Temuan ini mengindikasikan bahwa sebagian besar opini publik yang dikumpulkan dari media sosial Twitter dapat diklasifikasikan secara efektif ke dalam kategori sentimen positif maupun negatif. Penelitian ini juga mencatat bahwa opini masyarakat cenderung terbagi antara kepuasan terhadap kecepatan pengiriman dan ketidakpuasan terhadap pelayanan kurir atau kondisi paket. Hasil-hasil tersebut menunjukkan bahwa media sosial dapat menjadi sumber data yang kaya dan bermanfaat untuk mengukur persepsi masyarakat secara real-time, serta bahwa algoritma Naïve Bayes merupakan alat yang tepat untuk mengolah data tersebut dalam konteks analisis sentimen.

2.2.3.4 Kesimpulan Penelitian

Kesimpulan dari penelitian yang dilakukan oleh Ansyahry dan Al Amin (2023) menyatakan bahwa algoritma Naïve Bayes mampu digunakan secara efektif untuk mengklasifikasikan opini masyarakat terhadap layanan J&T Express berdasarkan data yang diambil dari media sosial Twitter. Dengan akurasi mencapai 84%, serta nilai presisi dan recall yang cukup tinggi, penelitian ini membuktikan bahwa Naïve Bayes merupakan metode yang andal dalam pemrosesan teks berbahasa Indonesia, khususnya dalam konteks analisis sentimen dua kelas (positif dan negatif). Selain itu, penelitian ini menegaskan bahwa media sosial merupakan sumber data yang potensial dan representatif dalam mengukur persepsi publik terhadap suatu layanan secara real-time. Hasil penelitian ini dapat dimanfaatkan

oleh perusahaan ekspedisi sebagai bahan evaluasi dalam meningkatkan mutu pelayanan berdasarkan umpan balik langsung dari pengguna layanan di media sosial.

