

BAB III

METODOLOGI PENELITIAN

3.1 Gambaran Umum Objek Penelitian

3.1.1 Objek Penelitian

Objek penelitian dalam penelitian ini adalah untuk melakukan analisis sentimen terhadap aplikasi Twitter sebelum dan sesudah terjadinya *rebranding* menjadi "X". Penelitian ini akan memfokuskan analisis terhadap ulasan pengguna diantaranya sebagai berikut:

1) Analisis Sentimen sebelum terjadinya *rebranding* Twitter menjadi "X" (Oktober 26, 2021 - Oktober 26, 2022): Penelitian akan menganalisis sentimen dalam ulasan pengguna aplikasi Twitter di Indonesia selama periode satu tahun sebelum terjadinya *rebranding*. Contoh ulasan dari periode ini yaitu:

a) *"Apk ini juga bagus kalian wajib dwoland Twiter karena kalian bisa melihat idol k-pop kalian di twitter dwoland sekarang."* (Sentimen positif).

b) *"Tolong dong, kok pas buka aplikasi ga bisa malah pas buka langsung keluar sendiri udah coba uninstal dan instal berulang kali tapi tetap sama aja. Mohon perbaikannya!"* (Sentimen negatif).

2) Analisis Sentimen sesudah terjadinya *rebranding* Twitter menjadi "X" (Oktober 27, 2022 - Oktober 27, 2023): Penelitian juga akan menganalisis sentimen dalam ulasan pengguna aplikasi Twitter di Indonesia selama periode satu tahun sesudah terjadinya *rebranding*. Selanjutnya contoh dari ulasan periode ini yaitu:

a. *"Bagus sih gua suka, semoga kedepannya mkin bagus deh app nya."* (Sentimen positif).

b. *"Twitter pindah ke Elon Musk jadi banyak masalah."* (Sentimen negatif).

Pada dalam penelitian ini akan memahami dampak sentimen dalam ulasan Google Play Store Indonesia sebelum dan sesudah terjadinya *rebranding* Twitter menjadi "X".

3.2 Metode Penelitian

3.2.1 Metode Penyelesaian

Penelitian ini mengacu pada studi terdahulu yang telah dibahas dalam Tabel 2.1 di Bab 2, yang memberikan dasar bagi pemilihan metode dan pendekatan yang digunakan. Sebagai framework utama dalam proses analisis data, penelitian ini menggunakan CRISP-DM (Cross-Industry Standard Process for Data Mining). Pemilihan CRISP-DM didasarkan pada hasil perbandingan framework yang telah diuraikan dalam Tabel 2.2 di Bab 2, yang menunjukkan fleksibilitas dan sistematika kerja CRISP-DM dalam menangani analisis data hingga tahap deployment.

Selain itu, penelitian ini menerapkan tiga metode pembelajaran mesin untuk klasifikasi sentimen, yaitu Naïve Bayes, Support Vector Machine (SVM), dan Logistic Regression. Pemilihan metode ini didasarkan pada kajian literatur yang telah dibandingkan dalam Tabel 2.3 di Bab 2, yang menunjukkan efektivitas ketiga algoritma tersebut dalam tugas analisis sentimen, khususnya pada data dari Google Play Store yang memiliki karakteristik serupa dengan dataset penelitian ini.

Untuk memperjelas perbedaan fungsi dan peran masing-masing pendekatan, Tabel 3.1 dan Tabel 3.2 berikut merangkum deskripsi framework dan algoritma yang digunakan dalam penelitian ini:

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 tahapan: Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Modeling, Evaluasi, dan Deployment.	Menyediakan struktur kerja yang fleksibel dan iteratif, memastikan implementasi hasil penelitian dapat diterapkan.

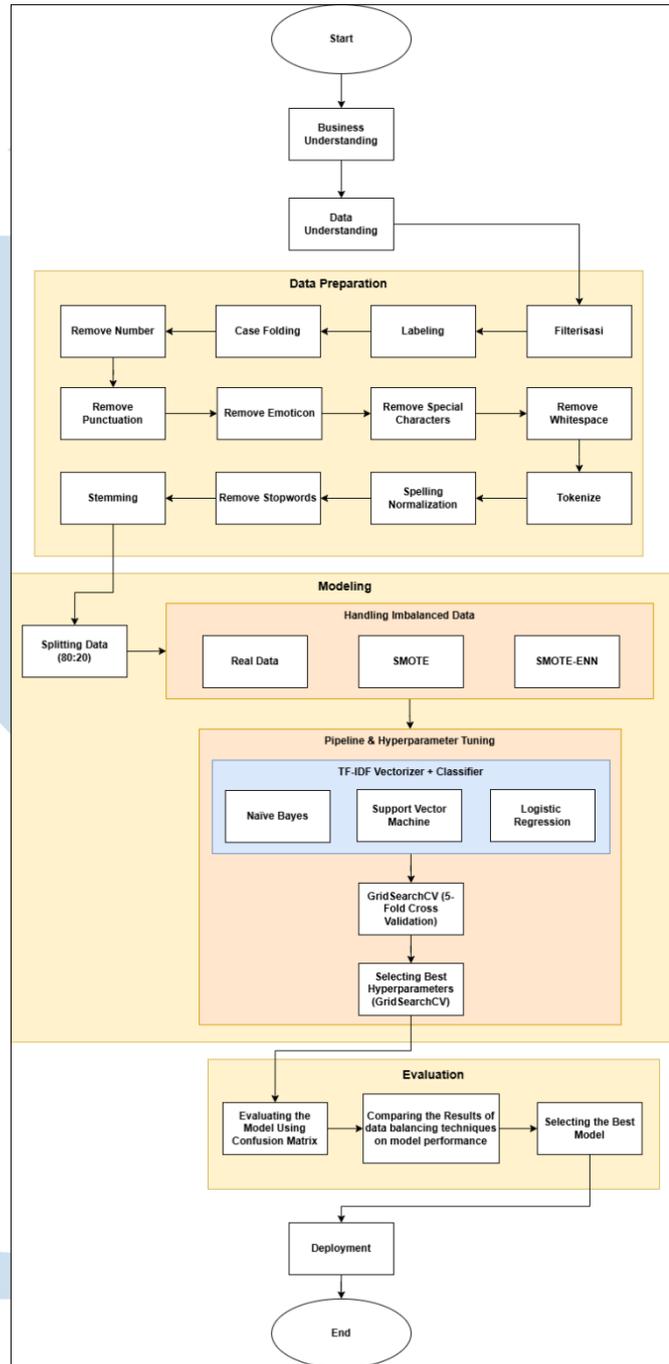
Tabel 3.2 Algoritma Klasifikasi yang Digunakan dalam Penelitian

Algoritma	Fungsi	Manfaat
Naïve Bayes	Algoritma klasifikasi berbasis teori probabilitas yang bekerja dengan mempertimbangkan keberadaan suatu kata dalam kelas target tertentu.	Cocok untuk klasifikasi teks dengan dataset besar, cepat dalam proses klasifikasi, serta menghasilkan probabilitas yang dapat digunakan untuk mengukur tingkat kepastian 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 satu kelas dibanding kelas lainnya.
Logistic Regression	Algoritma yang digunakan untuk klasifikasi biner maupun multikelas. Menggunakan fungsi sigmoid untuk mengubah output menjadi probabilitas antara 0 dan 1.	Dalam klasifikasi sentimen, memberikan interpretasi yang lebih mudah dibandingkan metode lain, serta dapat digunakan untuk menentukan apakah suatu ulasan bersifat positif atau negatif.

Framework CRISP-DM digunakan sebagai kerangka kerja utama dalam penelitian ini untuk memastikan proses analisis data berjalan secara sistematis. Sementara itu, tiga algoritma pembelajaran mesin, yaitu Naïve Bayes, SVM, dan Logistic Regression, diterapkan dalam klasifikasi sentimen ulasan pengguna aplikasi X (Twitter) sebelum dan sesudah *rebranding*.



3.2.2 Alur Penelitian



Gambar 3.1 Diagram Alur Penelitian

Gambar 3.1 merupakan *flowchart* atau diagram alur yang menggambarkan proses yang dilakukan pada penelitian ini. Dengan mengadopsi CRISP – DM sebagai kerangka pengerjaan alur diagram, maka 6 tahapan yang akan diaplikasikan sebagai berikut:

1) *Business Understanding*

Pada tahap ini, memiliki fokus utama dengan pemahaman akan permasalahan seperti mengapa penelitian harus dilakukan, dan harus diselesaikan. Pada penelitian ini dilakukan untuk menyelesaikan permasalahan terkait dengan dampak sentimen para pengguna aplikasi Twitter menjadi “X” di Indonesia sesudah terjadinya *rebranding*. Untuk mengetahui dampaknya, dilakukan analisis sentimen dengan 2 kali iterasi pada data ulasan para pengguna Twitter atau “X” 1 tahun sebelum dan sesudah terjadinya *rebranding* dari Google Play Store.

2) *Data Understanding*

Data understanding memiliki tujuan untuk melakukan pengambilan dan pengumpulan data, serta diakhiri dengan memahami data yang telah didapatkan. Dengan tujuan awal pada penelitian ini, maka data ulasan para pengguna Twitter atau “X” didapatkan dengan menggunakan bantuan library *google-play-scraper* pada bahasa Python. Karakteristik data yang diharapkan pada penelitian ini yaitu, terdapat data ulasan para pengguna yang memiliki tipe data teks. Alasan ini dihadirkan untuk menyelesaikan permasalahan analisis sentimen yang membutuhkan data teks.

3) *Data Preparation*

Setelah tahap *data understanding* terselesaikan, tahap *data preparation* fokus utamanya mempersiapkan data agar dapat menyelesaikan tujuan awal pada penelitian. Mempersiapkan ini mencakup merubah dari data mentah menjadi data bersih. Langkah-langkahnya sebagai berikut:

a) **Filterisasi**

Langkah pertama dalam *data preparation* adalah melakukan filterisasi data berdasarkan tanggal dan waktu ulasan. Ulasan sebelum terjadinya *rebranding* akan diambil dalam rentang

waktu 1 tahun, yaitu dari tanggal 26 Oktober 2021 hingga 26 Oktober 2022.

Sementara itu, ulasan sesudah *rebranding* akan diambil dalam rentang waktu 27 Oktober 2022 hingga 27 Oktober 2023. Data yang telah difilter akan dipisahkan ke dalam dua dataframe yang berbeda, yaitu data sebelum *rebranding* dan data sesudah *rebranding*. Dengan demikian, seluruh tahapan analisis selanjutnya akan dilakukan dua kali pada masing-masing dataframe untuk memastikan perbandingan yang valid sesuai dengan tujuan penelitian.

b) ***Labeling***

Setelah proses filterisasi selesai, tahap selanjutnya adalah data *labeling*, yaitu proses pemberian label pada data ulasan pengguna. Pada penelitian ini, data akan dibagi menjadi dua label, yaitu sentimen positif dan sentimen negatif. Pelabelan ini dilakukan karena penelitian ini menerapkan perbandingan model pembelajaran mesin klasifikasi yang membutuhkan data berlabel. Untuk memastikan objektivitas dalam pelabelan data, proses ini akan dilakukan secara manual dengan melibatkan tiga orang annotator yang tidak termasuk peneliti. Pemilihan annotator ini didasarkan pada kriteria bahwa mereka adalah pengguna aktif aplikasi Twitter atau X, yang setidaknya telah menggunakan aplikasi tersebut selama 1 tahun sebelum dan sesudah *rebranding*.

Sebagai metode penentuan label akhir, digunakan pendekatan *majority voting*, di mana label yang dipilih oleh mayoritas annotator akan dianggap sebagai label sentimen akhir untuk suatu ulasan [22]. Adapun penelitian sebelumnya menunjukkan bahwa metode *majority voting* yang melibatkan pelabelan manual oleh beberapa annotator dapat meningkatkan

kualitas dan keandalan data pelatihan, serta mengurangi ambiguitas dalam pelabelan. Selain itu, sistem pelabelan otomatis hanya mencapai sekitar 80% akurasi dibandingkan pelabelan manual, sehingga justifikasi penggunaan pelabelan manual dalam penelitian ini menjadi semakin kuat [69].

c) ***Case Folding***

Case Folding atau mengonversi teks ulasan akan dilakukan apabila proses *labeling* telah terselesaikan. Tujuan dari tahap ini pada *data preparation* adalah mengonversi seluruh teks ulasan pengguna menjadi huruf kecil (lowercase). Proses ini dilakukan untuk mengubah seluruh teks menjadi huruf kecil menggunakan fungsi `casefold()` dari Python, sehingga membantu menciptakan konsistensi dalam analisis data teks dan mempermudah pemrosesan pada tahap selanjutnya [38].

d) ***Remove Number***

Setelah tahap *case folding*, beberapa proses pembersihan teks dilakukan untuk meningkatkan kualitas data. Langkah selanjutnya adalah *remove number*, yaitu proses penghapusan angka dalam teks agar tidak mengganggu analisis, dengan menggunakan pustaka `re` karena kemampuannya dalam menangani pola numerik secara baik [22].

e) ***Remove Punctuation***

Remove punctuation bertujuan untuk menghapus tanda baca seperti titik, koma, tanda seru, tanda tanya, dan sejenisnya. Tanda baca umumnya tidak memberikan kontribusi berarti terhadap analisis sentimen dan dapat mengganggu proses tokenisasi. Proses ini dilakukan dengan menggunakan pustaka `string` dalam Python untuk menghilangkan karakter tanda baca secara konsisten [22].

f) ***Remove Emoticon***

Selanjutnya, *remove emoticon* dilakukan untuk menghapus karakter emoji dan *emoticon* yang mungkin terdapat pada teks ulasan. Proses ini dilakukan menggunakan pustaka *re* dari Python untuk mendeteksi pola simbol grafis yang umum muncul dalam bentuk emoji [70].

g) ***Remove Special Characters***

Selain itu, *remove special characters* juga diterapkan untuk menghapus karakter non-ASCII seperti simbol '²' dan karakter khusus lainnya yang tidak diperlukan dalam pengolahan data. Tahap ini juga memanfaatkan pustaka *re*, yang memungkinkan penghapusan karakter di luar rentang ASCII secara efektif [70].

h) ***Remove Whitespace***

Remove whitespace dilakukan untuk mengganti spasi yang berlebih menjadi satu spasi saja, sehingga format teks menjadi lebih rapi dan mudah diolah dalam tahap selanjutnya. Proses ini dilakukan menggunakan pustaka *re* dari Python karena pustaka ini memungkinkan pengenalan pola spasi berlebih dan penyederhanaannya dengan cara yang sistematis [22].

i) ***Tokenize***

Apabila Tahap *keep single unique char* sudah dikerjakan, maka tahap *tokenize* bisa dijalankan. Tahapan *tokenize* atau tokenisasi, merupakan pembagian teks ulasan menjadi token-token atau kata-kata. Tokenisasi memungkinkan representasi kata-kata yang lebih baik untuk analisis selanjutnya. Pada pengerjaannya akan diperbantukan *library* dari python bernama NLTK *word_tokenize* [71]. *Library* tersebut dipilih karena merupakan salah satu pustaka paling populer dalam pemrosesan bahasa alami (NLP) [3].

j) ***Spelling Normalization***

Spelling normalization merupakan tahap penting dalam proses persiapan data pada penelitian ini. Proses ini bertujuan untuk mengubah teks ulasan yang mengandung bahasa tidak formal atau alay menjadi bentuk yang lebih baku dan sesuai dengan kaidah bahasa standar. Penggunaan kamus bahasa alay dipilih karena ulasan pengguna sering kali bersifat informal dan menggunakan bahasa sehari-hari yang tidak sepenuhnya mengikuti aturan tata bahasa yang baku [72].

Dalam implementasinya, proses normalisasi ini menggunakan kamus alay bahasa Indonesia yang tersedia secara publik melalui repository GitHub [73]. Kamus ini disusun berdasarkan hasil penelitian terdahulu dan telah banyak digunakan dalam berbagai studi terkait pemrosesan bahasa alami di konteks bahasa Indonesia. Keberadaan kamus ini memungkinkan penyesuaian kata-kata tidak baku menjadi bentuk yang lebih konsisten dan sesuai dengan kaidah bahasa Indonesia.

k) ***Remove Stop Word***

Remove stop word dilakukan untuk penghapusan kata yang tidak memiliki fungsi pada ulasan. Dalam penelitian ini, penghapusan *stopwords* akan dilakukan menggunakan daftar kata yang telah disediakan oleh Natural Language Toolkit (NLTK), sebuah pustaka Python yang banyak digunakan dalam pemrosesan bahasa alami (NLP). Penggunaan NLTK dipilih karena kelengkapan dan keandalannya dalam menangani berbagai bahasa, termasuk bahasa Indonesia, sehingga memungkinkan pembersihan teks yang lebih optimal serta meningkatkan akurasi analisis yang akan dilakukan [22].

l) ***Stemming***

	'karena', 'enggak', 'ada', 'jaringan', 'seluler', 'masih', 'minta', 'kode', 'verifikasi', 'dari', 'nomor', 'saya', 'bukan', 'main', 'enggak', 'nyaman', 'banget', 'pakai', 'aplikasi', 'ini', 'uninstall']	'bukan', 'main', 'nyaman', 'banget', 'pakai', 'aplikasi', 'uninstall']
<i>Stemming</i>	['masuk', 'twitter', 'pakai', 'alamat', 'email', 'jaringan', 'seluler', 'minta', 'kode', 'verifikasi', 'nomor', 'bukan', 'main', 'nyaman', 'banget', 'pakai', 'aplikasi', 'uninstall']	Tidak ada perubahan (kata tidak terjadi perubahan kedalam bentuk dasarnya, karena sudah dalam bentuk dasar)

4) **Modeling**

Modeling yang akan dilakukan dalam penelitian kali ini akan dikerjakan sebanyak 2 kali terkait dengan tujuan penelitian sebelumnya. Dalam pengulangan iterasi sebanyak 2 kali tersebut, pengerjaan modeling akan bergerak dalam cabang pembelajaran mesin klasifikasi analisis sentimen. Tahapan *modeling* diantaranya sebagai berikut:

a) **Splitting Data**

Data ulasan akan dibagi menjadi dua bagian, dengan 80% digunakan sebagai data pelatihan dan 20% untuk data pengujian. Pemilihan rasio 80:20 didasarkan pada temuan dari penelitian sebelumnya, yang menunjukkan bahwa pembagian data dengan rasio tersebut umumnya menghasilkan performa model yang lebih baik dibandingkan rasio lain seperti 70:30 atau 90:10 [15]. Oleh karena itu, rasio 80:20 dipilih untuk memaksimalkan efektivitas proses pelatihan dan evaluasi model dalam penelitian ini.

b) **Handling Imbalanced Data**

Pada tahap modeling ini, dilakukan beberapa eksperimen dengan pendekatan yang berbeda untuk menangani ketidakseimbangan data (*imbalanced data*). Pendekatan yang digunakan dalam penelitian ini meliputi:

- i. Menggunakan data asli

- ii. Menggunakan Synthetic Minority Over-sampling Technique (SMOTE)
- iii. Menggunakan kombinasi SMOTE dan Edited Nearest Neighbors (SMOTE-ENN)

Setiap pendekatan, digunakan tiga algoritma klasifikasi, yaitu Naïve Bayes (NB) dengan *MultinomialNB*, Support Vector Machine (SVM) dengan *SVC*, dan Logistic Regression. Proses klasifikasi diawali dengan membangun *pipeline* yang terdiri dari *TF-IDF Vectorizer* sebagai tahap untuk mengubah data teks menjadi bentuk numerik, kemudian dilanjutkan ke algoritma klasifikasi. Guna memperoleh performa optimal dalam proses pelatihan model klasifikasi sentimen, dilakukan eksplorasi terhadap kombinasi hyperparameter terbaik melalui pendekatan *GridSearchCV* dengan menerapkan validasi silang sebanyak lima lipatan (*5-fold cross-validation*) [51]. Dengan mencoba semua konfigurasi hyperparameter yang ditentukan, *GridSearchCV* membantu menemukan pengaturan hyperparameter terbaik untuk memaksimalkan metrik evaluasi seperti akurasi atau F1-score [53].

Seluruh rangkaian pengujian dikonstruksi dalam sebuah *pipeline*, yang dirancang untuk menyederhanakan proses integrasi mulai dari tahap praproses data hingga pemodelan akhir. Hyperparameter yang diuji mencakup dua aspek utama. Aspek pertama terletak pada proses representasi teks menggunakan *TF-IDF Vectorizer*, di mana dilakukan pengujian terhadap nilai *max_df* (0.25, 0.5, dan 0.75) serta variasi *n-gram* (*ngram_range*) yang mencakup (1,1), (1,2), dan (1,3). Penentuan rentang nilai ini merujuk pada pendekatan yang telah diterapkan dalam penelitian sebelumnya [22].

Adapun aspek kedua mencakup hyperparameter dari masing-masing algoritma klasifikasi yang diterapkan, yaitu Naïve Bayes, Support Vector Machine (SVM), dan Logistic Regression. Pemilihan nilai-nilai hyperparameter dilakukan berdasarkan konfigurasi yang umum digunakan dalam praktik klasifikasi teks. Rincian dari Hyperparameter yang digunakan dalam eksperimen ini disajikan pada Tabel 3.4 berikut:

Tabel 3.4 Pencarian Hyperparameter Terbaik Menggunakan Gridsearchcv

Metode	Hyperparameter	Nilai yang Dicoba
TF-IDF Vectorizer	max_df	0.25, 0.5, 0.75
	ngram_range	(1,1), (1,2), (1,3) (unigram, bigram, trigram)
Naïve Bayes (MultinomialNB)	alpha	0.1, 0.01, 0.001
Support Vector Machine (SVM)	kernel	linear, poly, rbf, sigmoid
	C	0.1, 1, 10
	gamma	scale, auto
Logistic Regression	C	0.1, 1, 10
	solver	lbfgs, liblinear

Kombinasi hyperparameter pada Tabel 3.4 menjadi fondasi dalam proses eksplorasi konfigurasi model yang digunakan dalam penelitian ini. GridSearchCV dikenal sebagai metode hyperparameter tuning yang cukup agresif secara komputasi karena mengevaluasi semua kemungkinan kombinasi yang tersedia dalam ruang pencarian. Namun, dalam konteks penelitian ini, GridSearchCV tetap digunakan karena kemampuannya dalam memberikan evaluasi yang menyeluruh untuk masing-masing konfigurasi, sehingga memungkinkan analisis yang komprehensif terhadap pengaruh setiap kombinasi hyperparameter terhadap performa model.

Untuk mengatasi tantangan dari eksplorasi yang terlalu luas, diterapkan strategi pembatasan ruang pencarian hyperparameter.

Seluruh kombinasi yang digunakan telah ditentukan secara

eksplisit dan dicantumkan pada Tabel 3.4. Nilai-nilai yang dipilih merupakan representasi dari variasi yang umum digunakan dalam studi serupa, sehingga tetap mencakup konfigurasi yang relevan namun tidak berlebihan. Dengan cara ini, jumlah total kombinasi yang diuji dapat tetap terkontrol.

Pendekatan pembatasan ini sejalan dengan panduan dalam studi hyperparameter tuning, yang menekankan pentingnya menyeimbangkan antara cakupan eksplorasi dan beban komputasi, terutama ketika menggunakan metode seperti GridSearchCV yang menjelajah seluruh ruang parameter secara menyeluruh [75]. Langkah pencarian ini menghasilkan kombinasi hyperparameter terbaik berdasarkan performa validasi silang pada data latih. Model dengan kombinasi hyperparameter terbaik kemudian dievaluasi menggunakan data pengujian untuk menilai performa aktualnya, dan hasil inilah yang menjadi dasar dalam pemilihan model terbaik pada penelitian ini.

5) *Evaluation*

Pada tahap evaluasi ini, model dengan hyperparameter terbaik yang diperoleh dari proses GridSearchCV pada data latih akan diuji menggunakan data pengujian. Evaluasi dilakukan dengan memanfaatkan Confusion Matrix untuk menghitung metrik performa seperti akurasi, precision, recall, dan F1-score.

Masing-masing metrik digunakan karena memberikan sudut pandang yang berbeda dalam mengukur performa model. Akurasi mengukur proporsi prediksi yang benar terhadap keseluruhan data dan berguna untuk melihat kinerja umum model. Namun, karena data ulasan pengguna dalam penelitian ini berpotensi tidak seimbang, diperlukan metrik tambahan [54]. Precision mengukur seberapa tepat prediksi model terhadap suatu kelas. Precision positif menunjukkan proporsi

ulasan yang benar-benar positif dari seluruh prediksi positif. Sebaliknya, precision negatif menunjukkan proporsi ulasan yang benar-benar negatif dari seluruh prediksi negatif. Kedua metrik ini digunakan untuk menilai ketepatan model dalam mengklasifikasikan ulasan secara benar, baik sebagai positif maupun negatif. Recall mengukur kemampuan model dalam mendeteksi seluruh data yang relevan, penting ketika kesalahan dalam mengabaikan suatu kelas harus diminimalkan. F1-score sebagai rata-rata harmonis antara precision dan recall memberikan keseimbangan ketika terjadi *trade-off* di antara keduanya. Oleh karena itu, keempat metrik ini digunakan secara bersamaan untuk mengevaluasi kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya [55].

Selanjutnya, dilakukan perbandingan hasil performa model berdasarkan berbagai teknik penanganan data tidak seimbang (data asli, SMOTE, dan SMOTE-ENN) serta algoritma klasifikasi yang digunakan (Naïve Bayes, SVM, dan Logistic Regression). Dalam penelitian ini, metode SMOTE dan SMOTE-ENN dipilih karena keduanya dirancang untuk mengatasi tantangan ketidakseimbangan kelas yang bisa saja terjadi pada data [18]. SMOTE bekerja dengan menghasilkan sampel sintetis untuk kelas minoritas melalui interpolasi, sedangkan SMOTE-ENN mengombinasikan proses tersebut dengan penghapusan data mayoritas yang tidak representatif menggunakan pendekatan Edited Nearest Neighbor (ENN). Pendekatan ini bertujuan tidak hanya untuk menyeimbangkan distribusi kelas, tetapi juga untuk mengurangi *noise* dalam data latih dan meningkatkan kualitas model [46].

Penerapan ketiga pendekatan, yaitu data asli, SMOTE, dan SMOTE-ENN, dimaksudkan untuk mengeksplorasi pengaruh teknik penyeimbangan data terhadap performa klasifikasi sentimen dalam konteks *rebranding* Twitter menjadi X. Perbandingan ini diharapkan mampu mengidentifikasi kombinasi metode penyeimbangan dan algoritma klasifikasi yang paling efektif. Hasil evaluasi juga mendukung

analisis deskriptif terkait perubahan persepsi pengguna sebelum dan sesudah *rebranding*, sehingga memberikan gambaran yang lebih komprehensif mengenai dinamika sentimen pengguna secara temporal. Model dengan performa terbaik kemudian dipilih untuk digunakan pada tahap *deployment* dan potensi pengembangan penelitian lanjutan.

Hasil evaluasi tersebut juga akan mendukung analisis deskriptif terkait perubahan persepsi pengguna sebelum dan sesudah *rebranding*, sehingga memberikan gambaran yang lebih komprehensif mengenai dinamika sentimen pengguna secara temporal. Dari keseluruhan proses ini, dilakukan pemilihan model terbaik berdasarkan metrik performa untuk digunakan pada tahap *deployment*.

6) **Deployment**

Berdasarkan hasil evaluasi, model dengan performa terbaik dari masing-masing periode (sebelum dan sesudah *rebranding*) dipilih untuk tahap *deployment*. Model-model tersebut kemudian diimplementasikan ke dalam sebuah website sederhana berbasis Streamlit, yang memungkinkan pengguna untuk menguji hasil analisis sentimen secara mandiri.

3.3 Teknik Pengumpulan Data

Penelitian ini akan menggunakan data bersifat data primer. Pengumpulan data tersebut diperoleh melalui cara *scraping* pada Google Play Store dengan menggunakan *library* dari bahasa pemrograman Python di Jupyter Notebook [59]. Pemilihan Google Play Store didasarkan pada kemudahan akses data melalui *library google_play_scraper* dalam Python, yang memungkinkan proses *scraping* dilakukan secara langsung di Jupyter Notebook tanpa memerlukan ketergantungan terhadap API resmi atau izin tambahan dari penyedia aplikasi [58]. Selain itu, ulasan yang terdapat di Google Play Store hanya dapat diberikan oleh pengguna yang telah mengunduh dan menggunakan aplikasi secara langsung, sehingga dinilai kredibel dan relevan untuk dianalisis dalam konteks sentimen pengguna [12].

Data ulasan aplikasi para pengguna Twitter atau "X" di Indonesia akan diambil dengan batasan 1 tahun sebelum *rebranding* dan 1 tahun sesudah *rebranding* terjadi pada aplikasi Twitter menjadi X. Pengambilan data dilakukan untuk mendapatkan gambaran perubahan sentimen pengguna terhadap aplikasi tersebut sebelum dan sesudah terjadinya *rebranding*.

3.4 Variabel Penelitian

3.4.1 Variabel Independent

Variabel independent dalam penelitian ini adalah faktor-faktor yang mempengaruhi hasil analisis sentimen pengguna terhadap *rebranding* Twitter menjadi "X". Adapun variabel independent yang digunakan dalam penelitian ini meliputi:

- **Teks Ulasan Pengguna** – Data ulasan yang diambil dari Google Play Store terkait aplikasi Twitter/X. Contoh ulasan pengguna “*Sangat bagus, gampang di mengerti, serta sangat mudah untuk berinteraksi ke semua pengguna.....*”

3.4.2 Variabel Dependent

Variabel dependent dalam penelitian ini adalah hasil yang diperoleh setelah menerapkan metode analisis sentimen terhadap teks ulasan pengguna. Variabel dependent yang diamati meliputi:

- **Klasifikasi Sentimen** – Hasil prediksi sentimen berdasarkan model yang digunakan, dengan label: positif atau negatif.

3.5 Teknik Analisis Data

Dalam penelitian ini, analisis data dilakukan melalui serangkaian tahap sistematis guna memastikan data siap untuk diproses dalam model pembelajaran mesin. Tahapan utama yang diterapkan mencakup pembersihan data (data cleansing), pemberian label (data labeling), transformasi data (data transformation), serta penerapan model pembelajaran mesin.

1) Pembersihan Data (Data Cleansing)

Proses pembersihan data bertujuan untuk meningkatkan kualitas data dengan menghilangkan elemen-elemen yang tidak relevan atau berpotensi mengganggu analisis sentimen. Langkah-langkah yang dilakukan meliputi:

- a) **Case Folding:** Mengonversi seluruh teks ke dalam huruf kecil untuk menjaga konsistensi analisis teks.
- b) **Penghapusan Angka (Remove Number):** Menghapus angka yang tidak memiliki makna dalam analisis sentimen.
- c) **Penghapusan Tanda Baca (Remove Punctuation):** Mengeliminasi tanda baca seperti titik, koma, dan tanda seru yang tidak diperlukan.
- d) **Penghapusan Emotikon & Karakter Khusus (Remove Emoticon & Special Characters):** Menghapus simbol atau karakter yang tidak diperlukan dalam analisis teks.
- e) **Penghapusan Spasi Berlebih (Remove Whitespace):** Menghapus spasi yang tidak dibutuhkan dalam analisis sentimen.

2) Pemberian Label (Data Labeling)

Ketika proses filterisasi berdasarkan rentang waktu dilakukan, data ulasan pengguna dikategorikan menjadi dua periode utama, yaitu sebelum dan sesudah *rebranding* Twitter menjadi X.

- a) Data diberikan label sentimen positif atau negatif.
- b) Labeling dilakukan secara manual oleh tiga annotator yang memiliki pengalaman menggunakan aplikasi Twitter/X minimal satu tahun sebelum dan sesudah *rebranding*.
- c) Metode *majority voting* diterapkan untuk menentukan label akhir berdasarkan keputusan mayoritas dari ketiga annotator.
- d) Para *annotator* diberikan panduan teknis awal yang bersifat umum untuk menyamakan persepsi dalam penilaian sentimen, tanpa

menyertakan contoh data utama penelitian, agar menjaga konsistensi dan objektivitas dalam pelabelan.

3) Transformasi Data (Data Transformation)

Agar data lebih terstruktur dan siap diproses oleh model pembelajaran mesin, dilakukan beberapa tahapan transformasi sebagai berikut:

- a) **Tokenisasi (Tokenize)**: Memecah teks ulasan menjadi token (kata-kata) untuk dianalisis lebih lanjut.
- b) **Normalisasi Ejaan (Spelling Normalization)**: Mengonversi kata tidak baku atau bahasa alay ke dalam bentuk standar menggunakan kamus bahasa alay.
- c) **Penghapusan Kata Umum (Remove Stopword)**: Menghapus kata-kata yang tidak berkontribusi signifikan dalam analisis sentimen dengan bantuan pustaka NLTK.
- d) **Stemming**: Mengubah kata ke dalam bentuk dasar menggunakan pustaka **Sastrawi** untuk bahasa Indonesia.

4) Penerapan Model Pembelajaran Mesin

Setelah data siap, dilakukan eksperimen dengan berbagai pendekatan dalam menangani ketidakseimbangan data (imbalanced data):

- a) Menggunakan data asli tanpa penyesuaian.
- b) Menerapkan **Synthetic Minority Over-sampling Technique (SMOTE)**.
- c) Mengombinasikan **SMOTE dengan Edited Nearest Neighbors (SMOTE-ENN)**.

Selanjutnya, data diklasifikasikan menggunakan tiga algoritma:

- a) **Naïve Bayes (MultinomialNB)**
- b) **Support Vector Machine (SVM) dengan SVC**
- c) **Logistic Regression**

Sebelum klasifikasi dilakukan, data ulasan dikonversi ke bentuk numerik menggunakan TF-IDF *vectorizer*. Untuk memperoleh hasil optimal, digunakan metode GridSearchCV dengan validasi silang sebanyak lima kali lipat (5-fold cross-validation) guna menemukan kombinasi hyperparameter terbaik. Hasil dari tahapan ini diharapkan dapat menghasilkan data yang lebih bersih, terstruktur, serta optimal untuk analisis sentimen menggunakan model pembelajaran mesin.

