

# ANALISIS SENTIMEN TANGGAPAN MASYARAKAT TERHADAP CYBERBULLYING DI MEDIA SOSIAL MENGUNAKAN ALGORITMA NAÏVE BAYES ( NB )

**Hanna Sabilla Rifai<sup>1</sup> , Syafira Febrianti<sup>2</sup> , Imam Santoso<sup>3</sup>**

<sup>1</sup> Program Studi Sarjana Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya  
Informatika, <sup>2</sup> Dosen Universitas Teknologi Muhammadiyah Jakarta

[hanna@stikomcki.ac.id](mailto:hanna@stikomcki.ac.id) , [syafira@stikom.ac.id](mailto:syafira@stikom.ac.id) , [imam.santoso@utmj.ac.id](mailto:imam.santoso@utmj.ac.id)

## ABSTRAK

Penelitian analisis sentimen ini dilakukan untuk melihat sebuah opini seseorang atau masyarakat tanggapan terhadap cyberbullying yang berada di media sosial. sehingga dihasilkan opini yang dapat masuk ke kategori opini negatif ataupun opini positif. Dengan adanya penilaian maupun pandangan masyarakat yang disampaikan pada media sosial, maka peneliti berinisiatif membuat penelitian ini. Penelitian menggunakan 10 Fold Cross Validation dalam melakukan validasi, dan untuk mengukur akurasi dilakukan dengan Confusion Matrix, yaitu membandingkan akurasi Naïve Bayes dengan Support Vector Machine tanpa dan dengan menggunakan fitur seleksi Particle Swarm Optimization (PSO). Kemudian Kurva ROC digunakan untuk mengukur nilai AUC. Penelitian ini menggunakan algoritma Naïve Bayes untuk melakukan analisis sentimen. Kemudian menggunakan Accuracy dan AUC (Area Under Curve). Tahapan - tahapan pengujian dilakukan dengan menggunakan dataset sebanyak 376 komentar opini pengguna Youtube ,facebook, Tiktok, Instagram, dan Snack video mengenai Tanggapan masyarakat terhadap cyberbullying maka hasil Akurasi Algoritma Naïve Bayes yaitu sebesar 77.29 % , sedangkan untuk nilai AUC sebesar 0.500. Berdasarkan dari penelitian yang sudah dilakukan, dapat disimpulkan bahwa hasil akurasi dan AUC dengan algoritma Naïve Bayes, hasil pengujian pengklasifikasian AUC nilai keakuratannya dapat dikategorikan sebagai Excellent Classification. algoritma tersebut juga sangat besar, terjadi peningkatan nilai akurasi dan AUC sangat signifikan.

***Kata Kunci : Cyberbullying, Analisis Sentimen, Naïve Bayes***

## ABSTRACT

This sentiment analysis research was conducted to see a person's opinion or public response to cyberbullying on social media. so that opinions are generated that can enter into the category of negative opinion or positive opinion. With the assessment and public views conveyed on social media, the researchers took the initiative to conduct this research. The research uses 10 Fold Cross Validation in conducting validation, and to measure accuracy it is carried out with Confusion Matrix, which compares the accuracy of

Naïve Bayes with Support Vector Machine without and by using the Particle Swarm Optimization (PSO) selection feature. Then the ROC curve is used to measure the AUC value. This study uses the Naive Bayes algorithm to perform sentiment analysis. Then use Accuracy and AUC (Area Under Curve). The stages of testing were carried out using a dataset of 376 comments from Youtube, Facebook, Tiktok, Instagram, and Snack video users regarding public responses to cyberbullying, the results of the accuracy of the Naïve Bayes Algorithm are 77.29%, while the AUC value is 0.500. Based on the research that has been done, it can be concluded that the results of accuracy and AUC with the Naïve Bayes algorithm, the results of the AUC classification test for accuracy can be categorized as Excellent Classification. the algorithm is also very large, there is a very significant increase in the accuracy and AUC values.

**Keywords :** *Cyberbullying, Analisis Sentimen, Naïve Bayes*

## 1. PENDAHULUAN

Teknologi dan internet pada saat ini sudah tidak dapat lagi dipisahkan antara satu dengan lainnya. Hal ini dapat kita simak dengan maraknya kolaborasi antara teknologi dengan internet yang memiliki keluaran yang bervariasi. (Siwi et al., 2018)

Salah satunya adalah media sosial. Media sosial adalah sebuah media online, dengan para penggunanya bisa dengan mudah berpartisipasi, berbagi, dan menciptakan isi meliputi blog, jejaring sosial, wiki, forum dan dunia virtual. Contoh dari media sosial tersebut ada terdapat dari youtube, tiktok, whatsapp, facebook, twitter, instagram, dan lain sebagainya.

(Rachmatan & Ayunizar, 2017)

Media sosial biasa digunakan sebagai sarana komunikasi, informasi serta sebagai media hiburan bagi penggunanya. Instagram merupakan salah satu platform media sosial yang sering digunakan oleh masyarakat terlebih lagi kalangan milenial khususnya di Indonesia. Banyak orang yang tertarik dengan keunggulan media sosial tersebut karena instagram menawarkan ruang untuk berbagi foto dan video serta kemampuan untuk mengaplikasikan filter digital. Akan tetapi dari banyaknya manfaat yang didapatkan masih banyak penggunanya yang tidak memahami etika dalam menggunakan media sosial yang justru menjadikan media sosial

sebagai sarana untuk mengintimidasi orang lain di belakang layar. (Yunita, 2023)

Lembaga riset asal Inggris Ditch The Label yang fokus pada bullying menetapkan Instagram sebagai media sosial dengan kasus cyberbullying paling banyak. Di Indonesia diatur dalam Pasal 27 (3) Undang-Undang Informasi Elektronik Nomor 11 Tahun 2008, yang meliputi penyebaran atau pencemaran nama baik Cyberbullying adalah intimidasi yang terjadi melalui perangkat digital seperti ponsel, komputer, dan tablet. Cyberbullying dapat terjadi melalui SMS, dan aplikasi, atau online di media sosial, forum, atau game tempat orang dapat melihat, berpartisipasi, atau berbagi konten. Cyberbullying termasuk mengirim, memposting, atau membagikan konten negatif, berbahaya, palsu, atau jahat tentang orang lain. Ini dapat mencakup berbagi informasi pribadi atau pribadi tentang orang lain yang menyebabkan rasa malu atau terhina. Beberapa cyberbullying melewati batas menjadi perilaku yang melanggar hukum atau kriminal. Tempat paling umum di mana

cyberbullying terjadi adalah: Media Sosial, seperti youtube Facebook, Instagram, Snapchat, dan Tik Tok Perpesanan teks dan aplikasi perpesanan di perangkat seluler atau tablet Pesan instan, pesan langsung, dan obrolan online melalui internet Forum online, ruang obrolan, dan papan pesan, seperti RedditSurel Komunitas game online.

Menurut penelitian berjudul *A Majority of Teens Have Experienced Some Form of Cyberbullying*, ditemukan bahwa 59% remaja yang menggunakan internet pernah menjadi korban *cyberbullying*. Angka ini lebih besar dari korban berusia dewasa sebesar 33 persen.

Cyberbullying merupakan perilaku berulang yang ditujukan untuk menakuti, membuat marah, atau mempermalukan mereka yang menjadi sasaran. Contohnya termasuk:

- Menyebarkan kebohongan tentang seseorang atau memposting foto memalukan tentang seseorang di media sosial

- Mengirim pesan atau ancaman yang menyakitkan melalui platform *chatting*, menuliskan kata-kata menyakitkan pada kolom komentar media sosial, atau memposting sesuatu yang memalukan/menyakitkan
- Meniru atau mengatasnamakan seseorang (misalnya dengan akun palsu atau masuk melalui akun seseorang) dan mengirim pesan jahat kepada orang lain atas nama mereka.
- Trolling - pengiriman pesan yang mengancam atau menjengkelkan di jejaring sosial, ruang obrolan, atau game online
- Mengucilkan, mengecualikan, anak-anak dari game online, aktivitas, atau grup pertemanan
- Menyiapkan/membuat situs atau grup (group chat, room chat) yang berisi kebencian tentang seseorang atau dengan tujuan untuk menebar kebencian terhadap seseorang
- Menghasut anak-anak atau remaja lainnya untuk mempermalukan seseorang
- Memberikan suara untuk atau menentang seseorang dalam jajak pendapat yang melecehkan
- Membuat akun palsu, membajak, atau mencuri identitas online untuk mempermalukan seseorang atau menyebabkan masalah dalam menggunakan nama mereka
- Memaksa anak-anak agar mengirimkan gambar sensual atau terlibat dalam percakapan seksual.

Bullying secara langsung atau tatap muka dan cyberbullying seringkali dapat terjadi secara bersamaan. Namun cyberbullying meninggalkan jejak digital – sebuah rekaman atau catatan yang dapat berguna dan memberikan bukti ketika membantu menghentikan perilaku salah ini.

Pada penelitian analisis sentimen mengenai komentar pengguna youtube, instagram, tiktok dan facebook dengan menggunakan *Naive Bayes*. Hasil yang didapat dari penelitian tersebut *Naive Bayes* adalah 77.29 % untuk *Accuracy* dan 0.500 untuk *AUC*.

Pada penelitian analisis sentimen ini dilakukan untuk melihat sebuah opini seseorang atau masyarakat tanggapan terhadap cyberbullying yang berada di media sosial ini. sehingga

dihasilkan opini yang dapat masuk ke kategori opini negatif ataupun opini positif. Dengan adanya penilaian maupun pandangan masyarakat yang disampaikan pada media sosial, maka peneliti berinisiatif membuat penelitian ini, peneliti melihat belum ada penelitian tentang opinion mining sebelumnya yang mengangkat permasalahan tersebut, sehingga kedepannya diharapkan dapat bermanfaat membantu untuk melakukan riset atas opini masyarakat yang mengandung sentimen positif, negatif ataupun netral lainnya. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

## 2. LITERATUR DAN METODE

### 2.1 Analisis Sentimen.

*Sentiment Analysis*, atau yang biasa disebut juga *Opinion Mining*, adalah bidang studi yang bertujuan untuk menganalisis opini, sentimen, penilaian, evaluasi, sikap dan emosi publik terhadap suatu entitas dari produk, pelayanan, suatu permasalahan, organisasi, peristiwa tertentu, topik yang hangat dibicarakan dan atributnya.

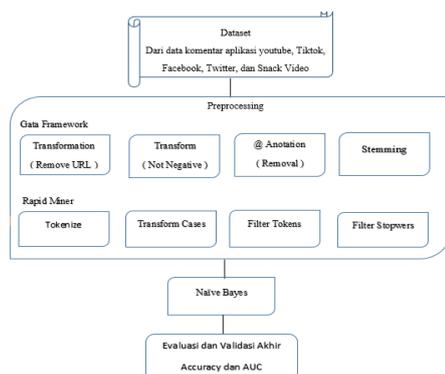
Ada beberapa macam metode yang bisa digunakan pada analisis sentimen, yaitu antara lain, NB (*Naive Bayes*), *Decision Tree*, KNN (*K-Nearest Neighbor*), *Neural Networks* dan juga SVM (*Support Vector Machines*). Dan metode yang saya gunakan yaitu metode Naive Bayes (NB)(*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

### 2.2 Algoritma Naive Bayes

Metode ini menggunakan teorema Bayes, yang ditemukan oleh Thomas Bayes abad ke-18 [5]. Klasifikasi Naive Bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Menurut Wu dan Kumar bahwa Naive Bayes merupakan metode klasifikasi populer dan masuk dalam sepuluh algoritma terbaik dalam data mining. Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan cara melihat frekuensi tiap klasifikasi pada data training.(*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

### 3. METODE

Berikut adalah langkah – langkah metode penelitian yang dijelaskan pada gambar dibawah ini



#### 3.1 Pengumpulan dan Pengolahan Data

Penelitian ini menggunakan data yang didapat dari komentar video youtube, Tiktok, Facebook, Snack Video, dan Instagram milik akun resmi dari 30 maret 2021

<https://youtu.be/sj9NWYFUKhM>,  
<https://vt.tiktok.com/ZSLkwon15/>,  
[https://m.facebook.com/video.php/?video\\_id=948546569300049](https://m.facebook.com/video.php/?video_id=948546569300049),  
<https://sck.io/p/1oKKczjn>, dan  
<https://www.instagram.com/p/CtdE8tqvRsG/?igshid=MTI1ZDU5ODQ3YW==>

dilakukan secara manual satu persatu pada tanggal 20 Juni 2022, data yang digunakan dalam penelitian ini sebanyak 376 data yang terdiri dari 57 data *review* positif dan 319 data *review* negatif

Setelah melakukan pengumpulan data, dilakukan proses pengolahan data atau *preprocessing*, tahapan ini mencakup kegiatan membangun data dan dilanjutkan kegiatan pembersihan data agar dapat dilakukan pengelolaan ketahap selanjutnya. Berikut tahapan *preprocessing* data, antara lain :

##### 1. Remove URL.

Yaitu proses menghilangkan *Uniform Resource Locator* yang ada pada dataset hasil pengumpulan data. Biasanya digunakan untuk merujuk pada *source* atau sumber dari suatu berita/informasi.

##### 2. Transform (Not Negative).

Proses mengubah kata – kata yang bermakna negatif yang akan disatukan dengan tanda garis bawah ( \_ ).

##### 3. Anotation Removal.

Menghapus tanda Mention (@) beserta teks yang ada

dibelakangnya. Pada media sosial biasanya berfungsi untuk memanggil user/pengguna lainnya.

#### 4. *Stemming*.

Tahapan ini berfungsi mengubah kata perkata menjadi sebuah kata dasar, dengan cara menghilangkan imbuhan, baik awalan ataupun akhiran.

#### 5. *Tokenize*.

Memecah sekumpulan karakter atau kalimat menjadi sebuah potongan karakter atau kata – kata sesuai dengan kebutuhan, biasa juga disebut tokenisasi.

#### 6. *Transforms Cases*.

Mengubah huruf kapital yang masih ada di dataset menjadi huruf - huruf kecil. Hal ini bertujuan agar terjadi keseragaman text pada model klasifikasi dan tidak terjadi kesalahan pada proses *tokenize*.

#### 7. *Filter Tokens (By Length)*.

Menghilangkan kata – kata dengan panjang karakter tertentu, biasanya kata yang memiliki hanya 2 karakter tidak memiliki arti.

#### 8. Filter Stopword.

Membuang kata – kata yang diabaikan pada sentimen analisis, biasanya yang berupa kata sambung dan kata keterangan.

*Preprocessing* pertama dilakukan menggunakan gataframework, yaitu pengolahan data teks berbasis *web*, dengan menggunakan bahasa pemrograman *php*. Gataframework dapat diakses di alamat [link](http://www.gataframework.com/textmining) [www.gataframework.com/textmining](http://www.gataframework.com/textmining)

. Banyak yang tersedia di Gataframework, antara lain *annotation removal*, *remove url* dan lainnya. Peneliti menggunakan Gataframework karena memiliki keunggulan dalam melakukan proses *Stemming* bahasa Indonesia. Pengolahan data selanjutnya dilakukan menggunakan tools *Rapidminer*(*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

### 3.2 Seleksi Fitur dan Algoritma

Peneliti mengusulkan menggunakan metode algoritma Naïve Bayes karena algoritma tersebut merupakan algoritma terbaik dalam data mining, dan sangat baik dalam melakukan proses

pengklasifikasian data.(*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

### 3.3 Evaluasi dan Validasi

Pada tahap pengujian metode, peneliti menggunakan *RapidMiner* dalam melakukan proses eksperimen, data *training* yang digunakan adalah komentar video *youtube*, *Tiktok*, *facebook*, *Snack Video* dan *Instagram* mengenai pro dan kontra seseorang atau masyarakat tanggapan terhadap *cyberbullying* yang berada di media sosial Kemudian *dataset* dikelompokkan menjadi 2 bagian yaitu opini positif dan negatif.

Penelitian menggunakan *10 Fold Cross Validation* dalam melakukan validasi, dan untuk mengukur akurasi dilakukan dengan *Confusion Matrix*, yaitu membandingkan akurasi *Naive Bayes* dengan *Support Vector Machine* tanpa dan dengan menggunakan fitur seleksi *Particle Swarm Optimization (PSO)*. Kemudian Kurva *ROC* digunakan untuk mengukur nilai *AUC*.

#### 1. Confusion Matrix

Matriks yang menginformasikan hasil prediksi secara keseluruhan dari nilai

akurasi dan untuk melihat kinerja pengklasifikasi, yaitu seberapa sering kasus *class X* yang benar diklasifikasikan sebagai *class X* atau kesalahan klasifikasi *class* yang lainnya. Ketika *dataset* hanya memiliki dua kelas, yaitu *class* positif dan *class* negatif, maka dapat dibuatkan seperti gambar dibawah ini :

Tabel 1. *Confusion Matrix*

Correct Classification	Classified As	
	+	-
+	True Positive	False Negative
-	False Positive	True Negative

*True* positif (TP) merupakan jumlah *record* positif dalam *dataset* yang diklasifikasikan positif. *True* negatif (TN) merupakan jumlah *record* negatif dalam *dataset* yang diklasifikasikan negatif. *False* positif merupakan jumlah *record* negatif dalam *dataset* yang diklasifikasikan positif. *False* negatif (fn) merupakan

jumlah *record* positif dalam *dataset* yang diklasifikasikan negatif.

Kurva ROC menunjukkan akurasi dan membandingkan klasifikasi secara visual. Kurva ROC mengekspresikan *confusion matrix*. ROC adalah grafik dua dimensi dengan *false* positif sebagai garis horizontal dan *true* positif sebagai garis vertikal. Pedoman umum untuk mengklasifikasikan keakuratan pengujian menggunakan AUC [9].

0.90 - 1.00 = *Excellent Classification*;

0.80 - 0.90 = *Good Classification*;

0.70 - 0.80 = *Fair Classification*;

0.60 - 0.70 = *Poor Classification*;

0.50- 0.60 = *Failure* (*Jurnal Jisamar*

*Kiki Setiawan (1), n.d.*)

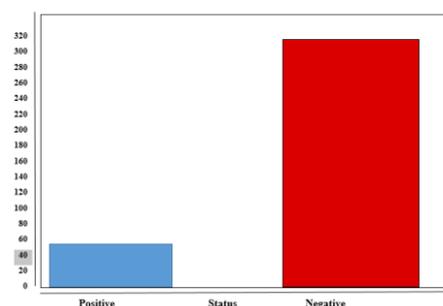
## 4. HASIL DAN PEMBAHASAN

### 4.1 Pengumpulan Data

Hasil pengumpulan data yang didapat dari komentar video youtube, Tiktok, Facebook, Snack Video dan Instagram mengenai Tanggapan Masyarakat terhadap cyberbullying di media sosial menggunakan algoritma Naïve Bayes, peneliti mengumpulkan

376 data opini pengguna youtube, Tiktok, facebook, Snack Video dan Instagram Kemudian data tersebut kumpulkan dalam format file excel, setelah itu data dikelompokkan menjadi 2 bagian yaitu data positif dan data negatif.

Pengelompokan data ini dilakukan oleh beberapa koresponden, sehingga kemudian dihasilkan 319 opini Positif dan 57 opini negatif. (*Jurnal Jisamar Kiki Setiawan (1), n.d.*)



Gambar 2. Statistik hasil pengumpulan data.

### 4.2 Pengolahan Data

Pengolahan data atau *preprocessing*. Berikut ini adalah tahapan *preprocessing* :

1. *Remove URL*.

Yaitu proses menghilangkan *Uniform Resource Locator* yang ada pada dataset hasil pengumpulan data. Berikut contoh perbedaan sebelum dan Setelah pengumpulan data dan membagi data menjadi komentar positif dan komentar negatif, selanjutnya dilakukan proses sesudah proses. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Biar tidak peduli sama omongan orang lain punya dua tangan pake buat nutup telinga karena kalau nutup mulut mereka tangan tidak_cukup	Biar tidak peduli sama bicara orang lain punya dua tangan pake buat nutup telinga karena kalau nutup mulut mereka tangan tidak cukup
---	--

### 3. Anotation Removal.

Menghapus tanda *Mention* (@) beserta teks yang ada dibelakangnya. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Tabel 4. Hasil Anotation Removal.

Sebelum	Sesudah
@m_yyyyy_g Bro semua sifat dan kebiasaan lu disebutin ini	semua sifat kebiasaan kamu disebutin ini

Tabel 2. Hasil Remove URL.

Sebelum	Sesudah
Buzzer malah lebih kejam daripada bullying yang telah membodohi NKRI sampai bangsa terbodoh se asian dan sedunia	Buzzer malah lebih kejam daripada bullying telah membodohi NKRI sampai bangsa bodoh asia dunia

### 2. Transform (Not Negative).

Proses mengubah kata – kata yang bermakna negatif yang akan disatukan dengan tanda garis bawah ( \_ ). (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Tabel 3. Hasil Transform Negative.

Sebelum	Sesudah
---------	---------

### 4. Stemming.

Tahapan ini berfungsi mengubah kata perkata menjadi sebuah kata dasar, dengan cara menghilangkan imbuhan, baik awalan ataupun akhiran. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Tabel 5. Hasil Stemming.

Sebelum	Sesudah
Buzzer malah lebih kejam daripada bullying yang telah membodohi NKRI sampai bangsa terbodoh se asian dan sedunia	Buzzer malah lebih kejam daripada bullying telah membodohi NKRI sampai bangsa bodoh asia dunia

### 5. Tokenize.

Memecah sekumpulan karakter atau kalimat menjadi sebuah potongan karakter atau kata – kata sesuai dengan kebutuhan, biasa juga disebut tokenisasi. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

6. *Transforms Cases.*

Mengubah huruf kapital yang masih ada di dataset menjadi huruf - huruf kecil. Hal ini bertujuan agar terjadi keseragaman text pada model klasifikasi dan tidak terjadi kesalahan pada proses *tokenize*. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Tabel 6. Hasil *Transforms Cases*.

Sebelum	Sesudah
Sudah ATTITUDE	sudah attitude
rendah pada individualisme	rendah pada individualisme
primitif terancam	primitif terancam
punah juga	punah juga

7. *Filter Tokens (By Lenght).*

Menghilangkan kata – kata dengan panjang karakter tertentu, biasanya kata yang memiliki hanya 2 karakter tidak memiliki arti. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

8. *Filter Stopword.*

Membuang kata – kata yang diabaikan pada sentimen analisis,

biasanya yang berupa kata sambung dan kata keterangan. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

Tabel 7. Hasil *Filter Stopword*.

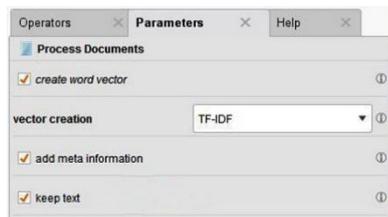
Sebelum	Sesudah
dikasih tau malah ' ini itu	kasih tau malah
Cyber BullyIng Tau	cyberbullying tau kesan
kesannya malah kayak	malah seperti ejek
ngejek	

Berikut gambar proses pengolahan data pada *tools rapidminer*.



Gambar 3. Proses Pengolahan Data Rapidminer.

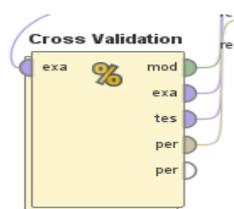
Data hasil *preprocessing* yang berupa kata akan diubah ke dalam bentuk angka dengan dilakukan proses pembobotan kata yang bertujuan untuk menghitung bobot pada masing-masing kata yang akan digunakan sebagai fitur, semakin banyak dokumen yang akan diproses maka semakin banyak fitur. Pada tahapan ini terdapat dua bagian proses yaitu *TF (Term Frequency)* dan *IDF (Inverse Document Frequency)*, (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)



Gambar 4. Proses *Tf Idf*.

### 5. Proses Validasi Algoritma.

Selanjutnya adalah penggunaan algoritma *Naive Bayes* untuk melakukan klasifikasi data yang dihubungkan dengan pengujian *10-fold cross validation* dimana proses ini untuk mengevaluasi proses kerja algoritma tersebut dengan membagi data secara acak ke dalam *10 fold* untuk mendapatkan 10 data yang sama, kemudian data tersebut digunakan *9 fold* untuk data *training* dan *1 fold* untuk data *testing*.

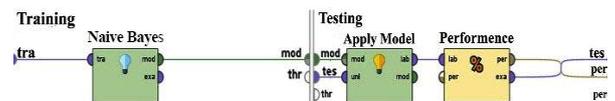


Gambar 5. Proses *Cross Validation* dengan *Rapidminer*.

Tahapan selanjutnya adalah proses Evaluasi, ialah mengevaluasi kinerja terhadap permodelan dan perhitungan yang telah digunakan dengan menerapkan algoritma *Naive*

*Bayes*. dihasilkan peningkatan akurasi pada penghitungan. Untuk mengestimasi performa dari model algoritma yang telah dipilih digunakan *Cross Validation* sehingga dapat menghasilkan nilai akurasi (Proses *Cross Validation* dapat dilihat pada Gambar 5).

Dibawah ini adalah model pengujian *Naive Bayes* dengan menggunakan *tools Rapidminer*. (Jurnal *Jisamar Kiki Setiawan (1)*, n.d.)



Gambar 6. Model Pengujian *Naive Bayes* dengan *RapidMiner*

### 6. Evaluasi

Penelitian ini menggunakan algoritma *Naive Bayes* untuk melakukan analisis sentimen. Kemudian dalam menentukan evaluasinya peneliti menggunakan *Accuracy* dan *AUC (Area Under Curve)*. Tahapan - tahapan pengujian yang sudah dilakukan dengan menggunakan *dataset* sebanyak 376 komentar opini pengguna *Youtube* dan *facebook* mengenai video *Indonesia stop sementara kirim TKI*

ke malaysia maka hasil Akurasi Algoritma *Naive Bayes* yaitu sebesar 77.29 %, sedangkan untuk nilai *AUC* sebesar 0.500.

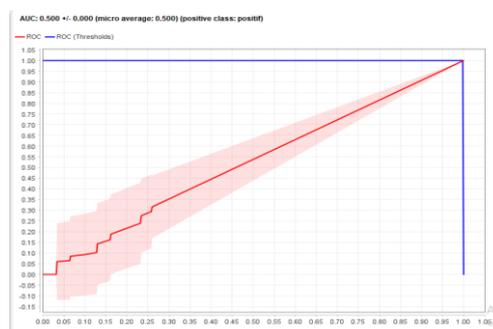
Berikut ini adalah Tabel *Confusion Matrix* Algoritma *Naive Bayes*

Tabel 7. *Confusion Matrix* Algoritma *NB*.

accuracy: 77.29% +/- 7.29% (micro average: 77.25%)			
	true negatif	true positif	class precision
pred. negatif	265	42	86.32%
pred. positif	39	10	20.41%
class recall	67.17%	19.23%	

Pada tabel 8 *Confusion Matrix* dapat dilihat, sebanyak 39 data diprediksi *class* positif ternyata sesuai, yaitu masuk kedalam *class* positif, sebanyak 265 data yang diprediksi *class* negatif ternyata termasuk kedalam prediksi *class* positif.

Dan sebanyak 10 data yang diprediksi *class* positif ternyata masuk dalam *class* negatif, kemudian 42 data di prediksi *class* negatif sesuai yaitu termasuk kedalam prediksi *class* negatif.



Gambar 8. Grafik *Area Under Curve* (*AUC*) *NB*.

Inilah tahapan - tahapan yang sudah dilakukan dengan menggunakan dataset sebanyak 376 komentar opini pengguna *Youtube*, *Tiktok*, facebook, *Snack Video*, dan *Instagram* mengenai video Tanggapan Masyarakat terhadap *Cyberbullying* di media sosial, maka hasil Akurasi Algoritma *Naive Bayes* dari *Rapidminer*. (*Jurnal Jisamar Kiki Setiawan (1)*, n.d.)

## 5. KESIMPULAN

Berdasarkan dari penelitian yang sudah dilakukan, dapat disimpulkan bahwa hasil akurasi dan *AUC* dengan algoritma *Naive Bayes*, hasil pengujian pengklasifikasian *AUC* nilai keakuratannya dapat dikategorikan sebagai *Excellent Classification*. algoritma tersebut

juga sangat besar, dapat dilihat dari hasil pengujian diatas, terjadi peningkatan nilai akurasi dan *AUC* sangat signifikan. Dan pada penelitian ini juga, hasil akurasi dan *AUC* dengan menggunakan algoritma *Naive Bayes* memiliki hasil yang lebih baik. Untuk mendukung penelitian ini, peneliti akan mengembangkan aplikasi opini dari komentar youtube, Tiktok, facebook, Snack Video, dan Instagram mengenai Tanggapan Masyarakat terhadap Cyberbullying di media sosial untuk mengklarifikasikan opini negatif dan positif menggunakan bahasa pemrograman php.

- 18(2), 257–262.  
<http://ejournal.bsi.ac.id/ejurnal/index.php/cakrawala%0APengaruh>
- 4) Yunita, R. (2023). Perundungan Maya (Cyber Bullying) Pada Remaja Awal. *Muhafadzah*, 1(2), 93–110.  
<https://doi.org/10.53888/muhafadzah.v1i2.430>

## DAFTAR PUSTAKA

- 1) *jurnal jisamar Kiki Setiawan (1)*. (n.d.).
- 2) Rachmatan, R., & Ayunizar, S. R. (2017). PADA REMAJA SMA DI BANDA ACEH Risana Rachmatan. *Jurnal UNMUH Jember*, 13(2), 67–79.  
<http://jurnal.unmuhjember.ac.id/index.php/INSIGHT/article/view/811>
- 3) Siwi, A., Utami, F., & Baiti, N. (2018). *Pengaruh Media Sosial Terhadap Perilaku Cyber Bullying Pada Kalangan R*Siwi, A., Utami, F., & Baiti, N. (2018). *Pengaruh Media Sosial Terhadap Perilaku Cyber Bullying Pada Kalangan Remaja*. 18(2), 257–262.emaja.