

BAB II

TINJAUAN PUSTAKA

2.1 Sentiment Analysis

Sentimen Analysis adalah proses penggunaan *text analytics* untuk mendapatkan sumber data dari internet dan beragam *platform* media social. Tujuannya untuk memperoleh opini dari pengguna yang terdapat pada *platform* tersebut. *Sentiment Analysis* berperan sebagai alat yang dapat menghubungkan seluruh data tersebut. Dengan begitu perusahaan dapat memperoleh masukan inti dari pengguna atau konsumen secara efisien.

Sentimen Analysis merupakan salah satu bidang dari *Natural Language Processing (NLP)* yang membangun system untuk mengenali dan mengekstraksi opini dalam bentuk teks. Informasi berbentuk teks saat ini banyak terdapat di internet dalam format forum, blog, media social, serta situs berisi *review*. Dengan bantuan *system analysis*, informasi yang tadinya tidak terstruktur dapat diubah menjadi data yang lebih terstruktur.

Sentiment Analysis saat ini dibagi menjadi penggunaan klasifikasi berdasarkan aturan, metode pembelajaran mesin menggunakan kata kata emosi sebagai klasifikasi fitur dan kamus emosi dapat digunakan untuk mewujudkan pemilihan karakteristik sentiment dengan cepat dan efisien (Isnain et al. 2021)

2.2 TikTok

Tiktok merupakan salah satu *platform* media social yang memberikan kemungkinan bagi para penggunanya untuk dapat membuat video pendek yang didukung dengan fitur musik, filter dan berbagai fitur kreatif lainnya. Pada awalnya tiktok diluncurkan bukan dengan nama Tiktok, pada bulan September tahun 2016, *ByteDance*, sebuah perusahaan yang berbasis di China, meluncurkan sebuah aplikasi video pendek yang memiliki pengguna sebanyak 100 juta pengguna dan tayangan video sebanyak 1 miliar tayangan setiap harinya hanya dalam jangka waktu 1 tahun. Oleh karena

kepopuleritasannya yang meningkat dengan pesat, *ByteDance* memutuskan untuk memperjelas jangkauan Douyin hingga ke luar China dengan nama baru yaitu TikTok (Ggu et al. 2021).

Dengan adanya media social TikTok di era globalisasi saat ini, seluruh masyarakat mendapatkan dampak negatifnya, seperti melanggar moral dan etika. Perilaku amoral dan tidak etis yang telah dilakukan oleh kaum milenial, bukan hanya mengomentari video, maupun juga saat membuat video ataupun membagikannya, sehingga banyak orang yang menyaksikannya. Tapi selain itu, TikTok juga saat ini dapat menjadi media informasi bagi para pengguna untuk memperoleh informasi yang dibutuhkan. Tidak sedikit kalangan remaja yang menjadi *Content Creator* dan memanfaatkan TikTok sebagai sumber penghasilan (Kustiawan 2022)

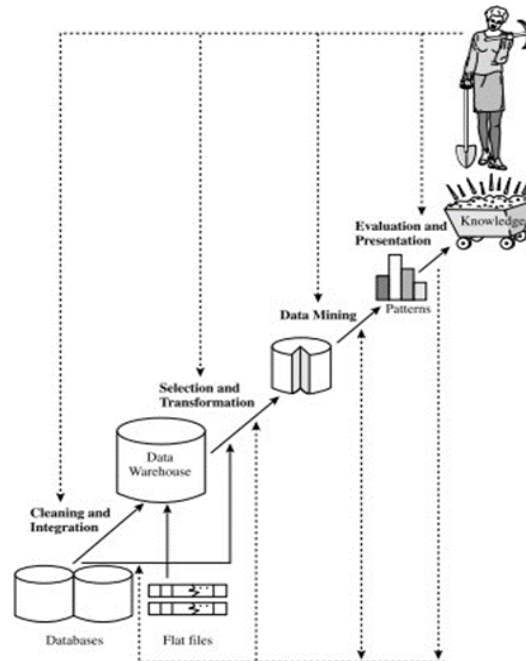
2.3 Data Mining

2.3.1 Pengertian Data Mining

Data Mining adalah proses untuk mendapatkan informasi yang berguna dari Gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah data mining kadang disebut juga *Knowledge Discovery*.

Salah satu Teknik yang dibuat dalam data mining adalah bagaimana menelusuri data agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. Kebutuhan untuk prediksi juga dapat memanfaatkan Teknik ini. Dalam data mining, pengelompokan data juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui pola *universal* data-data yang ada. Anomali data transaksi juga perlu di deteksi untuk dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional perusahaan sehingga tujuan akhir perusahaan diharapkan dapat tercapai.

Data mining merupakan bagian dari proses *Knowledge Discovery From Data (KDD)* (Ahmad, Samsugi, and Irawan 2022) Dibawah ini digambarkan skema dari proses KDD.



Gambar 2.1 Proses knowledge discovery

Gambar 2.1 menunjukkan proses penjelajahan pengetahuan dimulai dari beberapa database dilakukan proses *cleaning* dan *integration* sehingga menghasilkan data *warehouse*. Dilakukan proses *selection* dan *transformation* yang kemudian disebut sebagai data mining hingga menemukan pola dan memperoleh pengetahuan dari data (*knowledge*).

2.3.2 Sejarah Data Mining

Membahas mengenai Data Mining dan artinya, pastinya kita juga harus mengetahui apa sejarahnya bukan? Salah satu artikel pertama yang menggunakan frasa “Data Mining” diterbitkan oleh *Michael C. Lovell*.

Pada saat itu, *Lovell* dan banyak ekonom lain mengambil pandangan yang cukup negative dari praktik tersebut, percaya bahwa statistic dapat mengarah pada kesimpulan yang salah ketika tidak di

informasikan oleh pengetahuan tentang materi pelajaran. Tetapi pada tahun 1990-an, gagasan untuk mengekstraksi nilai dari data dengan mengidentifikasi *pattern* atau pola telah menjadi jauh lebih populer. Basis data dan vendor data *Warehouse* mulai menggunakan kata kunci untuk memasarkan perangkat lunak mereka dan perusahaan mulai menyadari manfaat potensial dari praktik tersebut.

Pada tahun 1996, sekelompok perusahaan yang termasuk terdata dan *NCR* memimpin proyek untuk mengstandarisasi dan mengformalkan metodologi penambangan data. Pekerjaan mereka menghasilkan proses standar lintas industri untuk penambangan data atau *Cross-Industry Standard Process for Data Mining* (CRISP-DM) (Tamba et al. 2023)

Standar terbuka ini memecah proses data mining menjadi 6 (enam) fase :

1. *Business Understanding* (Pemahaman bisnis)
2. *Data Understanding* (Pemahaman data)
3. *Data Preparation* (Persiapan Data)
4. *Modelling* (Pemodelan)
5. *Evaluation* (Evaluasi)
6. *Deployment* (Penyebaran)

Popularitas data mining terus meningkat secara signifikan pada tahun 1990-an, terutama dengan pelaksanaan sejumlah konferensi yang dikhususkan pada kajian tersebut, seperti: *ACM SIGKDD annual conference* tahun 1995, *European PKDD and Pacific/Asia (PAKDD)* pada tahun 1997. *IEEE ICDM* mulai diadakan pada tahun 2001 sebagai konferensi *SIAM* yang pertama. Hal ini meningkatkan popularitas data mining yang secara bersamaan didukung oleh kemajuan teknologi, kemampuan *CPU* dan media yang menyimpan data dalam jumlah besar dan mengolahnya dalam waktu yang lebih cepat. Menjadi hal biasa untuk perusahaan komersial untuk memelihara data dalam berbagai bentuk yang dapat dibaca komputer, dan dalam kebanyakan kasus digunakan untuk mendukung aktivitas bisnis dan pemikiran bahwa data tersebut

dapat “ditambang” sering masih dianggap sebagai prioritas kedua. Era tahun 1990-an juga ditandai dengan penggunaan kartu kesetiaan pelanggan yang memungkinkan perusahaan untuk mencatat belanjaan dari setiap pelanggan. Data yang tersimpan dalam volume besar tersebut dapat “ditambang” untuk mengetahui pola/perilaku pembelian oleh masing-masing pelanggan. Popularitas data mining terus berkembang sampai hari ini terlebih dengan kemungkinan menambang data dari data yang tidak standar (non tabular)

Pada awal tahun 2000 an, perusahaan web mulai melihat kekuatan dari data mining dan praktiknya benar-benar berjalan bahkan sampai sekarang ini. Sementara fase “data mining” sejak saat itu telah dikalahkan oleh kata kunci lain seperti “*data analytics*” atau analisis data “*big data*” dan “*machine learning*” prosesnya tetap menjadi bagian integral dari praktik bisnis. Bahkan wajar untuk mengatakan bahwa data mining telah menjadi bagian dari *de facto* dari menjalankan bisnis modern seperti seperti sekarang ini.

2.3.3 Fungsi Data Mining

Data mining memiliki beberapa fungsi, yaitu (Ahmad, Samsugi, and Irawan 2022) :

1. *Classification*

Classification adalah proses untuk mencari model atau fungsi yang menggambarkan dan membedakan kelas-kelas atau konsep data.

2. *Clustering*

Berfungsi untuk mencari pengelompokkan atribut ke dalam segmentasi-segmentasi berdasarkan similaritas.

3. *Association*

Berfungsi untuk mencari keterkaitan antara atribut atau item set, berdasarkan item yang muncul dan rule association yang ada.

4. *Regression*

Bertujuan untuk mencari prediksi dari suatu pola yang ada.

5. *Forecasting*

Berfungsi untuk meramalkan waktu yang akan datang berdasarkan trend yang telah terjadi di waktu sebelumnya

6. *Sequence Analysis*

Berfungsi untuk mencari pola urutan dari rangkaian kejadian.

7. *Deviation Analysis*

Berfungsi untuk mencari kejadian langka yang sangat berbeda dari keadaan normal (kejadian abnormal)

2.3.4 Pengelompokkan Data Mining

Pengelompokkan *data mining* dibagi menjadi beberapa kelompok, berdasarkan tugas yang dapat dilakukan, yaitu (Annisa, Ginting, and Syar 2022) :

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

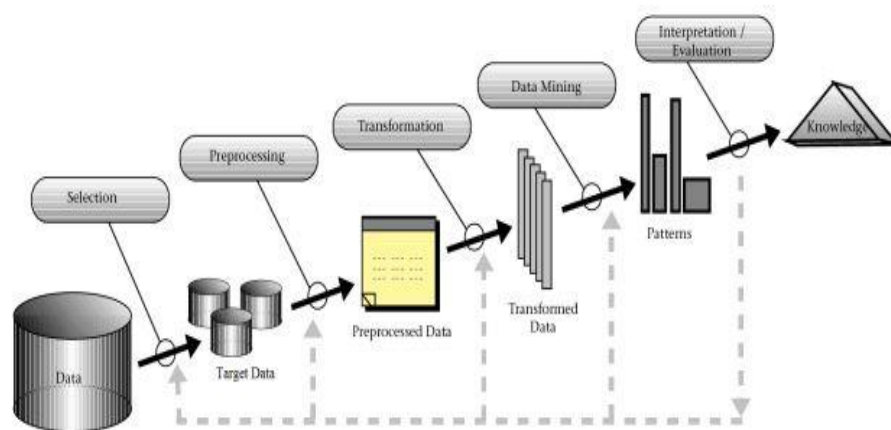
Merupakan pengelompokkan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki

kemiripan.

6. Asosiasi

Asosiasi bertugas menemukan atribut muncul dalam 1 waktu.

2.3.5 Tahapan Data Mining



Gambar 2.2 Tahapan Data Mining

Gambar 2.2 menunjukkan tahapan data mining, penjelasan tahapan data mining ialah sebagai berikut (Naldy and Andri 2021) :

1. *Data Selection*

Data yang ada akan dilakukan seleksi data dan atribut yang akan digunakan untuk proses selanjutnya. Seleksi data dari sekumpulan data operasional dilakukan sebelum tahap penggalian informasi. Data hasil seleksi akan digunakan untuk proses data mining dan di simpan dalam berkas yang terpisah dari data operasional.

2. *Pre-Processing (Cleaning)*

Tahapn cleaning merupakan hasil seleksi yang akan dilakukan proses pendahuluan dan pembersihan data. Sebelum proses data mining dilakukan, perlunya dilakukan cleaning data dengan tujuan untuk membuang duplikasi data, menghilangkan data yang tidak konsisten.

3. *Transformation*

Merupakan proses transformasi pada data yang telah dipilih untuk di ubah menjadi bentuk yang sesuai untuk di olah nantinya.

4. Data mining

Data mining adalah proses mencari suatu informasi atau pola yang menarik pada data yang telah dipilih menggunakan Teknik atau metode tertentu.

5. *Interpretation (Evaluation)*

Pada tahap ini, dilakukan identifikasi pola-pola yang dihasilkan dari data mining. Pola yang dihasilkan akan ditampilkan kedalam bentuk yang mudah di mengerti.

2.4 Naïve Bayes

Metode klasifikasi *Naïve Bayes* merupakan metode dengan algoritma yang sederhana namun sangat cepat dan akurat. Dalam penelitian ini, *Naïve Bayes* mampu bekerja dengan baik meskipun terdapat fitur yang sangat bergantung pada kumpulan data. Jumlah komentar yang diambil dari TikTok untuk digunakan sebagai data pelatihan rendah karena sulitnya mendapatkan komentar yang sesuai. Untuk permasalahan tersebut *Naïve Bayes* sangat cocok digunakan karena *Naïve Bayes* tetap dapat bekerja dengan baik dengan ukuran data training yang kecil. Selain itu, kesederhanaan algoritma *Naïve Bayes* dan kecepatannya yang tinggi selama pelatihan dan klasifikasi membuatnya menarik untuk digunakan sebagai metode klasifikasi. Dengan karakteristik tersebut di atas, maka metode klasifikasi *Naïve Bayes* sesuai jika digunakan pada penelitian ini (Darwis, Siskawati, and Abidin 2021). Berikut persamaan dari teorema Bayes:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

$$P(X)$$

Keterangan:

X : Data dengan class yang belum diketahui

C : Hipotesis data X merupakan suatu kelas spesifik

$P(C|X)$: Probabilitas hipotesis C

berdasarkan kondisi X (posterior)

probability)

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

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

$P(X)$: Probabilitas X (predictor prior probability)

2.4.1 Gaussian Naïve Bayes

Gaussian Naïve Bayes merupakan sebuah Teknik klasifikasi yang digunakan dalam machine Learning dengan menggunakan metode probabilitas dan Distribusi Gaussian atau Distribusi Normal. Distribusi Gaussian mengasumsikan bahwa setiap fitur pada data memiliki pengaruh yang independent dalam memprediksi target. Kombinasi prediksi dari seluruh parameter adalah prediksi akhir dengan probabilitas dari variable target yang diklasifikasikan ke dalam dua kelas. Klasifikasi akhirnya adalah hasil probabilitas yang lebih tinggi dari target group.

2.4.2 Bernoulli Naïve Bayes

Bernoulli Naïve Bayes adalah bagian dari keluarga *Naïve Bayes*, ini hanya membutuhkan nilai biner. Contoh paling umum adalah ketika kita memeriksa apakah setiap nilai akan menjadi kata yang muncul dalam dokumen atau tidak. Itu adalah model yang sangat disederhanakan. Dalam kasus dimana perhitungan frekuensi kata kurang penting, Bernoulli mungkin memberikan hasil yang lebih baik. Dengan kata sederhana kita harus menghitung setiap nilai fitur kemunculan istilah biner, yaitu suatu kata muncul dalam dokumen atau tidak. Fitur-fitur ini digunakan daripada mencari frekuensi suatu kata dalam dokumen.

2.4.3 Multinomial Naïve Bayes

Model multinomial memberikan kemampuan untuk mengklasifikasikan data, yang tidak dapat direpresentasikan secara numerik. Keuntungan utamanya adalah kompleksitasnya berkurang

secara signifikan. Ini memberikan kemampuan untuk melakukan klasifikasi, menggunakan set pelatihan kecil, tidak perlu dilatih ulang secara terus menerus.

Multinomial naïve Bayes banyak digunakan untuk menugaskan dokumen ke kelas berdasarkan analisis statistik isinya. Ini memberikan alternatif terhadap analisis semantik berbasis AI yang "berat" dan secara drastis menyederhanakan klasifikasi data tekstual. Klasifikasi bertujuan untuk menetapkan fragmen teks (yaitu dokumen) ke dalam kelas dengan menentukan probabilitas bahwa suatu dokumen termasuk dalam kelas dokumen lain, yang memiliki subjek yang sama. Setiap dokumen terdiri dari beberapa kata (yaitu istilah), yang berkontribusi pada pemahaman tentang isi dokumen. Kelas adalah tag dari satu atau beberapa dokumen, yang mengacu pada subjek yang sama

2.5 Rangkaian Penelitian

Rangkaian penelitian pada penelitian ini ialah sebagai berikut :

1. Studi Literatur

Untuk tahapan yang pertama peneliti melakukan proses studi literatur, yaitu mencari, membahas dan memahami penelitian sebelum yang berkaitan dengan objek yang akan dibahas.

2. Pengumpulan Data/*Crawling Data*

Pada penelitian ini pengumpulan data menggunakan Data sekunder yang diambil dari komentar pengguna Tiktok. Data yang diambil akan dijadikan sebagai data training untuk membuat kamus opini, pembuatan kamus data opini tersebut dengan mengumpulkan data opini melalui media online. Kamus data opini dikumpulkan di klasifikasikan menjadi kamus opini negative dan positif.

3. *Preprocessing*

Tahapan yang selanjutnya adalah preprocessing, dimana diharapkan dapat menghasilkan informasi yang dapat dianalisis, Informasi tersebut harus diproses terlebih dahulu untuk memastikan kualitas yang baik sebelum digunakan untuk analisis. Tujuan dari Prapemrosesan ini adalah untuk membersihkan dan menyiapkan data teks yang siap untuk di ekstraksi.

Tahapan Preprocessing :

1. *Cleaning*

Seringnya muncul symbol tanda baca dan bilangan angka pada komentar, membuat data menjadi tidak efektif dan tidak memiliki arti. Pada tahapan ini akan menghapus symbol, bilangan angka, dan tanda baca tersebut.

2. *Case Folding*

Data yang digunakan ialah dari komentar di Tiktok, Case Folding adalah proses merubah huruf besar menjadi huruf kecil atau huruf standar. Proses ini mempermudah pencarian karena tidak semua dokumen teks memiliki kapitalisasi yang konsisten.

3. *Tokenization*

Proses pengamanan data set dengan cara mengganti data-data pribadi sensitive dengan data informasi non-sensitif yang setara yang disebut dengan token dengan cara memisahkan kata satu persatu dalam bentuk koma.

4. *Stopwords Removal*

Stopwords Removal adalah proses menghilangkan kata-kata yang kurang penting atau menyesatkan dari kosakata yang tidak mempunyai arti.

5. *Stemming*

Stemming adalah mengurangi jumlah indikator yang berbeda dari informasi setelah ucapan yang diberi akhiran atau awalan ke bentuk dasarnya.

6. Labeling Data

Langkah selanjutnya adalah pelabelan pada kelas sentimen menggunakan kategori positif, kategori Netral dan kategori negatif. Tanggapan tanggapan positif pengguna Tiktok pada ChatGPT. Pada proses pelabelan untuk mempermudah klasifikasi, kelas sentimen positif diganti dengan label 2, kelas sentimen netral diganti dengan 1 sedangkan untuk kelas sentimen negatif 0

7. *Splitting Dataset*

Data *splitting* merupakan proses membagi data menjadi data training (data latih) untuk mengajari sistem dan data *testing* (data uji) untuk menguji sistem. Algoritma *Naïve Bayes* yang digunakan dalam penelitian ini, termasuk ke dalam supervised learning. Sehingga tiap data yang digunakan untuk Training model adalah yang telah diberikan label

8. Wordcloud

Pada Tahap ini, tampilan bobot kata yang dominan muncul, divisualisasikan dalam bentuk wordcloud.

4. Pembobotan Data

Setelah melakukan tahapan *pre-processing* langkah selanjutnya adalah menghitung bobot dari setiap term atau kata berdasarkan frekuensi kemunculan term tersebut dalam dokumen menggunakan metode TF-IDF. TF-IDF sebuah statistic numerik yang dapat menunjukkan kata kunci dengan kata tertentu. Selain dari itu, TF-IDF

juga dikenal efisien, sederhana dan akurat. Adapun rumus pembobotan data TF-IDF adalah sebagai berikut :

$$W_{t,d} = t_{f,t} \times d_{f,t} = \frac{t_{f,t}}{d_{f,t}} \times \log N$$

Keterangan :

$W_{t,d}$: Bobot TF-IDF
$T_{f,t}$: Jumlah frekuensi data
$d_{f,t}$: Jumlah inverse frekuensi dokumen tiap kata
$d_{f,t}$: Jumlah frekuensi dokumen tiap kata
N	: Jumlah total dokumen

5. Evaluasi

Pengujian pada penelitian ini untuk mengevaluasi hasil dari Algoritma yang digunakan. Metode evaluasi yang digunakan *Confusion Matrix*, metode ini cukup membantu untuk proses analisis kualitas *Classifier*. Pengujian ini akan menghitung nilai *accuracy*, *recall*, *precision*, dan *f1-score* yang akan ditampilkan dalam bentuk persentase. Perhitungan confusion matrix untuk menghitung akurasi pengujian, sebagai berikut :

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

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

$$\text{F1 - Score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Keterangan :

TP	: nilai <i>True Positive</i>
FP	: nilai <i>False Positive</i>
FN	: nilai <i>False Negative</i>
TN	: nilai <i>True Negative</i>

2.6 Penelitian Terdahulu

Berikut adalah beberapa ringkasan dari beberapa penelitian sebelumnya yang terkait dengan penelitian ini

Tabel 2.1 Penelitian Terdahulu

No	Judul	Metode	Tujuan	Peneliti	Tahun Penelitian	Hasil
1	“Analisis Senti ment Tingkat Kepua san Pelang gan Terha dap Layan an Kurir J&T Expre ss Mengunak an Algori tma Suppo	Support Vector Machin e (SVM)	Tujuan peneliti an ini adalah untuk menget ahui akurasi dari algorit ma SVM, perform a algorit ma SVM, serta menera pkan analisis sentime nt J&T Express	Indar Suwarsi h, Suzuki Sofian, Joko Tri Handok o, MS Hasibu an, Magiste r IIB Darmaj aya	2023	Hasil analisis sentiment terhadap aplikasi J&T Express, terdapat lebih banyak polaritas negative daripada positif.

	rt Vector Machi ne (SVM) Berda sarkan Ulasa n Pengg una Di Googl e Playst ore”		pada sebuah web dengan model yang sudah dilatih			
2	“Anali sis Senti men Objek Wisat a Bali di Googl e Maps mengg unaka n	Algorit ma Naïve Bayes	Menga nalisis ulasan wisata wan di Google Maps tentang tempat wisata yang direko mendas ikan di	Dian Siti Utami, Adhitia Erfina	Maret 2022	Hasil penelitian ini adalah Nusa Pedina dengan nilai akurasi 94,64% merupaka n objek wisata bali yang paling

	Algoritma Naïve Bayes”		Bali menggunakan Algoritma Naïve Bayes			banyak dikunjungi dan nilai akurasi nya lebih unggul dari Garuda Wisnu Kencana dengan nilai akurasi 82,86% maka Wisata Nusa pedina adalah rekomendasi tempat wisata di Bali berdasarkan hasil akhir dari penelitian ini.
--	------------------------	--	--	--	--	--

3	“Analisis Sentimen terhadap opini masyarakat tentang vaksin Covid-19 menggunakan Algoritma Naïve Bayes Classifier”	Naïve Bayes Classifier	Penelitian ini bertujuan untuk menganalisis pendapat tentang vaksinasi Covid-19 di Indonesia	Winda Yulita, Eko Dwi Nugroho, Muhammad Habib Algirafi	2021	Penelitian ini dilakukan terhadap data 3780 tweets, berdasarkan hasil analisis sebagian besar tweets memiliki sikap positif (60,3 %), sementara jumlah tweets yang netral (34,4%) melebihi jumlah tweets yang menentang (5,4 %). Nilai akurasi yang dihasilkan
---	--	------------------------	--	--	------	--

						sebesar 0,93 (93 %)
4	“Analisis Sentiment pembelajaran Darling pada Twitter di masa Pandemi Covid-19 menggunakan Metode Naïve Bayes	Naïve Bayes	Penelitian ini bertujuan menganalisis opini public terhadap pembelajaran daring pada masa pandemi covid-19 di indonesia	Samsir, Ambiyar, Unung Verwardina, Firman Edi, Ronal Wariant hos	Januari 2021	Hasil dari penelitian ini yaitu memiliki 30 % sentiment positif, 69 % sentiment negative dan 1% sentiment netral
5	“Analisis sentiment Review	Naïve Bayes, K-Nearest	Penelitian ini bertujuan untuk	Dian Ardiansyah, Atang Saepudi	Juli, 2023	Hasil dari penelitian ini yaitu mendapat nilai

	w pada aplika si media social Tiktok mengg unaka n Algori tma KNN dan SVM berbas is PSO”	Neighb or	mengan alisis sentime nt penggu na media social pada suatu produk dan mempa ndingka n kedua metode tersebut	n, Riska Aryanti , Royadi		akurasi KNN sebesar 83,40% dan Naïve bayes 69,20 %
6	“Anali sis sentim ent aplika si tiktok mengg unaka n algorit ma	Naïve Bayes dan SVM	Peneliti an ini bertuju an untuk mengan alisis koment ar masyar akat terhada	Friska Aditia Indriya ni, Ahmad Fauzi, Sutan Faisal	Juli 2023	Hasil dari penelitian ini mendapat akurasi sebesar 76.7% berlabel positif, 23.3% negatif, dan

	naïve bayes dan SVM		p aplikasi Tiktok			accuracy sebesar 79% untuk metode Naïve Bayes dan metode SVM sebesar 84 % lebih tinggi dibanding kan Naïve Bayes
7	“Analisis Senti ment tanggan Masya rakat Aplika si Tiktok menggunaka n Metode	Naïve Bayes, Categori Propotio nal Differe nce	Tujuan dari peneliti an ini adalah untuk memba ntu para orangtu a dalam pemilih an platfor m yang baik untuk	Junda Alfiah Zulqorn ain, Indriati, Putra Pandu Adikara	Juli 2021	Hasil dari penelitian ini adalah mendapat nilai accuracy 72 % dan 74%

	Naïve Bayes dan CPD		anak-anakny a			
8	“Analisis Senti ment pengg una Media Sosial terhad ap aplika si M-Health peduli lindun gi denga n metod e lexico n based dan naïve bayes	Lexico n Based dan Naïve Bayes	Tujuan peneliti an ini ialah mengan alisis opini public terhad ap aplikasi m-health untuk mening katkan kepercayaan pada Masyarakat	Riky Iskanda r Syah, Horiya h, Miftahu l Wahid	Juni, 2023	Hasil penelitian ini yaitu masyarakat cenderung memiliki opini positif sedangkan untuk perbandin gan metode nya Naïve Bayes kurang mampu mengklasi fikasi kan data tersebut

9	“Analisis Senti ment tanggan pan masya rakat terhad ap Tiktok shop dan Shope e di Twitte er mengg unaka n metod e Naïve Bayes dan K-Neare st Neigh bor”	Naïve Bayes dan K-Nearest Neighbor	Peneliti an ini bertuju an untuk menent ukan classifi er mana yang membe rikan hasil accurac y terbaik	Sevira Chairun nisa Nandar esta, Cecep Warman	Agustus, 2023	Hasil dari penelitian ini yaitu nilai akurasi Naïve Bayes pada aplikasi shopee yaitu 99.06% dan pada Tiktok yaitu 92.94% sedangkan pada KNN mendapat nilai akurasi shopee sebanyak 83 % dan Tiktok sebanyak 81 %
---	---	------------------------------------	--	---	---------------	--

10	“Analisis Sentiment Review penggunaan Tiktok melalui pendekatan Algoritma Naïve Bayes”	Naïve Bayes	Penelitian ini bertujuan menganalisis terhadap p tinjauan pengguna Tiktok dengan menggunakan pendekatan Algoritma	Nur Aprilia Ni, Nana Suarna, Willy Priharto no	Desember 2023	Hasil dari penelitian ini yaitu memperoleh nilai akurasi sebesar 80%, Precision sebesar 79% Recall sebesar 74% dan F1-Score sebesar 75%
----	--	-------------	---	--	---------------	---