

Penerapan Analisis Sentimen Terhadap Kekerasan atau Tindak *Bullying* pada Media Sosial Twitter

Bayu Nur Rahman Eswantoro, Handaru Jati

Universitas Negeri Yogyakarta

Email: bayunur.2020@student.uny.ac.id

ABSTRAK

Penelitian ini bertujuan untuk: (1) untuk menguji serta menganalisis bagaimana anonimitas dapat memengaruhi perilaku *cyberbullying* di media sosial Twitter; (2) untuk mengetahui bagaimana pemanfaatan teknik analisis sentimen di media sosial Twitter menggunakan klasifikasi *Support Vector Machine* (SVM) dapat digunakan secara efektif dalam mengidentifikasi *cyberbullying*. Penelitian ini menggunakan metode penelitian kuantitatif. Pengambilan data menggunakan alat Tweet Harvest. Subjek penelitian yaitu seluruh cuitan pengguna di Twitter berbahasa Indonesia yang berkaitan dengan *cyberbullying* atau tindakan kekerasan verbal lainnya. Proses analisis sentimen yang dilakukan memuat *crawling data, preprocessing, labeling, feature extraction and modeling, evaluating, dan visualization*. Teknik analisis yang akan digunakan dalam penelitian ini adalah analisis kinerja model menggunakan *Confusion Matrix* untuk menentukan seberapa baik model dalam menentukan, memprediksi dan mengklasifikasikan sentimen. Hasil Penelitian menunjukkan bahwa: (1) Anonimitas terbukti dapat memengaruhi perilaku *cyberbullying* di media sosial Twitter. Hal ini terlihat pada tahap identifikasi anonimitas, yang menunjukkan bahwa 811 dari 1.365 pengguna akun Twitter, teridentifikasi sebagai akun anonim. Selain itu, berdasarkan hasil analisis sentimen terhadap 1.282 data cuitan, diketahui bahwa 68,8% di antaranya mengandung unsur negatif. Hasil penelitian ini juga sejalan dengan teori *disinhibition effect*, di mana pengguna dapat dengan bebas mengutarakan pikirannya di media sosial tanpa ada yang mengawasi dan menganggap status seseorang di media sosial itu setara; (2) model klasifikasi *Support Vector Machine* terbukti efektif dalam mengidentifikasi *cyberbullying* di bandingkan *Random Forest* dan *Naïve Bayes* dengan nilai akurasi $81\% > 80\% > 78\%$.

Kata kunci: Analisis sentimen, *Cyberbullying*, Perbandingan Model, *Support Vector Machine*, Twitter.

ABSTRACT

This study aims to (1) test and analyze how anonymity affects cyberbullying behaviour on Twitter social media; (2) investigate how sentiment analysis techniques, utilizing Support Vector Machine (SVM) classification, can be effectively applied to identify cyberbullying on Twitter social media. This study uses quantitative research methods. Data collection uses the Tweet Harvest tool. The subjects of the study were all tweets from Indonesian-language Twitter users related to cyberbullying or other verbal violence. The sentiment analysis process carried out includes data crawling, preprocessing, labelling, feature extraction and modelling, evaluating, and visualization. The analysis technique used in this study is model performance analysis, utilizing a Confusion Matrix to determine how well the model performs in determining, predicting, and classifying sentiment. The results of the study show that (1) Anonymity is proven to affect cyberbullying behaviour on Twitter, a social media platform. This can be seen in the anonymity identification stage, which shows that 811 out of 1,365 Twitter account users are identified as anonymous accounts. In addition, based on the results of sentiment analysis of 1,282 tweets, it is known that 68.8% of them contain harmful elements. The results of this study are also in line with the theory of the disinhibition effect, where users can freely express their thoughts on social media without anyone supervising and considering a person's status on social media to be equal; (2) the Support Vector Machine classification model is proven to be effective in identifying cyberbullying compared to Random Forest and Naïve Bayes with an accuracy value of $81\% > 80\% > 78\%$.

Keywords: *Cyberbullying*, Model comparison, Sentiment analysis, *Support Vector Machine*, Twitter

PENDAHULUAN

Perkembangan teknologi informasi telah membawa perubahan besar dalam cara masyarakat berinteraksi dan berkomunikasi. Salah satu dampak yang paling menonjol adalah meningkatnya penggunaan media sosial sebagai sumber utama informasi bagi masyarakat serta sebagai sarana untuk mengekspresikan opini, berdiskusi, dan mengkritisi berbagai isu sosial. Twitter, sebagai salah satu *platform* media sosial yang paling populer, telah menjadi medium utama dalam menyampaikan informasi dan pandangan secara cepat, terbuka, dan masif (Humam & Laksito, 2023). Menurut Taufik dan Pamungkas dalam Salsabila (2022), dengan adanya fitur seperti tagar (*hashtag*), pengguna dapat mengikuti dan berpartisipasi dalam percakapan yang sedang hangat, sehingga membentuk berbagai macam dinamika opini publik. Namun, kebebasan berekspresi di media sosial tidak jarang disalahgunakan. Fenomena *cyberbullying* atau perundungan di dunia maya menjadi salah satu konsekuensi negatif dari interaksi digital yang minim regulasi dan cenderung anonim. Pengguna media sosial sering kali secara tidak sadar terlibat dalam tindakan *cyberbullying*, baik bagi pelaku, korban, maupun pengamat yang pasif. Hal ini dipengaruhi beberapa faktor yang di antaranya anonimitas di media sosial, kurangnya pemahaman tentang *cyberbullying*, kurangnya empati, pengaruh kelompok, serta kurangnya literasi digital menjadi pemicu utama maraknya kekerasan verbal dalam *platform* seperti Twitter (Rizkyanti et al.,

2021). Menurut (Jariyah, 2022), *cyberbullying* merupakan salah satu bagian dari perundungan (*bullying*) melalui internet oleh seseorang atau sekelompok orang yang dilakukan secara sengaja untuk menyakiti atau melukai keadaan emosi dan psikologi seseorang dengan melontarkan kata-kata kasar atau menghina, mempermalukan hingga mengancam korban yang menyebabkan korban mengalami stres, depresi, dan merasa terisolasi. Menurut (Munawaroh & Ninik Setiyowati, 2024), *cyberbullying* adalah salah satu perilaku penyimpangan yang dilakukan remaja menggunakan internet yang memberikan dampak negatif bagi korban dan pelaku.

Menurut riset *polling* Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) terhadap 5.900 sampel, 49% netizen Indonesia pernah mengalami *bullying* di media sosial, 47% tidak pernah mengalami *bullying*, dan sisanya tidak menanggapi, dengan 31,6% dari 49% tersebut membiarkan perlakuan *bullying* dan tidak melakukan apa-apa (Pratomo & Yusuf, 2019). Berikut merupakan beberapa kasus *bullying* dan tindak kekerasan yang pernah dibahas dan ramai diperbincangkan di media sosial Twitter. Di antaranya, kasus yang menimpa Natalius Pigai, atau sering dikenal dengan sebutan “Putra Papua”, yang mengalami serangan kebencian berunsur SARA oleh Ambroncius Nababan selaku ketua relawan Pro Jokowi-Amin, karena melakukan kritik terhadap pemerintah melalui media sosial Facebook (Humas, 2021). Selain itu, Komisi Perlindungan Anak Indonesia

(KPAI) menyatakan bahwa Luluk Nuril telah melakukan *cyberbullying* di media sosial TikTok kepada murid SMK yang sedang melakukan Praktik Kerja Lapangan (PKL), sehingga murid tersebut kehilangan kepercayaan diri dan berniat untuk berhenti melakukan PKL (Ikhsanudin, 2023). Selanjutnya, seorang anak laki-laki berusia 11 tahun di Tasikmalaya mengalami tindak perundungan siber yang dilakukan oleh temannya ketika bermain, korban dipaksa untuk menyetubuhi kucing lalu direkam menggunakan telepon seluler oleh para pelaku, sehingga korban diduga mengalami depresi dan meninggal dunia (CNN Indonesia, 2022). Adapun pengalaman buruk dialami oleh seorang remaja putri berinisial RDS (19), yang mendapat ujaran kebencian saat melakukan siaran langsung di media sosial, hingga berujung pada penggeroyokan oleh pelaku berinisial HA dan 20 orang temannya (Adliyah, 2024). Serta Wakil Menteri Kesehatan Dante Saksono Harbuwono, menerima laporan sebanyak 1.000 kasus perundungan, dengan 30% di antaranya telah teridentifikasi sebagai kasus perundungan. Di antara kasus tersebut, terdapat dugaan kasus bunuh diri akibat perundungan yang dialami oleh almarhumah dr. Aulia Risma Lestari saat menjalani Program Pendidikan Dokter Spesialis (PPDS) Anestesi di RS Kariadi (Novianto, 2024). Berdasarkan data statistik yang telah dilakukan oleh Jaringan Pemantau Pendidikan Indonesia (JPPI) dalam Zuhriyah (2024), pada tahun 2024 tercatat 573 kasus kekerasan yang

dilakukan di lingkungan pendidikan, selain itu kasus perundungan dan tindak kekerasan ini mengalami peningkatan yang cukup signifikan jika dibandingkan pada tahun sebelumnya yang mencatat 91 kasus pada tahun 2020, 142 kasus pada tahun 2021, 194 kasus pada tahun 2022, dan 285 kasus pada tahun 2023. Meningkatnya kasus perundungan di kalangan masyarakat Indonesia menjadi cerminan bahwa pemahaman dan kesadaran masyarakat modern saat ini masih tergolong rendah terkait dampak buruk dari tindak perundungan. Hal ini juga didukung oleh data yang diperoleh dari Microsoft (2021) dalam Rizkyanti et al. (2021), yang menyatakan bahwa skor *Digital Civility Indeks* (DCI) di Indonesia merupakan yang tertinggi di Asia Tenggara, yaitu sebesar 76% pada tahun 2020 sehingga Indonesia mendapat sebutan sebagai "Negara dengan Warga Netizen Paling Tidak Beradab di Asia Tenggara".

Berbagai kasus perundungan di media sosial menunjukkan bahwa *cyberbullying* telah menjadi masalah serius lintas usia dan profesi. Mulai dari ujaran kebencian berbau SARA, perundungan terhadap siswa magang oleh tokoh publik, hingga kasus dugaan bunuh diri akibat tekanan sosial daring, semuanya memperlihatkan urgensi untuk memahami dan menanggulangi perilaku ini secara sistematis berbasis data. Meskipun dengan kemajuan teknologi dan informasi yang seharusnya membawa peningkatan kesadaran sosial, kenyataannya banyak individu yang belum sepenuhnya memahami

konsekuensi serius terhadap tindak perundungan. Menurut Tazkiyah et al. (2021), korban perundungan sering kali mengalami trauma fisik maupun psikologis yang berkepanjangan, yang tidak hanya mengganggu kesejahteraan mereka, tetapi juga mempengaruhi kualitas hidup dan perkembangan mental mereka. Kondisi ini menjadi peringatan bagi seluruh bangsa Indonesia untuk lebih memperhatikan, memahami, dan bertindak tegas dalam menanggulangi perundungan demi menciptakan masyarakat yang lebih beradab dan harmonis.

Salah satu pendekatan yang dapat digunakan untuk mengidentifikasi pola perilaku cyberbullying adalah analisis sentimen terhadap unggahan di media sosial. Penggunaan Twitter untuk melakukan analisis sentimen sudah banyak digunakan oleh peneliti lain seperti Khaira et al. (2020), Andrianus (2023), Madyatmadja et al. (2024), Syahputra & Rosnita (2024), dan peneliti lainnya, dengan struktur datanya yang terbuka dan masif, menjadi sumber data yang potensial untuk memahami persepsi publik terhadap berbagai isu, termasuk perundungan siber. Dengan bantuan teknologi seperti *machine learning*, analisis sentimen dapat digunakan untuk mengklasifikasikan opini publik ke dalam kategori positif, negatif, atau netral. Selain itu, analisis sentimen Twitter juga dapat membantu masyarakat dan lembaga terkait dalam memantau dinamika sosial dan psikologis serta mengambil tindakan yang tepat untuk mencegah dan menangani kasus perundungan

(Tazkiyah et al., 2021). Menurut Zein et al. (2022) analisis sentimen merupakan sebuah metode yang digunakan untuk mengetahui sifat dari opini atau pemikiran publik terhadap suatu topik maupun isu tertentu yang sedang beredar sehingga menghasilkan dua kategori yaitu pendapat positif dan pendapat negatif. Untuk melakukan analisis sentimen dibutuhkan metode klasifikasi *supervised learning*, hal ini disebabkan metode klasifikasi *supervised learning* menggunakan data berlabel dalam pekerjaannya, salah satu metode klasifikasi yang sering digunakan untuk analisis sentimen adalah *Support Vector Machine* (SVM), yang mampu memaksimalkan jarak antar kelas (*margin*) tanpa adanya *overfitting* dengan menggunakan *hyperplane* untuk mendapatkan hasil akurasi yang tinggi (Ning Tan et al., 2019). Menurut Hilmi Zain et al. (2023), penggunaan analisis sentimen menggunakan Twitter menjadi sumber informasi alternatif yang bermanfaat bagi para peneliti dan akademisi untuk memahami pandangan dan opini masyarakat terhadap *cyberbullying* di Indonesia. Dalam beberapa tahun terakhir, penelitian-penelitian tersebut telah dilakukan oleh banyak peneliti dan akademisi untuk menggali informasi yang lebih mendalam dan akurat tentang persepsi masyarakat terhadap banyaknya kasus *cyberbullying* yang terjadi di media sosial. Dengan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan teknik analisis sentimen menggunakan SVM terhadap data Twitter yang berkaitan dengan tindak

perundungan atau *cyberbullying*. Penelitian ini juga membandingkan efektivitas SVM dengan metode klasifikasi lainnya untuk menentukan model terbaik dalam mengidentifikasi sentimen negatif sebagai indikator perilaku *bullying*.

METODE

Jenis Penelitian

Penelitian ini menggunakan metode kuantitatif yang berfokus kepada jumlah data yang diambil melalui Twitter menggunakan Tweet Harvest, selain itu penelitian ini bertujuan untuk menguji efektivitas penerapan teknik analisis sentimen menggunakan SVM dalam mengklasifikasikan sentimen terhadap pendapat masyarakat di media sosial Twitter. Untuk menguji efektivitas dari SVM akan dilakukan perbandingan menggunakan metode *Naives Bayes* (NB) serta *Random Forest* (RF).

Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan pada bulan November 2024 hingga April 2025 yang berlokasi di Jakarta, mulai dari pengumpulan data hingga analisis data.

Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh tweet berbahasa Indonesia yang ada di Twitter yang memiliki kata kunci kekerasan, *bullying*, dan *cyberbullying* dalam kurun waktu satu tahun terakhir (1365 tweet). Sedangkan sampel penelitian ini akan menggunakan teknik *purposive sampling* jika peneliti tidak memiliki waktu dan kemampuan yang cukup untuk menyelesaikan seluruh

rangkaian proses, di mana peneliti dapat menggunakan sebagian dari keseluruhan data yang mengandung indikasi *bullying* atau *cyberbullying*.

Teknik Pengumpulan Data

Penelitian ini menggunakan metode pengumpulan data dengan cara *crawling* menggunakan Tweet Harverst yang memanfaatkan *credentials token* pada Twitter.

Instrumen Penelitian

Instrumen yang digunakan untuk penelitian ini meliputi:

- 1) Tweet Harvest, mengumpulkan data mentah dari Twitter yang sesuai dengan topik penelitian (Satria, 2024).
- 2) *Natural Language Toolkit* (NLTK), memproses data mentah menjadi data siap pakai untuk di analisis seperti penyederhanaan teks ke bentuk dasar, memecah kalimat menjadi kata-kata (Bird et al., 2009).
- 3) *Valance Aware Dictionary and sEntiment Reasoner* (VADER), memberikan label sentimen pada tweet (Hutto, 2014).
- 4) *Term Frequency – Inverse Document Frequency* (TF-IDF), mengubah atau merepresentasikan teks menjadi sebuah nilai numerik berdasarkan frekuensi (Haj, 2019).
- 5) SVM, NB, dan RF, membuat sebuah prediksi sentimen berdasarkan dataset yang telah di ekstraksi (Ning Tan et al., 2019; Nugraha et al., 2022; Breiman, 2001).
- 6) *Confusion Matrix*, mengukur kinerja dari sebuah model klasifikasi dalam

- memprediksi sentimen berdasarkan nilai *accuracy*, *precision*, *recall*, *F1-score* (Salsabila, 2022).
- 7) *WordCloud*, memvisualisasikan hasil klasifikasi dan persebaran sentimen sehingga membantu peneliti untuk menganalisis data lebih lanjut Madyatmadja et al. (2024).

Teknik Analisis Data

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis kinerja model untuk menentukan seberapa baik model dalam menentukan, mengklasifikasikan, dan memprediksi sentimen menggunakan Confusion Matrix.

HASIL DAN PEMBAHASAN

Data Collection

Seperti yang terlihat pada **gambar 1**, pengumpulan data menggunakan Tweet Harvest akan menggunakan *credentials token* pada Twitter, pengumpulan data (\pm 1400 tweet) tersebut dilakukan pada bulan november 2024 dengan rentang waktu satu tahun terakhir yang kemudian disimpan ke dalam sebuah file dengan format excel (.csv) untuk memudahkan analisis data.

A	B	C	D	E	F	G	H	I	J	K	M	N	O
1	created_at		full_text		image_url		in_reply_to_screen_name						
2	VWed Jun 12 08:54	film	pembakaran	sekolah	Kok isi nya keban	in							
3	VWed Jun 05 09:45	in	utuh	bukan	bullying	ta	tiga	lindir	kekerasan	in			
4	Sun May 26 09:43	in	ada	ada	bully	nya	oleh	kekerasan	atau	ta	in		
5	Tue May 29 05:45	extremisme	Bers	intoleransi	nasionalisme	dan	toleransi	ber	toleransi	nasionalisme	dan		
6	Fri May 17 12:12	12	Kira-kira	kita	halan	ada	kerasan	ya	asereyebin	in			
7	Thu May 16 08:01	@bizarakanator	Tempat	kerja	seba	kerasakan	in						
8	Thu May 16 06:47	Bullying	dan	kekerasan	di	satuan	kominfo	kor	kominfo	kor			
9	Thu Apr 25 14:22	Pertaku	bullying	akan	menghancur	para	perkara	diri	in				
10	Sun Apr 21 10:55	@stumpung	nitar	giv	lapson	bk	jncsl	in					

Gambar 1. Hasil pengumpulan data menggunakan Tweet Harvest

Identifikasi Anonimitas

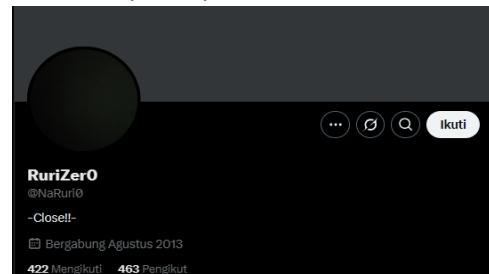
Pada tabel 1, terlihat hasil identifikasi akun pengguna anonim dan

bukan anonim pada dataset yang digunakan yang memperlihatkan sebanyak 811 dari 1365 akun teridentifikasi sebagai akun anonim.

Tabel 1. Identifikasi akun

Label	Jumlah Data
Anonim	811
Bukan Anonim	554
Jumlah	1365

Pada **gambar 2**, diperlihatkan akun pengguna Twitter berstatus anonim, yang menyembunyikan foto profil, alamat, dan tidak menggunakan nama asli. Dapat dilihat pada **gambar 3**, bahwa pengguna akun anonim lebih berani dan bebas dalam berpendapat



Gambar 2. Akun Anonim

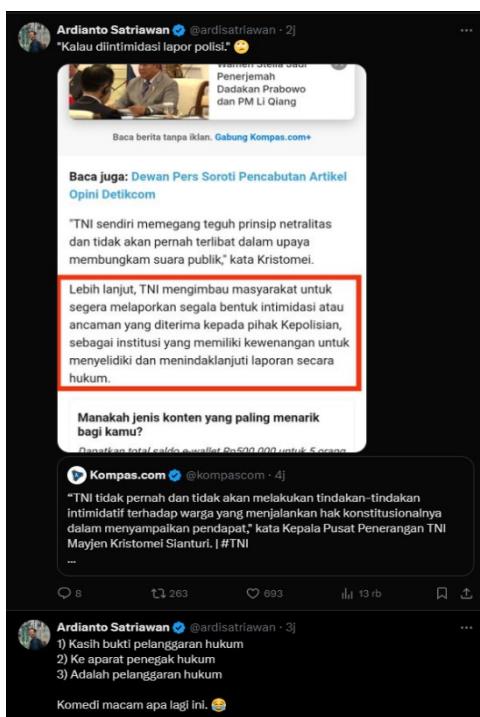


Gambar 3. Tweet akun anonim

Pada gambar 4, terlihat bahwa pengguna yang bukan anonim akan memperlihatkan foto profil, penggunaan nama yang sebenarnya, alamat tinggal, dan sudah diverifikasi oleh Twitter. Serta pengguna yang bukan anonim lebih sopan dalam menyampaikan pendapatnya seperti yang terlihat pada gambar 5.



Gambar 4. Akun bukan anonim



Gambar 5. Tweet akun bukan anonim

Preprocessing

Tahapan pertama dalam proses analisis sentimen adalah pembersihan data yang meliputi *case folding*, *slang dict*, kata tidak baku menjadi baku menurut KBBI, penghapusan data duplikat, dan lain-lain sehingga mendapat data bersih sebesar 1282 data. Pada tabel 2, terlihat data mentah yang akan diolah menjadi data bersih tanpa adanya noise.

Tabel 2. Data Cleaning

Full text	Cleaned text
film pendek bertemakan sekolah kok isi nya kebanyakan bullying dan aksi balas dendam Sangat wajar jika apa yang terjadi bakalan dicontoh oleh yang melihatnya Jangan kaget jika banyak aksi tindak kekerasan yang mungkin bisa berakibat hilangnya nyawa seseorang Faham ...!?	film pendek bertemakan sekolah mengapa isi nya kebanyakan bully dan aksi balas dendam sangat wajar jika apa yang terjadi bakalan dicontoh oleh yang melihatnya jangan terkejut jika banyak aksi tindak kekerasan yang mungkin bisa berakibat hilangnya nyawa seseorang paham
ini udah bukan bullying lg tapi tindak kekerasan	ini sudah bukan bully lagi tetapi tindak kekerasan

Tahapan kedua, yaitu memecah kalimat menjadi kata-kata yang dapat dimengerti oleh model atau *machine learning (tokenizing)* yang dapat dilihat pada tabel 3.

Tabel 3. Tokenizing

Cleaned text	Tokenizing
film pendek bertemakan sekolah mengapa isi nya kebanyakan	[film, pendek, bertemakan, sekolah, mengapa, isi, kebanyakan,

bully dan aksi balas dendam sangat wajar jika apa yang terjadi bakalan dicontoh oleh yang melihatnya jangan terkejut jika banyak aksi tindak kekerasan yang mungkin bisa berakibat hilangnya nyawa seseorang paham	kebanyakan, bully, dan, aksi, balas, dendam, sangat, wajar, jika, apa, yang, terjadi, bakalan, dicontoh, oleh, yang, melihatnya, jangan, terkejut, jika, banyak, aksi, tindak, kekerasan, yang, mungkin, bisa, berakibat, hilangnya, nyawa, seseorang, paham]	[ini, sudah, bukan, bully, lagi, tetapi, tindak, kekerasan]	[bully, tindak, kekerasan]
ini sudah bukan bully lagi tetapi tindak kekerasan	[ini, sudah, bukan, bully, lagi, tetapi, tindak, kekerasan]		

Tahapan ketiga, menghilangkan kata-kata yang tidak bermakna untuk membantu meningkatkan efisiensi proses selanjutnya (*Remove Stopword*) yang dapat dilihat pada **tabel 4**.

Tabel 4. Remove Stopword

Tokenizing	Remove Stopword
[film, pendek, bertemakan, sekolah, mengapa, isi, nya, kebanyakan, bully, dan, aksi, balas, dendam, sangat, wajar, jika, apa, yang, terjadi, bakalan, dicontoh, oleh, yang, melihatnya, jangan, terkejut, jika, banyak, aksi, tindak, kekerasan, yang, mungkin, bisa, berakibat, hilangnya, nyawa, seseorang, paham]	[film, pendek, bertemakan, sekolah, isi, nya, kebanyakan, bully, aksi, balas, dendam, wajar, dicontoh, terkejut, aksi, tindak, kekerasan, berakibat, hilangnya, nyawa, paham]

Tahapan keempat, *stemming* yang merupakan penyederhanaan kata-kata yang sudah di proses sebelumnya menjadi sebuah kata dasar atau tanpa imbuhan, seperti yang terlihat pada **tabel 5**.

Tabel 5. Stemming

Remove Stopword	Stemming
[film, pendek, bertemakan, sekolah, isi, nya, kebanyakan, bully, aksi, balas, dendam, wajar, dicontoh, terkejut, aksi, tindak, kekerasan, berakibat, hilangnya, nyawa, paham]	[film, pendek, bertemakan, sekolah, isi, nya, kebanyakan, bully, aksi, balas, dendam, wajar, dicontoh, terkejut, aksi, tindak, kekerasan, berakibat, hilangnya, nyawa, paham]
[bully, tindak, kekerasan]	[bully, tindak, kekerasan]

Labeling

Setelah seluruh rangkaian pada pemrosesan data selesai, langkah berikutnya adalah memberikan label kepada data secara otomatis menggunakan VADER seperti yang terlihat pada **tabel 6**. Selain itu, pada tahapan ini peneliti melakukan validasi menggunakan pelabelan secara manual untuk meningkatkan hasil dari model klasifikasi seperti yang ditunjukkan pada **tabel 7**.

Tabel 6. Labeling VADER

Teks	Sentimen
film pendek bertemakan sekolah mengapa	Negative

isi nya kebanyakan bully dan aksi balas dendam sangat wajar jika apa yang terjadi bakalan dicontoh oleh yang melihatnya jangan terkejut jika banyak aksi tindak kekerasan yang mungkin bisa berakibat hilangnya nyawa seseorang paham		yang masuk di kategorinya chat aku	
ini sudah bukan bully lagi tetapi tindak kekerasan	Negative		
satuan pembinaan masyarakat bersama unit pelayanan perempuan dan anak satuan reserse kriminal kepolisian resort kabupaten kotawaringin timur telah melaksanakan giat sebagai narasumber dalam rangka sosialisasi pencegahan dan penanganan tindak kekerasan bully di satuan pendidikan	Negative		
kira kira kalau kalian ada kerabat yang masuk dalam kategori boleh chat aku ya kategorinya penyandang disabilitas lebih diutamakan fisik sebelumnya pengguna narkoba korban bencana alam korban tindak kekerasan bisa bully apabila ada	Negative		
film pendek bertemakan sekolah mengapa isi nya kebanyakan bully dan aksi balas dendam sangat wajar jika apa yang terjadi bakalan dicontoh oleh yang melihatnya jangan terkejut jika banyak aksi tindak kekerasan yang mungkin bisa berakibat hilangnya nyawa seseorang paham		Negative	
ini sudah bukan bully lagi tetapi tindak kekerasan		Negative	
satuan pembinaan masyarakat bersama unit pelayanan perempuan dan anak satuan reserse kriminal kepolisian resort kabupaten kotawaringin timur telah melaksanakan giat sebagai narasumber dalam rangka sosialisasi pencegahan dan penanganan tindak kekerasan bully di satuan pendidikan		Positive	
kira kira kalau kalian ada kerabat yang masuk dalam kategori boleh chat aku ya kategorinya		Positive	

Tabel 7. Pelabelan Manual

penyandang disabilitas lebih diutamakan fisik sebelumnya pengguna narkoba korban bencana alam korban tindak kekerasan bisa bully apabila ada yang masuk di kategorinya chat aku	
---	--

Feature Extraction dan Modeling

Pada tahapan ini peneliti akan menggunakan TF-IDF sebagai *feature extraction* seperti yang sudah dijelaskan sebelumnya dan penggunaan Support Vector Machine, Naive Bayes, dan Random Forest sebagai model klasifikasi. Pada **tabel 8** akan terlihat pembagian data yang digunakan antara data latih dan data uji dengan rasio 70:30 dan pada **tabel 9** yang menunjukkan pembagian data dengan rasio 80:20, hal tersebut dilakukan untuk mengetahui apakah jumlah data memiliki pengaruh terhadap hasil model klasifikasi atau tidak.

Tabel 8. Split Data 70%

Label	Jumlah Data
Data Latih	897
Data Uji	385

Tabel 9. Split Data 80%

Label	Jumlah Data
Data Latih	1025
Data Uji	257

Penggunaan *Cross-validation* yang ditunjukkan pada **tabel 10** dan **tabel 11**, bertujuan untuk mengurangi atau menghindari *overfitting* data yang disebabkan oleh ketidakseimbangan data

antara label positif (400) dengan label negatif (882).

Tabel 10. Cross-validation 70%

Label	SVM	NB	RF
Rata-rata Akurasi	0.77	0.77	0.75
St. Deviasi Akurasi	0.02	0.03	0.02
Rata-rata F1-Weighted	0.75	0.73	0.72
St. Deviasi F1-Weighted	0.03	0.04	0.03
Akurasi Model	0.81	0.78	0.80

Tabel 11. Cross-validation 80%

Label	SVM	NB	RF
Rata-rata Akurasi	0.79	0.79	0.79
St. Deviasi Akurasi	0.06	0.04	0.04
Rata-rata F1-Weighted	0.78	0.77	0.77
St. Deviasi F1-Weighted	0.06	0.04	0.05
Akurasi Model	0.81	0.78	0.76

Evaluasi dan Perbandingan Support Vector Machine

Support Vector Machine merupakan *machine learning* yang mampu mengatasi permasalahan *linear* maupun *non linear* dan dapat memaksimalkan *margin* tanpa adanya *overfitting* dengan menggunakan *hyperplane*, hal tersebut terungkap pada tabel confusion matrix (**tabel 12**) yang memperlihatkan 68 *true positive*, 24 *false positive*, 49 *false negative*, 244 *true negative*. Hasil evaluasi matriks pada **tabel 13** menunjukkan rata-rata *precision* 79%, rata-rata *recall* 75%, rata-rata *f1-score* 76%, dan akurasi model 81%.

Tabel 12. Confusion Matrix SVM

	Positive	Negative
Act. Pos	68	49
Act. Neg	24	244

Tabel 13. Evaluasi SVM

	Presisi	Recall	F1-score	Supp
Neg	0.83	0.91	0.87	268
Pos	0.74	0.58	0.65	117
Acc			0.81	385
Macro	0.79	0.75	0.76	385
Weight	0.80	0.81	0.80	385

Naive Bayes

Naive Bayes (NB) memanfaatkan perhitungan probabilitas dan statistik untuk melakukan prediksi masa depan menggunakan data saat ini, dalam analisis sentimen ini naive bayes memiliki keunggulan dalam kecepatan dan keakuratannya dalam memproses data. Pada **tabel 14** memperlihatkan 47 *true positive*, 15 *false positive*, 70 *false negative*, 254 *true negative*. Sementara itu untuk hasil evaluasinya adalah rata-rata *precision* 77%, rata-rata *recall* 67%, rata-rata *f1-score* 69%, dan akurasi model 78% yang dapat dilihat pada **tabel 15**.

Tabel 14. Confusion Matrix NB

	Positive	Negative
Act. Pos	47	70
Act. Neg	15	253

Tabel 15. Evaluasi NB

	Presisi	Recall	F1-score	Supp
Neg	0.78	0.94	0.86	268
Pos	0.76	0.40	0.53	117
Acc			0.78	385
Macro	0.77	0.67	0.69	385
Weight	0.78	0.78	0.76	385

Random Forest

Random Forest (RF) merupakan sebuah klasifikasi yang terdiri dari sekumpulan pohon keputusan yang terdiri dari vektor acak yang terbagi identik dan independen, cara kerja dari model ini adalah untuk setiap keputusan diberikan satu suara untuk menentukan kelas yang paling dominan berdasarkan input. Pada **tabel 16** menunjukkan 63 *true positive*, 23 *false positive*, 54 *false negative*, 245 *true negative* dan pada **tabel 17** menunjukkan hasil evaluasinya dengan rata-rata *precision* 78%, rata-rata *recall* 73%, rata-rata *f1-score* 74%, dan akurasi model 80%.

Tabel 16. Confusion Matrix RF

	Positive	Negative
Act. Pos	63	54
Act. Neg	23	245

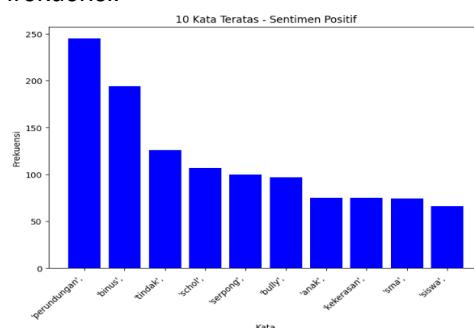
Tabel 17. Evaluasi RF

	Presisi	Recall	F1-score	Supp
Neg	0.82	0.91	0.86	268
Pos	0.73	0.54	0.62	117
Acc			0.80	385
Macro	0.78	0.73	0.74	385
Weight	0.79	0.80	0.79	385

Visualisasi

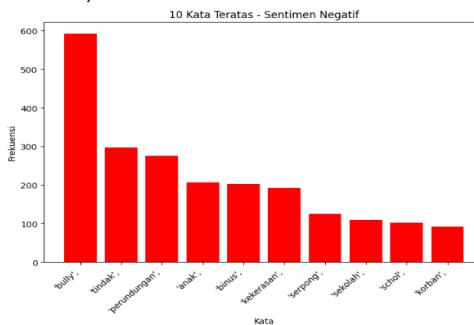
Setelah evaluasi, langkah visualisasi data cukup penting karena hal tersebut dapat membantu untuk memahami analisis lebih jauh lagi dan peneliti dapat mengidentifikasi sentimen positif dan negatif terhadap tindak *bullying* di media sosial Twitter. Langkah visualisasi sebenarnya dapat dilakukan di berbagai tahapan sesuai dengan

kebutuhan peneliti yang dapat dimulai dari setelah *preprocessing*, *labeling*, maupun setelah evaluasi. Penampilan diagram batang pada **gambar 6**, memvisualisasikan 10 kata teratas untuk label positif di antaranya “perundungan”, “binus”, “tindak”, “schol”, “serpong”, “bully”, “anak”, “kekerasan”, “sma”, dan “siswa”. Selain itu, kata “perundungan” merupakan kata yang paling banyak digunakan untuk menentukan sentimen positif yang hampir mencapai 250 frekuensi.



Gambar 6. 10 kata teratas label positif

Pada gambar 7, yang menampilkan diagram batang 10 kata teratas untuk label negatif yang diantaranya “bully” dengan frekuensi yang hampir 600, diikuti dengan kata “tindak”, “perundungan”, “anak”, “binus”, “kekerasan”, “serpong”, “sekolah”, “schol”, “korban”.



Gambar 7. 10 kata teratas label negatif

Visualisasi data menggunakan *WordCloud* dengan label positif pada **gambar 8**, memperlihatkan kata *bully*, perundungan, binus, dan kata lainnya yang sama pada *WordCloud* negatif. Hal tersebut dikarenakan sudah adanya kemajuan pada kasus perundungan di Binus school Serpong dan adanya kegiatan sosialisasi yang diselenggarakan oleh pihak kepolisian untuk mencegah tindak kekerasan dan *bullying*.



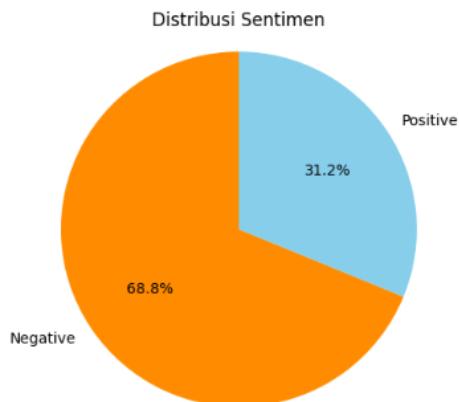
Gambar 8. WordCloud positif

Selain itu untuk **gambar 9**, terdapat juga kata-kata yang sama yang ditampilkan pada *WordCloud* positif seperti *bully*. tindak kekerasan, binus, serpong, dan lain-lain. Hal tersebut dikarenakan kasus *bullying* yang terjadi di Binus school Serpong yang dilakukan oleh geng tai sedang ramai diperbincangkan di media sosial Twitter.



Gambar 9. *WordCloud* negatif

Pada **gambar 10**, terlihat sebuah diagram lingkaran yang memperlihatkan distribusi sentimen dari *dataset* yang digunakan dalam penelitian. Dari hasil tersebut sebanyak 68,8% data bersifat negatif dan 32,2% data bersifat positif



Gambar 10. Distribusi sentimen

KESIMPULAN

Anonimitas terbukti memengaruhi perilaku *cyberbullying* di media sosial Twitter. Hal ini terlihat pada tahap identifikasi anonimitas yang menunjukkan bahwa 1.365 pengguna akun Twitter, sebanyak 811 akun teridentifikasi sebagai akun anonim. Selain itu, berdasarkan hasil analisis sentimen yang telah melalui seluruh tahapan, ditemukan 1.282 data cuitan dan diketahui bahwa 68,8% di antaranya mengandung unsur negatif. Hasil penelitian ini juga sejalan dengan teori *disinhibition effect*, yang mencakup beberapa aspek, di antaranya *Dissociative anonymity*, yaitu pemisahan identitas diri antara dunia maya dan dunia nyata; *Invisibility*, yaitu kebebasan seseorang dalam melakukan suatu hal di media sosial tanpa adanya yang mengawasi; *Solipsistic introjection*, yaitu kurangnya visual dan verbal pengguna terhadap lawan bicara sehingga pengguna seakan-akan berbicara terhadap diri sendiri pada media sosial dibandingkan berbicara secara langsung dengan seseorang dari dunia nyata; serta *Minimization of status*, yaitu status seseorang di dunia nyata menjadi tidak penting ketika berada di dunia maya, yang menyebabkan seseorang menjadi lebih bebas dan berani dalam mengekspresikan diri. Serta

Pemanfaatan teknik analisis sentimen menggunakan model klasifikasi *Support Vector Machine* terbukti efektif dalam mengidentifikasi *cyberbullying* dibandingkan *Random Forest* dan *Naïve Bayes*. Hasil dari klasifikasi *Support Vector Machine* unggul tipis akan *Random Forest* baik dari berbagai macam aspek nilai, dan cukup unggul jika dibandingkan dengan *Naïve Bayes* dengan hasil akhir nilai akurasi $81\% > 80\% > 78\%$ (*Support Vector Machine* > *Random Forest* > *Naïve Bayes*).

DAFTAR RUJUKAN

- Adliyah, S. (2024). *Remaja Putri di Palembang Dikeroyok, Pelaku Sering Hate Comment.* <https://www.detik.com/sumbagsel/hukum-dan-kriminal/d-7540667/remaja-putri-di-palembang-dikeroyok-pelaku-sering-hate-comment>
- Andrianus, Y. (2023). *IMPLEMENTASI ALGORITMA K-MEANS TERHADAP OPINI MASYARAKAT MENGENAI PERKIRAAN PEMILU 2024 PADA TWITTER.pdf.* 8, 305–308. <https://doi.org/https://doi.org/10.51876/simtek.v8i2.271>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python* (J. Steele (ed.); 1 ed.). O'Reilly Media, Inc. <https://www.nltk.org/#natural-language-toolkit>
- Breiman, L. (2001). RANDOM FOREST. *Statistics Department, University of California, Berkeley, CA 94720, 45(October 2001), 5–32.* <https://doi.org/https://doi.org/10.1023/A:1010933404324>
- CNN Indonesia. (2022). *Kasus Bullying Anak Setubuhi Kucing di Tasikmalaya Naik ke Penyidikan.*

- Jakarta.
<https://www.cnnindonesia.com/nasional/20220725173918-12-825930/kasus-bullying-anak-setubuhi-kucing-di-tasikmalaya-naik-ke-penyidikan>
- Haj, A. S. A. (2019). Analisis sentimen masyarakat terhadap kinerja kpu pada pemilu 2019 menggunakan algoritma k-means dengan confix stripping stemmer. *Repository.Uinjkt.Ac.Id.*
<https://repository.uinjkt.ac.id/dspace/handle/123456789/55857>
- Hilmi Zain, H., Maulana Awannga, R., & Isti Rahayu, W. (2023). Perbandingan Model Svm, Knn Dan Naïve Bayes Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024. *JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, 8(3), 2083–2093. <https://jim.usk.ac.id/sejarah>
- Humam, C., & Laksito, A. D. (2023). Implementasi Aplikasi Sentimen Pada Data Twitter Jelang Pemilu 2024. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(2), 105–112. <https://doi.org/10.30591/jpit.v8i2.5051>
- Humas. (2021). *Putra Papua Jadi Korban Cyber Bullying, Rasisme di Indonesia Makin Liar!* Universitas Al Azhar Indonesia. <https://uai.ac.id/en/putra-papua-jadi-korban-cyber-bullying-rasisme-di-indonesia-makin-liar/>
- Hutto, C. J. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text.* <https://github.com/cjhutto/vaderSentiment>
- Ikhsanudin, A. (2023). *KPAI: Luluk Nuril Lakukan Cyberbullying, Korban Hilang Percaya Diri.* [detiknews.](https://news.detik.com/berita/d-6918297/kpai-luluk-nuril-lakukan-cyberbullying-korban-hilang-percaya-diri) <https://news.detik.com/berita/d-6918297/kpai-luluk-nuril-lakukan-cyberbullying-korban-hilang-percaya-diri>
- Jariyah, S. A. (2022). *Fenomena Cyberbullying Dan Penanganannya (Studi Kasus Pada Dua Siswa Di Smp Negeri 1 Sungguminasa Kabupaten Gowa).* <http://eprints.unm.ac.id/25128/%0Ahttp://eprints.unm.ac.id/25128/1/SYANIAINUNJARIYAH.pdf>
- Khaira, U., Johanda, R., Utomo, P. E. P., & Suratno, T. (2020). Sentiment Analysis Of Cyberbullying On Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*, 3(1), 21. <https://doi.org/10.24014/ijaidm.v3i1.9145>
- Madyatmadja, E. D., Felix, Edrick, I. G. K., Indarto, J. W., & Sembiring, D. J. M. (2024). Harmonizing sentiments: Analyzing user reviews of Spotify through sentiment analysis. *Journal of Infrastructure, Policy and Development*, 8(9), 23. <https://doi.org/10.24294/jipd.v8i9.7101>
- Munawaroh, & Ninik Setiyowati. (2024). Systematic Literature Review Using Big Data Tentang Emotional Intelligence dan Perilaku Cyberbullying Remaja. *G-Couns: Jurnal Bimbingan dan Konseling*, 9(1), 506–517. <https://doi.org/10.31316/gcouns.v9i1.15991>
- Ning Tan, P., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *INTRODUCTION TO DATA MINING* (Second Edi).
- Novianto, R. D. (2024). *Wamenkes Ungkap 300 Kasus Perundungan di Program Pendidikan Dokter*

- Spesialis. sindonews. <https://nasional.sindonews.com/read/1448697/15/wamenkes-ungkap-300-kasus-perundungan-di-program-pendidikan-dokter-spesialis-1725365386>
- Nugraha, T., Purwantoro, & Umaidah, Y. (2022). *Analisis Sentimen terhadap Perpanjangan Masa Jabatan Presiden Indonesia Menggunakan Algoritma Naïve Bayes.* 4, 1349–1358.
- Pratomo, Y., & Yusuf, O. (2019). 49 Persen Netizen di Indonesia Pernah Mengalami “Bullying” di Medsos. Kompas.com. <https://tekno.kompas.com/read/2019/05/16/08290047/49-persen-netizen-di-indonesia-pernah-mengalami-bullying-di-medsos>
- Rizkyanti, C. A., Hesti Cahyani, A., Salsabilla, S., & Aulia, A. (2021). EMPATI DAN PERAN BYSTANDER DALAM CYBERBULLYING: FAMILY COMMUNICATION PATTERN SEBAGAI MEDIATOR. *Jurnal Psikohumanika*, 13(2), 10–24. <http://ejurnal.setiabudi.ac.id/ojs/index.php/psikohumanika>
- Salsabila, N. A. (2022). ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP TOKOH GUS DUR MENGGUNAKAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM).
- Satria, H. (2024). *Tweet Harvest.* <https://github.com/helmisatria/tweet-harvest>
- Syahputra, M. O., & Rosnita, L. (2024). *Analysis of Public Sentiment Toward Celebrity Endorsement On Media Social Using Support Vector Machine.* 4(3), 118–127.
- Tazkiyah, I., Fadillah, A. R., Kusuma, F. W., Siswantoro, M. F., & Cahyono, S. A. (2021). Peran Anonimitas terhadap Cyberbullying pada Media Sosial The Role of Anonymity to Cyberbullying on Social Media. *Seminar Nasional Teknologi dan Sistem Informasi*, 3(2), 77–83. <https://sitasi.upnjatim.ac.id/index.php/sitasi/article/download/74/11/>
- Zein, A., Farizy, S., & Suharyanto, E. (2022). SENTIMEN ANALISIS PADA KOMENTAR PENDEK EVALUASI DOSEN OLEH MAHASISWA (EDOM) PROGRAM STUDI SISTEM INFORMASI. V(01).
- Zuhriyah, U. (2024). *Data Kasus Bullying Terbaru 2024, Apakah Meningkat?* <https://tirto.id/data-kasus-bullying-terbaru-2024-apakah-meningkat-g621>