

Analisis Sentimen *Cyberbullying* Pada Komentar X Menggunakan Metode *Naïve Bayes*

Aprizal Machmud¹, Bany Wibisono², Nining Suryani^{3*}, Yunita⁴

^{1,2,3,4} Universitas Bina Sarana Informatika
Jl. Kramat Raya No.98 Kwitang, Kec. Senen, Jakarta Pusat, Indonesia

e-mail : aprizalmachmud1146@gmail.com¹, wibisonobany@gmail.com², ningning.nns@bsi.ac.id³,
yunita.ynt@bsi.ac.id⁴

(*) Corresponding Author

Artikel Info : Diterima : 23-08-2024 | Direvisi : 09-12-2024 | Disetujui : 16-12-2024

Abstrak - Kemajuan teknologi komunikasi dan media sosial mempermudah akses informasi global, namun juga meningkatkan kasus *cyberbullying* di platform seperti X. Dampak dari *cyberbullying* dapat mencakup gangguan fisik dan psikologis, seperti peningkatan kesepian, kecemasan, depresi, serta penurunan harga diri. Selain itu, korban *cyberbullying* mungkin merasakan tekanan yang dapat meningkatkan risiko keinginan untuk bunuh diri. Penelitian ini menggunakan metode *Naïve Bayes* untuk secara efektif dan efisien mengklasifikasikan komentar terkait *cyberbullying*. Model klasifikasi ini dikembangkan untuk mendeteksi *cyberbullying* pada komentar di X, menggunakan algoritma *Naïve Bayes* dan *dataset* dari Kaggle yang terdiri dari 650 data komentar yang mengandung ciri-ciri *cyberbullying* maupun yang tidak. Penelitian ini mencakup beberapa tahap *preprocessing* seperti tokenisasi, normalisasi, dan *stemming*. Data kemudian dibagi menjadi dua bagian: 80% untuk data *training* dan 20% untuk data *testing*. Hasil evaluasi menunjukkan akurasi model sebesar 80,77%, *precision* 81,25%, *recall* 70,91%, dan *AUC* 0,794. Inovasi dalam penelitian ini terletak pada penggunaan 2 (dua) operator *stemming* yaitu *stemming dictionary* dan *stemming snowball* yang dimana *stemming dictionary* menggunakan *file* khusus yang berisi kata-kata singkatan atau *slang*, yang sering digunakan dalam komentar di X sedangkan *stemming snowball* digunakan untuk mengonversikan kata-kata menjadi bentuk dasarnya. Model ini cenderung lebih akurat dalam mengklasifikasikan komentar sebagai *non-bullying* dibandingkan *bullying*. Saran peningkatan meliputi eksplorasi metode *preprocessing* dan algoritma lain, serta penggunaan *dataset* yang lebih besar dan bervariasi.

Kata Kunci : *Cyberbullying*, X, Klasifikasi, *Naive Bayes*, *Preprocessing*

Abstracts - Advances in communication technology and social media have made it easier to access global information, but have also increased cases of *cyberbullying* on platforms such as X. The impact of *cyberbullying* can include physical and psychological disorders, such as increased loneliness, anxiety, depression, and decreased self-esteem. In addition, victims of *cyberbullying* may feel distress that can increase the risk of suicidal ideation. This research utilizes the *Naïve Bayes* method to effectively and efficiently classify *cyberbullying*-related comments. This classification model was developed to detect *cyberbullying* in comments on X, using the *Naïve Bayes* algorithm and a dataset from Kaggle consisting of 650 comments that contain *cyberbullying* characteristics and those that do not. This research includes several preprocessing steps such as tokenization, normalization, and stemming. The data was then divided into two parts: 80% for training data and 20% for testing data. The evaluation results show a model accuracy of 80.77%, precision 81.25%, recall 70.91%, and AUC 0.794. The innovation in this research lies in the use of 2 (two) stemming operators, namely *stemming dictionary* and *stemming snowball*, where the *stemming dictionary* uses a special file containing abbreviations or slang words, which are often used in comments on the word becomes its basic form. This model tends to be more accurate in classifying comments as *non-bullying* than *bullying*. Suggestions for improvement include exploring other preprocessing methods and algorithms, as well as using larger and more varied datasets.

Keywords : *Cyberbullying*, X, Classification, *Naive Bayes*, *Preprocessing*

PENDAHULUAN

Di era globalisasi, perkembangan teknologi komunikasi memainkan peran sentral. Penggunaan media elektronik menjadi syarat utama, memungkinkan akses informasi dari seluruh dunia melalui internet dengan



mudah. Ini memberikan kemudahan dalam menjangkau berbagai informasi tanpa Batasan (Pakpahan, 2020). Informasi memainkan peran besar dalam kehidupan manusia dengan memungkinkan pengenalan lingkungan dan prediksi situasi yang dihadapi. Ini sejalan dengan perkembangan teknologi, di mana kedua hal tersebut berkembang bersama-sama. Majunya teknologi berarti kemajuan dalam bidang informasi (Yulieta F, 2021). Perkembangan teknologi telah menciptakan berbagai platform media sosial. Media sosial berperan penting dalam mengubah perilaku interaksi sosial para pengguna internet (Jubaidi, 2020).

Dengan kemajuan teknologi saat ini, banyak orang menggunakan media sosial sebagai sarana untuk berbagi informasi dan mengekspresikan diri, salah satunya adalah X. X adalah salah satu platform media sosial terbesar dan paling populer, dengan jutaan pengguna aktif setiap harinya (Cantika Samha A, 2023). X memiliki tingkat *anonimitas* yang relatif tinggi, yang sering digunakan penggunanya untuk mengungkapkan pendapatnya tanpa rasa takut. Namun, terdapat keganjilan berupa *cyberbullying* atau pelecehan online yang telah menjadi masalah yang sulit disangkal di lingkungan yang terkomputerisasi (Aldi, 2023). Tindakan intimidasi, mempermalukan, merendahkan, atau menyebabkan kerusakan emosional melalui teknologi digital, seperti komentar negatif di X, dikenal sebagai *cyberbullying*. Mengomentari di X adalah metode komunikasi yang sering digunakan (Nurlaely, 2023).

Cyberbullying dapat menggunakan komentar sebagai sarana untuk serangan verbal atau emosional terhadap korbannya. Contohnya adalah komentar yang merendahkan penampilan fisik, agama, etnis, orientasi seksual, atau karakteristik lainnya (Siroj M, 2024). Korban yang mengalami *cyberbullying* dapat menghadapi gangguan fisik dan psikologis, seperti kesepian, kecemasan, depresi yang lebih tinggi, serta penurunan harga diri. Selain itu, korban *cyberbullying* juga mungkin merasakan tekanan yang dapat meningkatkan keinginan untuk bunuh diri (Fitra Rizki, 2021).

Beberapa metode yang digunakan didalam menganalisis sentiment diantaranya, penelitian yang dilakukan oleh Firda Millennialita dkk yaitu melakukan komparasi Naïve Bayes dan Support Vector Machine untuk sentiment analisis terhadap respon bullying di Twitter dengan focus terhadap serangan bullying terhadap Kim Gara, dari LE SSERAFIM yang merupakan Idol dari Korea Selatan dimana didapatkan hasil algoritma Naïve Bayes dengan akurasi 97% sedangkan SVM menghasilkan akurasi sebesar 93% (Millennialita et al., 2024).

Penelitian yang dilakukan oleh Muhammad Iqbal dkk, menggunakan algoritma naïve bayes, dan Decision Tree terhadap ulasan aplikasi threads dan Twitter terhadap 899 komentar dengan 245 komentar negative dan 654 komentar positif, hasil akurasi yang didapatkan pada aplikasi twitter, naïve bayes sebesar 85.56%, sedangkan Decision Tree menghasilkan akurasi sebesar 72.78% sedangkan pada aplikasi threads naïve bayes menghasilkan akurasi 66.41% sedangkan Decision Tree menghasilkan akurasi sebesar 65.41% (Iqbal et al., 2023a).

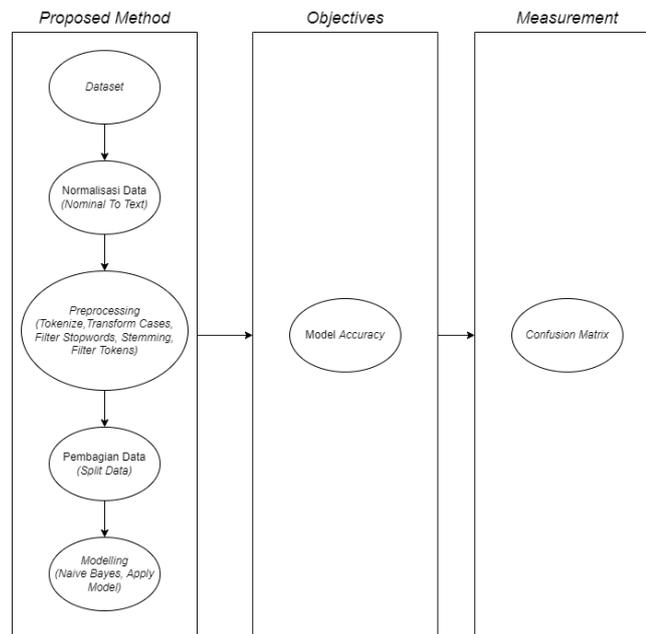
Metode *Naive Bayes* sangat cocok digunakan dalam penelitian ini karena sederhana dan efektif untuk tugas klasifikasi teks dengan menangani data teks secara baik melalui asumsi independensi antar fitur, memiliki kinerja yang baik, menawarkan pemrosesan yang cepat sehingga dapat mengolah dan mengklasifikasikan sejumlah besar data dengan efisien, dapat meningkatkan akurasi identifikasi komentar yang mengandung *cyberbullying*, serta mudah diterapkan dengan berbagai alat yang tersedia seperti *RapidMiner*, sehingga kombinasi dari kemudahan implementasi, efisiensi komputasi, dan kinerja yang solid menjadikannya pilihan yang tepat untuk penelitian ini (Rifai, 2023).

Inovasi dalam penelitian ini terletak pada penggunaan 2 (dua) operator *stemming* yaitu *stemming dictionary* dan *stemming snowball* yang dimana *stemming dictionary* menggunakan *file* khusus yang berisi kata-kata singkatan atau *slang*, yang sering digunakan dalam komentar di X sedangkan *stemming snowball* digunakan untuk mengonversikan kata-kata menjadi bentuk dasarnya.

Tujuan penelitian ini adalah untuk menerapkan dan menguji model algoritma *naive bayes* dalam mengklasifikasi terhadap komentar di X yang mengandung unsur *bullying* dan *non bullying*.

METODE PENELITIAN

2.1. Kerangka Penelitian



Sumber: Hasil Penelitian (2024)

Gambar 2. Kerangka Pemikiran

Pada gambar 2 diatas, penelitian ini dimulai dengan memilih *dataset* yang sesuai untuk analisis. Lalu *dataset* tersebut kemudian akan diolah menggunakan operator *nominal to text* untuk mengubah seluruh nilai numerik menjadi *text* yang dapat diproses lebih lanjut. Proses selanjutnya melibatkan tahapan *preprocessing*, yang mencakup tokenisasi untuk memecah teks menjadi bagian-bagian kata yang lebih kecil, transformasi cases untuk mengubah teks menjadi huruf kecil, filter *stopwords* untuk menghapus kata-kata umum yang tidak relevan, filter token untuk menghapus kata-kata yang panjangnya telah ditentukan, dan *stemming* untuk mengonversi kata-kata menjadi bentuk dasarnya dan mengatasi kata singkatan dan slang.

Setelah tahapan *preprocessing* selesai, *dataset* akan dibagi menggunakan operator *split data* untuk pembagian data menjadi subset data pelatihan (*training*) dan data pengujian (*testing*). Metode klasifikasi *Naive Bayes* akan diterapkan pada data pelatihan untuk membangun model klasifikasi.

Selanjutnya, model yang telah dibuat akan diterapkan pada data pengujian (*testing*) menggunakan operator *apply model* untuk mengklasifikasikan data pengujian. Terakhir, kinerja model akan dievaluasi menggunakan operator *performance* menggunakan aplikasi *RapidMiner* untuk menganalisis seberapa baik model *Naive Bayes* dapat mengklasifikasikan data pengujian

HASIL DAN PEMBAHASAN

Hasil dari Penelitian sedang dilakukan untuk menganalisis penggunaan algoritma *Naive Bayes* dalam mengklasifikasikan komentar apakah termasuk kategori *cyberbullying* atau *non-cyberbullying* pada komentar *X* dengan pengujian yang akurat. Berikut adalah langkah-langkah yang akan diambil dalam analisis sentimen pada penelitian ini.

3.1.1 Dataset

Dalam penelitian ini, pengumpulan data dilakukan dengan mengakses dan menggunakan *dataset* yang tersedia pada website resmi bernama *www.kaggle.com*. Data yang diambil berisi 650 data komentar media sosial *X* yang memiliki ciri ciri *cyberbullying* dan *non-cyberbullying*. Berikut adalah tampilan dataset komentar *twitter*.

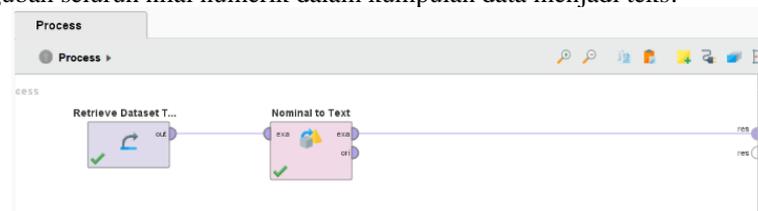
Row No.	Kategori	Komentar
1	Bullying	"Makin jelek aja anaknya, padahal ibu ayahnya cakep2"
2	Bullying	"jkk astagah sok bat cakep njs"
3	Non-bullying	"Kaya tulus ya mukanya klo dilas2"
4	Bullying	"Jadi diri sendiri, jangan jadi ular apalagi dugong"
5	Bullying	"Aduuh, terlalu halu. hahaha lebay editan nyaa lemak kemana lemak... hahaha. mimpi ketinggian."
6	Non-bullying	"nikmatnya bisa menghirup oksigen dengan bebas ya :)"
7	Bullying	"Mudah mudahan lu jadi gembel. Amin ya Allah!"
8	Non-bullying	"masih good vibes walaupun sedang lapar"
9	Bullying	"Kakinya gedong sayyy kek kaki jantan... Upp kan dia suka berbau ama wanwerwor. Mahluk astral. Yg kakinya gedong"
10	Non-bullying	"Kukira kehadiranmu akan menjadi sesuatu kebanggaan, ternyata tidak"

Sumber: Hasil Penelitian (2024)

Gambar 3. Dataset Komentar Twitter

3.1.2 Normalisasi Data

Setelah *dataset* dimasukkan ke dalam aplikasi *rapidminer*, data akan di normalisasikan agar algoritma dapat memproses data dengan lebih mudah. Normalisasi dilakukan menggunakan fitur *nominal to text* yang digunakan untuk mengubah seluruh nilai numerik dalam kumpulan data menjadi teks.



Sumber: Hasil Penelitian (2024)

Gambar 4. Proses *Normalisasi Data*

Komentar	Kategori
polynomial	binomial label
1 "Makin jelek aja anaknya, padahal ibu ayahnya cakep2"	1
2 "jkk astagah sok bat cakep njs"	1
3 "Kaya tulus ya mukanya klo dilas2"	0
4 "Jadi diri sendiri, jangan jadi ular apalagi dugong"	1
5 "Aduuh, terlalu halu. hahaha lebay editan nyaa lemak kemana lemak... hahaha. mimpi ketinggian."	1
6 "nikmatnya bisa menghirup oksigen dengan bebas ya :)"	0
7 "Mudah mudahan lu jadi gembel. Amin ya Allah!"	1
8 "masih good vibes walaupun sedang lapar"	0
9 "Kakinya gedong sayyy kek kaki jantan... Upp kan dia suka berbau ama wanwerwor. M."	1
10 "Kukira kehadiranmu akan menjadi sesuatu kebanggaan, ternyata tidak"	0

Sumber: Hasil Penelitian (2024)

Gambar 4. Data Sebelum Dinormalisasikan

Pada Gambar 4 adalah tampilan data sebelum dinormalisasikan, *attribute* pada data komentar yaitu kategori memiliki nilai nominal, nilai 0 untuk menunjukkan *non-cyberbullying* dan nilai 1 untuk *cyberbullying*.

Row No.	Kategori	Komentar
1	Bullying	"Makin jelek aja anaknya, padahal ibu ayahnya cakep2"
2	Bullying	"jkk astagah sok bat cakep njs"
3	Non-bullying	"Kaya tulus ya mukanya klo dilas2"
4	Bullying	"Jadi diri sendiri, jangan jadi ular apalagi dugong"
5	Bullying	"Aduuh, terlalu halu. hahaha lebay editan nyaa lemak kemana lemak... hahaha. mimpi ketinggian."
6	Non-bullying	"nikmatnya bisa menghirup oksigen dengan bebas ya :)"
7	Bullying	"Mudah mudahan lu jadi gembel. Amin ya Allah!"
8	Non-bullying	"masih good vibes walaupun sedang lapar"
9	Bullying	"Kakinya gedong sayyy kek kaki jantan... Upp kan dia suka berbau ama wanwerwor. Mahluk astral. Yg kakinya gedong"
10	Non-bullying	"Kukira kehadiranmu akan menjadi sesuatu kebanggaan, ternyata tidak"

Sumber: Hasil Penelitian (2024)

Gambar 5. Data Sesudah Dinormalisasikan

Gambar 5 adalah tampilan data sesudah dinormalisasikan, nilai *attribute* yang sebelumnya memiliki nilai nominal dengan angka 0 dan 1 telah di ubah menjadi teks menjadi *cyberbullying* dan *non-cyberbullying*.

3.1.3 Preprocessing

Preprocessing mempermudah proses berikutnya dengan menghapus informasi yang tidak penting atau mengubah data ke dalam format yang lebih sesuai. Berikut adalah proses *preprocessing*.

A. Tokenize

Langkah pertama yang dilakukan yaitu proses *tokenize*. Proses *Tokenize* digunakan untuk membagi kalimat menjadi kata-kata dengan cara memisahkan kalimat berdasarkan spasi di antara kata-kata. Hasil *tokenize* bisa dilihat pada *table 1*.

Tabel 1 Hasil Tokenize

Sebelum <i>Tokenize</i>	Sesudah <i>Tokenize</i>
Saya berharap kak @isyanasarasvati panjang umur. Sehingga bisa jadi inspirasi untuk musisi seluruh Indonesia. Saya suka musik indonesia dengan adanya kak Isyana	[Saya,berharap,kak,@isyanasarasvati,panjang,umur,..,Sehingga,bisa,jadi,inspirasi,untuk,musisi,seluruh,Indonesia,..,Saya,suka,musik,indonesia,dengan,adanya,kak,Isyana]
Hua seneng banget kemaren walaupun Cuma 35 menitan tapi seneng liat isyan, adaptasi kebiasaan baruku smga bisa lebih teratur maemnya, ditunggu karya selanjutnya kaka	[Hua,seneng,banget,kemaren,walaupun,Cuma,35,menitan,tapi,seneng,liat,isyana,adaptasi,kebiasaan,baru,suga,bisa,lebih,teratur,maemnya,ditunggu,karya,selanjutnya,kaka]

Sumber: Hasil Penelitian (2024)

B. Transform Cases

Langkah kedua yaitu *transform cases*. Tahap *Transform cases* digunakan untuk mengubah huruf pada kata-kata dalam dokumen menjadi huruf kecil secara keseluruhan. Hasil *transform cases* bisa dilihat pada *table 2*.

Tabel 2 Hasil Transform Cases

Sebelum <i>Transform cases</i>	Sesudah <i>Transform cases</i>
[Saya,berharap,kak,@isyanasarasvati,panjang,umur,..,Sehingga,bisa,jadi,inspirasi,untuk,musisi,seluruh,Indonesia,..,Saya,suka,musik,indonesia,dengan,adanya,kak,Isyana]	[saya,berharap,kak,@isyanasarasvati,panjang,umur,..,sehingga,bisa,jadi,inspirasi,untuk,musisi,seluruh,Indonesia,..,saya,suka,musik,indonesia,dengan,adanya,kak,isyana]
[Hua,seneng,banget,kemaren,walaupun,Cuma,35,menitan,tapi,seneng,liat,isyana,adaptasi,kebiasaan,baru,suga,bisa,lebih,teratur,maemnya,ditunggu,karya,selanjutnya,kaka]	[hua,seneng,banget,kemaren,walaupun,cuma,35,menitan,tapi,seneng,liat,isyana,adaptasi,kebiasaan,baru,suga,bisa,lebih,teratur,maemnya,ditunggu,karya,selanjutnya,kaka]

Sumber: Hasil Penelitian (2024)

C. Stopwords

Langkah ketiga yaitu *stopwords*. *Filter stopwords* berfungsi untuk menghilangkan kata-kata yang tidak relevan pada kalimat dengan daftar *stopwords* yang telah ditentukan. Hasil *stopwords* bisa dilihat pada *table 3*

Tabel 3 Hasil Stopwords

Sebelum <i>Stopwords</i>	Sesudah <i>Stopwords</i>
[saya,berharap,kak,@isyanasarasvati,panjang,umur,..,sehingga,bisa,jadi,inspirasi,untuk,musisi,seluruh,Indonesia,..,saya,suka,musik,indonesia,dengan,adanya,kak,isyana]	[berharap,kak,@isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,kak,isyana]
[hua,seneng,banget,kemaren,walaupun,cuma,35,menitan,tapi,seneng,liat,isyana,adaptasi,kebiasaan,baru,suga,bisa,lebih,teratur,maemnya,ditunggu,karya,selanjutnya,kaka]	[seneng,kemaren,35,menitan,seneng,liat,isyana,adaptasi,kebiasaan,suga,lebih,teratur,maemnya,ditunggu,karya,kaka]

Sumber: Hasil Penelitian (2024)

D. Filter Tokens By Length

Langkah keempat yaitu *tokens by length*. Pada tahapan *Filter tokens by length* dilakukan untuk membatasi kata-kata dengan cara menghapus kata-kata yang panjangnya melebihi atau kurang dari panjang yang telah ditentukan. Hasil *tokens by length* bisa dilihat pada *table 4*.

Tabel III. 4 Hasil Tokens by Length

Sebelum <i>Filter Token by Length</i>	Sesudah <i>Filter Token by Length</i>
[berharap,kak,@isyanasarasvati,panjang,umur,inspirasi,i,musisi,seluruh,Indonesia,suka,musik,indonesia,kak,isyana]	[berharap,isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,isyana]
[seneng,kemaren,35,menitan,seneng,liat,isyana,adaptasi,kebiasaan,smga,lebih,teratur,maennya,ditunggu,karya,kaka]	[seneng,kemaren,menitan,seneng,liat,isyana,adaptasi,kebiasaan,smga,lebih,teratur,maennya,ditunggu,karya,kaka]

Sumber: Hasil Penelitian (2024)

E. Stemming

Langkah kelima atau terakhir yaitu *stemming*. Pada tahapan ini, menggunakan 2 operator *stemming* yaitu *stemming dictionary* dan *stemming snowball* yang dimana *stemming dictionary* menggunakan file khusus yang berisi kata-kata singkatan atau slang, sedangkan *stemming snowball* digunakan untuk mengonversikan kata-kata menjadi bentuk dasarnya. Hasil *stemming* bisa dilihat pada *table 5* dan *6*.

Tabel 5 Hasil Stemming Dictionary

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
[berharap,isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,isyana]	[harap,isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,isyana]
[seneng,kemaren,menitan,seneng,liat,isyana,adaptasi,kebiasaan,smga,lebih,teratur,maennya,ditunggu,karya,kaka]	[seneng,kemaren,menit,seneng,liat,isyana,adaptasi,kebiasaan,semoga,lebih,atur,maen,tunggu,karya,kaka]

Sumber: Hasil Penelitian (2024)

Tabel 6 Hasil Stemming Snowball

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
[berharap,isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,isyana]	[harap,isyanasarasvati,panjang,umur,inspirasi,musisi,seluruh,Indonesia,suka,musik,indonesia,isyana]
[seneng,kemaren,menitan,seneng,liat,isyana,adaptasi,kebiasaan,semoga,lebih,teratur,maennya,ditunggu,karya,kaka]	[seneng,kemaren,menit,seneng,liat,isyana,adaptasi,kebiasaan,moga,lebih,atur,maen,tunggu,karya,kaka]

Sumber: Hasil Penelitian (2024)

3.1.4 Pembagian Data (*Split Data*)

Setelah proses *preprocessing* selesai, *dataset* akan dibagi menjadi dua subset yang berbeda yaitu data *training* dan data *testing* menggunakan *split data*. tahapan pembagian data di proporsikan dengan *ratio* 0.8 dan 0.2 atau 80% untuk data *training* dan 20% untuk data *testing*. Kami menguji *ratio* 60:40, 70:30, 80:20, dan 90:10 untuk membandingkan *ratio* mana yang lebih akurat dalam memprediksi dan memberikan hasil yang optimal.

Untuk *ratio* 60:40 mendapatkan nilai *accuracy* sebesar 76.92%, *precision* sebesar 74.56%, *recall* sebesar 73.28% dan *AUC* sebesar 0.717.

accuracy: 76.92%			
	true Bullying	true Non-bullying	class precision
pred. Bullying	115	31	78.77%
pred. Non-bullying	29	85	74.58%
class recall	79.86%	73.28%	

Sumber: Hasil Penelitian (2024)
 Gambar 6. Hasil Accuracy ratio 60:40

Untuk ratio 70:30 mendapatkan nilai *accuracy* sebesar 74.87%, *precision* sebesar 72.15%, *recall* sebesar 67.86% dan *AUC* sebesar 0.706.

accuracy: 74.87%			
	true Bullying	true Non-bullying	class precision
pred. Bullying	89	27	76.72%
pred. Non-bullying	22	57	72.15%
class recall	80.18%	67.86%	

Sumber: Hasil Penelitian (2024)
 Gambar 7. Hasil Accuracy ratio 70:30

Untuk ratio 80:20 mendapatkan nilai *accuracy* sebesar 80.77%, *precision* sebesar 81.25%, *recall* sebesar 70.91% dan *AUC* sebesar 0.794. ini bisa dilihat pada gambar 10.

Untuk ratio 90:10 mendapatkan nilai *accuracy* sebesar 78.46%, *precision* sebesar 85.71%, *recall* sebesar 62.07% dan *AUC* sebesar 0.794.

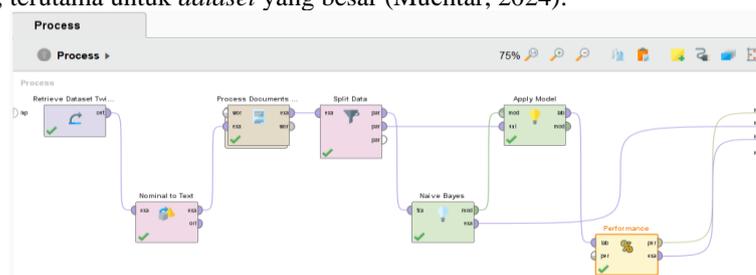
accuracy: 78.46%			
	true Bullying	true Non-bullying	class precision
pred. Bullying	33	11	75.00%
pred. Non-bullying	3	18	85.71%
class recall	91.67%	62.07%	

Sumber: Hasil Penelitian (2024)
 Gambar 8. Hasil Accuracy ratio 90:10

Pembagian data dengan rasio 80:20 menunjukkan bahwa model mampu memprediksi dengan akurat dan memberikan hasil kinerja yang optimal (Iqbal, 2023). Proses pembagian data (*split data*) melibatkan pembagian *dataset* menjadi dua bagian, yaitu subset pelatihan (*training*) dan subset pengujian (*testing*). Pembagian data dengan rasio 80:20 menunjukkan bahwa model mampu memprediksi dengan akurat dan memberikan hasil kinerja yang optimal. (Azrul, 2024). Berdasarkan percobaan yang dilakukan hasil akurasi terbaik didapatkan dengan pembagian data dengan rasio 80:20.

3.1.5 Modelling

Modelling dilakukan dengan menggunakan algoritma klasifikasi teks yaitu *naive bayes* dan operator *apply model*. Metode *naive bayes* dipilih untuk penelitian ini karena menawarkan keunggulan dibandingkan metode lain dalam tugas klasifikasi teks. Metode ini dikenal sederhana namun sangat efektif, mampu menangani data teks dengan baik melalui asumsi independensi antar fitur. Selain itu, *naive bayes* memiliki kinerja yang cepat, sehingga dapat mengolah dan mengklasifikasikan data dalam jumlah besar dengan efisien. Hal ini menjadikannya sangat cocok untuk penelitian yang berfokus pada pengolahan data teks dalam skala besar, seperti identifikasi komentar yang mengandung *cyberbullying*. Dibandingkan dengan metode lain seperti *support vector machine (SVM)* atau *random forest*, *naive bayes* lebih unggul dalam hal efisiensi komputasi dan kecepatan pemrosesan, terutama untuk *dataset* yang besar (Mughtar, 2024).



Sumber: Hasil Penelitian (2024)
 Gambar 9. Proses Modelling

3.1.6 Evaluasi

Dalam tahap evaluasi, dilakukan pengujian model *performance* dari model yang dihasilkan oleh algoritma *naïve bayes* menggunakan *confusion matrix*. Pengujian ini bertujuan untuk mengetahui seberapa baik model *naïve bayes* dalam melakukan klasifikasi komentar X. Untuk dapat melihat seberapa baik *naïve bayes* dalam melakukan klasifikasi, ini bisa dilihat dari penerapan *confusion matrix* yang menampilkan *accuracy*, *precision*, *recall*, dan *AUC (Area Under Curve)* sebagai berikut:

Pada gambar 10, 11 dan 12 terdapat total 650 data komentar twitter yang terbagi dalam dua kategori, yaitu 66 *true bullying* dan 39 *true non-bullying*. Komentar yang diprediksi sebagai *bullying* dan model memprediksinya sebagai *true bullying* berjumlah 66, sedangkan komentar yang diprediksi sebagai *bullying* tetapi model memprediksinya sebagai *true non-bullying* berjumlah 16.

Komentar yang diprediksi sebagai *non-bullying* sedangkan model memprediksinya sebagai *true bullying* berjumlah 9, lalu komentar yang diprediksi sebagai *non-bullying* dan model memprediksinya sebagai *true non-bullying* berjumlah 39.

Hasil *accuracy* mendapatkan hasil sebesar 80,77% dari total 650 data, hasil *precision* sebesar 81,25%, hasil *recall* sebesar 70,91% dan hasil nilai *AUC* sebesar 0,794 yang dimana *precision*, *recall*, dan *AUC* menunjukkan hasil klasifikasi *non-bullying*

accuracy: 80.77%

	true Bullying	true Non-bullying	class precision
pred. Bullying	66	16	80.49%
pred. Non-bullying	9	39	81.25%
class recall	88.00%	70.91%	

Sumber: Hasil Penelitian (2024)

Gambar 10. Hasil Accuracy

precision: 81.25% (positive class: Non-bullying)

	true Bullying	true Non-bullying	class precision
pred. Bullying	66	16	80.49%
pred. Non-bullying	9	39	81.25%
class recall	88.00%	70.91%	

Sumber: Hasil Penelitian (2024)

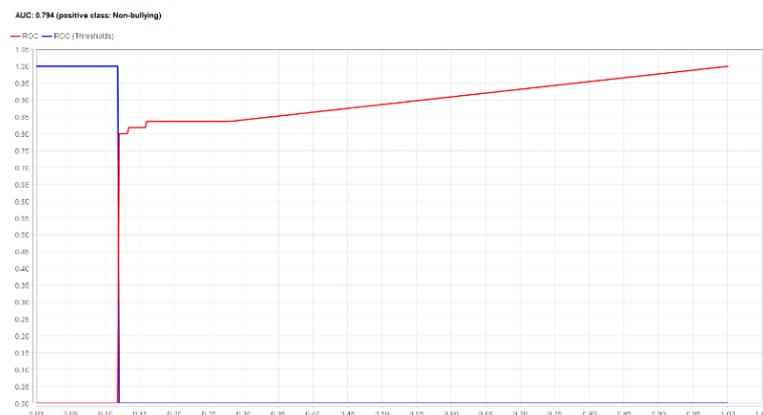
Gambar 11. Hasil Precision

recall: 70.91% (positive class: Non-bullying)

	true Bullying	true Non-bullying	class precision
pred. Bullying	66	16	80.49%
pred. Non-bullying	9	39	81.25%
class recall	88.00%	70.91%	

Sumber: Hasil Penelitian (2024)

Gambar 12. Hasil Recall



KESIMPULAN

Berdasarkan hasil penelitian, penerapan *naïve bayes* pada analisis sentimen *cyberbullying* dalam komentar X diperoleh nilai *accuracy* sebesar 80,77%, nilai *precision* sebesar 81,25%, nilai *recall* sebesar 70,91%, dan nilai *AUC* sebesar 0,794 ini menunjukkan bahwa metode *naïve bayes* efektif dalam mengklasifikasi komentar X sebagai *bullying* dan *non-bullying*. Dibandingkan dengan penelitian sebelumnya, seperti Millennialita (2024), yang menggunakan *naïve bayes* dan *support vector machine* untuk analisis *bullying* di Twitter dengan akurasi 97% atau Iqbal (2023) yang membandingkan *naïve bayes* dan *decision tree* dengan akurasi 85,56% pada *dataset twitter*

namun tanpa teknik khusus untuk menangani slang atau kata informal, penelitian ini memperkenalkan pembaruan berupa kombinasi teknik *stemming dictionary* dan *stemming snowball*. Teknik ini memungkinkan model untuk menangani komentar dengan kata-kata slang dan informal secara lebih baik.

Berdasarkan kesimpulan diatas, disarankan untuk menggunakan *dataset* yang lebih besar dan bervariasi, termasuk data dari berbagai platform media sosial seperti *tikTok*, *threads*, atau *youtube* untuk meningkatkan kemampuan model dalam mengenali pola yang lebih kompleks dan beragam. Selain itu, eksplorasi metode *preprocessing* lain, seperti *bert* atau *word2vec* dapat dilakukan untuk memberikan representasi teks yang lebih kaya dan meningkatkan akurasi model. Penelitian juga dapat diperluas dengan membandingkan algoritma lain seperti *decision tree*, *random forest*, *k-nearest neighbour*, dan *artificial neural network*. untuk mengevaluasi apakah hasil yang lebih baik dapat dicapai dalam klasifikasi *cyberbullying*.

REFERENSI

- Aldi, A. F., & Psikologi, J. (2023). *Kekerasan Dalam Rumah Tangga: Studi Etnografi Virtual Dengan Pendekatan Sentimen Analisis Twitter Terhadap Fenomena Kekerasan Dalam Rumah Tangga Selama Masa Pandemi Tahun 2020-2022 SKRIPSI Oleh.*
- Azrul, A., Purnamasari, A. I., & Ali, I. (2024). ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PERKEMBANGAN ARTIFICIAL INTELLIGENCE DENGAN PENERAPAN ALGORITMA LONG SHORT-TERM MEMORY (LSTM). In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 1).
- Cantika Samha, A., Ramadhona, Y., Zahara Putri, M., Effendi, F., Wahyu Puspaningrum, D., Fahlevi Arribath, A., Islam Negeri Raden Fatah Palembang, U., & Islam Negeri Raden Intan Lampung, U. (2023). UPAYA MENINGKATKAN KONSEP DIRI DALAM MENGATASI CYBERBULLYING PADA REMAJA DI ERA SOCIETY 5.0. In *Educate: Journal Of Education and Learning* (Vol. 1, Issue 1). <http://jurnal.dokicti.org/index.php/educate/index>
- Fitra Rizki, M., Auliasari, K., & Primaswara Prasetya, R. (2021). ANALISIS SENTIMENT CYBERBULLYING PADA SOSIAL MEDIA TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 5, Issue 2).
- Iqbal, M., Davy Wiranata, A., Suwito, R., & Faiz Ananda, R. (2023a). KLIK: Kajian Ilmiah Informatika dan Komputer Perbandingan Algoritma Naïve Bayes, KNN, dan Decision Tree terhadap Ulasan Aplikasi Threads dan Twitter. *Media Online*, 4(3), 1799–1807. <https://doi.org/10.30865/klik.v4i3.1402>
- Iqbal, M., Davy Wiranata, A., Suwito, R., & Faiz Ananda, R. (2023b). KLIK: Kajian Ilmiah Informatika dan Komputer Perbandingan Algoritma Naïve Bayes, KNN, dan Decision Tree terhadap Ulasan Aplikasi Threads dan Twitter. *Media Online*, 4(3), 1799–1807. <https://doi.org/10.30865/klik.v4i3.1402>
- Jubaidi, M., & Fadilla, N. (2020). PENGARUH FENOMENA CYBERBULLYING SEBAGAI CYBER-CRIME DI INSTAGRAM DAN DAMPAK NEGATIFNYA. *Shaut Al-Maktabah: Jurnal Perpustakaan, Arsip Dan Dokumentasi*, 12(2), 117–134. <https://doi.org/10.37108/shaut.v12i2.327>
- Millennianita, F., Athiyah, U., & Muhammad, A. W. (2024). Journal of Mechatronics and Artificial Intelligence Comparison of Naïve Bayes Classifier and Support Vector Machine Methods for Sentiment Classification of Responses to Bullying Cases on Twitter. In *Journal of Mechatronics and Artificial Intelligence* (Vol. 11, Issue 1). <http://ejournal.upi.edu/index.php/jmai/>
- Muchtar, M., Pasrun, Y. P., Rasyid, R., Miftachurohmah, N., & Mardawati, M. (2024). PENERAPAN METODE NAÏVE BAYES DALAM KLASIFIKASI KESEGARAN IKAN BERDASARKAN WARNA PADA CITRA AREA MATA. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(1). <https://doi.org/10.23960/jitet.v12i1.3879>
- Nurlaely, R., Sartika Simatupang, D., & Lucia Kharisma, I. (2023). *Analisis Sentimen Twitter Terhadap Cyberbullying Menggunakan Metode Support Vector Machine (SVM)*. 4(2), 376–384. <https://doi.org/10.37859/coscitech.v4i2.5161>
- Penulis, K., Pakpahan, H., & Luh Ayu Mondrisa Dwipayana, N. (2020). | 63 | CYBERBULLYING DI MEDIA SOSIAL. In *Bhirawa Law Journal* (Vol. 1, Issue 2).
- Rifai, H. S., Febrianti, S., & Santoso, I. (n.d.). ANALISIS SENTIMEN TANGGAPAN MASYARAKAT TERHADAP CYBERBULLYING DI MEDIA SOSIAL MENGGUNAKAN ALGORITMA NAÏVE BAYES (NB). <https://journals.upi-yai.ac.id/index.php/ikraith-informatika/issue/archive>
- Siroj, M., Zulfa, A., & Info, A. (n.d.). *Dampak Cyberbullying Pada Remaja di Media Sosial The Impact of Cyberbullying on Teenagers on Social Media*. <https://jicnusantara.com/index.php/jicn>
- Yulietta, F. T., Syafira, H. N. A., Alkautsar, M. H., Maharani, S., & Audrey, V. (2021). Pengaruh Cyberbullying di Media Sosial Terhadap Kesehatan Mental. *De Cive: Jurnal Penelitian Pendidikan Pancasila Dan Kewarganegaraan*, 1(8). <https://doi.org/10.56393/decive.v1i8.298>