

# Analisis Sentimen Penggunaan *Paylater* pada Generasi Z di Twitter menggunakan Metode *Naïve Bayes Classifier*

**M. BUCCI RYANDO<sup>1</sup>, MUCHAMAD IQBAL<sup>2</sup>, KURNIA SYAHIDAH<sup>3</sup>**

<sup>1</sup>Teknik Informatika, Institut Teknologi dan Bisnis Bina Sarana Global, Indonesia

<sup>2,3</sup>Sistem Informasi, Institut Teknologi dan Bisnis Bina Sarana Global, Indonesia

Email: bucci@global.ac.id

## ABSTRAK

Popularitas Paylater sebagai metode pembayaran e-commerce pasca-pandemi di Indonesia meningkat pesat, namun disertai risiko utang konsumtif, khususnya di kalangan Generasi Z. Penelitian ini bertujuan menganalisis sentimen dan opini publik guna memahami faktor yang memengaruhi keputusan adopsi Paylater. Dengan menggunakan metode analisis sentimen berbasis Naïve Bayes Classifier terhadap data Twitter (kini X), penelitian ini mengklasifikasikan tanggapan masyarakat terhadap penggunaan Paylater. Model yang dibangun divalidasi dengan nilai F1-score 0.432 dan Precision 0.508. Hasil analisis menunjukkan mayoritas sentimen (50,75%) bersifat netral atau ambigu, mencerminkan adanya keraguan publik terhadap penggunaan layanan ini. Selain itu, ditemukan dominasi sentimen negatif yang menyoroti isu peningkatan utang, kesulitan mengelola keuangan, serta ketergantungan terhadap fasilitas kredit konsumtif. Penelitian ini berkontribusi dalam pemanfaatan text mining untuk memetakan persepsi Generasi Z terhadap adopsi Paylater, sehingga hasilnya dapat menjadi dasar bagi perusahaan fintech dalam merumuskan strategi pemasaran yang lebih bijak dan bertanggung jawab.

**Kata kunci:** paylater, analisis sentimen, text mining, generasi z, naïve bayes classifier.

## ABSTRACT

*The popularity of Paylater as an e-commerce payment method in post-pandemic Indonesia has grown rapidly but is accompanied by the risk of consumptive debt, particularly among Generation Z. This study aims to analyze public sentiment and opinions to understand the factors influencing Paylater adoption decisions. Using a Naïve Bayes Classifier-based sentiment analysis method on Twitter (now X) data, this research classifies public responses toward Paylater usage. The developed model was validated with an F1-score of 0.432 and Precision of 0.508. The results indicate that the majority of sentiments (50.75%) are neutral or ambiguous, reflecting public uncertainty toward the service. In addition, dominant negative sentiments were identified, highlighting issues such as increasing debt, financial management difficulties, and dependency on credit facilities. This study contributes to the use of text mining in mapping Generation Z's perceptions of Paylater adoption, providing insights that can help fintech companies develop more responsible and ethical marketing strategies.*

**Keywords:** paylater, sentiment analysis, text mining, z generation, naïve bayes classifier.

## 1. PENDAHULUAN

Dalam era digital yang semakin maju, perkembangan teknologi keuangan telah memunculkan beragam inovasi untuk mempermudah dan meningkatkan kualitas layanan keuangan. Salah satu inovasi tersebut adalah layanan *Paylater*, yang memungkinkan pengguna untuk melakukan pembelian sekarang dan membayar nanti. Kepraktisan dan kemudahan dalam menggunakan layanan *Paylater* telah menarik minat masyarakat, baik individu maupun pelaku bisnis (**Hozairi dkk., 2021; Indriani & Wibowo, 2022**). Pada kenyataannya, kehadiran teknologi *Paylater* ini juga membawa beberapa permasalahan yang perlu diperhatikan secara mendalam. Salah satu permasalahan penting yang menjadi fokus dalam penelitian ini adalah sejauh mana opini masyarakat dapat mempengaruhi keputusan individu dalam menggunakan layanan *Paylater* (**Arisandy dkk., 2023**). Dengan meningkatnya popularitas media sosial dan platform berbagi opini, pengaruh masyarakat dalam pembuatan keputusan semakin penting untuk dipahami. Pengguna metode pembayaran *Paylater* di Indonesia meningkat selama pandemi. Sebanyak 55% dari konsumen yang menyatakan pernah menggunakan *Paylater*, baru menggunakannya saat pandemi. (**Wiguna & Rifai, 2021**)

Pada penelitian sebelumnya dengan judul penelitian Analisis Sentimen Masyarakat Terhadap *Paylater* menggunakan Metode *Naïve Bayes Classifier*, yang dimana dari hasil pengumpulan dataset yang didapat sebanyak 405 data mengenai *Paylater*, dapat disimpulkan bahwa masyarakat merasa kurang baik dengan penggunaan *Paylater* dibuktikan hasil analisis sentimen yang lebih banyak mendapatkan sentimen negatif dari pada positif. Dengan penggunaan algoritma *Naïve Bayes* dari pustaka *TextBlob* yang diterapkan pada pengujian model menggunakan *confussion matrix* menghasilkan bahwa tingkat akurasi yang lebih besar yaitu 95% jika dibandingkan dengan menggunakan pustaka *TextBlob* yang hanya senilai 61% (**Nugroho dkk., 2016**). Sedangkan pada penelitian ini memiliki batasan yang jelas dalam proses pengklasifikasian dampak dari penggunaan *Paylater*. Metode yang digunakan dalam penelitian melibatkan observasi terhadap perilaku pengguna, pengumpulan data melalui proses *crawling*, dan penerapan *Naïve Bayes Classifier* untuk mengklasifikasikan pandangan masyarakat terhadap layanan *Paylater*. Selain itu pada penelitian sebelumnya dengan judul penelitian Analisis Sentimen *Twitter* menggunakan *Text mining* dengan Algoritma *Naïve Bayes*. Dengan hasil dari penelitian ini yang menggunakan algoritma *Naïve Bayes Classifier* yakni sangat efektif untuk digunakan sebagai proses klasifikasi tweet yang dibutuhkan dalam sistem analisis sentimen ini dimana nilai yang di dapatkan dalam pengujian sampai 84%. Metode *Naïve Bayes Classifier* dapat digunakan untuk melakukan klasifikasi *tweet* dengan cukup baik pada sistem analisis sentimen. 300 data yang dibagi menjadi 2 yaitu data latih sebanyak 200 data dan data uji berjumlah 100 data. Hasil dari klasifikasi diketahui bahwa 100 data yang diuji masuk dalam kategori sentimen negatif (**Safira & Hasan, 2023**). Berdasarkan hasil uji pada penelitian sebelumnya, dengan judul penelitian *Analysis of Sentiment Adiraku App Reviews on Google Play Store Using Vector Machine Support Algorithm and Naïve Bayes*, aplikasi Adiraku di Google Play Store mendapatkan ulasan positif yang lebih banyak (1412 ulasan) dibanding ulasan negatif (588 ulasan) dari total 2000 ulasan yang diperoleh menggunakan metode Web Scraping. Hasil uji juga menunjukkan bahwa algoritma *Support Vector Machine* lebih unggul daripada *Naïve Bayes*, dengan mencapai nilai akurasi masing-masing sebesar 96% dan 85%. Dalam klasifikasi ini, digunakan 2000 titik data dari Web Scraping, yang dibagi menjadi 1400 data pelatihan dan 600 data pengujian. Dapat disimpulkan bahwa algoritma *Support Vector Machine* memiliki akurasi yang signifikan lebih baik dalam studi ini dibandingkan dengan *Naïve Bayes* (**Sudiantoro & Zuliarso, 2018**). Pada penelitian ini penulis juga menggunakan algoritma *Naïve Bayes Classifier*, guna mengklasifikasikan tweet pada sistem analisis sentimen (**Sammut & Webb, 2017**). Hasil klasifikasi pada penelitian *Sentiment Analysis of Presidential Candidates Anies Baswedan and Ganjar Pranowo Using Naïve Bayes Method* menggunakan

algoritma *Naïve Bayes* pada dataset Berita CNN Indonesia tentang Pencalonan Presiden dalam Pengambilan dan Pembersihan Data. Memberikan label sebagai berikut: 1 (sangat negatif), 2 (negatif), 3 (netral), 4 (positif), dan 5 (sangat positif). Label-label tersebut akan menentukan penggunaan tokenisasi yang menggunakan Quadgram yang memiliki akurasi 42,10%, presisi 42,70%, dan mempertimbangkan serta menghasilkan sentimen tertinggi, yaitu sentimen negative. (**Saputra dkk., 2022**)

Pada penelitian ini penulis menggunakan dua *software* yang digunakan untuk analisis *text mining* yaitu *RapidMiner* dan *Orange data mining*. Yang dimana kedua *software* tersebut menerapkan kerangka umum dalam menganalisis teks yang diperoleh dari media sosial (**Agrani & Rikumahu, 2020**). *Preprocessing data* merupakan tahapan yang dilakukan sebelum melaksanakan analisis teks, *preprocessing data* itu sendiri merupakan sebuah proses mempersiapkan data sebelum dioleh menjadi sebuah informasi. *Preprocessing* juga merupakan suatu proses mengolah data mentah menjadi data yang lebih berkualitas (**Padilah dkk., 2023**). Dalam analisis yang pertama pada penelitian ini penulis menggunakan *Rapidminer* sebagai *software* yang digunakan untuk *Preprocessing* data sekaligus pengolahan data hingga memperoleh *output* berupa visualisasi data analisis sentimen. Dalam analisis yang kedua pada penelitian ini penulis menggunakan *Orange Data Mining* sebagai *software* yang juga digunakan untuk *processing data* namun dengan langkah yang lebih singkat dan tetap relevan (**Rofiqoh dkk., 2017**). Sedangkan untuk metodologi pengembangan yang digunakan pada penelitian ini menggunakan model CRISP-DM (**Iqbal dkk., 2022**) dengan tahapan-tahapan yang relevan dalam analisis sentimen. Selain itu dalam penelitian ini digunakan metode *Naïve Bayes* didasarkan pada *Theorema Bayes*, yang menyatakan bahwa probabilitas suatu kejadian dapat dihitung berdasarkan probabilitas kondisional dari kejadian tersebut dan probabilitas dari kejadian yang terkait. *Theorema Bayes* sendiri merupakan model matematika dengan dasar statistik dan probabilitas. *Naïve* dalam metode ini merujuk pada asumsi yang dibuat bahwa setiap fitur atau atribut yang digunakan untuk klasifikasi adalah independen satu sama lainnya, meskipun kenyataannya tidak selalu demikian, metode *Naïve Bayes* masih sering digunakan karena memiliki keuntungan dalam pengolahan data yang besar dan dapat menghasilkan hasil klasifikasi yang cukup akurat dalam waktu yang singkat. (**Hadna dkk., 2016**)

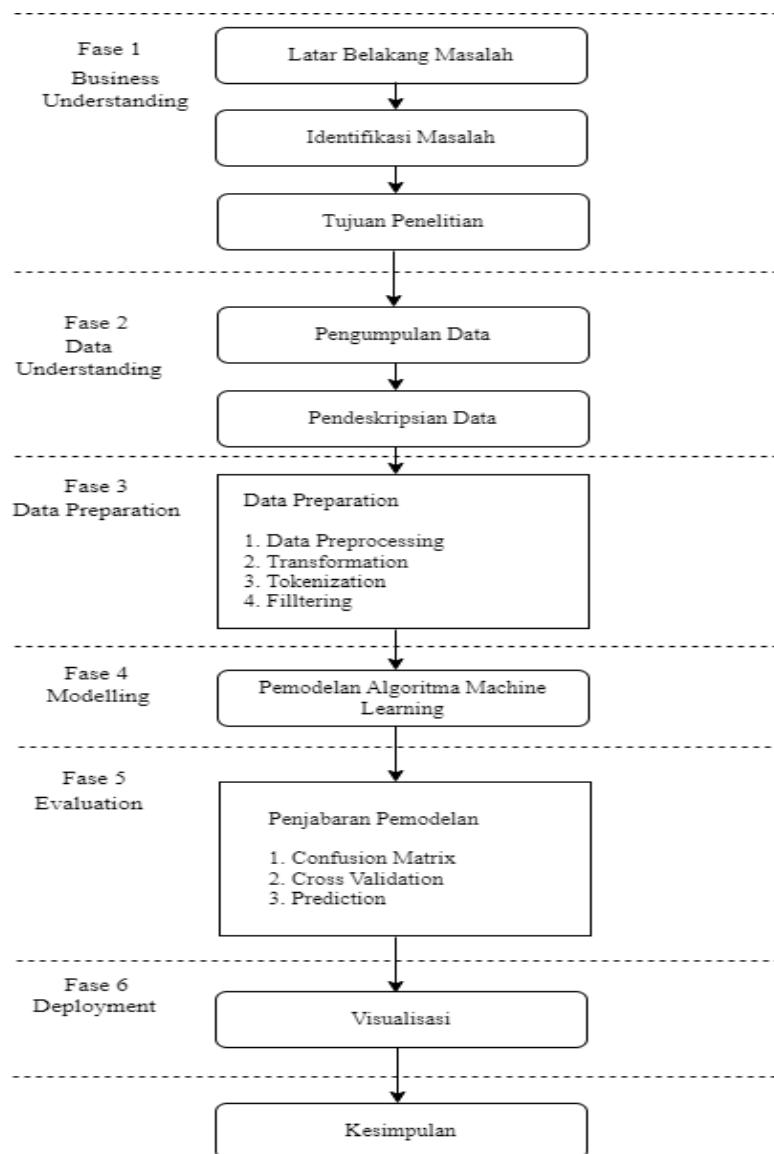
Sebagai kontribusi penting bagi penelitian-penelitian sebelumnya, penelitian ini menawarkan analisis mendalam, pengamatan yang cermat, serta proses klasifikasi data yang lebih terfokus. Proses *preprocessing* data dan data *cleaning* yang dilakukan juga akan memastikan integritas dan keakuratan hasil penelitian. Penelitian ini berupaya memberikan pemahaman yang lebih komprehensif tentang peran opini masyarakat dalam penggunaan teknologi keuangan yang inovatif, sehingga dapat memberikan panduan yang lebih baik bagi pengambilan keputusan di masa depan. Dengan demikian, diharapkan bahwa hasil dari penelitian ini dapat memberikan sumbangan berharga bagi perkembangan industri *fintech*, serta memberikan pemahaman yang lebih dalam tentang pola perilaku pengguna dalam menggunakan layanan *Paylater*. Selain itu, penelitian ini juga berpotensi memberikan wawasan baru dalam menerapkan metode analisis *text mining* untuk memahami pendapat dan sentimen masyarakat dalam konteks teknologi keuangan.

Penelitian ini bertujuan untuk menganalisis seberapa besar peran opini masyarakat dalam mempengaruhi keputusan individu dalam menggunakan layanan *Paylater*. Dalam konteks ini, pendekatan analisis *text mining* akan digunakan untuk mengidentifikasi dan memahami sentimen serta pandangan pengguna terkait dengan layanan ini (**Prasastio & Heriyanto, 2022**). Dengan adanya media sosial, banyaknya data yang ada berupa gambar, *comment*, *text* atau *emoticon*, video dan lainnya tersebut, masyarakat bebas beropini. Dengan adanya

analisis sentimen opini yang berkembang dan banyak di media sosial tersebut dapat menghasilkan data dan informasi yang bermanfaat. Analisis sentimen berperan sebagai alat yang dapat menghubungkan seluruh data tersebut. Dengan begitu, perusahaan dapat memperoleh masukan inti dari pengguna atau konsumen secara efisien. Saat ini, perusahaan memiliki data teks dalam volume besar seperti email, salinan obrolan dukungan pelanggan, komentar media sosial, dan ulasan. (**Hadianti & Yosep Tember, 2022**)

## 2. METODE PENELITIAN

Metodologi pengembangan yang digunakan pada penelitian ini menggunakan model CRISP-DM dengan tahapan-tahapan yang relevan dalam analisis sentimen.



**Gambar 1. Flowchart Analisis Sentimen Dampak Penggunaan Paylater**

Dalam tahapan-tahapan tersebut dijelaskan (**Ryando & Sutarmen, 2024**):

**Business understanding (Pemahaman Bisnis):** Pada tahap ini berfokus pada pemahaman terhadap tujuan bisnis atau masalah yang ingin diselesaikan melalui analisis data.

Tujuan dari tahap ini adalah untuk mengidentifikasi masalah atau peluang yang perlu diatasi melalui analisis data.

**Data Understanding (Pemahaman Data):** Pada tahap ini, fokusnya adalah untuk memahami dan mengumpulkan data yang tersedia. Pengumpulan data dilakukan dengan cara menggunakan metode *crawling* dan didapatkan sebanyak 7057. Selain itu, tahapan ini melibatkan pengeksplorasi data, evaluasi kualitas data, dan pemahaman lebih lanjut tentang variabel dan hubungan antar variabel dalam data.

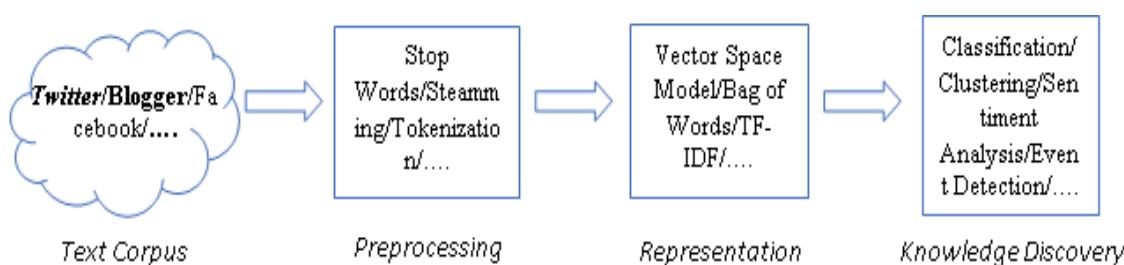
**Data Preparation (Persiapan Data):** Tahap ini melibatkan pra-pemrosesan data, transformasi, dan pembersihan data. Tugas-tugas dalam tahap ini termasuk pemilihan variabel yang relevan, penanganan *missing data*, penghapusan *outliers*, penggabungan data dari sumber yang berbeda, dan penyesuaian format data. (**Arisandy dkk., 2023**)

**Modeling (Pemodelan):** Pada tahap ini, model atau teknik analisis yang sesuai dipilih dan diterapkan pada data yang telah dipersiapkan sebelumnya. Ini melibatkan penggunaan algoritma analisis data untuk membangun model yang dapat memecahkan masalah bisnis yang diidentifikasi.

**Evaluation (Evaluasi):** Tahap ini bertujuan untuk mengevaluasi kualitas dan kinerja model yang telah dibangun. Evaluasi melibatkan pengujian model menggunakan data yang tidak digunakan dalam proses pembangunan model, serta penilaian terhadap hasil model berdasarkan kriteria bisnis yang ditentukan sebelumnya.

**Deployment (Implementasi):** Tahap terakhir adalah implementasi dan penggunaan model yang telah dievaluasi. Hasil dari analisis data diterapkan dalam lingkungan bisnis dan digunakan untuk mendukung pengambilan keputusan atau solusi dari masalah yang ditargetkan.

Pada penelitian ini penulis menggunakan dua *software* yang digunakan untuk analisis *text mining* yaitu *RapidMiner* dan *Orange data mining*. Yang dimana kedua *software* tersebut menerapkan kerangka umum dalam menganalisis teks yang diperoleh dari media sosial. *RapidMiner* dipilih karena menawarkan *visual workflow* yang kuat untuk manajemen proses *text mining* kompleks dan validasi model tingkat lanjut (F1-score, Precision), menjadikannya ideal untuk riset terstruktur. Sementara itu, *Orange Data Mining* digunakan karena antarmuka *user-friendly* dan fokusnya pada visualisasi data yang cepat dan eksploratif, memungkinkan pengujian dan eksperimen algoritma Naïve Bayes secara interaktif. Keduanya memfasilitasi implementasi analisis sentimen *no-code*.



**Gambar 2. Alur Analisis Text Mining**

**Acquiring Text:** atau dalam memperoleh data teks penulis menggunakan metode *Crawling*. Yang dimana *Crawling* sendiri merupakan sebuah cara untuk mengumpulkan data pada era big data saat ini. Biasanya *Crawling* dilakukan pada web atau sosial media (*Twitter*, *facebook*,

blog, dan yang lainnya). Biasanya teknik *Crawling* digunakan dengan menggunakan *Application Programming Interface (API)*, sebagai jalur komunikasi untuk mendapatkan data.

**Preprocessing:** Merupakan sebuah proses mempersiapkan data sebelum dioleh menjadi sebuah informasi. *Preprocessing* juga merupakan suatu proses mengolah data mentah menjadi data yang lebih berkualitas. Pada persiapan data, penulis beracuan pada tahapan-tahapan sebagai berikut:

- a. **Tokenization:** merupakan proses untuk mendapatkan dengan cara memilah semua kata yang ada pada rangkaian karakter. Kumpulan data tersebut dalam semua dokumen pada sebuah koleksi disebut *dictionary* atau *vocabulary*.
- b. **Normalization:** proses mentransformasi data teks kedalam kata pada umumnya.
- c. **Stop Word Removal:** proses menghilangkan kata tidak penting dalam data teks, biasanya berupa *spam*, atau kata yang tidak terlalu diperlukan.

Selain itu dalam penelitian ini digunakan metode *Naïve Bayes* didasarkan pada *Theorema Bayes*, yang menyatakan bahwa probabilitas suatu kejadian dapat dihitung berdasarkan probabilitas kondisional dari kejadian tersebut dan probabilitas dari kejadian yang terkait. *Theorema Bayes* sendiri merupakan model matematika dengan dasar statistik dan probabilitas. *Naïve* dalam metode ini merujuk pada asumsi yang dibuat bahwa setiap fitur atau atribut yang digunakan untuk klasifikasi adalah independen satu sama lainnya, meskipun kenyataannya tidak selalu demikian, metode *Naïve Bayes* masih sering digunakan karena memiliki keuntungan dalam pengolahan data yang besar dan dapat menghasilkan hasil klasifikasi yang cukup akurat dalam waktu yang singkat.

### 3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengetahui seberapa besar pengaruh opini masyarakat terhadap keputusan individu dalam menggunakan layanan *Paylater*. Proses yang digunakan. Proses pengolahan data teks atau analisis sentimen yang dilakukan pada penelitian ini menggunakan *Data Mining* proses yaitu CRISP-DM, yang mana alur yang dilakukan pada penelitian ini mengacu pada Gambar 1 yang dibagi menjadi beberapa fase untuk dapat mendeskripsikan secara lengkap proses yang terjadi dalam analisis *sentiment* terhadap penggunaan *Paylater*, mulai dari pemahaman bisnis, data yang didapatkan, pengolahan data, proses analisis sentimen menggunakan algoritma *Naïve Bayes Classifier*, mengevaluasi hasil pemodelan, sampai akhirnya proses *deployment* hasil dari seluruh fase yang dilakukan.

#### 3.1 Business Understanding

Pada tahap ini pemrosesan data dimulai dari pemahaman data atau menterjemahkan data sesuai sasaran dan pembatasan data dengan rumusan masalah dalam *Data Mining*. Setelahnya diikuti dengan mempersiapkan strategi awal untuk mencapai tujuan-tujuan penelitian. Dan terakhir merancang apa yang akan dibangun.

#### 3.2 Data Understanding

Pada tahap *data understanding* dimulai dengan pengumpulan data awal kemudian melakukan identifikasi dan eksplorasi data. Mengevaluasi kualitas data. Tahapan ini dimulai dari cara mendapatkan data:

**Pengumpulan Data:** Dalam hal ini penulis mendapatkan data dengan menggunakan metode *crawling*. *Crawling* didefinisikan sebagai sebuah teknik pengumpulan data pada sebuah

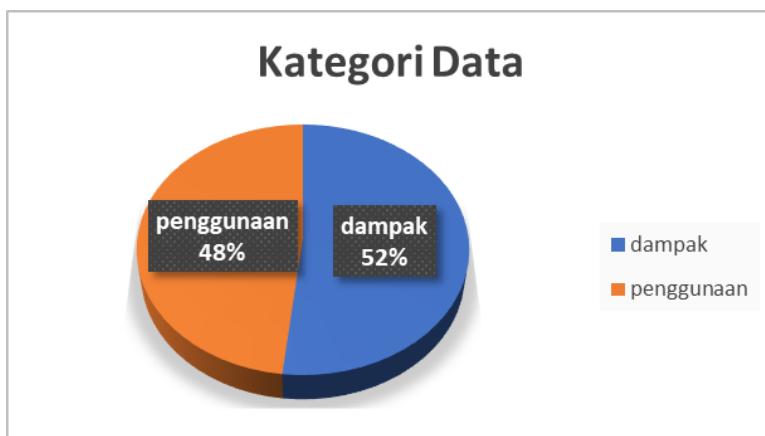
website dengan memasukkan *Uniform Resource Locator* (URL). *Crawling* data dapat dilakukan untuk mengumpulkan berbagai jenis data dari *internet*. Pada penelitian kali ini penulis menggunakan internet media sosial *Twitter* (sekarang *X*) sebagai sumber mendapatkan data teks.

"Created-At"
(ala-ala personal loan/cash advance CC)
melalui merchant yang offer servis ni. 10%-15% caj diaorang ni.
Teringat zaman student jual Baucar buku dulu sebab nak cash."
CAPCOM FIGURE BUILDER MONSTER HUNTER STANDARD MODEL P...
Available
Paylater ?
? <a href="https://t.co/fYDYYQ6Zfc">https://t.co/fYDYYQ6Zfc</a>
? Spaylater 7.223.000
? Lazada paylater 6.000.000
? Buns @solelggy open gestun:
— spaylater, gopaylater / cicil, kredivo, hci, ceria, lazada, akulaku, yup, dll
? <a href="https://t.co/5FtCcAG67w">https://t.co/5FtCcAG67w</a>

**Gambar 3. Data Teks**

Gambar 3 adalah representasi dari data Twitter yang mengandung Opini/Sentimen yaitu Pengalaman pengguna dan kritik terhadap biaya Paylater. Kedua, Konstruksi Sosial yaitu bukti adanya praktik "Gestun" (gesek tunai) menggunakan limit Paylater. Ketiga, Data Bisnis/Penawaran yaitu ketersediaan dan penerimaan berbagai merek Paylater sebagai metode pembayaran. Data ini sangat berguna untuk analisis sentimen karena menunjukkan sentimen negatif/waspada (terkait biaya/gestun) dan juga netral/informasi (terkait ketersediaan layanan).

Data yang diperoleh menggunakan metode *crawling* setelah dikumpulkan terdapat 7057 data. Data tersebut dibagi menjadi dua kategori data yaitu penggunaan dan dampak. Data penggunaan sebanyak 3401 data dan data dampak sebanyak 3655 data. Data tersebut akan melalui proses *Preprocessing* yang mana kemungkinan besar datanya akan berkurang untuk meminimalisir *duplicate* data. Persentase penggunaan data dapat dilihat pada gambar 4.



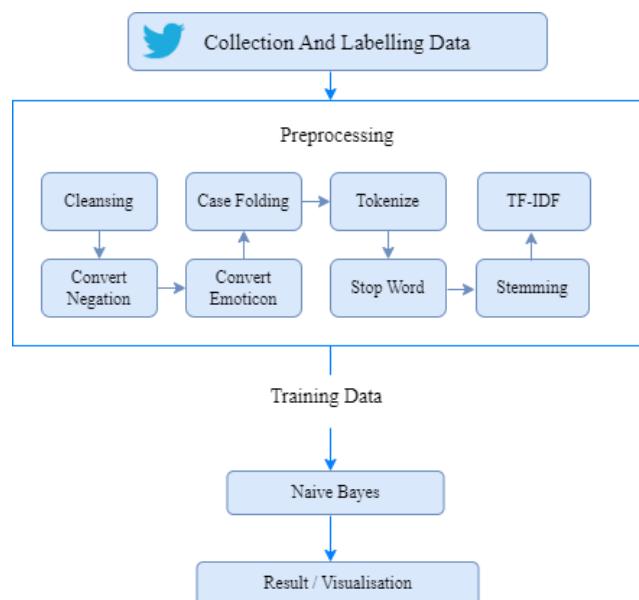
**Gambar 4. Persentase Penggunaan Data**

### 3.3 Data Preparation

Tahap ini melibatkan pra-pemrosesan data, transformasi, dan pembersihan data. Tugas-tugas dalam tahap ini termasuk pemilihan variabel yang relevan, penanganan *missing data*, penghapusan *outliers*, penggabungan data dari sumber yang berbeda, dan penyesuaian format data. *Preprocessing data* merupakan tahapan yang dilakukan sebelum melaksanakan analisis teks, dalam analisis yang pertama pada penelitian ini penulis menggunakan *Rapidminer* sebagai *software* yang digunakan untuk *Preprocessing* data sekaligus pengolahan data hingga memperoleh *output* berupa visualisasi data *sentiment* analisis.

### 3.4 Modelling (Pemodelan)

Alur analisis sentimen menggunakan RapidMiner dapat dilihat pada gambar 5.



Gambar 5. Diagram Sistem analisis sentimen RapidMiner

**Tokenize:** pada proses *Preprocessing* pada aplikasi *Rapidminer* ini berfungsi sebagai proses untuk membagi teks yang dapat berupa dokumen, kalimat atau paragraf, menjadi bagian-bagian tertentu.

Tabel 1. *Tokenize*

Kalimat Sebelum	Kalimat Sesudah
"Berharap banget sama gaji bulan ini untuk bayar Paylater"	"berharap", "banget", "sama", "gaji", "bulan", "ini", "untuk", "bayar", "Paylater"

**Transform cases:** Dalam *Preprocessing* penulis akan menggunakan operator Text Mining yang mana didalam nya ada operator dengan label *Transform cases* yang berfungsi untuk mengubah kapitalisasi huruf karakter dalam dokumen.

**Tabel 2. *Transform Cases***

Kalimat Sebelum	Kalimat Sesudah
HRD Mau <i>Resign</i> Tapi Masih Ada Tunggakan <i>GoPaylater</i>	hrd mau <i>resign</i> tapi masih ada tunggakan <i>goPaylater</i>

**Filter Stopword:** Dengan operator ini maka teks sebelum di klasifikasikan di hilangkan dulu teks yang tidak berhubungan dengan analisa *sentiment* sehingga dimensi teks akan berkurang tanpa mengurangi isi *sentiment* dari teks tersebut. *Filter stopword* bahasa Indonesia ini penulis ambil dari internet yang bersumber dari *Kaggle platform* pusat data yang biasa digunakan untuk penelitian: [Indonesian Stoplist | Kaggle](#)

**Filter Tokens by length:** Pada operator ini penulis melakukan filterisasi berdasarkan panjangnya kalimat sesuai dengan kriteria yang ditentukan.

### 3.5 Evaluation (*Evaluasi*)

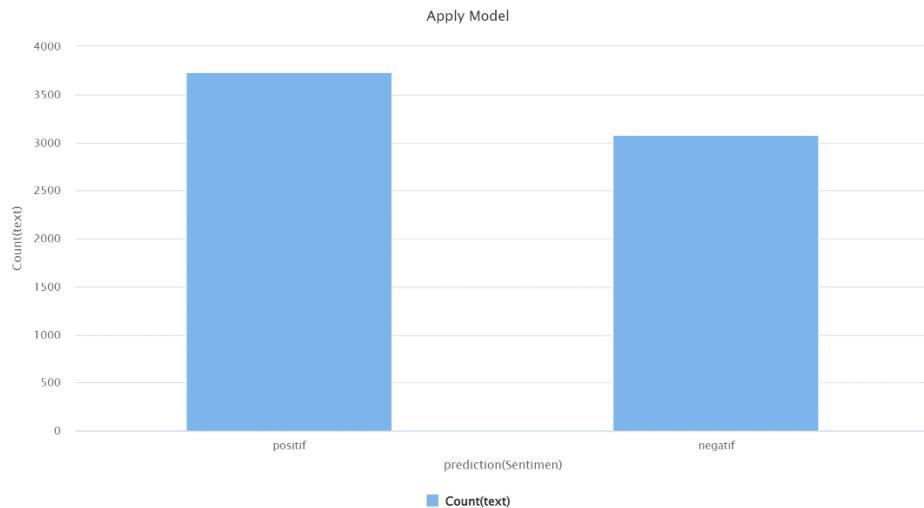


**Gambar 6. *WorldCloud Visualization***

Dalam dataset ini yang berisi ulasan pelanggan tentang penggunaan *Paylater*, *wordcloud* ini memberikan gambaran visual tentang kata-kata yang paling umum digunakan dalam ulasan tersebut. Ukuran kata dalam *wordcloud* mewakili frekuensi kemunculannya, dengan kata-kata yang lebih sering muncul memiliki ukuran yang lebih besar.

Dalam hal ini *wordcloud* tersebut memberikan gambaran visual awal dan tidak memberikan informasi tentang hubungan antara kata-kata atau konteks di mana kata-kata tersebut digunakan.

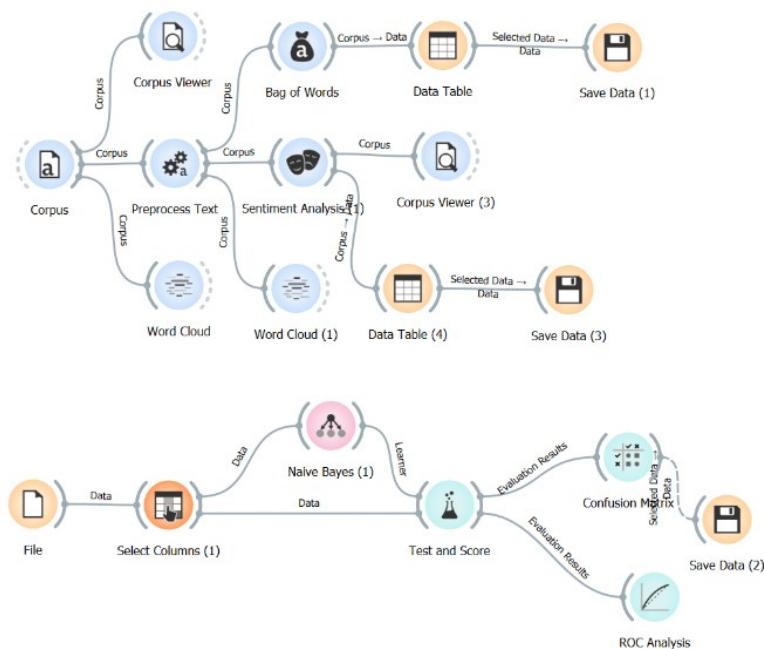
### 3.6 Deployment



**Gambar 7. Visualisasi hasil penggunaan Paylater**

Dalam analisis sentimen, menggunakan Rapidminer dengan metode *Naïve Bayes Classifier* yang digunakan untuk mengklasifikasikan teks ke dalam kategori sentimen positif atau negatif berdasarkan probabilitas masing-masing kata atau fitur dalam teks tersebut. Dalam konteks ini, hasil visualisasi diagram batang menunjukkan bahwa jumlah teks yang dianalisis dengan sentimen positif lebih besar daripada sentimen negatif.

Alur analisis sentimen *Orange Data Mining* dapat dilihat pada gambar.



**Gambar 8. Alur Proses Analisis Sentimen *Orange Data Mining***

**Transformation:** Hal pertama yang dilakukan dalam preprocess text adalah transformation yang merupakan proses mengubah data input untuk transformasi huruf kecil secara *default*. *Lowercase* berfungsi untuk mengubah semua huruf dalam teks menjadi huruf kecil. *Remove*

*Accents* untuk menghilangkan tanda akses dalam teks. *Parse html* untuk mengetahui tag html dan mem-parsing teks. *Remove url* untuk menghapus url dari teks.

**Tokenization:** Setelah data dibersihkan, langkah selanjutnya adalah memisahkan teks menjadi kata-kata individual atau token. Biasanya proses ini disebut tokenisasi. Pada bagian ini penulis menggunakan widget "*Tokenization*" pada Orange Data Mining untuk melakukan tokenisasi pada data teks tersebut.

**Filtering:** Dalam teks, ada kata-kata yang umum atau tidak memiliki makna yang signifikan, seperti kata hubung (seperti "atau", "dan") atau kata umum (seperti "dari", "di"), pada tahap ini widget yang penulis gunakan adalah "stopwords"



**Gambar 9. Wordcloud sebelum Preprocessing**

Deskripsi data sebelum *Preprocessing* ini memberikan pemahaman awal tentang karakteristik dan kondisi data mentah sebelum dilakukan proses pengolahan lebih lanjut untuk mempersiapkan data untuk dilakukan analisis sentimen.

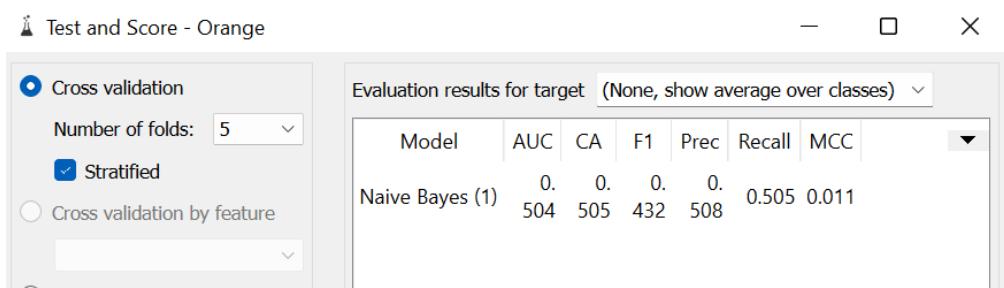
Setelah dilakukan *Preprocessing*, data mengalami beberapa perubahan dan pemrosesan untuk mempersiapkan data tersebut agar lebih siap digunakan. Terlihat perubahan dari data yang telah dilakukan *preprocessing*: penanganan, *missing values*, penanganan *outliers*, pembersihan data, penghapusan duplikat



**Gambar 10.** *WordCloud* setelah *Preprocessing*

*Preprocessing* akan sangat tergantung pada tujuan analisis dan jenis data yang dimiliki. Setelah dilakukan *Preprocessing*, data siap untuk digunakan dalam analisis lebih lanjut.

Hasil dari nilai akurasi berdasarkan *Test and Score* dari *Cross validation* atau bisa diartikan seberapa akurat algoritma yang digunakan dalam memprediksi data yang didapatkan.



**Gambar 11.** *Cross validation* menggunakan Orange Data Mining

**Tabel 3.** Hasil *Cross validation*

Model	AUC	CA	F1	Precision	Recall	MCC
Naïve Bayes	0.504	0.505	0.432	0.508	0.505	0.011

Berikut pengertian dari hasil metrik evaluasi yang digunakan dalam *cross-validation*:

**AUC:** *Area Under the Curve* (AUC) adalah metrik evaluasi yang mengukur sejauh mana model dapat membedakan antara kelas positif dan negatif dalam kurva *Receiver Operating Characteristic* (ROC).

**CA:** *Classification Accuracy* (CA) adalah metrik evaluasi yang mengukur akurasi secara keseluruhan dari model klasifikasi. Ini merupakan perbandingan jumlah prediksi yang benar dengan jumlah total sampel. CA memberikan gambaran tentang sejauh mana model dapat memprediksi dengan benar kelas-kelas target.

**F1:** *F1-score* adalah metrik evaluasi yang menggabungkan presisi (*precision*) dan recall (*recall*) menjadi satu angka tunggal. *F1-score* memberikan ukuran seimbang antara presisi dan recall, dan umumnya digunakan ketika kelas target tidak seimbang. *F1-score* memberikan informasi tentang sejauh mana model dapat mempertahankan keseimbangan antara mengklasifikasikan dengan benar kelas positif dan menghindari salah mengklasifikasikan kelas negatif.

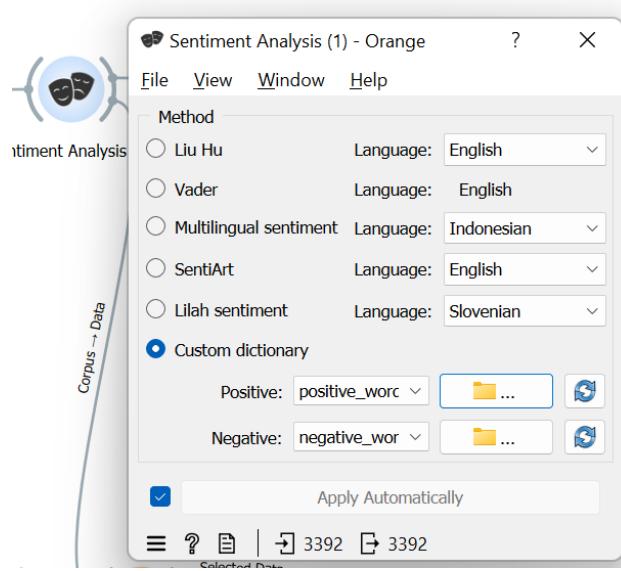
**Prec:** *Precision* (presisi) adalah metrik evaluasi yang mengukur sejauh mana prediksi positif yang dibuat oleh model adalah benar. Ini menggambarkan proporsi dari prediksi positif yang sebenarnya benar terhadap total prediksi positif yang dibuat oleh model.

**Recall:** *Recall* (recall) adalah metrik evaluasi yang mengukur sejauh mana model dapat mengidentifikasi dengan benar kelas positif. Ini menggambarkan proporsi dari kelas positif yang sebenarnya benar terhadap total jumlah sampel kelas positif dalam dataset.

**MCC:** *Matthews Correlation Coefficient* (MCC) adalah metrik evaluasi yang mengukur korelasi antara prediksi dan label aktual. MCC memberikan informasi tentang sejauh mana model dapat memprediksi dengan benar dan menghindari kesalahan prediksi palsu.

Dalam konteks *cross-validation*, metrik evaluasi ini dapat dihitung untuk setiap *fold* dalam *cross-validation* dan kemudian digunakan untuk menghasilkan rata-rata atau statistik lainnya yang merepresentasikan kinerja model secara keseluruhan.

Dengan menggunakan *custom dictionary* dalam analisis sentimen menggunakan *Orange Data Mining*, penulis dapat mengontrol kata-kata atau frasa yang digunakan dalam analisis sentimen dan sesuaikan dengan konteks. Ini memungkinkan untuk memperoleh pemahaman data yang lebih akurat dan relevan tentang sentimen yang terkandung dalam data teks yang digunakan.



**Gambar 12. Analisis Sentimen Orange Data Mining**

Melalui *Orange Data Mining*, analisis sentimen dapat dilakukan secara efisien dan intuitif dengan berbagai komponen dan alat yang tersedia. Dengan demikian, tujuan analisis sentimen menggunakan *Orange* adalah untuk menggali wawasan dan informasi yang berharga dari data teks yang dapat digunakan untuk pengambilan keputusan yang lebih baik dan pemahaman yang lebih baik tentang pandangan dan sentimen pengguna.

<b>Text:</b>	paylater membuat pengeluaran jadi nggak terkontrol membentuk habit kita untuk spending lebih mudah habit ini bahaya bila nggak terkontrol dari sekedar paylater jadi paylater yang nggak terbayar dan terpaksa harus ambil hutang
<b>sentimen:</b>	negatif
<b>kategori:</b>	dampak
<b>Text:</b>	padahal barang yg aku beli itu nominalnya ga gede bahkan ga sampe 5k trs bbrp kali reorder pake paylater sampe akhirnya udah mulai numpuk smpai cicilan paylaterribuanbulan nya jdi yg awalnya cicilannya 8paylaterk aja perbulan jadi ampe paylaterribuan
<b>sentimen:</b>	negatif
<b>kategori:</b>	dampak

**Gambar 13. Capture Data setelah dilakukan Analisis Sentimen**

## Analisis Sentimen Penggunaan Paylater pada Generasi Z di Twitter menggunakan Metode Naïve Bayes Classifier

	Text	sentimen	kategori	sentiment
1	paylater memb...	negatif	dampak	-5
2	tanyakanrl kok ...	negatif	dampak	-4.54545
3	ragu bisa paylat...	negatif	dampak	0
4	jenius paylater	positif	dampak	0
5	yup paylater	positif	dampak	0
6	akulaku via scan	positif	dampak	0
7	cv spay koin	positif	dampak	0
8	cv saldo prem ...	positif	dampak	0
9	zonauang zona...	positif	dampak	0
10	trusted dan pro...	positif	dampak	0
11	gestun instan c...	positif	dampak	0
12	fee terjangkau	positif	dampak	50
13	avail all paylater	positif	dampak	0
14	fee sta k aja bis...	positif	dampak	0
15	open gestun all...	positif	dampak	0
16	mau gestun ya ...	positif	dampak	0
17	gopaylater cicil	positif	dampak	0
18	fee murce	positif	dampak	0
19	pencairan instan	positif	dampak	0
20	spaylater akula...	positif	dampak	0
21	testie udah dan...	positif	dampak	0
22	hallo yang mau...	positif	dampak	-5.26316
23	trusted ooo	positif	dampak	0
24	fee rendah	positif	dampak	-50
25	paylater apa aia...	positif	dampak	0

**Gambar 14. Data kategori sentimen**

Dengan menggunakan *custom dictionary* dalam analisis sentimen menggunakan *Orange Data Mining*, penulis dapat mengontrol kata-kata atau frasa yang digunakan dalam analisis sentimen dan sesuaikan dengan konteks. Ini memungkinkan untuk memperoleh pemahaman data yang lebih akurat dan relevan tentang sentimen yang terkandung dalam data teks yang digunakan.

		Predicted		$\Sigma$
		dampak	penggunaan	
Actual	dampak	51.2 %	49.7 %	3392
	penggunaan	48.8 %	50.3 %	3408
		$\Sigma$	965	5835
				6800

**Gambar 15. Presentase Prediksi Opini Pengaruh Masyarakat**

Berdasarkan dari hasil analisis yang digunakan mengenai tanggapan masyarakat terhadap penggunaan *Paylater* maka diperoleh 51.2% opini masyarakat berdasarkan data dampak penggunaan *Paylater*, sedangkan berdasarkan penggunaan diperoleh nilai sebesar 50.3%. Berdasarkan presentase keduanya maka dapat disimpulkan bahwa sebesar 50.75% tanggapan masyarakat berpengaruh terhadap keputusan penggunaan *Paylater*.

## 4. KESIMPULAN

### 4.1 Kesimpulan

Penelitian ini menunjukkan bahwa persepsi publik terhadap penggunaan *Paylater* di Indonesia masih cenderung ambivalen, dengan sekitar separuh tanggapan bersifat netral. Namun, pada kalangan Generasi Z, muncul kecenderungan sentimen negatif yang menyoroti risiko peningkatan utang, kesulitan pengelolaan keuangan, dan ketergantungan terhadap kredit konsumtif. Temuan ini mengindikasikan adanya kekhawatiran terhadap dampak jangka

panjang penggunaan *Paylater*. Secara praktis, hasil penelitian ini dapat menjadi masukan bagi perusahaan *fintech* untuk merancang strategi edukasi dan pemasaran yang lebih bertanggung jawab, serta bagi regulator dalam menyusun kebijakan perlindungan konsumen dan pengelolaan risiko keuangan digital.

#### **4.2 Saran**

Penelitian selanjutnya disarankan untuk memperluas cakupan dataset agar representasi opini publik lebih komprehensif, serta membandingkan beragam metode klasifikasi, seperti *Support Vector Machine* (SVM) dan *K-Means*, guna memperoleh hasil prediksi yang lebih akurat. Selain itu, penggunaan sumber data lain di luar Twitter, seperti ulasan di *platform e-commerce* atau forum daring, dapat memberikan gambaran yang lebih luas mengenai persepsi masyarakat terhadap layanan *Paylater*.

### **DAFTAR RUJUKAN**

Agrani, A., & Rikumahu, B. (2020). Perbandingan Analisis Sentimen Terhadap Digital Payment 'Go-Pay'Dan 'OVO'Di Media Sosial Twitter Menggunakan Algoritma Naïve Bayes Dan Word Cloud. *EProceedings Manag*, 7(2), 2534.

Arisandy, Y., Dasril, Y. Bin, Salahudin, S. N. Bin, Muslim, M. A., Adnan, A., & Wen, G. K. (2023). Buy Now Pay Later Services on Generation Z: Exploratory Data Analysis Using Machine Learning. *Journal of Theoretical and Applied Information Technology*, 101(11), 4194–4204.  
[https://www.jatit.org/volumes/Vol101No11/7Vol101No11.pdf?utm\\_source=chatgpt.com](https://www.jatit.org/volumes/Vol101No11/7Vol101No11.pdf?utm_source=chatgpt.com)

Hadianti, S., & Yosep Tember, F. (2022). Analisis Sentimen COVID-19 di Twitter Menggunakan Metode Naïve Bayes dan SVM. *Jurnal Teknologi Informasi*, 6(1), 58–63.  
<https://doi.org/https://doi.org/10.36294/jurti.v6i1.2557>

Hadna, N. M. S., Paulus, I. S., & Winarno, W. (2016). Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter. *Seminar Nasional Teknologi Informasi Dan Komunikasi*, March, 1–8. <https://id.scribd.com/document/540279328/95>

Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, 6(2), 133–144.  
<https://doi.org/10.21107/nero.v6i2.237>

Indriani, F., & Wibowo, W. (2022). Sentiment analysis data by ShopeePayLater Twitter's opinion using naive Bayes classifier. *THE 3RD INTERNATIONAL CONFERENCE ON MATHEMATICS AND SCIENCES (THE 3RD ICMS): A Brighter Future with Tropical Innovation in the Application of Industry 4.0*, 070018.  
<https://pubs.aip.org/aip/acp/article-abstract/2668/1/070018/2832001/Sentiment-analysis-data-by-ShopeePayLater-Twitter>

Iqbal, M., Ryando, M. B., Triono, T., & Nurmaesah, N. (2022). Clustering of Prospective New Students Using Agglomerative Hierarchical Clustering. *International Proceeding Conference on Information Technology, Multimedia, Architecture, Design, and E-Business (IMADE)*, 2(August), 183–192.

Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2016). Analisis Sentimen pada Jasa Ojek Online. *Prosiding SNST Fakultas Teknik*, 156–161. [https://publikasiilmiah.unwahas.ac.id/PROSIDING\\_SNST\\_FT/article/view/1526](https://publikasiilmiah.unwahas.ac.id/PROSIDING_SNST_FT/article/view/1526) [https://publikasiilmiah.unwahas.ac.id/index.php/PROSIDING\\_SNST\\_FT/article/download/1526/1609](https://publikasiilmiah.unwahas.ac.id/index.php/PROSIDING_SNST_FT/article/download/1526/1609)

Padilah, B., Pratama, A. R., & Juwita, A. R. (2023). Analysis of Sentiment Adiraku App Reviews on Google Play Store Using Vector Machine Support Algorithm and Naïve Bayes. *Jurnal Sisfotek Global*, 13(1), 8. <https://doi.org/10.38101/sisfotek.v13i1.2943>

Prasastio, F. R., & Heriyanto, W. K. (2022). *Sentiment Analysis of the Covid-19 Vaccine Using the Naïve Bayes Algorithm and Levenshtein Distance Word Correction*.

Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1725–1732. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>

Ryando, M. B., & Sutarman, S. (2024). Clustering of Eligibility and Characteristics of Smart Indonesia Card Recipients for College using Agglomerative Hierarchical Clustering. *Jurnal Sisfotek Global*, 14(2), 137–146. <https://doi.org/http://dx.doi.org/10.38101/sisfotek.v14i2.15707>

Safira, A., & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *ZONAsi: Jurnal Sistem Informasi*, 5(1), 59–70. <https://doi.org/10.31849/zn.v5i1.12856>

Sammut, C., & Webb, G. I. (2017). *Encyclopedia of machine learning and data mining*. Springer Publishing Company, Incorporated.

Saputra, N., Nurbagja, K., & Turiyan, T. (2022). Sentiment Analysis of Presidential Candidates Anies Baswedan and Ganjar Pranowo Using Naïve Bayes Method. *Jurnal Sisfotek Global*, 12(2), 114. <https://doi.org/10.38101/sisfotek.v12i2.552>

Sudiantoro, A. V., & Zuliarso, E. (2018). ANALISIS SENTIMEN TWITTER MENGGUNAKAN TEXT MINING DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER. *Dinamika Informatika*, 10(2), 398–401.

Wiguna, R. A. raffaidy, & Rifai, A. I. (2021). Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining. *Journal of Information Systems and Informatics*, 3(1), 1–12. <https://doi.org/10.33557/journalisi.v3i1.78>