

BAB III

METODOLOGI PENELITIAN

3.1. Gambaran Umum Objek Penelitian

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap penerapan teknologi *AI* dalam sektor *e-commerce* berdasarkan data *tweet* yang diposting di Platform X (dulu Twitter). Dataset digunakan dalam penelitian ini untuk menggambarkan sentimen pengguna di Platform X

Data ini diperoleh melalui platform X (Twitter) yang difilter berdasarkan lang id. Dataset ini mengandung *tweet-tweet* yang dibagikan oleh pengguna Indonesia mengenai pengalaman mereka dengan teknologi *AI* dalam *e-commerce*. Dataset ini terdiri dari kolom-kolom seperti Username, Handle, Tanggal, dan Isi *Tweet*, yang berisi komentar-komentar asli dari pengguna Indonesia. Total *tweet* yang dianalisis dalam dataset ini adalah 5741 *tweet*.

[illegible]

Gambar 3. 1 Dataset dengan Filter lang:id

Contoh *Tweet* terkait:

- *"AI Tumbuhkan Lapangan Kerja Mantan Menteri Pariwisata: bagaimana AI menciptakan lapangan pekerjaan baru untuk sektor pariwisata?"(Sentimen Positif)*
- *"Masa depan sudah tiba! Bayangkan satu alat AI bisa menyelesaikan masalah besar, AI seharusnya bisa memberikan solusi untuk ekonomi kita."(Sentimen Positif)*
- *"Meskipun menggunakan AI, sistem ini tidak lebih baik dalam mengenali kebutuhan konsumen. Seringkali hasilnya salah."(Sentimen Negatif)*

- "AI dalam e-commerce masih jauh dari sempurna, sering memberikan rekomendasi produk yang tidak relevan." (Sentimen Negatif)

3.2. Metode Penelitian

3.2.1. Metode Penyelesaian

Penelitian ini menggunakan pendekatan Data Mining dengan fokus pada analisis sentimen menggunakan data teks. Sebagai kerangka kerja utama dalam analisis data, penelitian ini mengadopsi CRISP-DM (Cross-Industry Standard Process for Data Mining). Pemilihan CRISP-DM didasarkan pada fleksibilitas dan pendekatan sistematisnya dalam menangani proses analisis data, mulai dari tahap pemahaman bisnis, persiapan data, pemodelan, evaluasi, hingga deployment.

Untuk menganalisis sentimen terkait penerapan teknologi AI dalam e-commerce, penelitian ini menggunakan tiga algoritma *machine learning* yang telah terbukti efektif dalam klasifikasi teks. *Naïve Bayes* dipilih karena kemampuannya dalam menangani dataset besar secara cepat, sementara *Support Vector Machine (SVM)* diandalkan untuk menghasilkan akurasi tinggi, terutama pada data berdimensi besar. *Random Forest* digunakan untuk meningkatkan akurasi model dengan menggabungkan banyak pohon keputusan yang bekerja bersama. Penjelasan lebih lanjut mengenai framework dan algoritma yang digunakan dalam penelitian ini dapat dilihat pada Tabel 3.1 dan Tabel 3.2 berikut.

Tabel 3. 1 Framework Penelitian yang Digunakan

Framework	Fungsi	Manfaat
CRISP-DM (Cross-Industry Standard Process for Data Mining)	Framework utama yang digunakan untuk memastikan proses analisis berjalan secara sistematis melalui enam tahap: Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Modeling, Evaluasi, dan Deployment.	Menyediakan struktur kerja yang fleksibel dan iteratif, memastikan implementasi hasil penelitian dapat diterapkan.

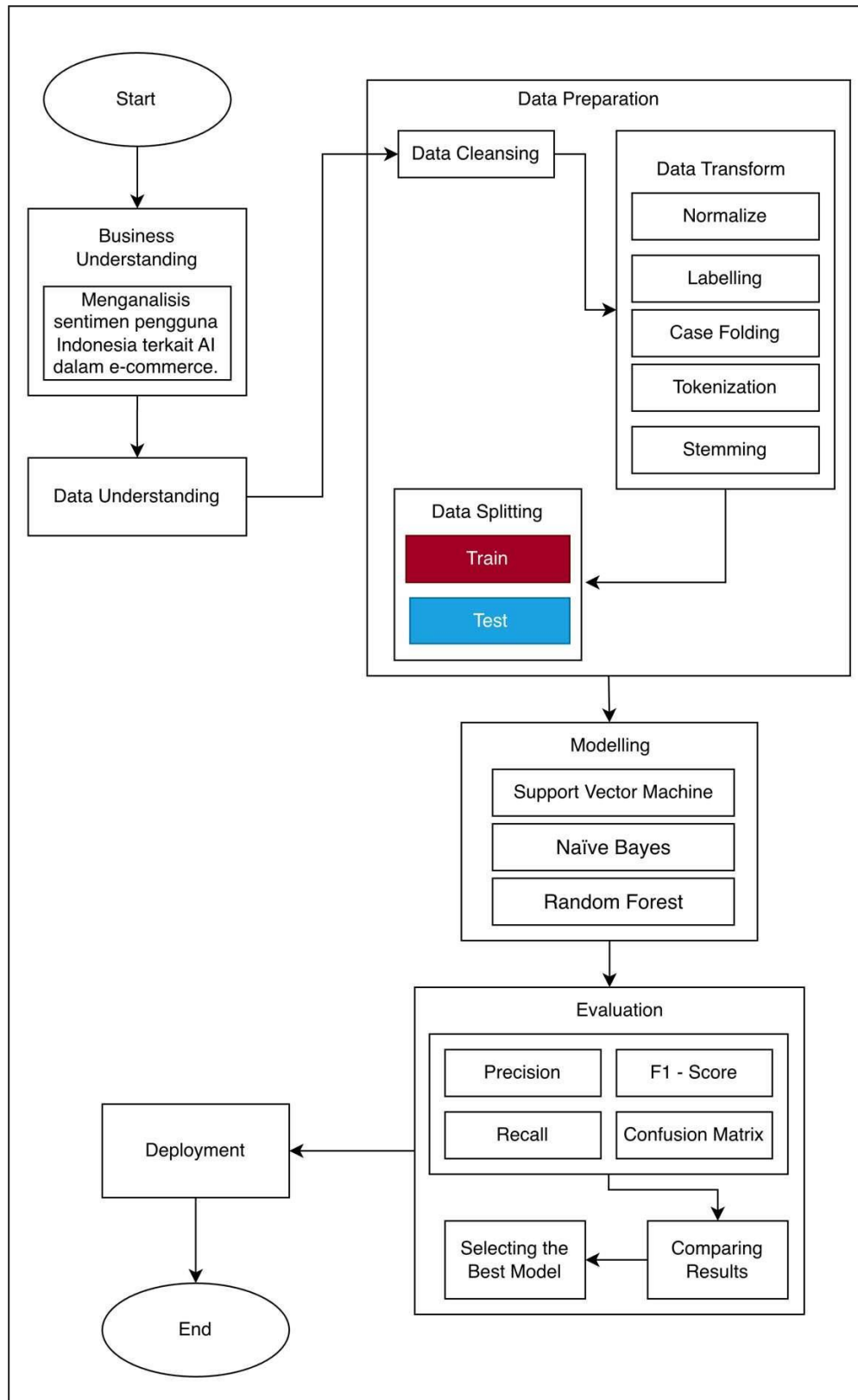
Tabel 3. 2 Algoritma yang Digunakan dalam Penelitian

Algoritma	Fungsi	Manfaat
Naïve Bayes	Algoritma klasifikasi berbasis teori probabilitas yang bekerja dengan mempertimbangkan	Cocok untuk klasifikasi teks dengan dataset besar, cepat dalam proses klasifikasi, serta menghasilkan

Algoritma	Fungsi	Manfaat
	keberadaan suatu kata dalam kelas target tertentu.	probabilitas yang dapat digunakan untuk mengukur tingkat kepercayaan suatu teks masuk ke kelas tertentu.
Support Vector Machine(SVM)	Algoritma yang bekerja dengan menemukan hyperplane optimal untuk memisahkan kelas sentimen dalam dataset ulasan pengguna aplikasi X.	Memberikan akurasi tinggi dalam analisis sentimen, terutama untuk data berdimensi besar, serta menghasilkan skor keputusan (decision score) untuk menentukan sejauh mana suatu sampel berada dalam kelas yang dibandingkan.
Random Forest	Algoritma ensemble yang menggunakan banyak pohon keputusan untuk meningkatkan akurasi dalam klasifikasi data.	Memperbaiki akurasi model dengan mengkombinasikan hasil dari banyak pohon keputusan, sehingga menghasilkan prediksi yang lebih stabil dan lebih tepat.

Framework CRISP-DM digunakan sebagai kerangka kerja utama dalam penelitian ini untuk memastikan proses analisis sentimen berjalan dengan sistematis dan terstruktur. Proses yang terdiri dari enam tahapan utama ini memberikan fleksibilitas dalam penyesuaian dan pengoptimalan analisis pada setiap tahap, dari pemahaman bisnis hingga implementasi model. Selain itu, tiga algoritma *machine learning* yang digunakan, yaitu *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest*, diterapkan untuk mengklasifikasikan sentimen dari *tweet* pengguna mengenai implementasi *AI* dalam *e-commerce*. Masing-masing algoritma ini dipilih berdasarkan keunggulannya dalam menangani data teks yang tidak terstruktur dan mengidentifikasi pola sentimen yang relevan dalam konteks penerapan *AI* di platform X.

3.2.2. Alur Penelitian



Gambar 3. 2 Diagram Alur Penelitian

Gambar 3.2 Diagram Alur Penelitian menggambarkan seluruh proses yang dilakukan dalam menganalisis sentimen pengguna Indonesia terhadap penerapan teknologi *AI* dalam *e-commerce* menggunakan data *tweet* yang diambil dari Platform X (Twitter). Alur ini mengikuti *framework* CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang terdiri dari enam tahapan utama, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* [65], [66]. Berikut penjelasan rinci mengenai setiap tahapan dalam diagram alur penelitian:

1. Business Understanding

Pada tahap *Business Understanding*, penelitian ini bertujuan untuk menganalisis sentimen pengguna Indonesia terhadap penerapan teknologi *AI* dalam dunia bisnis dan *e-commerce* melalui data *tweet* yang diperoleh dari Platform X (Twitter). Fokus utama penelitian ini adalah untuk memahami bagaimana pengguna Indonesia merespons teknologi *AI* dalam *e-commerce*, khususnya pada aspek penerapan *AI* untuk meningkatkan pengalaman pelanggan, personalisasi produk, serta efisiensi operasional. Dengan menggunakan analisis sentimen, penelitian ini berusaha memberikan wawasan yang mendalam mengenai bagaimana persepsi konsumen terhadap teknologi *AI* yang diterapkan oleh perusahaan-perusahaan *e-commerce*, serta dampaknya terhadap strategi bisnis yang dijalankan. Hasil penelitian ini diharapkan dapat membantu perusahaan dalam merancang fitur dan layanan berbasis *AI* yang lebih efektif, berdasarkan umpan balik langsung dari pengguna yang tercermin dalam komentar mereka di Platform X. Analisis sentimen ini tidak hanya memberikan gambaran umum mengenai penerimaan teknologi *AI*, tetapi juga menyoroti potensi tantangan, seperti masalah privasi, ketidakakuratan rekomendasi produk, atau kekhawatiran terkait penggunaan data pribadi yang

sering kali mempengaruhi respons pengguna terhadap implementasi *AI* di *e-commerce*.

2. Data Understanding

Data Understanding bertujuan untuk mengumpulkan dan memahami data yang diperoleh, serta memastikan bahwa data yang ada relevan dan siap digunakan dalam penelitian ini. Dalam penelitian ini, data berupa *tweet* dari pengguna Platform X (Twitter) diperoleh melalui web scraping menggunakan bantuan library *Tweet Harvest* dalam bahasa *Python*. Data yang dikumpulkan terdiri dari berbagai informasi, termasuk teks *tweet* yang berisi opini dan tanggapan dari pengguna terkait penerapan teknologi *AI* dalam *e-commerce*. Karakteristik data yang diharapkan adalah adanya data teks yang dapat digunakan untuk analisis sentimen, yang sangat bergantung pada pemahaman teks. Proses ini memastikan bahwa data yang dikumpulkan memiliki kualitas yang cukup untuk dianalisis lebih lanjut dalam tahapan *Data Preparation*, di mana data akan diproses untuk pemodelan dan analisis sentimen lebih lanjut.

3. Data Preparation

Pada tahap ini, data yang telah dikumpulkan melalui *scraping* diproses dan dipersiapkan untuk analisis lebih lanjut dalam modeling. Langkah-langkah yang dilakukan meliputi:

1) Data Cleansing

Data cleansing dilakukan untuk menghapus elemen-elemen yang tidak relevan atau gangguan dalam data, seperti angka, tanda baca, dan emotikon yang tidak memberikan kontribusi terhadap analisis sentimen.

2) Case Folding

Proses *Case Folding* dilakukan untuk mengubah seluruh teks menjadi huruf kecil. Hal ini bertujuan untuk mengurangi variasi kata akibat perbedaan kapitalisasi, sehingga meningkatkan konsistensi dan memudahkan proses analisis tanpa membedakan antara huruf besar dan kecil.

3) Normalization

Langkah ini bertujuan untuk menyamakan format teks agar konsisten, seperti mengubah huruf kapital menjadi huruf kecil dan menghapus spasi yang tidak perlu. Ini membantu untuk mengurangi variasi dalam data dan meningkatkan akurasi model.

4) Tokenization

Tokenization adalah proses di mana teks *tweet* dipecah menjadi kata-kata atau token, yang memungkinkan model pembelajaran mesin untuk memahami dan mengolah data lebih lanjut.

5) Stemming

Stemming adalah proses mengubah kata menjadi bentuk dasarnya. Misalnya, kata "memproses" menjadi "proses". Teknik ini digunakan untuk mengurangi variasi kata agar analisis lebih fokus pada inti kata tersebut.

6) Labelling

Labelling merupakan proses pemberian label atau kategori pada data berdasarkan kriteria tertentu. Dalam konteks analisis sentimen, pelabelan dilakukan untuk menentukan sentimen yang terkandung dalam data, seperti *tweet*, yang dapat dikategorikan menjadi positif, negatif, atau netral. Proses pelabelan ini dapat dilakukan dengan berbagai metode, salah satunya adalah pendekatan berbasis leksikon, di mana data dianalisis menggunakan sekumpulan kata-kata yang telah ditentukan untuk masing-masing kategori sentimen.

7) Data Splitting (Train-Test Split)

Setelah data dipersiapkan, langkah terakhir dalam tahap ini adalah membagi data menjadi dua bagian: data pelatihan (*Train*) yang akan digunakan untuk melatih model, dan data pengujian (*Test*) yang akan digunakan untuk menguji hasil model.

4. Modeling

Pada tahap *Modeling*, berbagai algoritma pembelajaran mesin diterapkan untuk melakukan klasifikasi sentimen pada *tweet* yang telah dipersiapkan. Tiga algoritma utama yang digunakan adalah:

1) Support Vector Machine(SVM)

SVM digunakan untuk mencari *hyperplane* terbaik yang memisahkan kelas *sentiment* positif, negatif, dan netral dengan margin terbesar, menghasilkan model yang sangat efektif dalam menangani data berdimensi tinggi.

2) Naïve Bayes

Algoritma ini digunakan untuk mengklasifikasikan *tweet* berdasarkan probabilitas kata yang ada dalam sentimen yang ditargetkan. *Naïve Bayes* bekerja dengan baik pada data teks yang besar dan merupakan pilihan yang baik untuk masalah klasifikasi teks.

3) Random Forest

Random Forest adalah algoritma *ensemble* yang menggunakan banyak pohon keputusan untuk melakukan klasifikasi. Dengan menggunakan pendekatan ini, model dapat mengurangi *overfitting* dan meningkatkan akurasi klasifikasi.

5. Evaluation

Pada tahap *Evaluation*, kinerja model yang telah dibangun dievaluasi dengan menggunakan beberapa metrik untuk memastikan bahwa model yang dipilih memberikan hasil yang optimal. Beberapa metrik yang digunakan antara lain:

1) *Precision*

Precision mengukur sejauh mana *tweet* yang diklasifikasikan sebagai positif benar-benar positif. Metrik ini penting untuk mengurangi jumlah *false positive*.

2) *Recall*

Recall mengukur sejauh mana *tweet* yang sebenarnya positif dapat terdeteksi oleh model. Ini penting untuk memastikan bahwa tidak ada *tweet* positif yang terlewatkan.

3) *F1-Score*

F1-Score adalah metrik yang menggabungkan *precision* dan *recall* dalam satu angka, memberikan gambaran umum tentang kinerja model secara keseluruhan.

4) *Confusion Matrix*

Confusion Matrix adalah alat untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi dengan label sebenarnya. Matriks ini menunjukkan jumlah prediksi yang benar dan salah, membantu menganalisis performa model dalam mengklasifikasikan data.

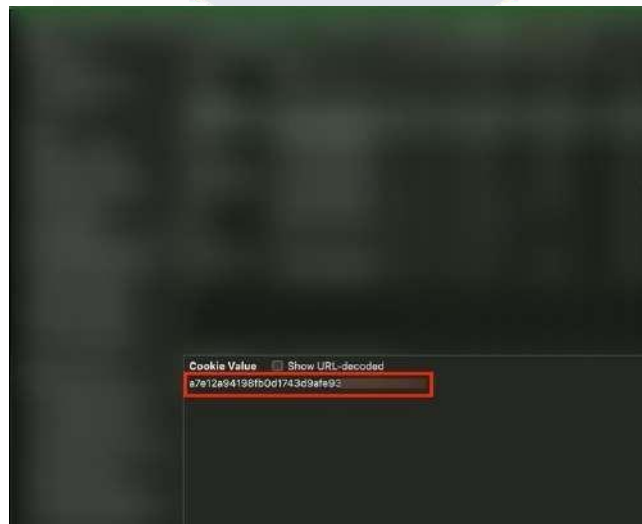
6. Deployment

Pada tahap *Deployment*, model yang telah dipilih akan diterapkan untuk menganalisis sentimen *tweet* terkait penerapan AI dalam *e-commerce* menggunakan platform *Streamlit*. Model yang telah diuji dan dievaluasi akan memproses data *tweet* secara *real-time* atau *batch*, memberikan wawasan tentang sentimen pengguna.

Hasil analisis ini membantu perusahaan *e-commerce* menyesuaikan strategi pemasaran atau pengembangan produk berdasarkan umpan balik akurat dari pengguna. *Streamlit* memungkinkan implementasi model dalam aplikasi web interaktif, memudahkan pengguna untuk melakukan analisis sentimen dengan memilih dataset, algoritma, dan memasukkan *tweet* untuk analisis.

3.3 Teknik Pengumpulan Data

Pada penelitian ini, pengumpulan data dilakukan dengan menggunakan teknik *scraping* yang mengandalkan *Tweet Harvest*. Data dikumpulkan dengan menggunakan lima kata kunci utama yang relevan dengan topik penelitian, yaitu "*AI in E-commerce*", "*AI recommendations*", "*AI Customer experience*", "*Live Commerce AI*", dan "*AI in Business*". *Tweet Harvest* digunakan setelah autentikasi dengan *Twitter Auth Token* untuk mengakses API Twitter. Pencarian *tweet* dilakukan dengan filter berdasarkan kata kunci yang ditentukan serta rentang waktu tertentu. Hasil pengumpulan data disimpan dalam format CSV untuk memudahkan analisis sentimen lebih lanjut.



Gambar 3. 3 API X (Privacy)

```
filename = 'scrapex.csv'  
search_keyword = 'AI in E-Commerce since:2024-08-01 until:2025-08-31 lang:id'  
limit = 700
```

Gambar 3. 4 *Tweet Harvest*

3.4 Teknik Analisis Data

Dalam penelitian ini, teknik analisis sentimen diterapkan untuk menganalisis sentimen pengguna mengenai penerapan *AI* dalam *e-commerce* dengan mengembangkan dan membandingkan tiga model klasifikasi berbasis *machine learning*, yaitu *Naïve Bayes*, *Support Vector Machine(SVM)*, dan *Random Forest* [9], [10], [11]. Ketiga model ini dipilih karena kemampuan mereka dalam menangani dataset teks yang besar dan kompleks serta efektivitasnya dalam mengklasifikasikan sentimen yang terkandung dalam *tweet*. Setiap model dilatih menggunakan data yang telah melalui proses *preprocessing* yang mencakup pembersihan data, tokenisasi, dan *labelling* teks. Setelah data diproses, dataset yang sama kemudian dibagi ke dalam tiga skenario pembagian data latih dan data uji, yaitu 80:20, 60:40, dan 50:50, sehingga performa setiap model dapat dievaluasi pada proporsi data pelatihan yang berbeda-beda dan terlihat konsistensi hasil klasifikasinya [67], [68].

Selain itu, penelitian ini menerapkan metode *data augmentation* sederhana untuk memperkaya representasi faktor eksternal yang berkaitan dengan privasi, keamanan data, dan kepercayaan pengguna terhadap penerapan *AI* pada bisnis dan *e-commerce*. Pada tahap ini, *tweet* yang mengandung kata kunci yang merepresentasikan faktor eksternal terlebih dahulu diidentifikasi melalui proses *keyword-based tagging*. *Tweet* yang telah teridentifikasi akan digabungkan kembali ke dalam dataset utama tanpa mengubah label sentimen yang sudah diberikan, sehingga jumlah data pada kategori faktor eksternal menjadi lebih seimbang dibandingkan kelas lain. Dengan pendekatan *data augmentation* ini, model diharapkan dapat belajar pola sentimen dalam konteks faktor eksternal dengan lebih baik dan menghasilkan analisis yang lebih stabil ketika mengevaluasi hubungan antara privasi, keamanan data, kepercayaan, dan sentimen pengguna terhadap penerapan *AI* dalam bisnis dan *e-commerce*.

Evaluasi kinerja masing-masing model dilakukan dengan menggunakan beberapa metrik, yaitu Akurasi, *Precision*, *Recall*, dan *F1-Score*. Metrik-metrik ini digunakan untuk mengukur seberapa baik model dalam mengklasifikasikan

sentimen *tweet* pengguna dan memastikan bahwa model yang terpilih memberikan hasil yang akurat. Dengan menggunakan pendekatan ini, penelitian bertujuan untuk menentukan model terbaik dalam menganalisis sentimen yang berkaitan dengan penerapan *AI* dalam *e-commerce*.

Tabel 3. 3 Metode Evaluasi Model Klasifikasi Sentimen

Metrik Evaluasi	Deskripsi	Tujuan
Akurasi	Persentase <i>tweet</i> yang diklasifikasikan dengan benar dibandingkan dengan total <i>tweet</i> yang diuji.	Mengukur seberapa baik model dapat mengklasifikasikan <i>tweet</i> secara keseluruhan.
Precision	Persentase <i>tweet</i> yang diklasifikasikan sebagai positif yang benar-benar positif.	Menilai seberapa tepat model dalam mengklasifikasikan <i>tweet</i> sebagai positif.
Recall	Persentase <i>tweet</i> positif yang berhasil ditemukan oleh model.	Mengukur kemampuan model dalam mendeteksi <i>tweet</i> positif yang sebenarnya.
F1-Score	Rata-rata harmonis dari <i>Precision</i> dan <i>Recall</i> , memberikan gambaran yang lebih seimbang tentang kinerja model.	Menyediakan ukuran kinerja yang lebih baik ketika ada ketidakseimbangan antara <i>Precision</i> dan <i>Recall</i> .