

ANALISIS SENTIMEN PADA TWITTER TENTANG ISU PERILAKU ANTISOSIAL DENGAN ALGORITMA NAÏVE BAYES

Retika Nur Fadila; Andri Triyono, S.T., M.T.; Dhika Malita P.A, M.Kom.

ABSTRACT

In 2023, around 78.19% of the 275.77% or 215.63 million Indonesian population will be connected to the internet, with positive impacts such as fast communication, entertainment and new knowledge. The internet makes non-cash transactions easier and has negative impacts such as addiction and antisocial behavior such as indifference to people around you. Teenagers often access social media, especially Twitter, to express opinions and vent both positive and negative. Sentiment analysis is used to determine opinions about antisocial behavior on Twitter by using text mining techniques to analyze teenagers' opinions. Naive Bayes and SVM algorithms are used in sentiment analysis on the Twitter dataset to analyze antisocial behavior. Actions to evaluate the Naive Bayes algorithm in assessing antisocial behavior sentiments had the best accuracy results of 59.71% with $k=7$ without n -grams. The Naive Bayes algorithm with $k=5$ and n -gram $n=2$ has the best precision of 33.76% and the best recall of 33.45%. Future research can try to use other classification algorithms such as KNN, SVM, etc. To find the best accuracy of the antisocial behavior dataset.

Keywords: Internet; Twitter; antisocial behaviour; Sentiment analysis;

Correspondence:

Retika Nur Fadila

Universitas An Nuur Purwodadi, Email; retikanurfadila98@gmail.com

PENDAHULUAN

Teknologi dan internet berkembang di berbagai bidang dan terus menunjukkan kemajuan dari waktu ke waktu (Syah & Hermawati, 2018). Berdasarkan data dari dataindonesia.id, pada tahun 2023 sekitar 78,19% dari 275,77% atau 215,63 juta penduduk Indonesia telah terhubung dengan internet. Perkembangan teknologi informasi berdampak positif bagi masyarakat Indonesia dengan adanya internet memungkinkan komunikasi yang cepat dan mudah antar individu serta menawarkan hiburan dan pengetahuan baru. Selain itu, internet juga mempermudah transaksi non tunai. Namun, penggunaan yang tidak bijak dapat menyebabkan dampak negatif seperti kecanduan dan perilaku antisosial, terutama jika terlalu sering menggunakan media sosial. Contoh perilaku antisosial ialah tidak memedulikan orang di sekitar

kita. Kebiasaan remaja yang sering mengakses media sosial adalah salah satu penyebab perilaku antisosial. Media sosial *Twitter* sering digunakan untuk beropini dan curhat, baik positif maupun negatif (Sadewo et al., 2021). Untuk mengetahui jumlah individu yang beropini tentang perilaku antisosial di *Twitter*, diperlukan analisis sentimen. Analisis sentimen adalah pemahaman dan pemrosesan data teks untuk mendapatkan informasi tentang suatu topik (Rahman, 2020). Teknik yang digunakan untuk analisis sentimen adalah *text mining* untuk mengambil opini dan menganalisis opini remaja. Ada beberapa algoritma atau metode untuk analisis sentimen, seperti *Naive Bayes* dan SVM. Penelitian ini menggunakan algoritma *Naive Bayes* pada dataset publik dari *Twitter* untuk menganalisis sentimen terkait isu perilaku antisosial.

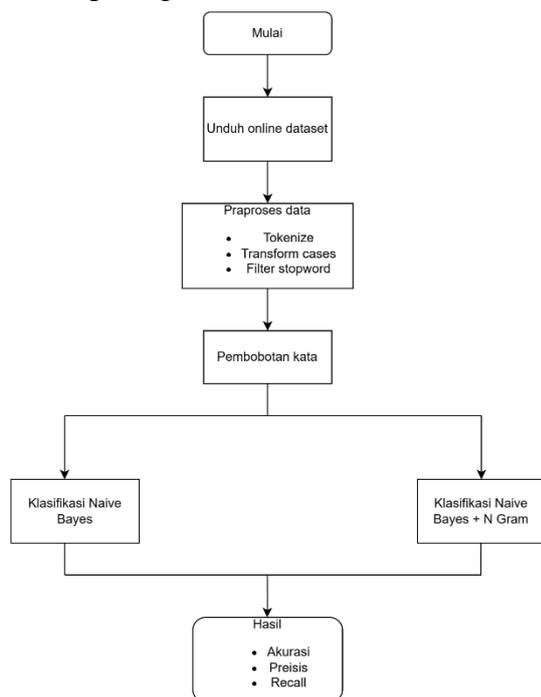
yang gegabah untuk keamanan, 4) Kurangnya penyesalan, 5) Non-antisosial atau Umum.

METODE PENELITIAN

Model penelitian yang digunakan dalam penelitian ini adalah model eksperimen dan evaluasi.

Metode Yang Diusulkan

Metode yang diusulkan merupakan gambaran dari tahapan pekerjaan dari awal penelitian itu dimulai hingga berakhir. Proses penelitian dimulai dari awal pengambilan data kemudian tahapan pra proses data, pembobotan kata, dan klasifikasi. Berikut langkah-langkah dapat dilihat pada gambar 1.



Gambar 1 metode yang diusulkan

Pengumpulan Dataset

Dataset didapatkan dari hasil unduh dataset publik dari [Antisocial Behaviour Public Twitter Indonesia | Kaggle](#) yang diunggah oleh Fitri Andri Astuti berjumlah 1.251 data. Dataset tersebut memiliki 5 nama kelas yaitu: 1) Kegagalan untuk menyesuaikan diri dengan norma-norma sosial tentang perilaku yang sah, 2) Iritabilitas dan agresivitas, 3) Pengabaian

Praproses Data

Pada tahap pra proses data yang dilakukan dalam penelitian ini ada 3 langkah, yaitu:

1. Tokenize

tokenize adalah tahap pemisahan kata, simbol, frasa, dan entitas penting lainnya (disebut token) dari teks (Ridwansyah, 2022).

2. Transform cases

Transform Cases proses mengubah seluruh karakter pada data sesuai keinginan, seperti mengubah huruf besar menjadi huruf kecil dan sebaliknya (Rahman, 2020).

3. Filter stopwords

Pada tahap ini, semua kata diperiksa, jika dalam dokumen terdapat kata-kata yang tidak memberikan kontribusi banyak seperti kata sambung, kata depan, atau kata ganti, maka kata-kata tersebut akan dihilangkan (Maulana et al., 2020).

4. N-gram

N-gram menjelaskan bahwa bahasa tidak terdiri dari kata-kata tunggal, melainkan rangkaian kata-kata tunggal dan frase dua, tiga atau lebih kata, masing-masing berisi informasi unik (Fahrur Rozi et al., 2020).

Pembobotan Kata

Pembobotan TF-IDF adalah pembobotan kata atau term berdasarkan frekuensi kemunculannya (TF) dalam suatu dokumen. Persamaan frekuensi kata $tf(t,d)$ ditunjukkan pada Persamaan berikut:

$$tf(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \dots\dots\dots(1)$$

Keterangan:

$tf(t,d)$ = jumlah kemunculan term pada dokumen d.

$f_{t,d}$ = pencacahan mentah
 $\frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$ = frekuensi istilah
 Untuk menghitung IDF dapat menggunakan persamaan berikut:
 $idf = \log\left(\frac{N}{df}\right)$ (2)

Keterangan:
 idf = ukuran informasi yang diberikan oleh istilah t
 N = Jumlah total dokumen yang digunakan.
 df = Jumlah dokumen di mana term yang dipilih muncul.

Untuk menghitung TF-IDF dapat menggunakan persamaan berikut:
 $TF-IDF = tf(t,d) \times idf$ (3)

Keterangan:
 $tf(t,d)$: jumlah kemunculan termpada dokumen d .
 idf : ukuran informasi yang diberikan oleh istilah t

Klasifikasi Naïve Bayes

Klasifikasi adalah teknik menguji objek dan mengidentifikasinya dalam satu kelas tertentu dari kumpulan kelas lain. Klasifikasi *Naïve Bayes* adalah Memprediksi peluang masa depan berdasarkan pengalaman masa lalu dikenal dengan teorema Bayes (Nugroho, 2018). Keuntungan menggunakan algoritma *Naïve Bayes* yaitu algoritma ini hanya membutuhkan data pelatihan kecil (Imron, n.d.).

Hasil (Evaluasi dan Validasi)

Validasi yang digunakan pada penelitian ini adalah *k-fold validation*, *k-fold validation* adalah metode memilah data menjadi *train* data dan *test* data (Ridwansyah, 2022). Misal k yang digunakan adalah $k=7$ maka dataset dipecah menjadi 7 bagian, 6 bagian digunakan untuk data latih, dan 1 bagian digunakan untuk data uji. Proses ini dilakukan sebanyak 7 kali (Wahono et al., 2014). Untuk mendapatkan hasil akurasi, presisi, dan *recall* diperlukan evaluasi atau pengukuran performa atau biasa disebut

dengan *confussion matrix*, *confussion matrix* adalah tabel yang menampilkan hitungan jawaban benar dan salah pada data tes (Normawati & Prayogi, 2021). *Confussion matrix* dapat disajikan pada tabel 1.

Tabel 1 confusion matrix

Aktual Kelas	Terklasifikasi Positif (1)	Terklasifikasi Negatif (0)
Positif (1)	TP	FP
Negatif (0)	FN	TN

Keterangan:

- TP adalah *True Positive*, yaitu jumlah dokumen dari kelas 1 yang benar terklasifikasikan sebagai kelas 1
- TN adalah *True Negative*, yaitu jumlah dokumen dari kelas 0 yang benar terklasifikasikan sebagai kelas 0
- FN adalah *False Negative*, yaitu jumlah dokumen dari kelas 0 yang salah terklasifikasikan sebagai kelas 1
- FP adalah *False Positive*, yaitu jumlah dokumen dari kelas 1 yang salah terklasifikasikan sebagai kelas 0

HASIL DAN PEMBAHASAN

Setelah mendapatkan kephahaman pada landasan teori maka selanjutnya tahap yang berurutan yaitu dari proses dan hasil praproses data, pembobotan kata, implementasi dan hasil klasifikasi *Naïve Bayes*, serta evaluasi performa.

Praproses data

Data penelitian sebagai contoh praproses data diambil data cuitan yang bersifat mentah dan terdapat karakter, simbol, singkatan, dan bahasa

asing yang perlu diolah terlebih dahulu seperti tabel 2.

Tabel 2 hasil praproses data

Data asli	Tokenize	Transform Cases	Filter stopwords
Saya selalu benar dan kamu selalu salah.. apabila saya salah.. kembali ke pasal 1. Saya selalu benar 😊	'Saya', 'selalu', 'benar', 'dan', 'kamu', 'selalu', 'salah', 'apabila', 'saya', 'salah', 'kembali', , 'ke', 'pasal', 'saya', 'selalu', 'benar'	saya selalu benar dan kamu selalu salah.. apabila saya salah kembali ke pasal 1. saya selalu benar	selalu benar kamu selalu salah apabila salah ke pasal selalu benar
Yah gimana ya, generasi boomer : orang tua selalu benar	'Yah', 'gimana', 'ya', 'generasi boomer', 'orang', 'tua', 'selalu', 'benar'	yah gimana ya, generasi boomer : orang tua selalu benar	yah gimana generasi boomer orang tua selalu benar

Selain *tokenize*, *transform cases*, dan *filter stopwords* ada satu langkah agar bisa melanjutkan pada pengklasifikasian *Naïve Bayes* dengan penambahan *n-gram*. Berikut ditampilkan pada tabel 3 hasil dari *n-gram*, $n=2$

Tabel 3 *n-gram*

Data Asli	<i>n-gram</i> ($n=2$)
Saya selalu benar dan kamu selalu salah.. apabila saya salah.. kembali ke pasal 1. Saya selalu benar 😊	"saya selalu", "selalu benar", "benar dan", "dan kamu", "kamu selalu", "selalu salah". "apabila saya", "saya salah".

	"kembali ke", "ke pasal", "pasal 1". "saya selalu", "selalu benar".
Yah gimana ya, generasi boomer : orang tua selalu benar	"yah gimana", "gimana ya", "generasi boomer", "orang tua", "tua selalu", "selalu benar"

Pembobotan Kata

Sebagai studi kasus diambil lima dokumen yang akan digunakan sebagai contoh pembobotan kata, *tools* yang digunakan untuk pembobotan kata adalah *Microsoft excel*. Dokumen yang akan digunakan untuk pembobotan kata dapat dilihat pada tabel 4.

a. Dokumen yang akan digunakan
Tabel 4 dokumen yang digunakan

D1	aku selalu marah bukan karena mereka salah tapi karna memang mereka benar
D2	emang gua selalu benar
D3	medusa aku selalu benar
D4	feeling tidak selalu benar
D5	mungkin memang ada hukum hukum yang gak akan bisa kita langgar prinsip kehidupan yang gak berubah

Tabel 4 merupakan dokumen yang akan digunakan sebagai simulasi penghitungan TF-IDF dengan berjumlah 5 dokumen menggunakan *Microsoft excel*.

b. Menghitung *Term-Frequency* (TF)

Tabel 5 menghitung TF

Keterangan:

Term	D1	D2	D3	D4	D5	DF
aku	1	0	1	0	0	2
selalu	1	1	1	1	0	4
marah	1	0	0	0	0	1
bukan	1	0	0	0	0	1
karena	1	0	0	0	0	1
mereka	2	0	0	0	0	1
salah	1	0	0	0	0	1
tapi	1	0	0	0	0	1
karna	1	0	0	0	0	1
memang	1	0	0	0	1	2
benar	1	1	1	1	0	4
emang	0	1	0	0	0	1
gua	0	1	0	0	0	1
medusa	0	0	1	0	0	1
feeling	0	0	0	1	0	1
tidak	0	0	0	1	0	1
mungkin	0	0	0	0	1	1
ada	0	0	0	0	1	1
hukum	0	0	0	0	2	1
yang	0	0	0	0	2	1
gak	0	0	0	0	1	1
akan	0	0	0	0	1	1
bisa	0	0	0	0	1	1
kita	0	0	0	0	1	1
langgar	0	0	0	0	1	1
prinsip	0	0	0	0	1	1
kehidupan	0	0	0	0	1	1
berubah	0	0	0	0	1	1
Jumlah Dokumen(Dn)	12	4	4	4	15	

TF = *Term Frequency* (banyaknya kata pada sebuah dokumen)

Dn = Jumlah dokumen ke-n

DF = *Document Frequency* (banyaknya dokumen yang mengandung kata tersebut)

c. Menghitung TF normalisasi

Pada penghitungan TF normalisasi Dalam bentuk frekuensi mentah, tf hanyalah frekuensi

istilah tersebut dalam setiap dokumen. Kedua istilah ini muncul satu kali dalam setiap dokumen. Namun, dokumen 2 berisi lebih banyak kata, sehingga frekuensi relatifnya

TF Normalisasi ($\frac{TF}{Dn}$)					
Term	d1	d2	d3	d4	d5
aku	0,083	0	0,25	0	0
selalu	0,083	0,25	0,25	0,25	0
marah	0,083	0	0	0	0
bukan	0,083	0	0	0	0
karena	0,083	0	0	0	0
mereka	0,167	0	0	0	0
salah	0,083	0	0	0	0
tapi	0,083	0	0	0	0
karna	0,083	0	0	0	0
memang	0,083	0	0	0	0,067
benar	0,083	0,25	0,25	0,25	0
emang	0	0,25	0	0	0
gua	0	0,25	0	0	0
medusa	0	0	0,25	0	0
feeling	0	0	0	0,25	0
tidak	0	0	0	0,25	0
mungkin	0	0	0	0	0,067
ada	0	0	0	0	0,067
hukum	0	0	0	0	0,133
yang	0	0	0	0	0,133
gak	0	0	0	0	0,067
akan	0	0	0	0	0,067
bisa	0	0	0	0	0,067
kita	0	0	0	0	0,067
langgar	0	0	0	0	0,067
prinsip	0	0	0	0	0,067
kehidupan	0	0	0	0	0,067
berubah	0	0	0	0	0,067

lebih rendah.

Tabel 6 menghitung TF normalisasi Contoh penghitungan TF normalisasi misal *term* aku pada dokumen 1 terdapat 1 *term* maka untuk menghasilkan TF normalisasi dibutuhkan rumus $tf("aku", d_1) = \frac{1}{12} = 0,083$.

d. Menghitung IDF
Tabel 7 Menghitung IDF

Term	DF	IDF $\left(\log \frac{N}{DF}\right)$
aku	2	0,398
selalu	4	0,097
marah	1	0,699
bukan	1	0,699
karena	1	0,699
mereka	1	0,699
salah	1	0,699
tapi	1	0,699
karna	1	0,699
memang	2	0,398
benar	4	0,097
emang	1	0,699
gua	1	0,699
medusa	1	0,699
feeling	1	0,699
tidak	1	0,699
mungkin	1	0,699
ada	1	0,699
hukum	1	0,699
yang	1	0,699
gak	1	0,699
akan	1	0,699
bisa	1	0,699
kita	1	0,699
langgar	1	0,699
prinsip	1	0,699
kehidupan	1	0,699
berubah	1	0,699

Keterangan :

TF = *Term Frequency* (banyaknya kata pada sebuah dokumen)

DF = *Document Frequency* (banyaknya dokumen yang mengandung kata tersebut)

N = total Dokumen

IDF = *Invers Document Frequency*

e. Menghitung TF-IDF
Tabel 8 menghitung TF-IDF

TF-IDF (TF normalisasi*IDF)				
d1	d2	d3	d4	d5
0,033	0	0,1	0	0
0,008	0,024	0,024	0,024	0
0,058	0	0	0	0
0,058	0	0	0	0
0,058	0	0	0	0
0,117	0	0	0	0
0,058	0	0	0	0
0,058	0	0	0	0
0,058	0	0	0	0
0,033	0	0	0	0,027
0,008	0,024	0,024	0,024	0
0	0,175	0	0	0
0	0,175	0	0	0
0	0	0,175	0	0
0	0	0	0,175	0
0	0	0	0,175	0
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,093
0	0	0	0	0,093
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,047
0	0	0	0	0,047
0,547	0,398	0,323	0,398	0,683

Hasil pengujian algoritma Naïve Bayes

Pengujian algoritma *Naïve Bayes* akan dilakukan sebanyak dua kali yaitu pertama pengujian algoritma *Naïve Bayes* tanpa *n-gram*, berikut hasil pengujian algoritma *Naïve Bayes* tanpa *n-gram*:

Tabel 9 hasil pengujian algoritma *Naïve Bayes* tanpa *n-gram*

k-fold validation	akurasi	presisi	recall

5	59,23%	33,45%	32,23%
6	58,67%	33,46%	31,30%
7	59,71%	33,98%	32,18%
8	59,00%	33,72%	31,65%
9	58,11%	32,87%	30,65%
10	58,20%	32,67%	30,49%

Tabel 9 merupakan hasil dari pengujian *Naïve Bayes* tanpa fitur *n-gram* dengan menggunakan *k-fold validation* $k=5$, $k=6$, $k=7$, $k=8$, $k=9$, dan $k=10$. Akurasi tertinggi diperoleh dari *fold* ke-7 yaitu 59,71%, presisi tertinggi diperoleh dari *fold* ke-8 yaitu 33,72%, *recall* tertinggi diperoleh dari *fold* ke-1 yaitu 32,23%.

Pengujian algoritma yang kedua yaitu pengujian algoritma *Naïve bayes* dengan *n-gram*, berikut hasil pengujian algoritma *Naïve Bayes* dengan *n-gram*:

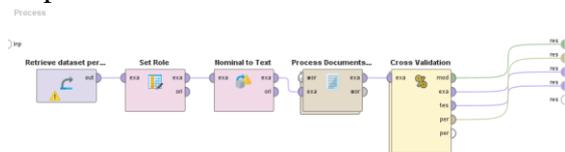
Tabel 10 hasil pengujian algoritma *Naïve Bayes* dengan *n-gram*

k-fold validation	akurasi	presisi	recall
5	58,52%	33,76%	33,45%
6	57,63%	31,33%	31,12%
7	58,04%	33,20%	32,90%
8	57,79%	32,50%	32,78%
9	57,47%	33,24%	32,70%
10	58,11%	33,55%	32,63%

Tabel 10 merupakan hasil dari pengujian *Naïve Bayes* dengan fitur *n-gram* dengan menggunakan *k-fold validation* $k=5$, $k=6$, $k=7$, $k=8$, $k=9$, dan $k=10$. Akurasi tertinggi diperoleh dari *fold* ke-5 yaitu 58,52%, presisi tertinggi diperoleh dari *fold* ke-5 yaitu 33,76%, *recall* tertinggi diperoleh dari *fold* ke-5 yaitu 33,45%.

Design pengujian algoritma *Naïve Bayes*

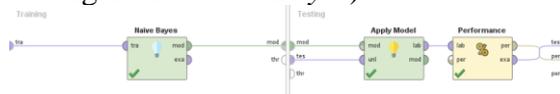
Pada pengujian algoritma *Naïve Bayes* menggunakan tools RapidMiner, berikut gambarana design pengujian algoritma *Naïve Bayes* menggunakan RapidMiner:



Gambar 2 Pengujian algoritma *Naïve Bayes*

Gambar 2 merupakan *design* dari pengujian algoritma *Naïve Bayes* yang memiliki beberapa langkah atau tahapan yaitu:

1. *Import* dataset yang akan digunakan untuk penelitian
2. Operator *set role* untuk menentukan label pada dataset
3. Operator *nominal to text* berfungsi untuk merubah atribut yang berbentuk nominal atau angka ke teks
4. Operator *Process document from data*(tahap praproses data) pada tahap praproses data ada dua pengujian yaitu pertama pengujian tanpa menggunakan fitur *n-gram* kedua pengujian menambahkan fitur *n-gram*.
5. *Cross validation* (evaluasi dan validasi algoritma *Naïve Bayes*).



Gambar 3 cross validation

Pada gambar 3 *cross validation* dibagi menjadi dua bagian yaitu bagian pertama *training* yang berisi operator algoritma *Naïve Bayes* dan bagian kedua *testing* yang berisi operator *apply model* dan *performance*.

SIMPULAN DAN SARAN

Kesimpulan

Berdasarkan pengujian algoritma *Naïve Bayes* yang ditambahkan fitur *n-gram* dan tidak menambahkan fitur *n-gram* terhadap perilaku antisosial pada twitter maka dapat ditarik kesimpulan, yaitu:

1. Algoritma *Naïve Bayes* dapat digunakan untuk mengklasifikasikan sentimen publik yang ada pada *twitter* terutama *tweet* mengenai perilaku antisosial.
2. Hasil akurasi algoritma *Naïve Bayes* terbaik didapatkan dengan k terbaik $k=7$ tanpa *n-gram*(*unigram*) yaitu 59,71%.
3. Hasil *presisi* algoritma *Naïve Bayes* terbaik didapatkan pada $k=5$ dengan *n-gram* $n=2$ (*bigram*) yaitu 33,76%.

4. Hasil *recall* algoritma *Naïve Bayes* terbaik didapatkan pada $k=5$ dengan *n-gram* $n=2$ (*bigram*) yaitu 33,45%.

Saran

Diakhir penelitian ini, penulis memberikan beberapa saran, berikut saran yang disampaikan oleh penulis:

1. Peneliti berpendapat dengan hasil akurasi 59,71% penelitian ini hanya dapat digunakan sebagai referensi bukan acuan untuk dinas terkait atau masyarakat umum apabila ingin meninjau perilaku antisosial diinternet khususnya media sosial *twitter*.
2. Pada penelitian selanjutnya dengan dataset perilaku antisosial dapat menggunakan algoritma klasifikasi yang lain misal *K-Nearest Neighbors (KNN)*, *Support Vector Machine (SVM)*, dll yang bertujuan untuk mencari nilai akurasi terbaik.

DAFTAR PUSTAKA

- Fahrur Rozi, I., Taufika Firdausi, A., & Islamiyah, K. (2020). Analisis Sentimen Pada Twitter Mengenai Pasca Bencana Menggunakan Metode *Naïve Bayes* Dengan Fitur N-Gram. *Jurnal Informatika Polinema*, 6(2), 33–39.
<https://doi.org/10.33795/jip.v6i2.316>
- Imron, A. (n.d.). *KABUPATEN REMBANG MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER*.
- Maulana, F. A., Ernawati, I., Labu, P., & Selatan, J. (2020). Analisa sentimen cyberbullying di jejaring sosial twitter dengan algoritma *naïve bayes*. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 529–538.
<https://conference.upnvj.ac.id/index.php/senamika/article/view/619>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi *Naïve Bayes Classifier* Dan *Confusion Matrix* Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
<http://ejournal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- Nugroho, A. (2018). Analisis Sentimen Pada Media Sosial Twitter Menggunakan *Naive Bayes Classifier* Dengan Ekstrasi Fitur N-Gram. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 2(2), 200.
<https://doi.org/10.30645/j-sakti.v2i2.83>
- Rahman, A. (2020). *Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier*. 8(2), 50–58.
- Ridwansyah, T. (2022). Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan *Naïve Bayes Classifier*. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 2(5), 178–185.
<https://doi.org/10.30865/klik.v2i5.362>
- Sadewo, T. A., Kusuma, P. D., Setianingsih, C., & Telkom, U. (2021). *CLUSTERING PADA DATA SENTIMEN BPJS KESEHATAN MENGGUNAKAN ALGORITMA AGGLOMERATIVE HIERARCHICAL CLUSTERING AVERAGE LINKAGE*. 8(5), 6078–6088.
- Syah, R., & Hermawati, I. (2018). The Prevention Efforts on Cyberbullying Case for Indonesian Adolescent Social Media Users. *Jurnal Penelitian Kesejahteraan Sosial*, 17(2), 131–146.
- Wahono, R. S., Herman, N. S., & Ahmad, S. (2014). A comparison framework of classification models for software defect prediction. *Advanced Science Letters*, 20(10–12), 1945–1950.
<https://doi.org/10.1166/asl.2014.5640>