

**PENERAPAN ALGORITMA KLASIFIKASI *NAIVE BAYES*
DAN *SUPPORT VECTOR MACHINE* UNTUK ANALISIS
SENTIMEN *CYBERBULLYING BILINGUAL* DI APLIKASI X**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer pada
Program Studi Sistem Informasi



Oleh:

NOVITA SARI
12050322233



UIN SUSKA RIAU
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2025

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERSETUJUAN

**PENERAPAN ALGORITMA KLASIFIKASI NAIVE BAYES
DAN SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN CYBERBULLYING BILINGUAL DI APLIKASI X**

TUGAS AKHIR

Oleh:

NOVITA SARI
12050322233

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 14 Januari 2025

Ketua Program Studi

E. Saputra, S.Kom., M.Kom.
NIP. 198307162011011008

Pembimbing

M. Jazman, S.Kom., M.InfoSys.
NIP. 198206042015031004

© Hak cipta milik UIN Suska Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Site Islamia University of Sultan Syarif Kasim R

UIN SUSKA RIAU



LEMBAR PENGESAHAN

**PENERAPAN ALGORITMA KLASIFIKASI NAIVE BAYES
DAN SUPPORT VECTOR MACHINE UNTUK ANALISIS
SENTIMEN CYBERBULLYING BILINGUAL DI APLIKASI X**

TUGAS AKHIR

Oleh:

NOVITA SARI

12050322233

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 06 Januari 2025

Pekanbaru, 06 Januari 2025
Mengesahkan,

Ketua Program Studi

Eki Saputra, S.Kom., M.Kom.
NIP. 198307162011011008

Dekan
Dr. Hartono, M.Pd.
NIP. 196403011992031003

DEWAN PENGUJI:

Ketua : Arif Marsal, Lc., MA.

Sekretaris : M. Jazman, S.Kom., M.InfoSys.

Anggota 1 : T. Khairil Ahsyar, S.Kom., M.Kom.

Anggota 2 : Syaifullah, SE., M.Sc.

Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : NOVITA SARI

NIM : 12050322233

Tempat/Tgl. Lahir : MANDAH / 23 FEBRUARI 2000

Fakultas/Pascasarjana : SAINS DAN TEKNOLOGI

Prodi : SISTEM INFORMASI

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:

PENERAPAN ALGORITMA KLASIFIKASI NAIVE BAYES DAN SUPPORT VECTOR
MACHINE UNTUK ANALISIS SENTIMEN CYBERBULLYING BILINGUAL DI
APLIKASI X

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya pyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesua peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 17 JANUARI 2025

membuat pernyataan



NOVITA SARI

NIM : 12050322233

*pilih salah satu sesuai jenis karya tulis

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin penulis dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



UIN SUSKA RIAU



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 06 Januari 2025
Yang membuat pernyataan,



NOVITA SARI
NIM. 12050322233

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang

Assalamu 'alaikum Warahmatullahi Wabarakatuh.

Alhamdulillah Rabbil 'Alamin, segala puji bagi Allah Subhanahu Wa Ta 'ala sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad Shallallahu 'Alaihi Wa Sallam dengan mengucapkan Allahumma Sholli'ala Sayyidina Muhammad Wa'ala Ali Sayyidina Muhammad. Semoga kita semua selalu senantiasa mendapat syafa'at-Nya di dunia maupun di akhirat, Aamiin Ya Rabbal'alamiin.

Dengan penuh rasa hormat dan cinta, peneliti mempersembahkan Tugas Akhir yang telah diselesaikan ini kepada Ibunda Suryani tercinta dan Bapak Abdul Murat yang selalu memberikan inspirasi, kekuatan, motivasi, serta dukungan semangat. Berkat do'a dan kasih sayangmu, anakmu telah berhasil memperoleh gelar sarjana seperti yang engkau harapkan. Tiada apapun di dunia ini yang dapat membalas semua jasa dan pengorbananmu. Untuk itu, saya anakmu ini selalu mendo'akan yang terbaik dan bahagia dunia akhirat serta diberikan tempat istimewa di sisi-Nya kelak sehingga kita bisa berkumpul kembali bersama-sama di Jannah-Nya.

Peneliti ingin mengucapkan terima kasih yang mendalam juga kepada Dosen Pembimbing peneliti yaitu bapak Muhammad Jazman, S.Kom., M.InfoSys yang dengan sabar memberikan masukan, arahan, ilmu yang berharga, dan bersedia menyempatkan waktu untuk mengajari dan membimbing peneliti dalam semua hal proses penelitian sehingga peneliti dapat menyelesaikan Tugas Akhir ini. Tidak lupa pula peneliti ingin mengucapkan terima kasih kepada Bapak dan Ibu Dosen Sistem Informasi yang telah mewariskan ilmu, memberikan motivasi, dan arahan untuk menyelesaikan studi di Program Studi Sistem Informasi ini. Terima kasih untuk sahabat terdekat dan teman-teman seperjuangan yang tidak dapat peneliti sebutkan satu persatu bersama kalian masa perkuliahan menjadi lebih bermakna. Semoga dimasa mendatang kita dapat bertemu lagi dalam keadaan yang lebih baik.

Wassalamu 'alaikum Warahmatullahi Wabarakaatuh.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Alhamdulillah Rabbil 'Alamin, bersyukur kehadiran Allah *Subhanahu Wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini. *Shalawat* serta salam kita ucapkan kepada Nabi Muhammad *Shalallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli'Ala Sayyidina Muhammad Wa'Ala Ali Sayyidina Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer di Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Pada penulisan Tugas Akhir ini terdapat beberapa pihak yang sudah berkontribusi dan mendukung peneliti baik berupa materi, moril, dan motivasi. Peneliti ingin mengucapkan banyak terima kasih kepada:

1. Bapak Prof. Dr. H. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom sebagai Sekretaris Program Studi Sistem Informasi.
5. Bapak Arif Marsal, Lc., MA sebagai Ketua Sidang sekaligus Dosen Pembimbing Akademik peneliti yang telah memberikan arahan dari Semester 1 hingga Semester 9.
6. Bapak Muhammad Jazman, S.Kom., M.InfoSys sebagai Dosen Pembimbing Tugas Akhir ini yang telah banyak memberikan arahan, masukan, nasihat, serta motivasinya baik dalam penyelesaian Tugas Akhir maupun juga dalam perkuliahan dan kehidupan sehari-hari. Setiap motivasi yang diberikan akan selalu peneliti ingat dan dijadikan sebagai pelajaran hidup.
7. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Penguji I yang telah dapat meluangkan waktunya untuk memberikan banyak arahan, masukan, nasihat, dan motivasi dalam penyelesaian Tugas Akhir ini.
8. Bapak Syaifullah, SE., M.Sc sebagai Penguji II peneliti yang bersedia meluangkan waktunya untuk membantu memberikan masukan dan saran terhadap peneliti sehingga Tugas Akhir ini dapat terselesaikan dengan baik.
9. Seluruh Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti. Semoga ilmu yang diberikan dapat peneliti amalkan dan menjadi amal jariyah.
10. Seluruh Pegawai dan Staf Fakultas Sains dan Teknologi Universitas Islam

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Negeri Sultan Syarif Kasim Riau yang telah membantu dan mempermudah proses administrasi selama perkuliahan ini.

11. Ibunda tercinta Suryani dan Bapak Abdul Murat yang tanpa lelah selalu memberikan semangat, dukungan, serta do'a terbaiknya, dan selalu menjadi motivasi peneliti dalam menyelesaikan Strata Satu (S1) ini. Terima kasih atas segala keringat, jerih payah pengorbanan, dan kerja keras yang telah kalian berikan dengan penuh keikhlasan demi menuju kesuksesan anakmu ini. Semoga Allah *Subhanahu Wa Ta'ala* selalu menjaga dan melindungi ibu dan bapak dimanapun kalian berada.
12. Teman seperjuangan yaitu Rahma Yanti dan Nova Ariani, S.I.Kom yang berjuang bersama dalam menyelesaikan Tugas Akhir ini dan memberikan dukungan kepada peneliti.
13. Semua pihak yang tidak dapat peneliti sebutkan satu-persatu yang terlibat dalam perjuangan penyelesaian pendidikan Strata Satu (S1) yang telah memberikan bantuan, dukungan, dan semangat kepada peneliti baik dalam pengumpulan data maupun penyusunan Tugas Akhir ini.

Semoga segala do'a dan dorongan yang telah diberikan selama ini menjadi amal kebajikan dan mendapat balasan setimpal dari Allah *Subhanahu Wa Ta'ala*. Peneliti menyadari bahwa penulisan Tugas Akhir ini masih banyak terdapat kekurangan dan jauh dari kata sempurna. Untuk itu, kritik dan saran yang membangun sangat diharapkan untuk kesempurnaan Tugas Akhir ini yang dapat dikirim melalui *email* 12050322233@students.uin-suska.ac.id. Semoga Laporan Tugas Akhir ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terima kasih.

Pekanbaru, 14 Januari 2025

Peneliti,



NOVITA SARI
NIM. 12050322233

Letter of Acceptance

21 November 2024

hak cipta milik UIN Suska Riau

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ipta Dilindungi Undang-Undang

arang mengutip sebagian atau seluruhnya karya tulis ini tanpa mengutip sumbernya.

ngutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

State Islamic University of Sultan

Kasim Riau

2. Dilarang mengutip sebagian atau seluruhnya karya tulis ini tanpa mengutip sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

Dear Novita Sari,
Congratulations, We are pleased to inform you that your following manuscript has been accepted and will be published in SISTEMASI, Vol 14 No. 1, January, 2025 pISSN: 2302-8149 eISSN: 2540-9719.

Penerapan Algoritma Klasifikasi Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Cyberbullying Bilingual di Aplikasi X

: Application of Naive Bayes Classification Algoritma and Support Vector Machine for Bilingual Cyberbullying Sentimen Analysis in X App

Authors : Novita Sari, Muhammad Jazman, Tengku Khairil Ahsyar & Syaifullah

Email : 12050322233@students.uin-suska.ac.id

Received on : 12 November 2024

Revised on : 18 November 2024

Accepted on : 20 November 2024

Thank you very much for submitting your article to "SISTEMASI"



Best Regards,

Prof. Dr. Abdullah Husin, S.Si., M.kom
Chief Editor

UIN SUSKA RIAU



Penerapan Algoritma Klasifikasi *Naive Bayes* dan *Support Vector Machine* untuk Analisis Sentimen *Cyberbullying Bilingual* di Aplikasi X

Application of Naive Bayes Classification Algoritma and Support Vector Machine for Bilingual Cyberbullying Sentimen Analysis in X App

¹Novita Sari *, ²Muhammad Jazman, ³Tengku Khairil Ahsyar, ⁴Syaifullah, ⁵Arif Marsal
^{1,2,3,4,5}Information Systems, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

*e-mail: 12050322233@students.uin-suska.ac.id

(received: ?, revised: ?, accepted: ? diisi oleh editor)

Abstrak

Peningkatan penggunaan media sosial yang signifikan turut mempengaruhi peningkatan insiden *Cyberbullying*, khususnya dalam konteks keanekaragaman penggunaan bahasa. Penelitian ini akan melakukan analisis sentimen guna mendeteksi konten potensial *Cyberbullying* pada aplikasi X dengan pendekatan *bilingual* (bahasa Indonesia dan Inggris) menggunakan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM). Data *tweet* terkumpul dan diproses melalui tahap pra-pemrosesan untuk mengekstraksi fitur yang relevan bagi analisis sentimen. Kedua algoritma kemudian diterapkan dalam mengklasifikasikan *tweet* menjadi kategori positif, negatif, atau netral serta mengidentifikasi indikasi *Cyberbullying*. Hasil uji coba menunjukkan bahwa algoritma NB memberikan kinerja yang lebih unggul dibandingkan SVM dengan tingkat akurasi sebesar 87%. Sedangkan dalam mengidentifikasi pola *Cyberbullying* dalam teks *bilingual* NB mencapai tingkat akurasi tertinggi pada bahasa Indonesia yaitu sebesar 87%. Dengan hasil tersebut, diharapkan penelitian ini dapat menjadi referensi untuk pengembangan sistem deteksi *Cyberbullying* yang lebih akurat dan responsif pada platform media sosial *bilingual*.

Kata kunci: Aplikasi X, Analisis Sentimen *Bilingual*, *Cyberbullying*, NB, SVM.

Abstract

The significant increase in the use of social media has contributed to the increase in cyberbullying incidents, especially in the context of language diversity. This research will conduct sentiment analysis to detect potential *Cyberbullying* content on X application with a *bilingual* approach (Indonesian and English) using *Naive Bayes* and *Support Vector Machine* (SVM) algorithms. The *tweet* data was collected and processed through a pre-processing stage to extract relevant features for sentiment analysis. Both algorithms were then applied to classify tweets into positive, negative or neutral categories and identify indications of cyberbullying. Experimental results show that the NB algorithm provides superior performance compared to SVM with an accuracy rate of 87%. Meanwhile, in identifying *Cyberbullying* patterns in *bilingual* text, NB achieved the highest accuracy rate in Indonesian, which was 87%. With these results, it is hoped that this research can be a reference for the development of a more accurate and responsive *Cyberbullying* detection system on *bilingual* social media platforms.

Keywords: *Bilingual Sentiment Analysis*, *Cyberbullying*, NB, SVM, X app.

Pendahuluan

Salah satu bentuk komunikasi yang sepenuhnya berbasis internet adalah media sosial [1]. Jaringan sosial ahli dalam menghubungkan orang-orang yang memiliki minat dan pemikiran yang sama satu sama lain secara teratur [2]. Salah satu media sosial yang cukup banyak peminatnya yaitu (dulunya *Twitter*) [3]. Perkiraan analisis menurut riset dari website Data Reportal yang dilakukan oleh tim Kepios menyebutkan jumlah pengguna X pada tahun 2024 berada pada angka 335,7 juta pengguna dengan berada di peringkat 12 sebagai jejaring sosial media paling populer [4].

Pengguna platform X dapat berinteraksi dan mengekspresikan diri melalui berbagai *tweet* seperti teks, gambar dan video [5]. Platform ini memungkinkan penggunanya untuk berbagi informasi, berita, opini, dan berinteraksi dengan orang lain secara real-time [6]. Sayangnya sekali, komunikasi ini dapat disalahgunakan untuk *Cyberbullying*, yang merupakan bentuk pelecehan elektronik di mana orang menyerang dan menyinggung orang lain [7].

Konten *Cyberbullying* tidak hanya memicu perilaku berbahaya, tetapi juga membahayakan komunitas online dan kesehatan mental individu [8]. Perilaku ini ditandai dengan tindakan menindas yang dilakukan secara berulang, seringkali didorong oleh berbagai motif seperti kemarahan, balas dendam, atau bahkan hanya untuk mencari sensasi [9]. Fenomena *Cyberbullying* merupakan manifestasi modern dari penindasan tradisional yang memanfaatkan anonimitas dan jangkauan luas internet. Ancaman *Cyberbullying* yang bersifat non-spasial dan berkelanjutan dapat memicu kecemasan, depresi, dan bahkan ide bunuh diri pada korban [10].

Cyberbullying telah menjadi perhatian serius di masyarakat saat ini, seiring dengan meluasnya penggunaan platform media sosial dan saluran komunikasi online lainnya [11]. Dalam penelitian ini, peneliti mengeksplorasi penggunaan algoritma pembelajaran mesin untuk menganalisis konten pesan dan mengidentifikasi pesan yang berisi bahasa yang kasar atau melecehkan [12]. Studi komputasi mengenai opini yang populer dan efisien untuk menganalisis data besar, yang dapat menghasilkan pengambilan keputusan yang lebih baik yaitu menggunakan sentimen analisis [13].

Analisis sentimen merupakan teknik otomatis untuk mengidentifikasi opini positif, negatif, dan netral berdasarkan ulasan pengguna [14]. Analisis sentimen adalah metode populer untuk memanfaatkan keputusan berdasarkan data, yang dapat dibuat dengan mengekstraksi wawasan dari komentar media sosial, tanggapan survei, dan ulasan produk [15]. Teknik dan metode yang sering digunakan dalam klasifikasi sentimen diantaranya *Naive Bayes*(NB), *Support Vector Machine*(SVM), KNN berbasis *machine learning*, *Random Forest* dll [16].

Penelitian ini mengimplementasikan dua algoritma klasifikasi untuk sentimen analisis yaitu *Naive Bayes*(NB) dan *Support Vector Machine*(SVM) [17]. NB merupakan salah satu teknik untuk mengevaluasi atau memvalidasi akurasi model yang dibangun berdasarkan dataset yang digunakan [18]. Sedangkan Salah satu fungsi utama dari SVM adalah mencari garis (*hyperplane*) terbaik yang memisahkan kelas-kelas dengan jarak sejauh mungkin di dalam ruang input [19][20].

Tujuan penelitian ini adalah untuk menemukan algoritma terbaik dari NB dan SVM dalam mengklasifikasi sentimen *Cyberbullying bilingual* di X. Selain itu penelitian ini juga ingin melihat apakah algoritma *Naive Bayes* dan SVM masih bisa digunakan untuk analisis sentimen.

2 Tinjauan Literatur

Ada beberapa penelitian sebelumnya yang telah menerapkan algoritma NB dan SVM untuk menganalisis sentimen pada kasus *Cyberbullying* antara lain analisis perbandingan teknik Machine Learning untuk mendeteksi *Cyberbullying* di Twitter dengan membandingkan 7 model klasifikasi Machine Learning diantaranya NB, SVM, LR, LGBM, SGD, RF dan ADB, berdasarkan hasil klasifikasi diperoleh hasil eksperimen menunjukkan keunggulan LR, yang mencapai akurasi median sekitar 90,57%. Di antara pengklasifikasi, regresi logistik mencapai skor F1 terbaik (0,928), SGD mencapai presisi terbaik (0,968), dan SVM mencapai recall terbaik (1,00) [21].

Perbandingan Metode *Support Vector Machine* dan *Naive Bayes* dalam Klasifikasi *Cyberbullying* di Twitter, hasil penelitian ini adalah akurasi klasifikasi *Naive Bayes* sebesar 97,99% dan akurasi klasifikasi SVM sebesar 99,60% [22]. Kemudian deteksi tingkat keparahan *Cyberbullying*: Pendekatan pembelajaran mesin, dalam penelitian ini kami menerapkan fitur Embedding, Sentiment, dan Leicon beserta orientasi semantik PMI. Fitur yang diekstraksi diterapkan dengan algoritma *Naive Bayes*, KNN, Decision Tree, Random Forest, dan *Support Vector Machine*. Hasil dari



eksperimen dengan kerangka kerja yang kami usulkan dalam pengaturan multikelas menjanjikan baik berkenaan dengan Kappa, akurasi pengklasifikasi dan metrik f-measure, maupun dalam pengaturan ner [23]. Sistem deteksi *Cyberbullying* di media sosial menggunakan supervised machine learning dengan membandingkan model *Support Vector Machine* (SVM), Naïve Bayes, dan Logistic Regression (LR), didapatkan hasil jika dibandingkan dengan algoritma lain, metode SVM memiliki tingkat akurasi yang tinggi yaitu 75,5% [24].

Berdasarkan hasil penelitian terdahulu sebelumnya yang telah saya cantumkan, maka dapat disimpulkan bahwa algoritma NB dan SVM terbukti dapat diimplementasikan untuk sentimen analisis *Cyberbullying* di aplikasi X. Selain itu, dalam penelitian ini juga akan melakukan penelitian sentimen dalam konteks *Cyberbullying bilingual*, yaitu perbandingan akurasi sentimen bahasa Indonesia dengan sentimen bahasa *Inggris* di X.

Metode Penelitian

Bagian 'Metode Penelitian' merupakan bagian penting dari penelitian karena memberikan penjelasan rinci tentang bagaimana penelitian dilakukan. Peneliti menguraikan pendekatan penelitian yang digunakan, populasi dan sampel yang digunakan, variabel yang diteliti, alat yang digunakan untuk pengumpulan data, dan metode yang digunakan untuk analisis data. Tujuan utama bagian ini adalah untuk memberikan gambaran yang jelas dan terbuka mengenai proses penelitian, sehingga pembaca dapat memahami dan menilai kualitas penelitian.

3.1 Pendekatan Penelitian

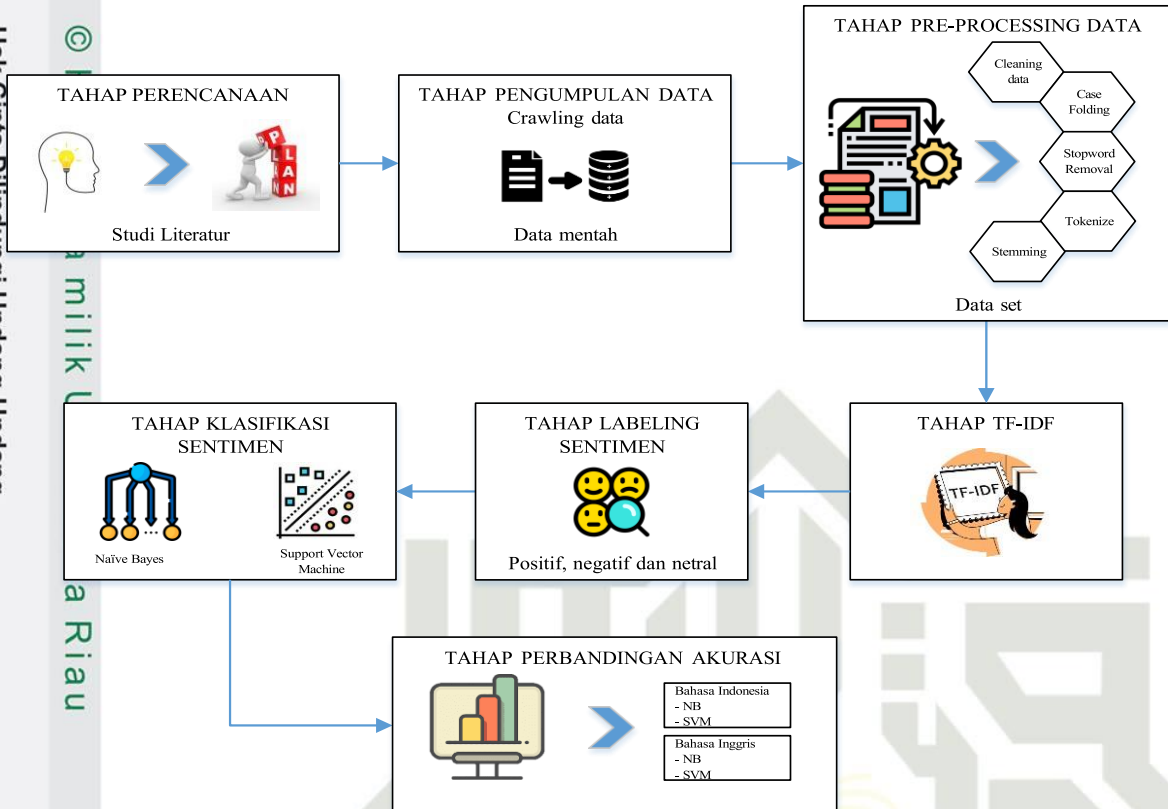
Penelitian ini menggunakan metode klasifikasi sentimen *bilingual* untuk menentukan *tweet* dengan kata kunci "*Cyberbullying*" di X sebagai positif, negatif, atau netral berdasarkan sentimen yang terkandung di dalamnya. Platform *Google Colab* dan bahasa pemrograman *Python* digunakan dalam penelitian ini untuk mengolah data *tweet* yang dikumpulkan oleh X. *Google Colab* menyediakan lingkungan komputasi awan di mana peneliti dapat menjalankan kode *Python* tanpa menginstal aplikasi tambahan. *Python* memiliki banyak perpustakaan kuat seperti NLTK dan spaCy untuk analisis data teks, yang membuatnya pilihan yang tepat. Untuk mendukung pengelolaan literatur dalam penelitian ini, aplikasi Zotero digunakan.

3.2 Populasi dan Sampel

Penelitian ini mengumpulkan semua *tweet* terbaru berbahasa Indonesia dan berbahasa *Inggris* yang diposting di X mengandung kata kunci "*Cyberbullying*". Populasi ini mewakili keseluruhan *tweet* berbahasa Indonesia dan berbahasa *Inggris* yang membahas tentang *Cyberbullying* di X tersebut. Metode pengambilan data atau sampel menggunakan teknik *Crawling* didapatkan 502 *tweet* data berbahasa Indonesia dan 492 *tweet* data berbahasa *Inggris* dari populasi. Ini memastikan bahwa setiap *tweet* di X yang memenuhi kriteria populasi memiliki peluang yang sama untuk masuk dalam sampel.

3.3 Metodologi Penelitian

Dalam penelitian ini, pertama-tama peneliti akan menentukan arah dan pondasi penelitian agar terlaksana dan efektif dan efisien, kemudian peneliti mengumpulkan data *tweet* tentang *Cyberbullying* dari X. Setelah itu dilakukan pra-pemrosesan data *tweet* yang telah di dapatkan. Langkah selanjutnya adalah melakukan proses *TF-IDF* untuk memberikan bobot pada kata yang penting. Kemudian kata tersebut diberikan label sentimen berupa label positif, negatif dan netral. Selanjutnya peneliti membandingkan dua model algoritma klasifikasi *Naive Bayes* dan *Support Vector Machine*. Tahap terakhir yaitu melakukan perbandingan akurasi pada kedua algoritma sekaligus perbandingan antara penggunaan algoritma untuk perbandingan dua bahasa. Gambar 1 memberikan gambaran lengkap tentang metodologi penelitian.



Gambar 1 Metodologi Penelitian

3.3.1 Tahap perencanaan

Tahapan perencanaan berisikan tahap studi literatur untuk mencari informasi mengenai referensi dari penelitian sebelumnya yang telah melakukan riset berkaitan dengan analisis sentimen *Cyberbullying bilingual* menggunakan algoritma NB dan SVM.

3.3.2 Pengumpulan Data

Tujuan pengumpulan data adalah untuk mendapatkan dataset yang dapat dianalisis. Data dikumpulkan melalui teknologi *crawling* dengan memanfaatkan bahasa pemrograman *Python* dan platform *Google Colab*. Data yang diambil berasal dari postingan pengguna pada aplikasi *X* dengan kata kunci —*Cyberbullying*—.

Web crawler adalah program komputer yang dimaksudkan untuk mengunjungi halaman web, menyalin konten yang ada, dan menyimpannya untuk analisis kemudian [25]. Data atau sampel yang dikumpulkan dari *X* didapatkan sebanyak 349 *tweet* berbahasa Indonesia dan 349 *tweet* berbahasa Inggris dari populasi. Hal ini memastikan bahwa semua *tweet* yang memenuhi kriteria populasi memiliki peluang yang sama untuk dimasukkan dalam sampel.

3.3.3 Pre-Processing Data

Preprocessing teks adalah prosedur untuk menghilangkan kebisingan atau meningkatkan kualitas teks [26]. Pra-pemrosesan data adalah metode pengolahan data yang mengubah data mentah menjadi format yang mudah dipahami. Tahap pra-pemrosesan data diperlukan untuk menyelesaikan masalah seperti data berisik, redundansi, dan nilai yang hilang [27].

Tahap ini berfokus pada manipulasi teks itu sendiri, seperti:

3.3.3.1 Cleaning Data

Pembersihan data melibatkan pencarian dan koreksi atau penghapusan data yang salah atau bermasalah dari kumpulan data. Proses ini biasanya digunakan untuk mencari dan mengganti data atau bagian yang tidak lengkap, tidak akurat, tidak relevan, atau salah [28].



3.3.3.2 Case Folding

Case Folding adalah teknik yang mengubah setiap huruf dalam dokumen atau kalimat menjadi huruf yang lebih kecil yang dapat dicari dengan lebih mudah. Dalam hal penggunaan huruf kapital, tidak semua data konsisten [29].

3.3.3.3 Stopword Removal

Proses menghilangkan kata-kata seperti "di", "dari", dan "yang", serta kata-kata seperti "a", "the", dan "is" yang tidak dapat ditemukan di mesin pencari komputer. Menghapus stopwords meningkatkan efisiensi dan akurasi aplikasi penambangan teks, dan juga mengurangi kompleksitas waktu dan ruang aplikasi penambangan teks secara keseluruhan [30].

3.3.3.4 Tokenization

Tokenisasi adalah pembagian kumpulan karakter berdasarkan ruang. Pada saat yang sama, juga dapat menghapus karakter tertentu, seperti tanda baca [31]

3.3.3.5 Stemming

Periksa kata-kata yang telah diubah menjadi huruf kecil. Stemming mengurangi daftar kata dalam data pelatihan dengan membakukan kata [32].

3.3.4 TF-IDF

Frekuensi Term - Frekuensi Dokumen Terbalik (TF-IDF) untuk mengatasi masalah bahwa kata yang sering muncul dilebih-lebihkan oleh jumlah kata mentah yang dikaitkan dengan fitur, pendekatan TF-IDF digunakan untuk mengukur seberapa sering kata tersebut muncul dalam dokumen. Oleh karena itu, beberapa algoritme klasifikasi mungkin tidak menghitung jumlah kata secara optimal menggunakan ukuran frekuensi kata dalam dokumen [33].

Rumus untuk algoritma TF-IDF ditunjukkan pada (1) - (3), di mana TF_{ij} menunjukkan nilai TF dari kata ke- i untuk dokumen ke- j , IDF_i menunjukkan nilai IDF dari kata ke- i , dan $TF-IDF_{ij}$ menunjukkan nilai TF-IDF dari kata ke- i untuk dokumen ke- j . D adalah jumlah total semua dokumen, D_j adalah dokumen ke- j , n_{ij} adalah jumlah kemunculan kata ke- i pada dokumen ke- j , dan N_i adalah jumlah kemunculan kata ke- i pada semua dokumen.

$$TF_{ij} = \frac{n_{ij}}{D_j} \quad (1)$$

$$IDF_i = \log\left(\frac{D}{N_i}\right) \quad (2)$$

$$TF - IDF_{ij} = TF_{ij} \cdot IDF_i \quad (3)$$

Rumus TF-IDF yang telah diperbaiki ditunjukkan pada (4), dengan kata ke- i muncul pada segmen pertama dan terakhir ketika $f(i)$ jika tidak, $f(i) = 1$; $g(i)$ ketika kata ke- i adalah kata benda, dan $g(i) = 1$ ketika kata ke- i memiliki sifat leksikal yang berbeda. adalah angka yang lebih besar dari nol, dan suku-suku yang lain memiliki makna yang sama seperti pada persamaan (1) - (3) [34].

$$TF - IDF_{ij} = f(i) \cdot g(i) \cdot TF_{ij} \cdot IDF_i \quad (4)$$

3.3.5 Labeling Sentimen

Nilai sentimen diberikan pada teks, yang dapat berupa nilai positif, negatif, atau netral. Nilai sentimen dapat dilabelkan dengan berbagai cara, seperti secara manual dengan memasukkan nilai sentimen pada teks, menggunakan fungsi TF-IDF, atau dengan *Python* dengan menggunakan library *TextBlob*. Pada penelitian ini peneliti menggunakan library *TextBlob* dalam proses pemberian labeling sentimen pada data.



3.3.6 Klasifikasi Sentimen

Pada bagian ini dilakukan proses analisis sentimen pada data yang telah melewati tahap pemrosesan data. Tahap ini terdiri dari proses klasifikasi *Naive Bayes* dan klasifikasi *Support Vector Machine*.

3.3.6.1 Naive Bayes (NB)

Algoritma pembelajaran mesin *Naive Bayes* termasuk dalam kategori pembelajaran terawasi pembelajaran mesin yang memerlukan sampel berlabel sebagai data pelatihan) [18]. Algoritme ini sangat disukai karena kesederhanaannya, efisiensinya, dan skalabilitasnya, yang membuatnya menarik untuk berbagai tugas klasifikasi [35].

Naive Bayes hanya bisa mengenali teks dan angka, tapi tidak bisa mengenali gambar. Metode ini menggunakan teorema Bayes untuk menghitung probabilitas [36]. Dalam bahasa sederhana, teorema Bayes didefinisikan sebagai kemungkinan bahwa hubungan A akan terjadi jika hubungan B telah terjadi sebelumnya dan sebaliknya [37]. Persamaan bayes seperti persamaan (5).

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (5)$$

Di mana :

X = Data dengan class yang belum diketahui

H = Hipotesis data merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)

P(H) = Probabilitas hipotesis H (prior probabilitas)

P(X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H

P(X) = Probabilitas X

3.3.6.2 Support Vector Machine (SVM)

Dengan menggunakan fungsi kernel, SVM dapat menangani data linier dan nonlinier. Ini memetakan ruang masukan ke ruang fitur berdimensi lebih besar, memungkinkan pemisahan linier [38]. Sedangkan Salah satu fungsi utama dari SVM adalah mencari garis (hyperplane) terbaik yang memisahkan kelas-kelas dengan jarak sejauh mungkin di dalam ruang input [20].

Akurasi algoritma SVM bergantung pada klasifikasi linear, yang digunakan, dan fungsi kernelnya. SVM dapat menangani data dengan karakteristik non-linier dengan baik dengan menggunakan berbagai jenis fungsi kernel, seperti kernel linier, polinomial, atau Gauss. Fungsi kernel adalah fungsi yang mengubah data ke dimensi yang lebih besar dengan tujuan meningkatkan struktur data sehingga mempermudah proses pemisahan [39].

Rumus fungsi kernel ditunjukkan pada persamaan (6)-(8).

$$\text{Linier } K(x, y) = x, y \quad (6)$$

$$\text{Polynomial } K(x_i, x_j) = ((x_i \cdot x_j) + c)^d \quad (7)$$

$$\text{Gaussian } K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right) \quad (8)$$

Keterangan:

K(X_i, X_j) = Fungsi kernel

X_i = Data ke-i

X_j = Data ke-j

eX^p = Operasi eksponen

d, σ = Parameter kernel

3.3.7 Perbandingan Akurasi

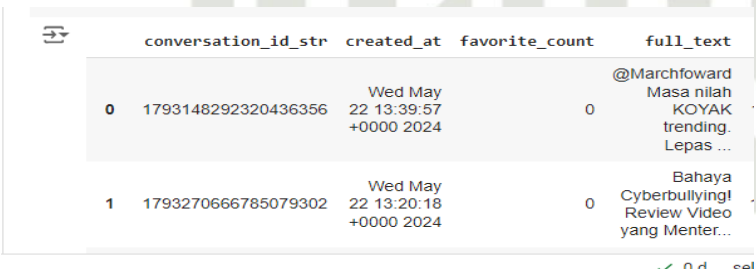
Perbandingan akurasi dilakukan untuk menentukan algoritma mana yang lebih baik dalam menghasilkan nilai pada data aplikasi X tentang *Cyberbullying* menggunakan algoritma *Naive Bayes* dan *Support Vector Machine*. Pada perbandingan akurasi juga dilakukan untuk menentukan bahasa apa yang memiliki akurasi yang lebih tinggi dalam melakukan analisis sentimen. Perbandingan yang dilakukan dengan membandingkan nilai akurasi, precision, recall, dan f1 score pada masing masing algoritma.

Hasil dan Pembahasan

Setelah mengetahui metodologi yang akan digunakan dalam penelitian ini, langkah berikutnya adalah menjelaskan hasilnya. Mula-mula, data dikumpulkan dari X; kemudian dibersihkan dari noise, duplikat, karakter atau simbol, dan pemberian bobot; kemudian, terlepas dari kata-kata yang digunakan, diberikan label positif, negatif, atau netral sesuai dengan perasaan yang digunakan. Kemudian data diklasifikasikan menurut algoritma NB dan SVM untuk mendapatkan nilai akurasi terbaik. Selanjutnya, perbandingan akurasi dilakukan untuk menentukan algoritma mana yang paling akurat dan untuk menentukan bahasa mana yang paling akurat dari kedua algoritma tersebut. Hasil dan pembahasan lengkap dapat dilihat pada tahapan 4.1 sampai 4.6.

4.1 Pengumpulan Data

Proses pengumpulan data dilakukan dengan teknik *Crawling Data* menggunakan *Google Colab* dan bahasa pemrograman *Python* di aplikasi *X* dengan kata kunci —*Cyberbullying*—. Data *tweet* yang dikumpulkan merupakan data terbaru. Untuk dapat mengakses data *X* diperlukannya *Auth Token X* tersebut sebagai kode akses agar bisa dilakukannya *Crawling Data* oleh *Google Colab*. Proses *Crawling Data* dilakukan pada dua bahasa yang berbeda, bahasa Indonesia dan bahasa *Inggris* dengan topik yang sama yaitu *Cyberbullying*. Hasil *Crawling Data* bahasa Indonesia dan bahasa *Inggris* dari *X* ditampilkan pada Gambar 2 dan 3.



	conversation_id_str	created_at	favorite_count	full_text
0	1793148292320436356	Wed May 22 13:39:57 +0000 2024	0	@Marchfoward Masa nih KOYAK trending. Lepas ...
1	1793270666785079302	Wed May 22 13:20:18 +0000 2024	0	Bahaya Cyberbullying! Review Video yang Menter...

Gambar 2 Hasil *Crawling data* bahasa Indonesia



	conversation_id_str	created_at	favorite_count	full_text
0	1793477405778305491	Thu May 23 03:01:49 +0000 2024	1	#Survivor #Survivor47 I will be cyberbullying ...
1	1793448307194544458	Thu May 23 02:59:03 +0000 2024	1	@MasteroftheTDS @Grumz Just a home address ? ...
2	1792577666794606948	Thu May 23 02:52:00 +0000 2024	7	@laylassong @acmecojim @AstroLamaBeans sittin...

Gambar 3 Hasil *Crawling data* bahasa Inggris.

Berdasarkan hasil *Crawling data* seperti terlihat pada gambar 2 dan gambar 3 diatas, didapatkan sebanyak 502 *tweet* berbahasa Indonesia dan 492 bahasa *Inggris* yang membahas tentang *Cyberbullying*.

4.2 Pre-Processing Data

Tahap ini bertujuan untuk membersihkan dan menstandarisasi data teks agar algoritma analisis sentimen menjadi lebih terstruktur dan lebih mudah diolah. Tabel 1 menunjukkan hasil *Pre-Processing* untuk bahasa Indonesia, dan Tabel 2 menunjukkan hasil untuk bahasa *Inggris*.

Tabel 1 Hasil *Pre-Processing* bahasa Indonesia

Alur <i>Pre-Processing</i>	Sebelum	Setelah
<i>Cleaning data</i>	Apa susahnya nyebut safa n si berflower itu saja alih-alih menyebut nama artis secara langsung untuk dijadiin lelucon? Sangat disayangkan pelaku cyberbullying (safa)	Apa susahnya nyebut safa n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon Sangat disayangkan

<p>© Hak cipta milik UIN Suska Riau</p>	<p>malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>	<p>pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>
<p><i>Case Folding</i></p>	<p>Apa susahnya nyebut safe n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon Sangat disayangkan pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>	<p>apa susahnya nyebut safe n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon sangat disayangkan pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>
<p><i>Stopword Removal</i></p>	<p>apa susahnya nyebut safe n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon sangat disayangkan pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>	<p>apa susahnya nyebut safe n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon sangat disayangkan pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>
<p><i>Tokenize</i></p>	<p>apa susahnya nyebut safe n si berflower itu saja alihalih menyebut nama artis secara langsung untuk dijadiin lelucon sangat disayangkan pelaku cyberbullying safe malah dijadikan seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak bersalah sllu disebut</p>	<p>[apa, susahnya, nyebut, safe, n, si, berflower, itu, saja, alihalih, menyebut, nama, artis, secara, langsung, untuk, dijadiin, lelucon, sangat, disayangkan, pelaku, cyberbullying, safe, malah, dijadikan, seperti, korban, karna, org, hanya, fokus, dgn, si, berflower, dan, nama, idol, yg, gak, bersalah, sllu, disebut]</p>
<p><i>Stemming</i></p>	<p>[apa, susahnya, nyebut, safe, n, si, berflower, itu, saja, alihalih, menyebut, nama, artis, secara, langsung, untuk, dijadiin, lelucon, sangat, disayangkan, pelaku, cyberbullying, safe, malah, dijadikan, seperti, korban, karna, org, hanya, fokus, dgn, si, berflower, dan, nama, idol, yg, gak, bersalah, sllu, disebut]</p>	<p>apa susah nyebut safe n si berflower itu saja alihalih sebut nama artis cara langsung untuk dijadiin lelucon sangat sayang laku cyberbullying safe malah jadi seperti korban karna org hanya fokus dgn si berflower dan nama idol yg gak salah sllu sebut</p>

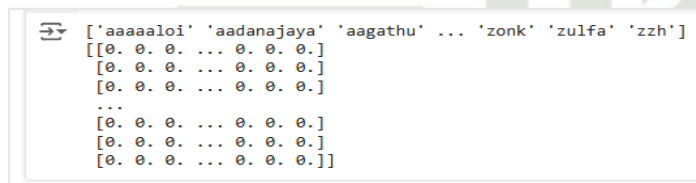
Tabel 2 Hasil Pre-Processing bahasa Inggris

Alur Pre-Processing	Sebelum	Sesudah
Cleaning data	@Griot2325 Most tame examples cyberbullying Ive ever seen in my life Twitter reactionaries cannot be real people	Griot Most tame examples cyberbullying Ive ever seen in my life Twitter reactionaries cannot be real people
Case Folding	Griot Most tame examples cyberbullying Ive ever seen in my life Twitter reactionaries cannot be real people	griot most tame examples cyberbullying ive ever seen in my life twitter reactionaries cannot be real people
Stopword Removal	griot most tame examples cyberbullying ive ever seen in my life twitter reactionaries cannot be real people	griot tame examples cyberbullying ive ever seen life twitter reactionaries cannot real people
Tokenize	griot tame examples cyberbullying ive ever seen life twitter reactionaries cannot real people	[griot, tame, examples, cyberbullying, ive, ever, seen, life, twitter, reactionaries, cannot, real, people]
Stemming	[griot, tame, examples, cyberbullying, ive, ever, seen, life, twitter, reactionaries, cannot, real, people]	griot tame examples cyberbullying ive ever seen life twitter reactionaries cannot real people

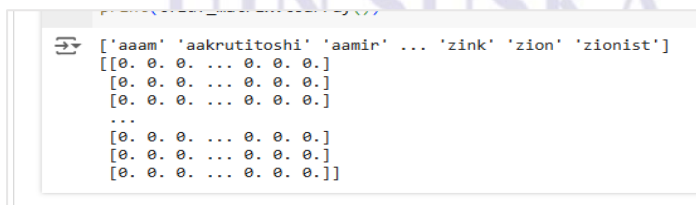
Seperti terlihat pada tabel di atas, kumpulan data yang lolos tahap prapemrosesan data bersifat seragam, terstruktur, dan tidak mengandung simbol di dalam teks. Sebab, ketika sudah sampai pada tahap selanjutnya yaitu tahap klasifikasi, komputasi menjadi lebih optimal.

4.3 TF-IDF

TF-IDF dapat mengekstrak fitur teks penting, yang membantu memahami makna dan memprediksi sentimen. Kata kunci yang sering digunakan tetapi jarang ditemukan di korpus lain akan menerima skor yang lebih tinggi, yang dapat digunakan untuk membedakan data. Sementara kata-kata umum yang sering digunakan akan dikurangi. Hasil pembobotan TF-IDF bahasa Indonesia dan bahasa Inggris dapat dilihat pada Gambar 4-5.



Gambar 4 Hasil TF-IDF data bahasa Indonesia

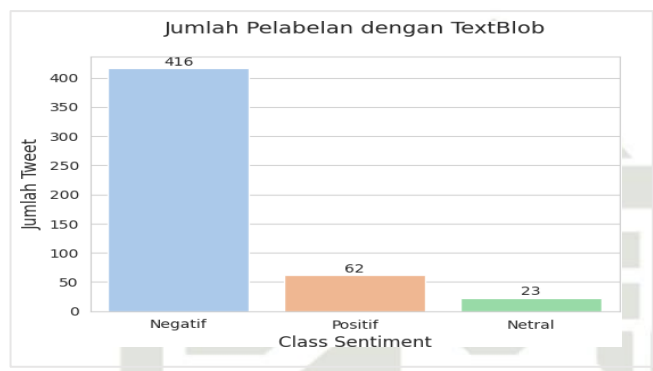


Gambar 5 Hasil TF-IDF data bahasa Inggris

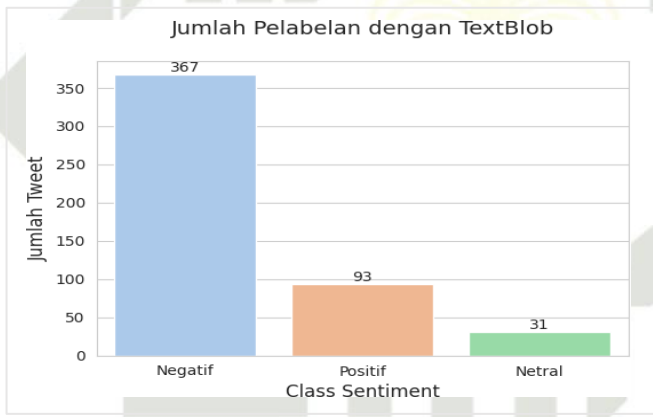
4.4 Labeling Data

Labeling adalah proses memberikan label sentimen pada data teks. Proses labeling menggunakan library pemrograman *TextBlob*. Label ini menunjukkan apakah teks tersebut memiliki sentimen positif, negatif atau netral. Pada data yang telah didapatkan sebanyak 502 data berbahasa Indonesia dan 492 data berbahasa *Inggris*, kedua data tersebut memiliki jumlah pelabelan yang berbeda.

Berdasarkan Gambar 6 dan 7 terlihat bahwa data dengan nilai sentimen negatif lebih tinggi dibandingkan dengan sentimen positif. Ini berarti bahwa sebagian besar diskusi atau komentar tentang cyberbullying cenderung mengekspresikan perasaan negatif seperti kemarahan, kesedihan, atau ketakutan. Topik ini kemungkinan besar dianggap sebagai masalah serius yang memiliki dampak negatif pada individu dan masyarakat.



Gambar 6 Hasil *labeling* sentimen pada data bahasa Indonesia



Gambar 7 Hasil *labeling* sentimen pada data bahasa *Inggris*

4.5 Klasifikasi Sentimen

Pada bagian ini akan menjelaskan bagaimana proses pengklasifikasian pada kedua algoritma yang digunakan yaitu *Naive Bayes* dan SVM. Setelah data melewati proses pra-pemrosesan dan *labeling* data, maka langkah selanjutnya mengklasifikasikan data menggunakan algoritma NB dan SVM untuk mendapatkan akurasi terbaik dari kedua algoritma tersebut. Selain itu pada bagian ini juga akan melihat akurasi algoritma pada bahasa mana yang tertinggi.

4.5.1 *Naive Bayes* (NB)

Naive Bayes adalah metode klasifikasi teks sederhana namun efektif yang digunakan untuk memprediksi kemungkinan suatu kategori berdasarkan bukti historis. Pada penelitian ini data set dibagi menjadi data latih dan data uji sebanyak 20% dari keseluruhan data. Hasil klasifikasi *Naive Bayes* dapat dilihat pada Tabel 3.

Tabel 3 Hasil klasifikasi *Naive Bayes*

Metriks klasifikasi	Bahasa Indonesia	Bahasa <i>Inggris</i>
Akurasi	69%	56%
<i>Precision</i>	76%	55%

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

<i>Recall</i>	69%	56%
<i>F1-score</i>	72%	54%

Precision merupakan proses mengukur seberapa sering prediksi positif model benar-benar positif. Tujuannya untuk meminimalkan *false positive* (memprediksi positif padahal sebenarnya negatif). Sedangkan *Recall* kebalikan dari *precision* dimana berfungsi untuk mengukur seberapa banyak dari semua contoh positif yang sebenarnya berhasil diidentifikasi oleh model. *F1-score* adalah nilai rata-rata harmonik untuk *precision* dan *recall*.

Ada sejumlah faktor yang dapat menyebabkan perbedaan hasil klasifikasi antara Bahasa Indonesia dan Bahasa Inggris di tabel tersebut. Struktur, ukuran, atau kompleksitas dataset yang digunakan dapat memengaruhi kinerja model. Aturan linguistik setiap bahasa memengaruhi kualitas data yang dihasilkan oleh proses preprocessing teks seperti tokenisasi dan stemming. Selain itu, mungkin model Naive Bayes sederhana lebih cocok untuk dataset atau bahasa tertentu.

5.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah salah satu algoritma yang dapat digunakan untuk klasifikasi dan regresi. Dalam konteks klasifikasi, SVM mencari *hyperplane* terbaik untuk memisahkan data menjadi dua kelas atau lebih. Pada penelitian ini data set dibagi menjadi data latih dan data uji sebanyak 20% dari keseluruhan data. Hasil klasifikasi SVM dapat dilihat pada Tabel 4.

Tabel 4 Hasil klasifikasi SVM

Metriks klasifikasi	Bahasa Indonesia	Bahasa Inggris
Akurasi	86%	66%
<i>Precision</i>	76%	43%
<i>Recall</i>	86%	66%
<i>F1-score</i>	81%	52%

Dari tabel 4 terlihat bahwa perbedaan hasil klasifikasi antara Bahasa Indonesia dan Bahasa Inggris mungkin disebabkan oleh variabel seperti perbedaan struktur bahasa, kualitas, dan kuantitas data latih. Mungkin karena data Bahasa Indonesia lebih representatif atau memiliki struktur yang lebih sederhana dibandingkan dengan Bahasa Inggris. Model SVM yang dilatih dengan data bahasa Indonesia lebih baik dalam mengklasifikasikan sentimen dibandingkan dengan model SVM yang dilatih dengan data bahasa Inggris. Ini menunjukkan bahwa kualitas dan kuantitas data latih, serta karakteristik bahasa, sangat penting untuk keberhasilan analisis sentimen.

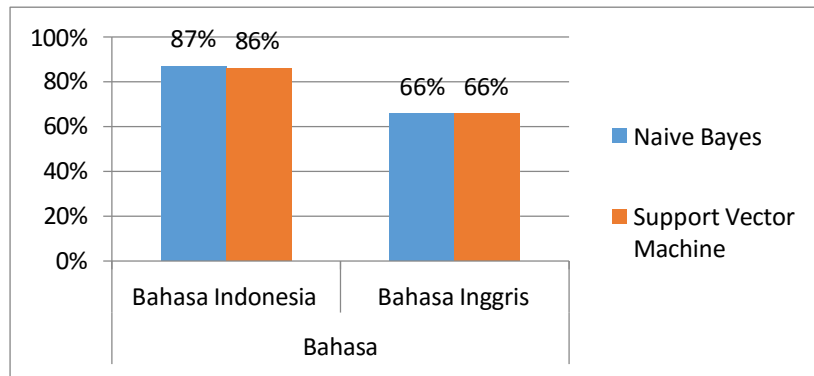
4.6 Perbandingan Akurasi

Perbandingan akurasi dilakukan untuk menentukan algoritma mana yang lebih baik dalam menghasilkan nilai pada data aplikasi X tentang *Cyberbullying* menggunakan algoritma Naive Bayes dan Support Vector Machine. Pada perbandingan akurasi juga dilakukan untuk menentukan bahasa apa yang memiliki akurasi yang lebih tinggi dalam melakukan analisis sentimen. Perbandingan yang dilakukan dengan membandingkan nilai akurasi, precision, recall, dan f1 score pada masing masing algoritma. Hasil perbandingan akurasi dapat dilihat pada Tabel 5 dan representasi hasil pada Gambar 8.

Tabel 5 Hasil perbandingan Akurasi

Algoritma	Bahasa	
	Bahasa Indonesia	Bahasa Inggris
<i>Naive Bayes</i>	87%	66%
<i>Support Vector Machine</i>	86%	66%

Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 8 Representasi hasil perbandingan akurasi

Berdasarkan Tabel 5, *Naive Bayes* memiliki akurasi tertinggi dari SVM dan bahasa Indonesia mendapatkan nilai akurasi tertinggi dibandingkan bahasa *Inggris*. Ini bisa terjadi karena perbedaan akurasi antara bahasa Indonesia dan bahasa Inggris dalam analisis sentimen dapat disebabkan oleh berbagai faktor, mulai dari kompleksitas bahasa hingga kualitas data latih. Karakteristik data, tujuan analisis, dan sumber daya komputasi menentukan pilihan antara Naive Bayes dan SVM. Dalam beberapa situasi, Naive Bayes mungkin memberikan hasil yang lebih baik daripada SVM, terutama dalam kasus di mana data yang digunakan berkualitas tinggi dan memiliki struktur yang sederhana.

5 Kesimpulan

Penelitian ini melakukan analisis sentimen pada ulasan pengguna *X* terhadap topik *Cyberbullying* dengan membandingkan 2 Algoritma yaitu NB dan SVM. Berdasarkan hasil klasifikasi yang diperoleh melalui tahapan *Pre-Processing*, algoritma NB terbukti lebih unggul dari algoritma SVM dengan nilai akurasi tertinggi sebesar 87%, sedangkan SVM 86%. Selain itu, perbandingan akurasi bahasa, tertinggi didapatkan oleh NB bahasa Indonesia yaitu sebesar 87%. Berarti hasil penelitian ini menunjukkan bahwa NB masih sangat relevan untuk algoritma klasifikasi sentimen *bilingual*. Pada penelitian berikutnya, kami sarankan untuk memperluas penggunaan algoritma terbaru dan penggunaan bahasa asing lainnya agar klasifikasi nya bisa lebih luas tidak hanya dalam ruang lingkup kecil saja.

Referensi

- [1] A. Lüders, A. Dinkelberg, and M. Quayle, 'Becoming —usl in digital spaces: How online users creatively and strategically exploit social media affordances to build up social identity', *Acta Psychologica*, vol. 228, p. 103643, Aug. 2022, doi: 10.1016/j.actpsy.2022.103643.
- [2] N. Sabermajidi, N. Valaei, M. S. Balaji, and S. K. Goh, 'Measuring brand-related content in social media: a socialization theory perspective', *Information Technology & People*, vol. 33, no. 4, pp. 1281–1302, Jan. 2020, doi: 10.1108/ITP-10-2018-0497.
- [3] L. Stracqualursi and P. Agati, 'Tweet topics and sentiments relating to distance learning among Italian Twitter users', *Sci Rep*, vol. 12, no. 1, p. 9163, Jun. 2022, doi: 10.1038/s41598-022-12915-w.
- [4] S. Matsa Sarah Naseer, Jacob Liedke and Katerina Eva, 'How Americans Get News on TikTok, X, Facebook and Instagram', Pew Research Center. Accessed: Aug. 11, 2024. [Online]. Available: <https://www.pewresearch.org/journalism/2024/06/12/how-americans-get-news-on-tiktok-x-facebook-and-instagram/>
- [5] R. Kullar, D. A. Goff, T. P. Gauthier, and T. C. Smith, 'To Tweet or Not to Tweet—a Review of the Viral Power of Twitter for Infectious Diseases', *Curr Infect Dis Rep*, vol. 22, no. 6, p. 14, Jun. 2020, doi: 10.1007/s11908-020-00723-0.
- [6] L. Stracqualursi and P. Agati, 'Twitter users perceptions of AI-based e-learning technologies', *Scientific Reports*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-56284-y.
- [7] O. A. Alismaiel, 'Digital Media Used in Education: The Influence on Cyberbullying Behaviors among Youth Students', *IJERPH*, vol. 20, no. 2, p. 1370, Jan. 2023, doi: 10.3390/ijerph20021370.

- Shak Cipta Diindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
2. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
- [8] D. Kim, 'Cyberbullying Behaviors in Online Travel Community: Members' Perceptions and Sustainability in Online Community', *Sustainability*, vol. 14, no. 9, p. 5220, Apr. 2022, doi: 10.3390/su14095220.
 - [9] Ahmad Mohamad Alomar and Hassan Sami Alabady, 'The Phenomenon of Cyber Bullying: Interpretation, Confrontation, and the Position of Islamic Law', *JNS*, vol. 34, May 2023, doi: 10.59670/jns.v34i.1123.
 - [10] Ahmad Mohammad Alomar et Al, 'Aspect and Special Distinct Nature of Cyberbullying', *Russian Law Journal*, vol. 11, no. 3, Art. no. 3, Apr. 2023, doi: 10.52783/rlj.v11i3.1817.
 - [11] L. H. Collantes, Y. Martafian, S. N. Khofifah, T. Kurnia Fajarwati, N. T. Lassela, and M. Khairunnisa, 'The Impact of Cyberbullying on Mental Health of the Victims', in *2020 4th International Conference on Vocational Education and Training (ICOVET)*, Malang, Indonesia: IEEE, Sep. 2020, pp. 30–35. doi: 10.1109/ICOVET50258.2020.9230008.
 - [12] M. P. Akhter, Z. Jiangbin, I. R. Naqvi, M. AbdelMajeed, and T. Zia, 'Correction to: Abusive language detection from social media comments using conventional machine learning and deep learning approaches', *Multimedia Systems*, vol. 29, no. 1, pp. 451–451, Feb. 2023, doi: 10.1007/s00530-021-00819-0.
 - [13] I. Awajan, M. Mohamad, and A. Al-Quran, 'Sentiment Analysis Technique and Neutrosophic Set Theory for Mining and Ranking Big Data From Online Reviews', *IEEE Access*, vol. 9, pp. 47338–47353, 2021, doi: 10.1109/ACCESS.2021.3067844.
 - [14] Fathurahman Bei and Sudin Saepudin, 'Analisis Sentimen Aplikasi Tiket Online di Play Store Menggunakan Metode Support Vector Machine (SVM)', 2021.
 - [15] H. Hertina et al., 'Data mining applied about polygamy using sentiment analysis on Twitters in Indonesian perception', *Bulletin EEI*, vol. 10, no. 4, pp. 2231–2236, Aug. 2021, doi: 10.11591/eei.v10i4.2325.
 - [16] B. AlBadani, R. Shi, and J. Dong, 'A Novel Machine Learning Approach for Sentiment Analysis on Twitter Incorporating the Universal Language Model Fine-Tuning and SVM', *ASI*, vol. 5, no. 1, p. 13, Jan. 2022, doi: 10.3390/asi5010013.
 - [17] M. R. Romadhon and F. Kurniawan, 'A Comparison of Naive Bayes Methods, Logistic Regression and KNN for Predicting Healing of Covid-19 Patients in Indonesia', in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, Surabaya, Indonesia: IEEE, Apr. 2021, pp. 41–44. doi: 10.1109/EIConCIT50028.2021.9431845.
 - [18] A. Roihan, P. A. Sunarya, and A. S. Rafika, 'Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper', *IJCIT*, vol. 5, no. 1, May 2020, doi: 10.31294/ijcit.v5i1.7951.
 - [19] 'Survey on Dietary Application through Image Processing for Calorie Management', *International Journal of Advanced Research in Science, Communication and Technology*, pp. 345–347, May 2022, doi: 10.48175/ijarsct-3666.
 - [20] M. Muhathir, M. H. Santoso, and D. A. Larasati, 'Wayang image classification using SVM method and GLCM feature extraction', *Journal Of Informatics And Telecommunication Engineering*, vol. 4, no. 2, pp. 373–382, 2021.
 - [21] A. Muneer and S. M. Fati, 'A comparative analysis of machine learning techniques for cyberbullying detection on twitter', *Future Internet*, vol. 12, no. 11, pp. 1–21, 2020, doi: 10.3390/fi12110187.
 - [22] N. Chamidah and R. Sahawaly, 'Comparison Support Vector Machine and Naive Bayes Methods for Classifying Cyberbullying in Twitter', *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 7, no. 2, p. 338, Sep. 2021, doi: 10.26555/jiteki.v7i2.21175.
 - [23] B. A. Talpur and D. O'Sullivan, 'Cyberbullying severity detection: A machine learning approach', *PLoS ONE*, vol. 15, no. 10, October, pp. 1–19, 2020, doi: 10.1371/journal.pone.0240924.
 - [24] A. Perera and P. Fernando, 'Cyberbullying Detection System on Social Media Using Supervised Machine Learning', *Procedia Computer Science*, vol. 239, pp. 506–516, 2024, doi: 10.1016/j.procs.2024.06.200.
 - [25] C. Muehlethaler and R. Albert, 'Collecting data on textiles from the internet using web crawling and web scraping tools', *Forensic Science International*, vol. 322, p. 110753, 2021, doi: 10.1016/j.forsciint.2021.110753.



[26] A. P. Natasuwarna, 'Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring', *Techno.Com*, vol. 19, no. 4, pp. 437–448, 2020, doi: 10.33633/tc.v19i4.4044.

[27] U. Naseem, I. Razzak, and P. W. Eklund, 'A survey of pre-processing techniques to improve short-text quality: a case study on hate speech detection on twitter', *Multimedia Tools and Applications*, vol. 80, no. 28–29, pp. 35239–35266, 2021, doi: 10.1007/s11042-020-10082-6.

[28] K. Maharana, S. Mondal, and B. Nemade, 'A review: Data pre-processing and data augmentation techniques', *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltip.2022.04.020.

[29] D. Alita and A. R. Isnain, 'Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier', *Jurnal Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.

[30] D. J. Ladani and N. P. Desai, 'Stopword Identification and Removal Techniques on TC and IR applications: A Survey', *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, pp. 466–472, 2020, doi: 10.1109/ICACCS48705.2020.9074166.

[31] Y. A. Singgalen, 'Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL', *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 4, pp. 1899–1914, 2023, doi: 10.47065/bits.v4i4.3231.

[32] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, 'Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations', *Organizational Research Methods*, vol. 25, no. 1, pp. 114–146, 2022, doi: 10.1177/1094428120971683.

[33] M. Kamyab, G. Liu, and M. Adjeisah, 'Attention-Based CNN and Bi-LSTM Model Based on TF-IDF and GloVe Word Embedding for Sentiment Analysis', *Applied Sciences (Switzerland)*, vol. 11, no. 23, 2021, doi: 10.3390/app112311255.

[34] M. Liang and T. Niu, 'Research on Text Classification Techniques Based on Improved TF-IDF Algorithm and LSTM Inputs', *Procedia Computer Science*, vol. 208, pp. 460–470, 2022, doi: 10.1016/j.procs.2022.10.064.

[35] I. Wickramasinghe and H. Kalutarage, 'Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation', *Soft Comput*, vol. 25, no. 3, pp. 2277–2293, Feb. 2021, doi: 10.1007/s00500-020-05297-6.

[36] W. A. Prabowo and C. Wiguna, 'Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM', *mib*, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.

[37] O. Baines, 'Naïve Bayes: Machine Learning and Text Classification Application of Bayes' Theorem'.

[38] J. Suzuki, 'Support Vector Machine', in *Statistical Learning with Math and R: 100 Exercises for Building Logic*, Singapore: Springer Nature Singapore, 2020, pp. 171–192. doi: 10.1007/978-981-15-7568-6_9.

[39] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, 'Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM', *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, 2023, doi: 10.57152/malcom.v3i2.897.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LoA Paper



SISTEMASI: JURNAL SISTEM INFORMASI
 PROGRAM STUDI SISTEM INFORMASI
 FAKULTAS TEKNIK DAN ILMU KOMPUTER
 UNIVERSITAS ISLAM INDRAGIRI

Letter of Acceptance

21 November 2024

Dear, Novita Sari

Congratulations, We are pleased to inform you that your following manuscript has been accepted and will be published in SISTEMASI, Vol 14 No. 1, January, 2025 pISSN: 2302-8149 eISSN: 2540-9719.

Title : **Penerapan Algoritma Klasifikasi Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Cyberbullying Bilingual di Aplikasi X**

: **Application of Naive Bayes Classification Algoritma and Support Vector Machine for Bilingual Cyberbullying Sentimen Analysis in X App**

Authors : Novita Sari, Muhammad Jazman, Tengku Khairil Ahsyar & Syaifullah

Email : 12050322233@students.uin-suska.ac.id

Received on : 12 November 2024

Revised on : 18 November 2024

Accepted on : 20 November 2024

Thank you very much for submitting your article to "SISTEMASI"



Best Regards,

Prof. Dr. Abdullah Husin, S.Si., M.kom
 Chief Editor



UIN SUSKA RIAU

Gambar A.1. LoA



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SERTIFIKAT JURNAL



Gambar B.1. Sertifikat Sinta Jurnal

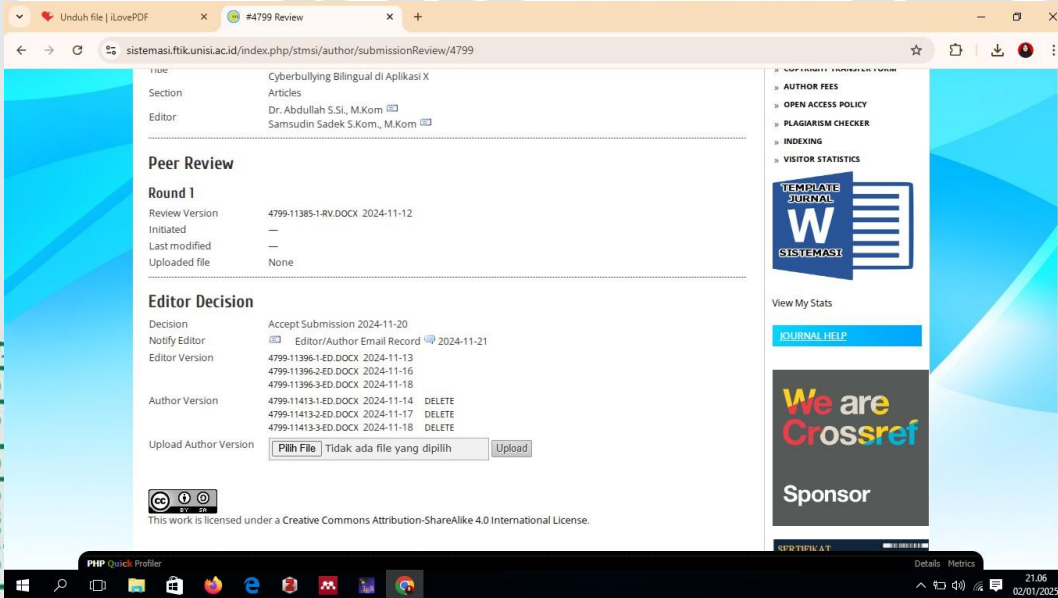


PROSES SUBMIT SAMPAI ACCEPTED

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar C.1. Submit Paper Aktif



Gambar C.2. Upload Revisi Paper

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

12/15/24, 5:26 AM

Email Universitas Islam Negeri Sultan Syarif Kasim Riau - [stmsi] Editor Decision



Novita Sari Sistem Informasi <12050322233@students.uin-suska.ac.id>

[stmsi] Editor Decision

1 pesan

Dr. Abdullah S.Si., M.Kom <sistemasi.ftik@gmail.com>
Kepada: Novita Sari <12050322233@students.uin-suska.ac.id>

13 November 2024 pukul 13.08

Novita Sari:

We have reached a decision regarding your submission to Sistemasi: Jurnal Sistem Informasi, "Penerapan Algoritma Klasifikasi Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Cyberbullying Bilingual di Aplikasi X".

Our decision is: Revisions Required

Dr. Abdullah S.Si., M.Kom
Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer,
Universitas Islam Indragiri, Indonesia
abdialam@gmail.com
Sistemasi: Jurnal Sistem Informasi

<https://mail.google.com/mail/u/0/?ik=b90d1530f2&view=pt&search=all&permthid=thread-f:1815586383297885288&simpl=msg-f:18155863832978...> 1/1

Gambar C.3. Email Revisi dari Editor

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

12/15/24, 5:21 AM

Email Universitas Islam Negeri Sultan Syarif Kasim Riau - [stmsi] Editor Decision



Novita Sari Sistem Informasi <12050322233@students.uin-suska.ac.id>

[stmsi] Editor Decision

1 pesan

Dr. Abdullah S.Si., M.Kom <sistemasi.ftik@gmail.com>
Kepada: Novita Sari <12050322233@students.uin-suska.ac.id>

20 November 2024 pukul 10.03

Yth. Sdri. Novita Sari:

Berkenaan dengan submission manuskrip ke Sistemasi: Jurnal Sistem Informasi, "Penerapan Algoritma Klasifikasi Naive Bayes dan Support Vector Machine untuk Analisis Sentimen Cyberbullying Bilingual di Aplikasi X".

Keputusan Editor: Diterima (Accepted)

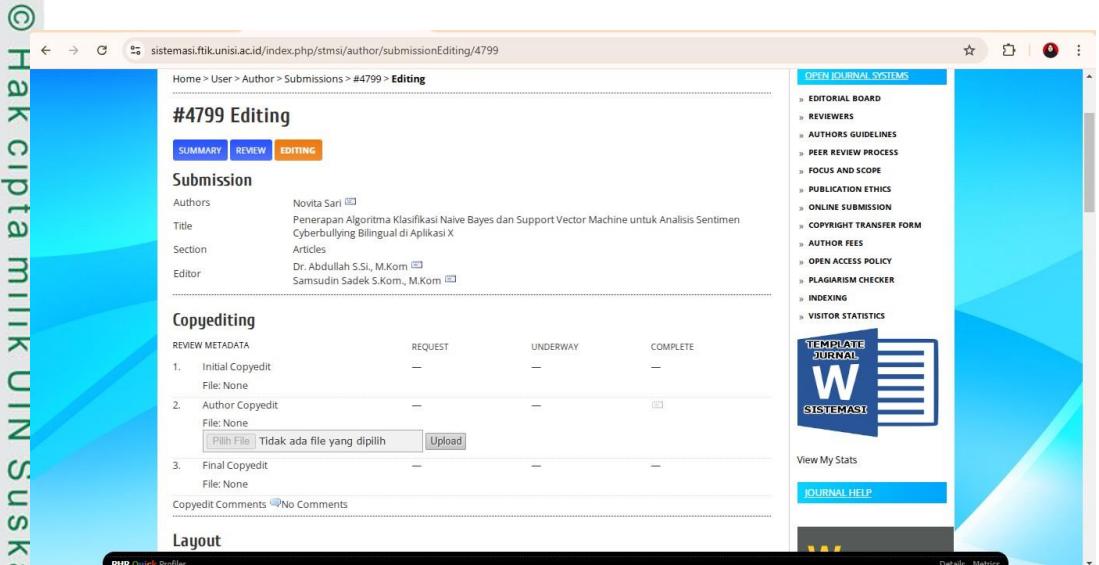
1. Agar mengisi (dengan rapi) Etika publikasi dan Copy Right Transfer (CTA) form (dapat di download pada web), dan mengirimkan kepada kami ke email: sistemasi.ftik@gmail.com
Pernyataan Etika Publikasi SISTEMASI.doc
https://drive.google.com/open?id=17xWEcXdbUd7QRVPS9b_nlzcqvmUPPo30
2. Silakan melakukan pengecekan similarity (Turnitin) secara mandiri maks 20% dan melaporkan dengan cara upload supplement file pada web atau mengirim ke kami email: sistemasi.ftik@gmail.com.
Jika ada perubahan naskah dapat diupload kembali pada tahap editing.
3. Artikel akan diterbitkan pada Vol 14 No 1 bulan Januari 2025
Donasi transfer ke rek BSI 7212506006 An. ABDULLAH
Terimakasih.
4. LoA akan dikirim jika ada permintaan dan telah memenuhi persyaratan.

Regards,

Dr. Abdullah Husin.
Chief Editor
Sistemasi: Jurnal Sistem Informasi

<https://mail.google.com/mail/u/0/?ik=b90d1530f2&view=pt&search=all&permthid=thread-f:1816208921099382428&simpl=msg-f:18162089210993...> 1/1

Gambar C.4. Email Accepted dari Editor



Gambar C.5. Editing Paper

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR RIWAYAT HIDUP



© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim I

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Peneliti bernama Novita Sari lahir di Mandah, Kabupaten Indragiri Hilir, Provinsi Riau pada tanggal 21 Februari 2000. Peneliti anak Pertama dari 2 bersaudara dari Ibu Suryani dan Bapak Abdul Murat. Peneliti Bertempat tinggal di Kelurahan Pangkalan Kerinci Kota Kecamatan Pangkalan Kerinci, Provinsi Riau. Pada tahun 2007 peneliti memulai pendidikan di SDN 016 Suak Apung Penjuru, Indragiri Hilir sampai tahun 2013. Selanjutnya peneliti melanjutkan pendidikan di SMPN 3 Kempas Jaya, Indragiri Hilir pada tahun 2013 sampai 2016. Setelah tamat dari tingkatan sebelumnya, peneliti melanjutkan pendidikan di SMAN Dharma Pendidikan Kempas Jaya, Indragiri Hilir dan lulus pada tahun 2019. Setelah menganggur selama 1 tahun, peneliti melanjutkan pendidikan Strata Satu (S1) di Universitas Islam Negeri Sultan Syarif Kasim Riau, tepatnya Fakultas Sains dan Teknologi, Program Studi Sistem Informasi pada Tahun 2020. Pada tahun 2022 peneliti melaksanakan Kerja Prakter (KP) di PT. RAPP Timas Suplindo Pangkalan Kerinci. pada tahun 2023 peneliti melaksanakan Kuliah Kerja Nyata (KKN) di Desa Telaga Sam-Sam, Kecamatan Kandis, Kabupaten Siak, Provinsi Riau selama kurang lebih 2 bulan. Terkait dengan pertanyaan kepada peneliti tentang yang dikerjakan dapat menghubungi kontak melalui *e-mail* 12050322233@students.uin-suska.ac.id untuk menjalin komunikasi yang lebih baik.

UIN SUSKA RIAU