

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **1.1 Analisis Sentimen**

Analisis sentimen adalah teknik pembersihan data dari kata dan simbol yang tidak relevan dan merubah data yang bernilai kualitatif menjadi kuantitatif, kemudian data ulasan pengguna akan diklasifikasikan untuk mendapatkan ulasan yang bersifat positif dan negatif [1]. Saat ini analisis sentimen merupakan salah satu topik yang bannyak digunakan.

Oleh para peneliti, yang memiliki tujuan untuk menyediakan informasi dari sebuah dataset yang tidak terstruktur [2]. Analisis sentimen ini juga bisa diterapkan pada opini semua bidang seperti ekonomi, politik, sosial, dan hukum. Analisis sentimen ini juga dapat mengelompokkan polaritas dari teks untuk mengetahui apakah opini pada dokumen termasuk positif atau negatif [3].

Cara kerja Analisis sentimen dalam mengambil data dapat dibagi menjadi tiga langkah, yakni klasifikasi, evaluasi, dan hasil visualisasi. [4].

#### **1. Klasifikasi**

Klasifikasi data yang dinilai sebagai opini dari sebuah teks, ada tiga klasifikasi dalam metode analisis sentimen yang dapat dilakukan, yakni:

a. Machine Learning

Fitur-fitur didalamnya dapat mengenali sentimen (sudut pandang seseorang) dalam sebuah teks. Metode Machine Learning kini semakin bertambah populer karena dapat dinilai representative

b. Lexicon based

Menggunakan berbagai kata yang dinilai dengan skor polaritas untuk mengetahui tanggapan masyarakat atau pengguna mengenai suatu topik. Keunggulannya adalah tidak memerlukan data pelatihan, tapi kelemahannya adalah banyak data yang belum termuat dalam leksikon

c. Campuran

Menggabungkan metode Machine Learning dan Leksikon, metode ini biasanya memberikan hasil yang lebih menjanjikan

## **2. Evaluasi**

Setelah data terklasifikasi, metode analisis sentimen berikutnya adalah menggunakan metrik evaluasi seperti Precision, Recall, F-score, dan Accuracy. Proses ini juga melibatkan pengukuran rata-rata seperti makro, mikro dan skor F1 tertimbang untuk menangani data yang masuk ke dalam dua klasifikasi atau lebih. Metrik yang digunakan didasarkan pada keseimbangan klasifikasi set data secara umum, skemanya adalah sebagai berikut : Tinjauan set data, processing, tokenizer, penghapusan stopwords, transformasi, klasifikasi dan evaluasi.

## **3. Visualisasi Data**

Langkah selanjutnya dalam metode analisis system adalah visualisasi data, Visualisasi data dilakukan menggunakan bagan sesuai kebutuhan perusahaan atau siapa saja yang memanfaatkan data-data ini. Sebagian besar orang biasanya menggunakan Teknik yang sudah dikenal seperti Grafik, histogram atau matriks.

Namun, hasil akhir dari sentiment analysis bisa sangat bervariasi, data yang ada dapat muncul disertai domain yang terlibat. Karena itu, Teknik visualisasi data berupa Wordcloud, peta interaktif dan gaya sparkline juga cukup efektif untuk menampilkan hasil analisis.

### **1.2 Tokopedia**

Tokopedia adalah salah satu Perusahaan jual beli berbasis digital terbesar di Indonesia. Sejak resmi diluncurkan, PT. Tokopedia berhasil menjadi salah satu perusahaan internet Indonesia dengan pertumbuhan yang pesat. Dengan mengusung model bisnis marketplace dan mall online, Tokopedia memungkinkan setiap individu, toko kecil dan brand untuk membuka dan mengelola toko online. Sejak diluncurkan sampai hingga akhir 2015, layanan dasar Tokopedia yang bisa digunakan oleh semua orang secara gratis. Tokopedia memiliki visi untuk “Membangun Indonesia yang Lebih Baik Lewat Internet”, Tokopedia mempunyai program untuk mendukung para pelaku Usaha Mikro Kecil dan Menengah (UMKM) dan perorangan untuk mengembangkan usaha mereka dengan memasarkan produk secara online.

Tokopedia merupakan perusahaan yang didirikan oleh William Tanuwijaya dan Leontinus Alpha Edison pada 6 Februari 2009, kemudian pada 17 Agustus 2009 secara resmi diluncurkan kepada publik. PT Tokopedia merupakan sebuah online marketplace yang mewadahi para pemilik usaha seperti UMKM di Indonesia untuk memiliki toko mereka secara online. Tokopedia menyediakan sarana bagi pemilik usaha ini untuk mengelola dan mengembangkan usaha mereka dengan mudah dan bebas biaya, selain itu juga memberi pengalaman yang baru dalam hal jual beli online.

Kemunculan Tokopedia bermula dari pengamatan William dan Leontinus yang melihat banyaknya pedagang di desa yang merantau ke kota untuk bisa memperluas jangkauan pasar dengan keuntungan yang lebih baik. Sementara itu, konsumen di daerah sulit mendapat produk yang diinginkan karena terbatasnya akses. Hal ini lah yang mendorong William dan Leontinus untuk mencari solusi guna memudahkan konsumen di daerah mendapatkan kebutuhannya yang mudah dijangkau. Sejak kemunculannya, Tokopedia menjadi salah satu perusahaan internet yang memiliki perkembangan cukup pesat di Indonesia. Tokopedia berusaha untuk selalu melakukan inovasi dan pengembangan pada perusahaannya.

Selama berjalannya PT Tokopedia ini selalu mendapatkan suntikan dana dari beberapa perusahaan besar, diantaranya adalah PT Indonusa Dwitama (2009), East Ventures (2010), Cyber Agent Ventures (2011), Netprice (2012), dan Softbank Ventures Korea (2013). Tokopedia menjadi perusahaan teknologi pertama di Asia Tenggara yang menerima investasi sebesar USD 100 juta dari Sequoia Capital dan Softbank Internet and Media Inc pada Oktober 2014. Kemudian pada April 2016, Tokopedia mendapat investasi kembali sebesar USD 147 juta. Pada Agustus 2017, Tokopedia kembali mendapat investasi dari Alibaba yaitu e-commerce asal Tiongkok yakni sebesar USD 1,1 miliar. Tokopedia kembali berhasil mendapatkan pendanaan sebesar USD 1,1 miliar dari sejumlah investor yang pendanaan tersebut dipimpin oleh Softbank Vision Fund dan Alibaba Group pada Desember 2018.

Dilihat dari kesuksesannya, Tokopedia berhasil menerima beberapa penghargaan antara lain Best Company in Consumer Industry dari Indonesia Digital Economy Award 2016, menjadi #3 Top Chart di Google Play, 'Fastest Value Growth' dalam acara BrandZ™ Top 50 Most Valueable Indonesian Brands, dan Best Companies to Work For dari HR Asia Award.

## 1.3 Data Mining

### 2.3.1 Pengertian Data Mining

Data mining adalah proses pengumpulan data dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada data. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika ataupun *teknologi Artificial Intelligence (AI)*. Data mining juga sering disebut *Knowlegde Discovery in Database (KDD)*.

Data mining adalah serangkaian proses untuk menambah serta mencari informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining biasanya digunakan untuk mencari pengetahuan dalam basis data yang besar sehingga sering disebut *Knowledge Discovery in Database (KDD)*

Data mining adalah pencarian pengetahuan dalam basis data pada proses Identifikasi pola-pola yang valid, yang berpotensi bermanfaat dan dapat dipahami dengan mudah. Data mining merupakan proses yang bertujuan untuk menemukan pengetahuan (*Knowledge Discovery*) yang ditambah dari sekumpulan data yang memiliki volume yang sangat besar.

Menurut **Larose**, definisi data mining adalah proses menemukan sesuatu yang bermakna oleh suatu korelasi baru, pola dan juga trend yang terdapat dengan cara memilah-milah data yang berukuran besar, dimana data tersebut disimpan dalam repository, menggunakan

teknologi sosialisasi pola serta statistic dan Teknik matematika pengertian penambangan data atau “Data Mining” merupakan analisis pengamatan database dalam menemukan hubungan yang tidak terduga serta juga untuk meringkas data dengan cara atau sebuah metode baru yang dapat di mengerti yang berguna pada pemilik data. Penggalan data juga dapat diartikan sebagai suatu proses ekstraksi sebuah informasi yang berguna serta potensial dari sekumpulan data yang secara implisit terdapat didalam suatu kumpulan data.

### 2.3.2 Pengelompokan Data Mining

Pengelompokan *data mining* dibagi menjadi beberapa kelompok, menurut [5], yaitu:

1. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali *variable* target estimasi lebih ke arah numerik daripada ke arah kategori. Model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.

3. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

4. Klasifikasi

Dalam klasifikasi terdapat target variabel kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

5. Pengklasteran

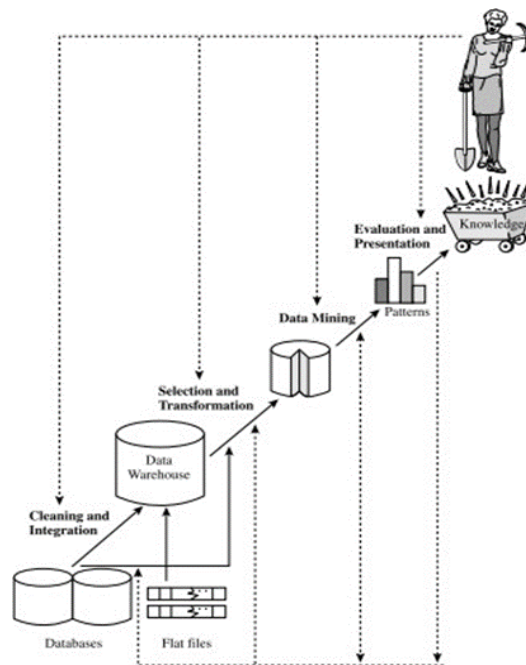
Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

## 6. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

### 2.3.3 Tahapan Data Mining

Tuntutan dari data mining salah satunya adalah ketika diterapkan pada data berskala besar adalah diperlukan metodologi sistematis tidak hanya ketika melakukan analisa saja tetapi juga ketika mempersiapkan data dan juga melakukan interpretasi dari hasilnya sehingga dapat menjadi aksi ataupun keputusan yang bermanfaat. Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada gambar 2. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan knowledge base



**Gambar 2.1 Tahapan Data Mining**

Tahap-tahap *data mining* adalah sebagai berikut:

1. Pembersihan data (*data cleaning*)

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan.

2. Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.

3. Seleksi data (*data selection*)

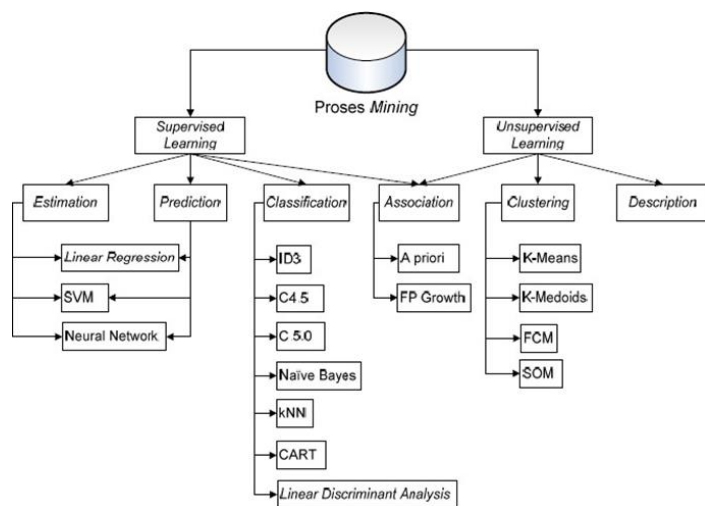
Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

4. Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

5. Proses *Mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan *data mining* dapat dilihat pada Gambar



**Gambar 2.2 Beberapa Metode Data Mining**

6. Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.

7. Presentasi pengetahuan (*knowledge presentation*)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. [6]

#### 1.4 Text Mining

Tujuan utama dari *text mining* adalah untuk mengolah dan menganalisis teks secara otomatis menggunakan teknik-teknik komputasi dan pembelajaran mesin. *Text mining* bertujuan untuk mengekstrak informasi yang bermanfaat dari teks, seperti menentukan tema atau topik, mengidentifikasi entitas atau kata kunci, atau mengukur sentiment atau emosi dari teks.

Dengan *text mining*, dapat dilakukan berbagai macam analisis teks, seperti klasifikasi teks, cluster atau pengelompokan teks, atau sumber daya linguistik yang terkait dengan teks. *Text mining* juga dapat membantu mengidentifikasi pola-pola yang tidak terlihat secara manual dari teks, sehingga dapat membantu mengambil keputusan atau menemukan pola-pola yang bermanfaat dari teks.

Selain itu, *text mining* juga dapat membantu mengoptimalkan pencarian dan navigasi informasi di dalam teks, sehingga dapat memudahkan pengguna untuk menemukan informasi yang diinginkan dengan cepat dan tepat. Dengan demikian, *text mining* memiliki banyak manfaat dan tujuan yang bermanfaat dalam mengelola dan menganalisis teks secara otomatis.

#### 1.5 Klasifikasi

Klasifikasi (*Classification*) data mining merupakan proses mendapatkan penjelasan kesamaan karakteristik pada suatu kelas atau kelompok dengan tujuan untuk memperkirakan kelas dari suatu objek yang belum diketahui labelnya. Dan metode klasifikasi menjadi salah satu yang paling sering digunakan pada data mining.



Klasifikasi pada data mining melakukan prosesnya dengan cara mempelajari data yang ada sebelumnya, selanjutnya mengklasifikasikan data baru, metode ini menghasilkan kategorikal (ordinal ataupun nominal). Untuk mengetahui apakah perkiraan akurasi yang dihasilkan benar, maka dapat diketahui melalui confusion matrix.

Klasifikasi memakai data uji guna menentukan keakuratan model. Umumnya kumpulan data yang digunakan selanjutnya dibagi menjadi dua bagian, bagian pertama adalah data latih dan bagian kedua adalah data uji. Model yang diharapkan dibentuk menggunakan data latih kemudian proses pengujian menggunakan data uji.

Proses pembersihan data dapat mempengaruhi kinerja metode klasifikasi. Tentunya jika data yang akan digunakan tidak bersih, memiliki banyak anomali, missing value dan masalah lainnya, maka hasil klasifikasi yang didapatkan tidak akan optimal. Confusion matrix juga dapat digunakan untuk menghitung kualitas metode klasifikasi yang digunakan dengan menghitung akurasi, spesifisitas, F-score dan sensitivitas.

## 1.6 Naïve Bayes

Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat pada Teknik klasifikasi. Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistic yang ditemukan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naïve dimana diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lain (Friska Aditia Indriyani et al., 2023a).

Persamaan dari teorema Bayes adalah :

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

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$  : Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$  : Probabilitas hipotesis H (prior probability)

$P(X|H)$  : Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  : Probabilitas X

Untuk menjelaskan teorema Naïve Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah penduduk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema bayes diatas disesuaikan sebagai berikut :

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)}$$

Dimana variable C mempresentasikan kelas, sementara variable  $F_1 \dots F_n$  mempresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik-karakteristik secara global (disebut juga evidence). Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut :

$$Posterior = \frac{Prir \times likelihood}{evidence}$$

Nilai evidence selalu tetap untuk setiap kelas pada satu sampel. Nilai dari posterior tersebut nantinya akan dibandingkan dengan nilai-nilai posterior kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan menjabarkan  $(C|F_1 \dots F_n)$  menggunakan aturan perkalian sebagai berikut :

$$\begin{aligned}
P(C|F_1, \dots, F_n) &= P(C) P(F_1, \dots, F_n|C) \\
&= P(C)P(F_1|C) P(F_2, \dots, F_n|C_1F_1) \\
&= P(C)P(F_1|C) P(F_2, \dots, F_n|C_1F_1), P(F_n|C, F_1, F_2) \\
&= P(C)P(F_1|C) P(F_2, \dots, F_n|C_1F_1), P(F_n|C, F_1, F_2), P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\
&= P(C)P(F_1|C) P(F_2, \dots, F_n|C_1F_1), P(F_n|C, F_1, F_2), P(F_4, \dots, F_n|C, F_1, F_2, F_3, \dots, F_{n-1})
\end{aligned}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor syarat yang mempengaruhi nilai probabilitas, yang hamper mustahil untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi untuk dilakukan

### 1.7 RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

Fitur-fitur dalam RapidMiner :

1. Banyak algoritma penambangan data seperti pohon keputusan dan peta yang mengatur sendiri
2. Visualisasi yang canggih seperti diagram histogram yang tumpang tindih, diagram pohon, dan diagram sebar 3D.
3. Tersedia banyak pilihan plugin termasuk juga plugin teks yang bisa dimanfaatkan untuk melakukan analisis text mining.

4. Menyediakan teknologi data mining dan machine learning termasuk ETL (*Extract, Transform, Load*), pra-pemrosesan data, visualisasi data, pemodelan dan evaluasi.
5. Proses penambangan data ditulis dalam XML dan terdiri dari operator nestable, dideskripsikan dengan XML, dan dibuat dalam GUI.
6. Mengintegrasikan proyek data mining Weka dan statistika R

### 1.8 Penelitian Terdahulu

Peneliti mendapatkan 10 penelitian terdahulu yang berkaitan dengan penelitian terhadap analisis sentimen menggunakan Algoritma Naïve Bayes.

**Tabel 2.8 Penelitian Terdahulu**

No	Nama	Tahun	Judul	Metode	Hasil
1.	Putra Kurniawan, Wasilah, Sutedi, Handoyo Widi Nugroho	2024	Implementasi Data Mining Dalam Klasifikasi Tingkat Kesenjangan Kompetensi PNS Menggunakan Metode Naive Bayes	Pada penelitian ini dilakukan pengujian hasil kasifikasi menggunakan salah satu teknik pengklasifikasian data mining, yaitu metode Naïve Bayes. Tujuan dari penelitian ini, yaitu melakukan pengujian terhadap performa algoritma Naïve Bayes dalam melakukan klasifikasi tingkat kesenjangan kompetensi PNS.	Hasil ini menunjukkan bahwa model yang dibangun memiliki performa yang sangat baik, baik secara keseluruhan maupun ketika mempertimbangkan distribusi kelas.
2.	Melda Agarina, Sutedi	2020	Penerapan Data Mining dalam Perancangan Sistem Pendukung Keputusan Seleksi Penerimaan Beasiswa Menggunakan Naive Bayes Classifier (Studi Kasus: IIB Darmajaya)	Metode Naive Bayes merupakan suatu metode yang digunakan untuk memprediksi suatu keputusan berdasarkan dengan kriteria yang sudah ditentukan. Sistem ini dibangun menggunakan bahasa pemrograman Java	Hasil dari implementasi sistem ini ialah memberikan keterangan tentang informasi penerima beasiswa berdasarkan rengking yang dapat digunakan sebagai alat

No	Nama	Tahun	Judul	Metode	Hasil
				dengan menggunakan MySQL sebagai databasenya.	bantu dalam proses pengambilan keputusan. Dengan adanya sistem ini, proses perhitungan untuk menentukan penerima beasiswa dapat dilakukan dengan mudah, cepat dan akurat.
3.	Rahayu Dwi Pemasari	2020	Analisis Sentimen nasabah bank bri terhadap menggunakan Naive Bayes	metode Pengklasifikasian Naive Bayes. Dengan Naive Bayes informasi sentimen dari sebuah kalimat dapat diketahui secara cepat karena sifat dari Naive Bayes sendiri yang sederhana dan independen sehingga tidak memperdulikan keberadaan fitur lainnya. Adapun proses yang dilakukan dalam penelitian ini adalah preprocessing terhadap data sentimen dan pengklasifikasian kalimat sentimen ke dalam kelas positif dan negatif.	Berdasarkan hasil yang didapat dalam penelitian ini bahwa dengan sistem yang dibangun diketahui nilai error yang diberikan metode Naive Bayes dalam mengklasifikasikan kalimat sentimen ke dalam kelas positif dan negatif sebesar 24,26%. Dengan nilai error yang diberikan, dapat disimpulkan bahwa metode Naive Bayes sesuai untuk dipergunakan dalam melakukan pengklasifikasian terhadap kalimat sentimen
4.	Indar Swarsih, Suzuki Sofian, Joko Tri	2023	Analisis Sentimen Tingkat Kepuasan Pelanggan Terhadap	Dalam kerangka penelitian ini, dilakukan analisis sentimen	Hasil penelitian menunjukkan bahwa dari

No	Nama	Tahun	Judul	Metode	Hasil
	Handoko, MS Hasibuan.		Layanan Kurir J&T Express Menggunakan Algoritma Support Vector Machine (SVM) Berdasarkan Ulasan Pengguna Di Google Playstore	menggunakan algoritma Support Vector Machine, dengan menerapkan metodologi CRISP- DM.	tahap pemahaman bisnis hingga tahap pemodelan, performa Support Vector Machine dapat dianggap baik. Sebagai tambahan, penelitian ini juga menghasilkan implementasi yang dapat diakses melalui situs web dengan alamat jnt- sentiment.strea mlit.app.
5.	Khusnul Khotimah	2021	Teknik Data Mining menggunakan Algoritma Decision Tree (C4.5) untuk Prediksi Seleksi Beasiswa Jalur KIP pada Universitas Muhammadiyah Kotabumi	Algoritma Decision Tree (C4.5) merupakan salah satu algoritma data mining yang dapat digunakan untuk klasifikasi data membantu menyelesaikan permasalahan klasifikasi.	Berdasarkan hasil penelitian implementasi data mining menggunakan algoritma decision tree (C4.5) diperoleh nilai akurasi sebesar 100%. Uji akurasi juga dilakukan dengan algoritma Naïve Bayes untuk memperoleh perbandingan tingkat akurasi.
6.	Susilo Hartono	2021	Faktor-Faktor yang Mempengaruhi Kesuksesan Implementasi Sistem Informasi (Studi Kasus Sistem Informasi Akademik STKIP Muhammadiyah Pringsewu Lampung)	Penelitian dilakukan dengan menggunakan model kesuksesan sistem informasi dari DeLone & McLean, yang telah di modifikasi sesuai dengan kebutuhan dan keadaan yang ada di STKIP MPL.	Hasil dari penelitian menunjukkan bahwa dari 5 (lima) hipotesa yang dirumuskan, terdapat 4 (empat) hipotesa yang diterima. Hipotesis tersebut adalah

No	Nama	Tahun	Judul	Metode	Hasil
					kualitas informasi berpengaruh terhadap kepuasan pengguna, kualitas sistem berpengaruh terhadap kepuasan pengguna, kepercayaan berpengaruh terhadap kepuasan pengguna dan kepuasan pengguna berpengaruh kepada manfaat-manfaat bersih.
7.	sulastri	2022	Penerapan Data Mining untuk Prediksi Rating Penjualan Buku Menggunakan Metode <i>Naive Bayes</i>	Dalam penelitian ini ada beberapa metode yaitu analisis kebutuhan, metode pengumpulan data, pengembangan system.	Teknik yang digunakan dalam prediksi rating penjualan buku ini memakai algoritma <i>Naive Bayes</i> . <i>Naive Bayes</i> digunakan untuk mencari nilai probabilitas paling besar pada setiap variabel yang sudah ada.
8.	Kevin Moniarga Suharta	2023	Penerapan Algoritma <i>Naive Bayes Classifier</i> untuk Mengetahui Minat Beli Pelanggan Terhadap Sofa (Studi Kasus di Mebel Kelumer Bayau)	Dalam penelitian ini di gunakan beberapa metode yaitu: metode pengumpulan data, metode pengembangan perangkat lunak	Penelitian ini menggunakan metode Algoritma <i>Naive Bayes Classifier</i> yang mempunyai akurasi dan kecepatan untuk dilakukan penggalian pengetahuan yang sudah ada pada database.

No	Nama	Tahun	Judul	Metode	Hasil
9.	Dicky Norfiansyah	2019	Penerapan Data Mining Dengan Algoritma <i>Naive Bayes Clasifier</i> untuk Mengetahui Minat Beli Pelanggan Terhadap Kartu Internet XL (Studi Kasus di CV. Sumber Utama Telekomunikasi)	Dalam penelitian ini di gunakan beberapa metode yaitu: metode pengumpulan data, metode pengembangan perangkat lunak	Hasil dari penelitian ini bisa memprediksi laku atau tidak terjual kartu perdana, sehingga perusahaan bisa mendapatkan keputusan sehingga bisa menaikkan strategi penjualan.
10.	Alvina Felicia Watratan	2022	Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia	Metode penelitian yang di terapkan pada penelitian ini ialah analisis masalah dan studi literatur, mengumpulkan data dan implementasi. Penerapan metode naive bayes diharapkan mampu memprediksi tingkat penyebaran COVID-19 di Indonesia.	Hasil penelitian pada klasifikasi metode Naive Bayes menunjukkan bahwa 16 data dari 33 data yang di uji dalam kasus Covid-19 per provinsi dengan keakuratan sebesar 48,4848%, di mana dari 33 data yang di uji dalam kasus Covid-19 per provinsi yang diuji terdapat 16 data yang berhasil diklasifikasikan dengan benar.