

## Analisis Sentimen Terhadap Cyberbullying Pada Komentar di Instagram Menggunakan Algoritma Naïve Bayes

Fauzan Baehaqi<sup>1</sup>, Nuri Cahyono<sup>2</sup>

[fauzanbaehaqi@students.amikom.ac.id](mailto:fauzanbaehaqi@students.amikom.ac.id), [nuricahyono@amikom.ac.id](mailto:nuricahyono@amikom.ac.id)

Universitas Amikom Yogyakarta

---

### Informasi Artikel

Diterima : 25 Jul 2023  
Direview : 14 Sep 2023  
Disetujui : 20 Jan 2024

---

### Kata Kunci

Instagram,  
Cyberbullying,  
Klasifikasi, Naïve Bayes

---

### Abstrak

Cyberbullying merupakan tindakan yang merujuk pada penggunaan teknologi digital untuk sengaja menyakiti, merendahkan, atau membuli orang lain secara daring. Penelitian ini berfokus pada analisis cyberbullying di media sosial, khususnya dalam komentar Instagram, di mana ketika banyak pihak yang kemudian menjadi sekumpulan orang yang tidak menyukai suatu postingan akan bersatu untuk memberikan pendapat dan komentar negatif, yang bisa berakibat menurunkan kepercayaan diri dan dampak buruk lainnya bagi pemilik postingan. Oleh karena itu dilakukan analisis sentimen cyberbullying pada komentar Instagram dengan menggunakan algoritma Naive Bayes sebagai upaya pencegahan, dan mengevaluasi kelayakan algoritma yang digunakan. Metode analisis sentimen Naive Bayes digunakan untuk mengidentifikasi sentimen positif dan negatif dalam komentar. Data penelitian terdiri dari 400 komentar Instagram yang telah dilabeli secara manual. Hasil evaluasi menunjukkan tingkat akurasi yang baik setelah melalui tahap preprocessing, pembobotan TF-IDF, dan feature selection. Akurasi mencapai nilai precision 92,10%, recall 89,74%, f1-score 90,98%, dan akurasi 91,25%. Penelitian ini diharapkan mampu memberikan kontribusi signifikan dalam upaya pencegahan cyberbullying pada platform media sosial.

---

### Keywords

Instagram, Cyberbullying,  
Naïve Bayes Classification

---

### Abstract

*Cyberbullying is an action that refers to the use of digital technology to intentionally hurt, degrade, or bully other people online. This research focuses on the analysis of cyberbullying on social media, especially in Instagram comments, where when many parties then become a group of people who don't like a post will unite to provide negative opinions and comments, which can result in lowered self-confidence and other bad impacts for post owner. Therefore, cyberbullying sentiment analysis was carried out on Instagram comments using the Naive Bayes algorithm as a prevention effort, and evaluating the feasibility of the algorithm used. The Naive Bayes sentiment analysis method is used to identify positive and negative sentiments in comments. The research data consists of 400 Instagram comments that have been manually labeled. The evaluation results show a good level of accuracy after going through the preprocessing, TF-IDF weighting and feature selection stages. Accuracy reached a precision value of 92.10%, recall 89.74%, f1-score 90.98%, and accuracy 91.25%. It is hoped that this research will be able to make a significant contribution to efforts to prevent cyberbullying on social media platforms.*

## A. Pendahuluan

Seiring kemajuan teknologi dan informasi, Perkembangan internet telah membawa perubahan besar termasuk dalam komunikasi dan pertukaran informasi. Jumlah pengguna internet terus meningkat setiap tahun, terutama di Indonesia, mencapai sekitar 77% dari total jumlah penduduk per Januari 2023[1]. Salah satu fenomena yang muncul seiring dengan perkembangan internet adalah penggunaan media sosial yang semakin meluas. Media sosial seperti Facebook, Twitter, dan Instagram telah menjadi platform populer bagi pengguna internet untuk berinteraksi, berbagi konten, dan membangun jejaring sosial[2]. Penggunaan media sosial telah mengubah cara kita berkomunikasi, berbagi informasi, dan membentuk hubungan sosial dalam kehidupan sehari-hari. Berdasarkan laporan terkini dari We Are Social dan Hootsuite, jumlah pengguna sosial media secara global mencapai 4,76 miliar. Yaitu setara dengan 59,4% dari keseluruhan penduduk dunia saat ini. Laporan tersebut juga mencatat bahwa di Indonesia, jumlah penggunanya mencapai 60,4% dari total populasi negara pada bulan yang sama (databoks, Januari 2023).

Dalam konteks ini, Instagram menjadi salah satu sosial media yang sangat populer bagi pengguna internet terutama di kalangan remaja dan orang dewasa. Instagram menawarkan beragam fitur menarik seperti unggahan foto dan video, fitur cerita (stories), penggunaan hashtag untuk mencari dan mengikuti topik-topik tertentu[3]. Data dari databoks menyebutkan bahwa India menjadi negara dengan jumlah pengguna terbanyak, dengan 326 juta pengguna, diikuti Amerika Serikat dengan 168 juta pengguna, Brasil dengan 132 juta pengguna, dan Indonesia dengan 106 juta pengguna (databoks, April 2023). Namun, di balik popularitasnya terdapat masalah, Instagram juga memiliki dampak buruk, fitur komentar memungkinkan pengguna memberikan tanggapan pada suatu postingan. Meskipun sebagian besar komentar di Instagram dapat berisi dukungan, apresiasi atau kritik sayangnya platform ini juga memberi ruang bagi komentar negatif, ejekan atau bisa disebut juga cyberbullying[4].

Cyberbullying adalah tindakan yang merujuk pada penggunaan teknologi digital untuk sengaja menyakiti, merendahkan, atau membuli orang lain secara daring[5]. Bentuk-bentuk cyberbullying sangat beragam, mencakup pengiriman pesan kebencian, pelecehan verbal, penyebaran informasi pribadi yang tidak sah, dan ancaman fisik atau psikologis[6]. Masalah yang terjadi ketika ada banyak pengguna yang tidak setuju atau tidak menyukai suatu postingan akan bersatu dan memberikan komentar negatif untuk menyerang pemilik postingan, tidak jarang pula terjadi perdebatan karena perbedaan pendapat pada kolom komentar bahkan menggunakan kata-kata negatif dan hinaan. Fenomena ini sering kali menimbulkan dampak buruk pada korban, berupa gangguan kesehatan mental seperti kehilangan rasa percaya diri, stres, dan depresi, dalam beberapa kasus yang tragis, cyberbullying bahkan dapat menyebabkan berakhirnya kehidupan seseorang[7].

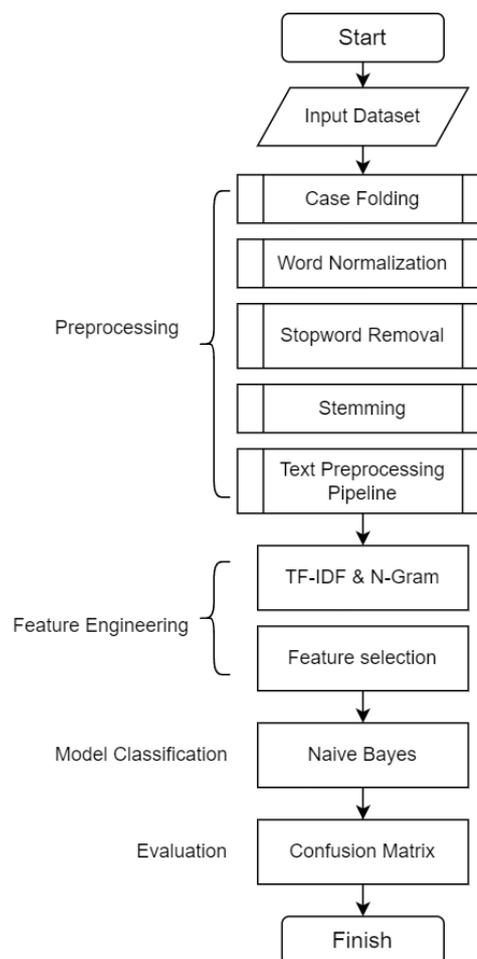
Melihat dampak serius yang ditimbulkan oleh cyberbullying, menjadi latar belakang dilakukannya penelitian untuk memahami secara mendalam fenomena ini dan mencari solusi yang efektif untuk menghadapinya. Oleh karena itu, penting untuk mengkaji permasalahan cyberbullying pada komentar di Instagram dalam konteks pemrosesan teks. Dalam upaya menyaring komentar-komentar di sosial media, analisis sentimen digunakan untuk mengidentifikasi komentar yang bersifat positif dan komentar yang bersifat negatif. Hasil analisis tersebut memungkinkan

adanya tindakan pencegahan yang dapat dilakukan, baik untuk membantu korban maupun pelaku.

Pada penelitian ini, untuk menganalisis sentimen pada komentar Instaram akan menggunakan algoritma Naïve Bayes. Pemilihan algoritma ini didasari oleh beberapa kelebihan yang relevan dan apakah algoritma ini dapat digunakan dalam menganalisis sentimen[8]. Naive Bayes adalah metode yang relatif sederhana dan mudah diimplementasikan, sehingga cocok untuk analisis sentimen pada komentar di Instagram, Terlebih, metode ini hanya membutuhkan jumlah data latih yang tidak terlalu besar dalam proses pengklasifikasian[9]. Tujuan utama dari penelitian ini adalah untuk mendalami dan memahami potensi penerapan algoritma Naive Bayes sebagai alat pencegahan cyberbullying dan meningkatkan kesadaran akan keamanan di platform Instagram.

## B. Metode Penelitian

Metode pada penelitian ini terdiri dari beberapa tahap. Alur penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Metode Penelitian

## **B.1. Input Dataset**

Dataset yang akan digunakan merupakan kumpulan data yang diperoleh dari [https://raw.githubusercontent.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/master/dataset\\_komentar\\_instagram\\_cyberbullying.csv/](https://raw.githubusercontent.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/master/dataset_komentar_instagram_cyberbullying.csv/) Dataset ini merupakan kumpulan data komentar Instagram yang diambil dari akun github milik rizalespe[10].

## **B.2. Preprocessing**

Tahapan preprocessing dalam analisis teks dengan langkah-langkah seperti case folding, word normalization, stopwords removal, stemming, dan text preprocessing pipeline berfungsi untuk menyiapkan data teks sebelum menjalani analisis lebih lanjut. Tahapan-tahapan yang dilaksanakan adalah sebagai berikut:

### **B.2.1. Case Folding**

Case folding akan mengonversi semua huruf menjadi huruf kecil, menghapus url, angka, tanda baca, dan karakter lain selain huruf abjad, dengan tujuan untuk menstandarisasi dokumen teks dalam data.

### **B.2.2. Word Normalization**

Proses Word normalization mengubah kata-kata dalam teks menjadi bentuk yang lebih standar atau normal yaitu dengan mengatasi variasi kata, mengatasi singkatan dan akronim. Tujuan dari word normalization adalah untuk mengurangi variasi dan menghasilkan representasi yang konsisten dengan mengurangi variasi kata dengan makna yang sama.

### **B.2.3. Stopword Removal**

Stopword removal berguna untuk menghapus kata-kata yang sering muncul dalam dokumen tetapi tidak memberikan informasi penting. Kata-kata tersebut umumnya termasuk dalam kategori kata bantu atau kata hubung yang sering muncul seperti "dan", "atau", "juga", "di", "dari", dan sebagainya. Hal ini bertujuan mengurangi dimensi data, meningkatkan efisiensi komputasi, dan memfokuskan analisis pada kata-kata yang lebih informatif dan penting.

### **B.2.4. Stemming**

Stemming adalah proses menghilangkan akhiran dan awalan kata. Misalnya, kata "berlari", "berlarian", dan "berlari-lari" akan diubah menjadi bentuk dasar "lari". Tujuan dari stemming adalah untuk mengurangi variasi kata-kata yang mungkin memiliki akar yang sama, sehingga dapat meningkatkan efisiensi dalam analisis teks dan pemrosesan bahasa alami.

### **B.2.5. Text Preprocessing Pipeline**

Text Preprocessing pipeline dapat digunakan untuk menyatukan beberapa langkah menjadi satu objek, dengan menyimpan data yang telah melalui proses preprocessing ke dalam file CSV. Ini bermanfaat agar tidak perlu menjalankan proses preprocessing dari awal setiap kali akan menggunakan data tersebut,

### **B.3. Feature Engineering**

Feature Engineering merupakan sebuah proses untuk mengembangkan dan memilih suatu fitur atau atribut yang akan digunakan untuk melakukan analisis data atau dalam melakukan pembuatan sebuah model machine learning. Dalam penelitian ini ada dua feature yang digunakan, yaitu TF-IDF dan Feature selection menggunakan chi square.

#### **B.3.1. TF-IDF**

Metode TF-IDF mengubah data teks menjadi vektor numerik dan membantu meningkatkan efisiensi dan akurasi dalam berbagai tugas pemrosesan teks[11]. Metode TF-IDF dikenal karena efisiensinya, kemudahan penggunaannya, dan hasil yang akurat. Pendekatan ini mengukur Term Frequency (TF) dan Inverse dari Document Frequency (IDF) untuk setiap kata di seluruh dokumen, membantu dalam menentukan frekuensi kemunculan suatu kata dalam dokumen. Hasil dari metode ini adalah matrix fitur TF-IDF yang menggambarkan representasi numerik dari kata-kata dalam dokumen. Ini dapat digunakan dalam tugas pemrosesan teks klasifikasi menggunakan algoritma naïve bayes. Matrix ini akan memberikan bobot untuk setiap kata dalam dokumen yang menggambarkan seberapa penting kata tersebut dalam konteks koleksi dokumen tersebut.

#### **B.3.2. Feature Selection**

Seleksi fitur merupakan suatu metode untuk mengurangi jumlah fitur atau variabel input dengan memilih fitur-fitur yang dianggap paling relevan terhadap model. yang bertujuan meningkatkan akurasi dan mengurangi kompleksitas algoritma sehingga bisa mempercepat waktu training[12].

### **B.4. Naïve Bayes**

Algoritma klasifikasi Naïve Bayes adalah model klasifikasi yang digunakan untuk melakukan klasifikasi data berdasarkan perhitungan probabilitas[13]. Dalam konteks analisis sentimen, model ini digunakan untuk memprediksi apakah sebuah ulasan memiliki sentimen positif atau negatif berdasarkan vektor TF-IDF dari kata-kata dalam ulasan[14]. Proses ini melibatkan perhitungan probabilitas untuk setiap kata dalam dokumen. Diperlukan untuk melakukan pembagian data menjadi dua bagian, yakni data pelatihan (train data) dan data uji (test data). Hal ini merupakan langkah penting dalam penerapan algoritma Naïve Bayes atau banyak algoritma klasifikasi lainnya. Tujuan utama pemisahan data adalah untuk melatih model pada data yang sudah ada, dan kemudian menguji kinerja model pada data yang belum pernah dilihat sebelumnya. Hal ini dilakukan untuk mengukur sejauh mana model dapat menggeneralisasi dari data yang telah dilihat. Data pelatihan digunakan untuk melatih model, yang berarti model akan "belajar" dari data tersebut, dan data pengujian digunakan untuk menguji sejauh mana model mampu mengklasifikasikan data baru yang tidak digunakan selama pelatihan. Dengan pemisahan ini, kita dapat mengevaluasi kinerja model dengan lebih baik dan memastikan bahwa model tersebut dapat digunakan untuk memprediksi sentimen atau klasifikasi lainnya pada data yang belum pernah dilihat sebelumnya.

## B.5. Confusion Matrix

Setelah proses pemodelan naïve bayes, hasilnya dievaluasi menggunakan confusion matrix. Untuk menentukan nilai akurasi data diatas akan dihitung dari jumlah data yang berhasil di prediksi model dan jumlah data yang salah di prediksi model. terdapat empat istilah yang digunakan sebagai representasi hasil dari proses klasifikasi, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

		Negative	Positive
Aktual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive
		Prediksi	

**Gambar 2.** Tabel Confusion Matrix

Akurasi (Accuracy): Mengukur sejauh mana model dapat mengklasifikasikan semua kelas dengan benar.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN})$$

Presisi (Precision): Mengukur sejauh mana hasil positif yang diprediksi oleh model adalah benar.

$$\text{Precision} = (\text{TP}) / (\text{TP} + \text{FP})$$

Recall (Sensitivitas): Mengukur sejauh mana model berhasil mengidentifikasi semua instance positif yang sebenarnya.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

F1 Score: Mengevaluasi kinerja model dengan mempertimbangkan keseimbangan antara presisi dan recall.

$$\text{F-1 Score} = (2 * \text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

Keterangan :

TP : Jumlah data yang diprediksi benar sebagai positif sesuai dengan nilai aktual

TN : Jumlah data yang diprediksi benar sebagai negatif sesuai dengan nilai aktual

FP : Jumlah data yang diprediksi positif tetapi negatif dari nilai aktual

FN : Jumlah data yang diprediksi negatif tetapi positif dari nilai aktual

## C. Hasil dan Pembahasan

### C.1. Input Dataset

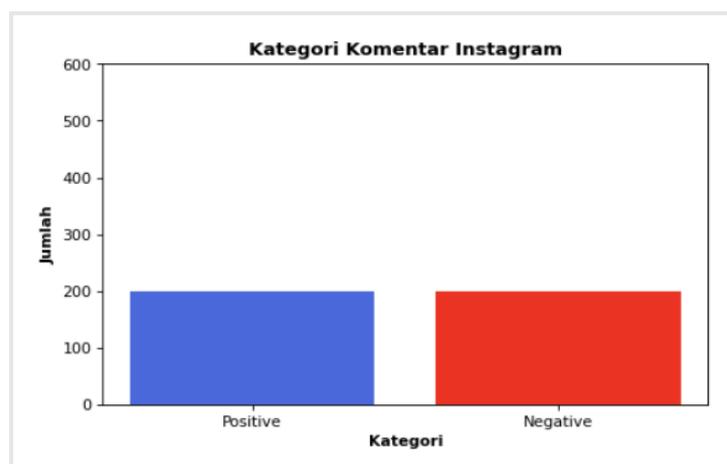
Dataset yang digunakan berasal dari [https://raw.githubusercontent.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/master/dataset\\_komentar\\_instagram\\_cyberbullying.csv/](https://raw.githubusercontent.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/master/dataset_komentar_instagram_cyberbullying.csv/) diambil dari akun github milik rizalespe[10]. Data ini dikumpulkan dari postingan komentar Instagram, lalu dilabeli secara manual oleh pemilik dataset. Dataset ini terdiri dari 400 komentar yang telah terbagi ke dalam dua kelas label, dengan 200 komentar yang dikategorikan sebagai positif dan 200 komentar lainnya dikategorikan sebagai negatif.

Id	Sentiment	Instagram Comment Text
0	1	negative <USERNAME> TOLOL!! Gak ada hubungan nya kegug...
1	2	negative Geblek lo tata ..cowo bgt dibela2in balikan .....
2	3	negative Kmn termewek2 skr lengket lg duhhh kok labil ...
3	4	negative Intinya kalau kesel dengan ATT nya, gausah ke ...
4	5	negative hadewwww perempuan itu lg!!!!sakit jiwa,knp ha...
...	...	...
395	396	positive Bangga sama suami yg selalu ingat istri disela...
396	397	positive Apaoun pekerjaannya yg penting halal u tuk men...
397	398	positive Gojek itu mayoritas peganguran yang lama gak ...
398	399	positive <USERNAME> aslinya cantik dan ayu loh mbak kr...
399	400	positive <USERNAME> suami saya seumuran sama saya mba,...

400 rows × 3 columns

**Gambar 3.** Input Dataset

Gambar dibawah ini adalah visualisasi kategori komentar dari dataset.



**Gambar 4.** Kategori Komentar Instagram

## C.2. Preprocessing

Pada tahap selanjutnya adalah melakukan preprocessing, preprocessing bertujuan untuk mengolah data menjadi efektif untuk digunakan. Pada tahap preprocessing ini ada empat proses yang perlu dilalui antara lain.

### C.2.1. Case Folding

Ini merupakan proses untuk mengonversi semua huruf menjadi huruf kecil, menghapus tanda baca, menghapus URL, menghapus angka, menghapus karakter tanda baca dan menghapus kata username. Hasil dari proses case folding bisa dilihat pada tabel dibawah ini:

**Tabel 1. Case Folding**

Sebelum Case Folding	Sesudah Case Folding
Geblek lo tata...cowo bgt dibela2in	geblek lo tatacowo bgt dibelain
balikan...hadeww...ntar ditinggal lg nyalahin	balikanhadewwntar ditinggal lg nyalahin tuh
tuh cowo...padahal kitenya yg oon.	cowopadahal kitenya yg oon

### C.2.2. Word Normalization

Pada proses word normalization kata-kata slang, gaul dan singkatan pada data akan dirubah menjadi kata yang baku. Data yang digunakan didownload dari corpus kumpulan slangwords pada link [https://raw.githubusercontent.com/ksnugroho/klasifikasi-spam-sms/master/data/key\\_norm.csv](https://raw.githubusercontent.com/ksnugroho/klasifikasi-spam-sms/master/data/key_norm.csv). Berikut adalah contoh data corpus yang diunduh:

**Tabel 2. Corpus Kumpulan Slangwords**

_id	Singkat	Hasil
1	abis	habis
2	accent	tekanan
3	accept	terima
4	accident	kecelakaan
5	achievement	prestasi

Kata-kata singkatan seperti “ktika akan diubah menjadi “ketika”, “gw” menjadi “saya”, “lbih” menjadi “lebih” dan “kece” menjadi “gaul” kata-kata diatas akan diubah menjadi kata yang baku.

**Tabel 3. Word Normalization**

Sebelum Word Normalization	Setelah Word Normalization
wkwk ini slg ktika gw ngedit fotonya biar lbh kece	wkwk ini slg ketika saya mengedit fotonya biar lebih gaul

### C.2.3. Stopword removal

Setelah itu kata-kata yang dianggap tidak memberikan informasi penting atau makna khusus akan dihapus. Kata-kata tersebut termasuk dalam kategori kata bantu atau kata hubung yang sering muncul seperti "dan", "atau", "juga", "di", "dari", dan sebagainya. Pada tahap ini, kata-kata yang ada dalam daftar stopwords diidentifikasi dan dihapus dari teks. Hasilnya bisa dilihat pada tabel 4:

**Tabel 4. Stopword Removal**

Sebelum Stopword Removal	Setelah Stopword Removal
geblek lo tatacowo bgt dibelain	geblek lo tatacowo bgt dibelain
balikanhadewwntar ditinggal lg nyalahin tuh	balikanhadewwntar ditinggal lg nyalahin tuh
cowopadahal kitenya yg oon	cowopadahal kitenya yg oon

#### C.2.4. Stemming

Setelah melewati beberapa langkah di atas, langkah terakhir dalam preprocessing pada penelitian ini adalah stemming. Stemming diperlukan untuk mengurangi variasi indeks dalam suatu data, sehingga kata-kata dengan awalan atau akhiran akan disederhanakan ke bentuk dasarnya. Hal ini memungkinkan pengelompokan kata-kata yang memiliki makna serupa, meskipun bentuknya berbeda karena adanya imbuhan yang berbeda. Rincian hasil stemming dapat dilihat seperti dibawah ini:

**Tabel 5. Stemming**

Sebelum Stemming	Setelah Stemming
geblek lo tatacowo bgt dibelain	geblek lo tatacowo bgt bain
balikanhadewwntar ditinggal lg nyalahin tuh	balikanhadewwntar tinggal lg nyalahin tuh
cowopadahal kitenya yg oon	cowopadahal kitenya yg oon

#### C.2.5. Text Preprocessing Pipeline

Semua Langkah pada tahap text preprocessing lalu digabungkan seperti gambar dibawah ini.

```

1 # Buat fungsi untuk menggabungkan seluruh langkah text preprocessing
2 def text_preprocessing_process(Text):
3     text = casefolding(Text)
4     text = text_normalize(Text)
5     text = remove_stop_words(Text)
6     text = stemming(Text)
7     return text

```

**Gambar 5. Text Preprocessing Pipeline**

Dibawah ini adalah data yang sudah melalui semua tahap preprocessing.

Id	Sentiment	Instagram Comment Text	clean_teks
0	1	negative <USERNAME> TOLOL!! Gak ada hubungan nya kegug...	username tolol gak ada hubung nya gugur dgn pa...
1	2	negative Geblek lo tata...cowo bgt dibela2in balikan.....	geblek lo tata cowo bgt dibela2in balik hadeww...
2	3	negative Kmm termewek2 skr lengket lg duhhh kok labil ...	kmm termewek2 skr lengket lg duhhh kok labil ...
3	4	negative Intinya kalau kesel dengan ATT nya, gausah ke ...	inti kalau kesel dengan att nya gausah ke anak...
4	5	negative hadewwww pempuan itu lg!!!sakit jiwa,knp ha...	hadewwww pempuan itu lg sakit jiwa knp harus...
...	...	...	...
395	396	positive Bangga sama suami yg selalu ingat istri disela...	bangga sama suami yg selalu ingat istri sela w...
396	397	positive Apaoun pekerjaannya yg penting halal u tuk men...	apaoun kerja yg penting halal u tuk nafkah kel...
397	398	positive Gojek itu mayoritas pegangguran yang lama gak ...	gojek itu mayoritas pegangguran yang lama gak ...
398	399	positive <USERNAME> aslinya cantik dan ayu loh mbak kr...	username asli cantik dan ayu loh mbak krn aku ...
399	400	positive <USERNAME> suami saya seumuran sama saya mba,...	username suami saya umur sama saya mba malah t...

400 rows x 4 columns

**Gambar 6. Hasil Data Preprocessing**

### C.3. Feature Engineering

#### C.3.1. TF-IDF

Setelah menyelesaikan preprocessing data, digunakan 'TfidfVectorizer' dari library scikit-learn untuk mengubah dokumen teks menjadi matrix fitur Term Frequency-Inverse Document Frequency (TF-IDF). matrix TF-IDF diubah menjadi sebuah DataFrame menggunakan pandas (data\_tf\_idf) dengan nama kolom yang sesuai dengan kata-kata dalam vektor TF-IDF. Ini memberikan kemudahan dalam analisis dan pemahaman hasil ekstraksi fitur. Terakhir, modul pickle digunakan untuk menyimpan matrix fitur TF-IDF ke dalam sebuah file biner (tf\_idf\_feature.pickle). Hal ini mempermudah penyimpanan dan berbagi data yang telah diolah, sehingga tidak perlu melakukan perhitungan ulang ketika akan menggunakannya. Dibawah ini adalah hasil jumlah token menggunakan TF-IDF.

	10	15	16	17	1996	20an	25	30	70	ामीन	...	yoo	you	youtube	yra	yuk	yutub	zaman	zholimi	zina	zzzz			
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
395	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
396	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
397	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
398	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
399	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

400 rows x 2715 columns

Gambar 7. TF-IDF

Selanjutnya akan dilakukan seleksi fitur menggunakan chi-square untuk mengurangi jumlah fitur yang kurang relevan.

#### C.3.2. Feature Selection

Seleksi fitur dilakukan pada data teks yang telah dilakukan TF-IDF menggunakan metode Chi-Square. Pustaka scikit-learn digunakan untuk mengimplementasikan metode Chi-Square. Hasilnya adalah 1000 fitur terbaik yang dipilih berdasarkan skor Chi-Square. Langkah pertama yang dilakukan adalah membuat sebuah dictionary kosong yang diberi nama kbest\_feature. Dictionary ini digunakan untuk menyimpan pasangan kata (k) dan indeks (v) dari vocab yang dihasilkan oleh TF-IDF. Selanjutnya, dilakukan iterasi melalui vocab yang dihasilkan oleh TF-IDF, dengan memeriksa setiap kata (k) apakah termasuk dalam daftar fitur yang telah diseleksi sebelumnya. Jika kata tersebut termasuk dalam fitur-fitur yang diseleksi, maka kata tersebut disimpan dalam dictionary kbest\_feature bersama dengan indeksnya. Dengan mengurangi jumlah fitur ini, diharapkan dapat meningkatkan efisiensi pemodelan, mengurangi dimensi data seperti mengurangi fitur yang kurang relevan. Hasil seleksi fitur ini selanjutnya yang akan digunakan pada pemodelan menggunakan naïve bayes. Hasil dari seleksi fitur menggunakan chi-square dapat dilihat pada gambar dibawah ini:

```
      Kata  Nilai
528   zina  2713
942   yra  2708
333  youtube 2707
653   you  2706
390   yng  2703
..     ...   ...
346  abiii  15
982  abbey  13
370  abal  11
941  aamiin 9
931  1996  4

[1000 rows x 2 columns]
```

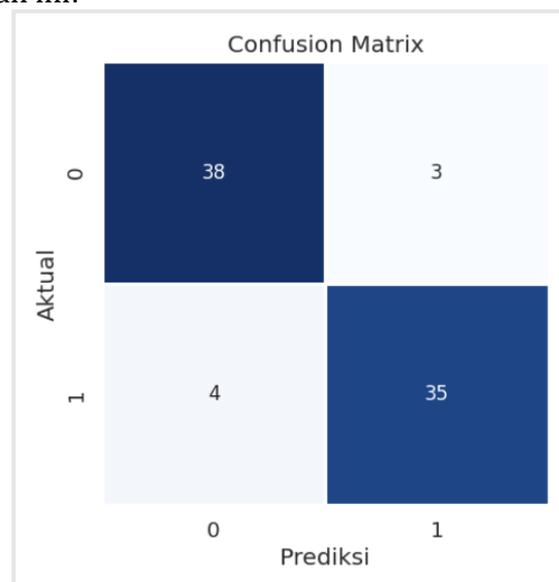
**Gambar 8.** Hasil Seleksi Fitur

#### C.4. Naïve Bayes

Setelah data siap, dilakukan pemisahan data latih dan data uji untuk mengukur kinerja model klasifikasi Naive Bayes. Proses pemisahan ini penting dilakukan untuk memastikan model dapat menggeneralisasi dengan baik. Dataset dibagi menjadi dua subset, data latih (80%) dan data uji (20%). Penggunaan fungsi `train_test_split` dari pustaka `scikit-learn` digunakan, di mana `X_kbest_features` adalah matrix fitur yang telah dipilih melalui seleksi fitur, dan `y` adalah vektor target yang berisi label kelas. Argumen `test_size=0.2` digunakan untuk menentukan bahwa 20% dari data akan menjadi data uji. Setelah pemisahan data, model klasifikasi Naive Bayes dilatih dengan menggunakan algoritma Multinomial Naive Bayes. Hasil model tersebut disimpan ke dalam sebuah file menggunakan `joblib` untuk digunakan dalam tahap evaluasi kinerja model dan prediksi.

##### C.4.1. Evaluasi

Hasil evaluasi yang diperoleh melalui confusion matrix menunjukkan bahwa model ini memiliki tingkat akurasi yang cukup baik. Confusion matrix bisa dilihat pada gambar dibawah ini:



**Gambar 9.** Hasil Evaluasi Confusion Matrix

Terdapat 38 kasus True Negative, yang menunjukkan klasifikasi data negatif dengan benar, 35 kasus True Positive dalam mengklasifikasikan data positif dengan benar, 4 kasus False Negative, yang berarti ada data positif yang salah diklasifikasikan sebagai negatif, dan 3 kasus False Positive, yang menunjukkan adanya data negatif yang salah diklasifikasikan sebagai positif.

Dari evaluasi confusion matrix di atas, perhitungan nilai akurasi dapat memberikan informasi mengenai persentase sentimen yang benar-benar diprediksi sebagai positif, negatif, serta sentimen yang salah diprediksi sebagai positif dan negatif. Sebagai cara mudah untuk mengingatkannya, jika dimulai dengan True, maka prediksinya benar; jika dimulai dengan false, maka prediksinya salah. Sebagai contoh, True negative berarti prediksi sentimen yang negatif adalah benar, di mana positive dan negative adalah hasil prediksi model, menunjukkan bahwa model memprediksi sentimen negatif dengan benar. Berikut adalah nilai matriks kinerja setelah dihitung menggunakan rumus:

**Tabel 6.** Hasil akurasi confusion matrix

Representasi Confusion Matrix	Nilai
multinomialNB precision	0.9210
multinomialNB recall	0.8974
multinomialNB f1-score	0.9098
multinomialNB accuracy	0.9125

#### D. Kesimpulan

Dalam penelitian ini, analisis sentimen komentar Instagram dilakukan dengan menggunakan dataset yang diambil dari akun GitHub milik RizalEspe, terdiri dari 400 data yang telah dilabeli secara manual, yaitu 200 data sentimen positif dan 200 sentimen negatif. Melalui berbagai tahap preprocessing, termasuk penggunaan TF-IDF dan seleksi fitur dengan Chi-Square, serta penerapan model klasifikasi Naive Bayes pada data training sebesar 80% dan data test sebesar 20%, hasil evaluasi menunjukkan kinerja yang sangat baik. Algoritma Naive Bayes mampu mengklasifikasikan sentimen positive dengan presisi sebesar 92,10%, recall sebesar 89,74%, f1-score sebesar 90,98%, dan akurasi mencapai 91.25%. Harapannya, hasil penelitian ini dapat memberikan kontribusi yang berarti dalam usaha pencegahan dan peningkatan keamanan serta kenyamanan dalam berinteraksi di sosial media, terutama dalam mendeteksi potensi cyberbullying. Dengan demikian, algoritma Naive Bayes terbukti menjadi alat yang sangat efektif dalam analisis sentimen komentar Instagram dan memiliki potensi untuk aplikasi lebih lanjut dalam pengembangan solusi terkait keamanan online yang diharapkan dapat meningkatkan kesadaran akan keamanan bersosial media seperti di Instagram.

#### E. Ucapan Terimakasih

Dalam penelitian ini, Peneliti ingin mengungkapkan rasa terima kasih yang kepada Universitas Amikom Yogyakarta yang telah memberikan fasilitas dan sumberdaya yang diperlukan dan seluruh tim pendukung yang telah memberikan kontribusi berharga dalam pelaksanaan penelitian ini. Tanpa dukungan dan

bantuan mereka, penelitian ini tidak akan mencapai hasil yang luar biasa. Terima kasih sekali lagi atas dukungan yang berharga ini.

## F. Referensi

- [1] R. Gunawan *et al.*, "ADIKSI MEDIA SOSIAL DAN GADGET BAGI PENGGUNA INTERNET DI INDONESIA," *Jurnal Techno-Socio Ekonomika*, vol. 14, no. 1, 2020.
- [2] P. Studi Ilmu Komunikasi, "TREN PENGGUNAAN MEDIA SOSIAL SELAMA PANDEMI DI INDONESIA," 2020.
- [3] A. Rachmayanti and Y. Candrasari, "PERILAKU CYBERBULLYING DI INSTAGRAM."
- [4] M. Jubaidi and N. Fadilla, "PENGARUH FENOMENA CYBERBULLYING SEBAGAI CYBER-CRIME DI INSTAGRAM DAN DAMPAK NEGATIFNYA," *Shaut Al-Maktabah : Jurnal Perpustakaan, Arsip dan Dokumentasi*, vol. 12, no. 2, pp. 117–134, Dec. 2020, doi: 10.37108/shaut.v12i2.327.
- [5] F. Syifa Mutma, "DESKRIPSI PEMAHAMAN CYBERBULLYING DI MEDIA SOSIAL PADA MAHASISWA," *Jurnal Common |*, vol. 4, no. 1, 2020, doi: 10.34010/common.
- [6] J. Penelitian *et al.*, "Agustus Tahun 2021 | Hal," vol. 1, no. 8, pp. 257–263, 2021, doi: 10.56393/decive.v1i8.298.
- [7] T. Sutini *et al.*, "DIRI PADA REMAJA : LITERATUR REVIEW."
- [8] R. Tsania, S. A. Putri, D. E. Ratnawati, and D. W. Brata, "Perbandingan Naïve Bayes dan K-Nearest Neighbor untuk Analisis Sentimen Aplikasi Gapura UB Berdasarkan Ulasan Pengguna pada Playstore," 2023. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] S. Saepudin, S. Widiastuti, and C. Irawan, "Sentiment Analysis of Social Media Platform Reviews Using the Naïve Bayes Classifier Algorithm," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 2, pp. 236–243, Jul. 2023, doi: 10.32736/sisfokom.v12i2.1650.
- [10] W. Athira Luqyana, I. Cholissodin, and R. S. Perdana, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine," 2018. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] "PENERAPAN METODE TF-IDF DAN N-GRAM".
- [12] T. Ernayanti, M. Mustafid, A. Rusgiyono, and A. R. Hakim, "PENGGUNAAN SELEKSI FITUR CHI-SQUARE DAN ALGORITMA MULTINOMIAL NAÏVE BAYES UNTUK ANALISIS SENTIMEN PELANGGAN TOKOPEDIA," *Jurnal Gaussian*, vol. 11, no. 4, pp. 562–571, Feb. 2023, doi: 10.14710/j.gauss.11.4.562-571.
- [13] D. Sepri, P. Algoritma, N. Bayes, U. Analisis, K. Penggunaan, and A. Bank, "media cetak," *Journal of Computer System and Informatics (JoSYC)*, vol. 2, no. 1, pp. 135–139, 2020.
- [14] S. Diajukan, S. Persyaratan, U. Memperoleh, G. S. Kom, G. H. Al, and R. Noer, "IMPLEMENTASI ALGORITMA NAÏVE BAYES DAN TF-IDF DALAM ANALISIS SENTIMEN DATA ULASAN (Studi Kasus: Ulasan Review Aplikasi E-Commerce Shopee Di Situs Google Playstore)," 2023.