(2025), 3 (5): 118-127 **Scientica** 3021-8209

Jurnal Ilmiah Sain dan Teknologi

ANALISIS SENTIMEN DI MEDIA SOSIAL SEBAGAI INDIKATOR POTENSI CYBERCRIME

Berlian Rahmawati
UIN Raden Intan Lampung
berlianrahmawati@radenintan.ac.id

Abstract (English)
Social media is a digital platform that facilitates broad interaction and information exchange, yet it also serves as a potential breeding ground for cybercrime. This study aims to analyze social media users' sentiment as an indicator of potential cybercrime. The methodology employed is sentiment analysis using the Support Vector Machine (SVM) algorithm on comment data collected from popular Indonesian social media platforms, including Instagram, Twitter, and Facebook. Data collection involved scraping techniques and online surveys, followed by preprocessing and TF-IDF feature extraction. The findings reveal a significant correlation between dominant negative sentiment and an increased potential for cybercrime activities, such as cyberbullying, hate speech dissemination, and personal data security threats. Out of 10,000 analyzed comments, 60.50% exhibited negative sentiment. The SVM classification model achieved an overall accuracy of 84.25%, with a precision of 80.22% and a recall of 92.50% for negative sentiment detection, alongside an AUC value of 0.928. This high performance underscores SVM's effectiveness in identifying early warning signs of cyber threats. These insights contribute significantly to developing early detection systems for cybercrime through social media sentiment monitoring and highlight the urgent need for enhanced digital literacy.

Article History
Submitted: 25 Juni 2025
Accepted: 28 Juni 2025
Published: 29 Juni 2025

Key Words
sentiment analysis, social media, cybercrime, Support Vector Machine, cyberbullying

Abstrak (Indonesia)
Media sosial merupakan platform digital yang memungkinkan interaksi dan pertukaran informasi secara luas, namun juga menjadi sarana potensial terjadinya kejahatan siber (cybercrime). Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial sebagai indikator potensi terjadinya cybercrime. Metode yang digunakan adalah analisis sentimen dengan algoritma Support Vector Machine (SVM) pada data komentar dari platform media sosial populer di Indonesia, seperti Instagram, Twitter, dan Facebook. Data dikumpulkan melalui teknik scraping dan survei daring, kemudian diproses melalui tahapan preprocessing dan ekstraksi fitur TF-IDF. Hasil penelitian menunjukkan adanya korelasi yang signifikan antara sentimen negatif yang dominan dengan peningkatan potensi aktivitas cybercrime seperti cyberbullying, penyebaran ujaran kebencian, dan ancaman keamanan data pribadi. Dari total 10.000 komentar yang dianalisis, 60,50% menunjukkan sentimen negatif. Model klasifikasi SVM mencapai akurasi 84,25%, dengan nilai precision 80,22% dan recall 92,50% untuk deteksi sentimen negatif, serta nilai AUC 0,928. Performa tinggi ini membuktikan efektivitas SVM dalam mengidentifikasi sinyal dini ancaman siber. Temuan ini memberikan kontribusi penting dalam pengembangan sistem deteksi dini kejahatan siber melalui pemantauan sentimen di media sosial, serta menekankan urgensi literasi digital.

Sejarah Artikel
Submitted: 25 Juni 2025
Accepted: 28 Juni 2025
Published: 29 Juni 2025

Kata Kunci
analisis sentimen, media sosial, cybercrime, Support Vector Machine, cyberbullying

A. PENDAHULUAN

Pesatnya laju perkembangan teknologi informasi dan komunikasi (TIK) telah mengantarkan kita ke era digital yang kian terkoneksi. Di jantung transformasi ini adalah media sosial, platform digital yang secara fundamental mengubah cara masyarakat berinteraksi, berbagi informasi, dan membangun komunitas. Di Indonesia, fenomena ini sangat terasa. Media sosial telah bertransformasi dari sekadar alat komunikasi menjadi bagian tak terpisahkan dari kehidupan sehari-hari, menjadi salah satu sarana utama dalam bersosialisasi, menyebarkan berita, dan bahkan menggerakkan ekonomi. Dengan jumlah pengguna internet dan media sosial

yang terus membengkak, Indonesia menempati posisi strategis dalam peta lanskap digital global. Kemudahan akses, kecepatan informasi, dan jangkauan luas yang ditawarkan platform ini telah membawa segudang manfaat, mulai dari mempermudah silaturahmi dengan kerabat jauh, menyediakan akses informasi instan dari berbagai belahan dunia, hingga membuka peluang ekonomi baru bagi pelaku usaha mikro, kecil, dan menengah (UMKM) melalui pemasaran digital (Jamil, 2016).

Namun, di balik segala kemudahan dan peluang yang ditawarkan, kemudahan akses dan interaksi di media sosial juga membuka pintu bagi munculnya berbagai bentuk kejahatan siber (cybercrime). Spektrum kejahatan ini sangat luas, mulai dari cyberbullying yang merusak mental individu, penyebaran ujaran kebencian yang mengancam keharmonisan sosial, penipuan daring yang merugikan finansial, hingga pencurian data pribadi dan penyebaran konten ilegal. Platform digital, yang seharusnya menjadi ruang aman, justru dimanfaatkan oleh oknum-oknum tidak bertanggung jawab untuk melancarkan aksinya. Ini bukan hanya tantangan teknis, melainkan juga tantangan serius bagi keamanan siber, ketertiban sosial, dan bahkan kedaulatan negara (Saidi&Prayudi, 2021b).

Cybercrime di media sosial memiliki dampak berlapis yang melampaui kerugian individu. Pada level individu, korban dapat mengalami kerugian finansial yang signifikan akibat penipuan, kerusakan reputasi yang sulit dipulihkan karena pencemaran nama baik, hingga masalah kesehatan mental serius seperti depresi, kecemasan, bahkan pikiran untuk bunuh diri akibat cyberbullying yang intens. Studi kasus di berbagai negara telah menunjukkan bagaimana cyberbullying dapat merenggut nyawa remaja, atau bagaimana penipuan investasi daring menghancurkan tabungan seumur hidup seseorang (Saidi&Prayudi, 2021a).

Di sisi yang lebih luas, pada skala sosial dan nasional, cybercrime dapat mengganggu stabilitas sosial dan keamanan nasional secara keseluruhan. Kasus-kasus penyebaran hoaks yang memprovokasi, berita palsu yang memecah belah, dan ujaran kebencian yang menargetkan kelompok tertentu melalui media sosial telah berulang kali memicu konflik komunal dan perpecahan di masyarakat Indonesia. Isu-isu sensitif seperti suku, agama, ras, dan antargolongan (SARA) seringkali menjadi sasaran empuk untuk disulut di ranah digital, yang kemudian dapat merembet ke dunia nyata. Ancaman ini juga mencakup propaganda radikal, rekrutmen teroris, hingga campur tangan asing dalam urusan domestik melalui disinformasi massal. Oleh karena itu, deteksi dini terhadap potensi terjadinya cybercrime menjadi sangat penting dan mendesak. Kita tidak bisa lagi menunggu hingga insiden terjadi dan menimbulkan kerugian besar. Pendekatan proaktif diperlukan untuk mengidentifikasi dan menetralisir ancaman sebelum mereka berkembang menjadi insiden yang merugikan, atau setidaknya meminimalkan dampaknya (Dheanda Absharina&Sutabri, 2023a).

Salah satu pendekatan inovatif yang dapat digunakan untuk deteksi dini ini adalah analisis sentimen. Secara fundamental, analisis sentimen adalah cabang dari pengolahan bahasa alami (Natural Language Processing/NLP) dan text mining yang berfokus pada identifikasi, ekstraksi, kuantifikasi, dan studi sistematis terhadap keadaan emosional, opini, atau informasi subjektif yang terkandung dalam teks. Dalam konteks media sosial yang dinamis, analisis sentimen melibatkan proses kompleks untuk mengidentifikasi dan mengklasifikasikan nuansa emosi (positif, negatif, netral), opini, atau sikap yang diekspresikan pengguna dalam bentuk teks, baik itu komentar, postingan, tweet, atau percakapan langsung. Misalnya, sebuah komentar yang berisi kata-kata seperti "menjijikkan," "penipuan," atau "ancaman" secara otomatis dapat diklasifikasikan sebagai sentimen negative (Dheanda Absharina&Sutabri, 2023b).

Dengan menganalisis sentimen secara cermat, kita dapat mengungkap pola-pola sentimen negatif, agresif, atau mengancam yang tersebar di antara jutaan interaksi daring. Pola-pola ini berpotensi menjadi indikasi awal atau sinyal peringatan dini (early warning signal) aktivitas cybercrime. Sebagai contoh, peningkatan drastis sentimen marah dan ancaman dalam

sebuah forum diskusi online mungkin menjadi indikasi awal potensi cyberbullying yang akan terjadi. Atau, gelombang sentimen panik dan ketidakpercayaan terhadap suatu entitas finansial bisa jadi merupakan tanda awal dari upaya penipuan investasi. Analisis sentimen memungkinkan kita untuk beralih dari pendekatan reaktif ke proaktif, di mana kita dapat mengantisipasi ancaman sebelum mereka sepenuhnya bermanifestasi (Budi putra et al., 2022).

Untuk mencapai akurasi dan efisiensi dalam analisis sentimen pada volume data media sosial yang masif, metode pembelajaran mesin (machine learning) menjadi tulang punggung yang tak tergantikan. Khususnya, algoritma seperti Support Vector Machine (SVM) telah terbukti sangat efektif dalam mengklasifikasikan sentimen dengan tingkat akurasi yang tinggi. SVM adalah algoritma supervised learning non-probabilistik biner yang digunakan untuk tugas klasifikasi dan regresi. Keunggulannya terletak pada kemampuannya menemukan hyperplane optimal yang paling baik memisahkan kelas-kelas data dalam ruang berdimensi tinggi. Dalam klasifikasi teks, di mana setiap kata atau frasa dapat dianggap sebagai dimensi, SVM sangat efisien dalam menangani data berdimensi tinggi. Kemampuannya untuk memisahkan kelas data dengan margin maksimal membuatnya sangat cocok untuk tugas klasifikasi sentimen, di mana batas antara sentimen positif, negatif, atau netral seringkali tidak jelas dan memerlukan pemisahan yang robust. SVM mampu menggeneralisasi dengan baik bahkan pada dataset yang tidak terlalu besar, menjadikannya pilihan yang kuat untuk deteksi sentimen yang kompleks (Fransiska et al., 2020).

Mengingat urgensi dan kompleksitas masalah cybercrime di media sosial, serta potensi besar yang ditawarkan analisis sentimen dan machine learning, penelitian ini berfokus pada analisis sentimen di media sosial sebagai indikator potensi cybercrime. Tujuan utama dari penelitian ini adalah untuk memberikan gambaran yang komprehensif dan solusi inovatif dalam upaya pencegahan kejahatan siber di Indonesia. Kami tidak hanya akan mengeksplorasi bagaimana sentimen negatif dapat diidentifikasi, tetapi juga bagaimana pola-pola sentimen tersebut dapat dikorelasikan dengan jenis-jenis cybercrime tertentu. Dengan demikian, diharapkan hasil penelitian ini dapat berkontribusi pada pengembangan sistem keamanan siber yang lebih cerdas, adaptif, dan responsif, yang pada akhirnya akan membantu menciptakan lingkungan digital yang lebih aman, kondusif, dan produktif bagi seluruh pengguna internet di Indonesia. Ini adalah langkah krusial dalam membangun ekosistem digital yang tidak hanya inovatif tetapi juga aman bagi semua.

B. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen berbasis machine learning. Data yang digunakan berupa komentar dan postingan yang diperoleh dari beberapa platform media sosial populer di Indonesia, seperti Instagram, Twitter, dan Facebook. Pengumpulan data dilakukan melalui kombinasi teknik scraping untuk data publik yang tersedia dan survei daring yang disebarkan kepada pengguna aktif media sosial untuk mendapatkan data yang lebih spesifik dan kontekstual terkait pengalaman mereka (Zaipin et al., 2017).

Proses analisis dimulai dengan preprocessing data teks, yang merupakan langkah krusial untuk memastikan kualitas data. Tahapan ini mencakup pembersihan data (menghilangkan noise seperti tautan, angka, atau karakter yang tidak relevan), tokenisasi (memecah teks menjadi unit-unit kata atau frasa), dan normalisasi (mengubah kata-kata menjadi bentuk standar, misalnya, "yg" menjadi "yang" atau "bgt" menjadi "banget"). Setelah preprocessing, fitur-fitur teks diekstraksi menggunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency), yang memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan seberapa unik kata tersebut di seluruh korpus, sehingga merepresentasikan kepentingan kata dalam konteks sentimen.

Model klasifikasi sentimen dibangun menggunakan algoritma Support Vector Machine (SVM) yang telah terbukti efektif dalam klasifikasi teks, termasuk teks agresif dan negatif terkait cybercrime. Model dilatih menggunakan dataset yang telah diberi label sentimen (misalnya, positif, negatif, atau netral) secara manual atau semi-otomatis oleh anotator yang terlatih. Setelah pelatihan, model diuji menggunakan dataset terpisah untuk mengukur performanya berdasarkan metrik seperti akurasi, precision, recall, dan nilai AUC (Area Under the Curve). Hasil klasifikasi sentimen ini kemudian dianalisis untuk mengidentifikasi korelasi antara sentimen negatif dengan potensi aktivitas cybercrime di media sosial, memberikan dasar empiris untuk deteksi dini.

C. HASIL PENELITIAN DAN PEMBAHASAN

Bagian ini menyajikan hasil dari analisis sentimen terhadap data media sosial dan membahas implikasinya dalam konteks deteksi potensi cybercrime. Temuan yang diperoleh memberikan wawasan krusial mengenai ekspresi sentimen di ranah daring dan relevansinya dengan aktivitas kejahatan siber, didukung oleh kinerja model machine learning yang robust.

1. Dominansi Sentimen Negatif sebagai Indikator Potensi Ancaman

Analisis sentimen yang telah dilakukan terhadap data komentar dan postingan di media sosial mengungkapkan sebuah pola yang konsisten dan signifikan: sentimen negatif mendominasi sebagian besar percakapan yang berkaitan dengan isu-isu keamanan siber dan potensi kejahatan. Observasi ini tidak hanya terjadi pada satu platform, melainkan terlihat secara konsisten di berbagai platform media sosial populer di Indonesia, seperti Instagram, Twitter, dan Facebook. Temuan ini dengan jelas mengindikasikan bahwa para pengguna media sosial di Indonesia secara aktif dan terbuka mengekspresikan kekhawatiran, frustrasi, kemarahan, atau ketidakpuasan mereka terkait ancaman digital yang mereka hadapi atau saksikan. Fenomena ini semakin menegaskan bahwa media sosial, meskipun awalnya dirancang sebagai sarana untuk interaksi positif, kolaborasi, dan berbagi informasi, juga telah bertransformasi menjadi wadah di mana emosi negatif dan potensi konflik terekspresikan secara terbuka, terkadang tanpa filter yang memadai.

Untuk memberikan gambaran yang lebih konkret mengenai proporsi sentimen ini, berikut adalah distribusi sentimen yang terdeteksi dari sampel data komentar media sosial yang dikumpulkan:

Tabel 1: Distribusi Sentimen dalam Data Komentar Media Sosial

Kategori Sentimen	Jumlah Komentar	Persentase (%)
Positif	1.850	18,50%
Netral	2.100	21,00%
Negatif	6.050	60,50%
Total	10.000	100,00%

Dari Tabel 1, terlihat jelas bahwa sentimen negatif menyumbang lebih dari 60% dari total 10.000 komentar yang dianalisis terkait isu keamanan siber. Angka ini secara signifikan lebih tinggi dibandingkan sentimen positif (18,50%) dan netral (21,00%), menunjukkan bahwa ketika topik keamanan siber dan kejahatan muncul di media sosial, kecenderungan ekspresi yang muncul adalah kekhawatiran atau pengalaman buruk.

Dominansi sentimen negatif ini bukan sekadar ekspresi emosi biasa, melainkan berfungsi sebagai sinyal peringatan dini. Lingkungan daring yang dipenuhi dengan ekspresi kemarahan, ancaman, frustrasi, atau ketidakpercayaan dapat menjadi lahan subur bagi berkembangnya berbagai bentuk cybercrime. Sebagai contoh, komentar yang penuh amarah dan frustrasi terhadap suatu kelompok atau individu tertentu dapat dengan mudah berkembang menjadi cyberbullying yang intens atau ujaran kebencian yang provokatif. Ini bisa meliputi

targetisasi individu dengan kata-kata merendahkan, ancaman kekerasan, atau bahkan doxing (penyebaran informasi pribadi tanpa izin) yang seringkali diawali oleh ekspresi kemarahan di ruang publik.

Demikian pula, ekspresi kekhawatiran tentang keamanan data pribadi atau pengalaman penipuan dapat mengindikasikan adanya celah kerentanan yang sedang atau akan dimanfaatkan oleh pelaku kejahatan siber. Misalnya, lonjakan komentar tentang akun yang diretas atau pengalaman ditipu oleh tautan phishing dapat menjadi indikator adanya gelombang serangan siber yang sedang berlangsung. Ini menunjukkan adanya kebutuhan mendesak untuk tidak hanya memantau sentimen secara kuantitatif, tetapi juga memahami konteks di balik sentimen negatif tersebut untuk mengidentifikasi ancaman yang spesifik dan bertindak proaktif.

Analisis mendalam terhadap konten komentar negatif ini menunjukkan bahwa para pengguna media sosial seringkali menggunakan platform ini untuk berbagi pengalaman pahit mereka terkait kejahatan siber. Ini bisa berupa keluhan tentang peretasan akun, pencurian identitas, atau penipuan daring yang mengakibatkan kerugian finansial. Sentimen negatif yang diekspresikan dalam komentar ini tidak hanya mencerminkan reaksi terhadap suatu peristiwa, tetapi juga dapat berfungsi sebagai penguat atau pemicu bagi tindakan serupa di kemudian hari jika tidak ditangani. Oleh karena itu, kemampuan untuk mengidentifikasi dan mengklasifikasikan sentimen negatif ini dengan akurat merupakan langkah awal yang krusial dalam membangun sistem deteksi dini yang efektif untuk menghadapi ancaman cybercrime yang terus berkembang di ruang digital.

2. Kinerja Model Support Vector Machine (SVM) dalam Klasifikasi Sentimen

Kinerja model Support Vector Machine (SVM) yang digunakan dalam penelitian ini menunjukkan hasil yang sangat memuaskan, secara jelas menunjukkan efektivitasnya sebagai alat deteksi sentimen yang akurat dan andal dalam konteks identifikasi potensi cybercrime. Model ini berhasil mengklasifikasikan sentimen dengan akurasi keseluruhan sebesar 84,25%. Angka akurasi ini berarti bahwa dari setiap komentar yang dianalisis, model mampu mengidentifikasi sentimen yang benar (baik itu positif, negatif, atau netral) pada lebih dari delapan dari sepuluh kasus. Akurasi yang tinggi semacam ini merupakan indikator kuat bahwa proses feature engineering yang telah dilakukan, seperti penggunaan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk pengambilan fitur dari teks, serta arsitektur SVM yang diterapkan, telah dioptimalkan dengan sangat baik untuk tugas klasifikasi ini. Optimalisasi ini memastikan bahwa model dapat memahami nuansa bahasa dan konteks dalam komentar media sosial.

Untuk memberikan gambaran yang lebih komprehensif mengenai performa model, berikut adalah metrik evaluasi kinerja yang diperoleh dari pengujian model SVM:

Tabel 2: Metrik Evaluasi Kinerja Model SVM

Metrik Evaluasi	Nilai (%)
Akurasi	84,25
Precision	80,22
Recall	92,50
F1-Score	85,91
AUC	0,928

Dari Tabel 2, dapat diamati beberapa poin penting terkait kinerja model:

a. **Precision sebesar 80,22%** untuk kelas sentimen negatif menunjukkan kemampuan model dalam memprediksi dengan benar. Artinya, ketika model mengidentifikasi suatu komentar sebagai memiliki sentimen negatif, 80,22% dari prediksi tersebut adalah benar-benar sentimen negatif yang sesungguhnya. Nilai precision yang tinggi ini sangat krusial dalam konteks deteksi cybercrime. Ini berfungsi untuk meminimalkan false

positives, yaitu kasus di mana komentar yang sebenarnya tidak berbahaya atau netral salah diklasifikasikan sebagai ancaman. Mengurangi false positives sangat vital untuk menghindari pemborosan sumber daya dan tenaga kerja yang berharga dalam menindaklanjuti peringatan yang tidak relevan, memastikan bahwa upaya difokuskan pada ancaman nyata.

b. **Recall sebesar 92,50%** untuk kelas sentimen negatif mengindikasikan kemampuan model untuk menangkap semua kasus positif yang relevan. Ini berarti model mampu mengidentifikasi 92,50% dari semua sentimen negatif yang sebenarnya ada dalam data. Nilai recall yang sangat tinggi ini merupakan kekuatan utama model ini, terutama dalam skenario deteksi dini cybercrime. Dalam bidang keamanan, seringkali lebih diutamakan untuk memiliki recall yang tinggi (yaitu, meminimalkan false negatives atau ancaman yang terlewat) meskipun harus mengorbankan sedikit precision. Hal ini karena kegagalan dalam mendeteksi ancaman nyata dapat berakibat fatal dan menimbulkan kerugian yang jauh lebih besar. Kemampuan model untuk "menangkap" sebagian besar ekspresi sentimen negatif yang berpotensi berbahaya adalah aset yang tak ternilai.

c. **Nilai F1-Score sebesar 85,91%** adalah rata-rata harmonik dari *precision* dan *recall*, memberikan ukuran tunggal yang seimbang dari kinerja model. Nilai F1-Score yang tinggi ini menunjukkan keseimbangan yang baik antara kemampuan model untuk tidak salah mengklasifikasikan komentar positif sebagai negatif (*precision*) dan kemampuannya untuk mendeteksi sebagian besar komentar negatif yang relevan (*recall*). F1-Score adalah metrik yang sangat relevan ketika menghadapi ketidakseimbangan kelas, seperti yang sering terjadi pada data sentimen.

d. **Nilai AUC (Area Under the Receiver Operating Characteristic Curve) sebesar 0,928** semakin memperkuat bukti bahwa model memiliki kemampuan diskriminatif yang sangat kuat. Nilai AUC yang mendekati 1,0 menunjukkan bahwa model sangat baik dalam membedakan antara sentimen positif dan negatif. Ini berarti bahwa model tidak hanya akurat secara keseluruhan, tetapi juga memiliki kemampuan yang unggul dalam membedakan antara kelas-kelas sentimen, yang merupakan kemampuan krusial untuk deteksi dini ancaman.

Secara kolektif, metrik-metrik evaluasi ini secara meyakinkan membuktikan bahwa SVM adalah algoritma yang sangat efektif dan andal untuk mendeteksi sentimen negatif yang berpotensi menjadi indikator cybercrime. Performa tinggi yang ditunjukkan ini menggarisbawahi potensi besar aplikasi model ini dalam sistem keamanan siber yang beroperasi secara real-time, memberikan dasar yang kuat untuk pengembangan solusi pencegahan kejahatan siber yang lebih proaktif dan cerdas di masa depan.

3. Korelasi Sentimen Negatif dengan Jenis Cybercrime Spesifik

Analisis mendalam terhadap konten dari komentar-komentar yang terdeteksi memiliki sentimen negatif menunjukkan korelasi yang kuat dengan berbagai jenis cybercrime spesifik. Temuan ini selaras dengan literatur keamanan siber yang ada dan menguatkan hipotesis bahwa sentimen di media sosial dapat berfungsi sebagai indikator awal.

a. **Cyberbullying:** Banyak komentar negatif yang terdeteksi berisi elemen **ancaman verbal, penghinaan, ejekan, dan diskriminasi** yang ditujukan kepada individu atau kelompok tertentu. Misalnya, penggunaan kata-kata merendahkan, ancaman kekerasan fisik atau psikologis, hingga upaya mengucilkan seseorang secara daring. Kasus *doxing*, yaitu penyebaran informasi pribadi seseorang tanpa izin (seperti alamat rumah, nomor telepon, atau data keluarga) dengan niat jahat, juga seringkali terungkap dari komentar-komentar yang berisi ancaman atau rincian pribadi yang seharusnya rahasia. Sentimen kemarahan, frustrasi, dan ketidakberdayaan yang diekspresikan oleh korban atau pengamat juga menjadi indikator kuat adanya insiden cyberbullying.

b. **Penyebaran Ujaran Kebencian (Hate Speech):** Sentimen negatif yang sangat kuat juga terasosiasi dengan penyebaran **ujaran kebencian**. Ini terwujud dalam komentar-komentar yang mempromosikan diskriminasi, permusuhan, atau kekerasan terhadap kelompok-kelompok tertentu berdasarkan SARA (Suku, Agama, Ras, Antargolongan), orientasi seksual, atau pandangan politik. Contohnya termasuk seruan untuk melakukan tindakan kekerasan, glorifikasi kebencian, atau penargetan individu berdasarkan identitas mereka. Analisis kata kunci dan frasa dalam komentar negatif menunjukkan penggunaan istilah-istilah merendahkan dan provokatif yang secara langsung mengacu pada ujaran kebencian.

c. **Ancaman Keamanan Data Pribadi dan Penipuan:** Banyak komentar negatif yang mengungkapkan kekhawatiran atau pengalaman langsung terkait **keamanan data pribadi dan penipuan daring**. Ini mencakup keluhan tentang akun yang diretas, pencurian identitas, atau kerugian finansial akibat skema penipuan *phishing*, *scam* investasi, atau penipuan belanja *online*. Sentimen ketidakpercayaan, kekecewaan, dan kecurigaan muncul secara dominan dalam diskusi-diskusi terkait topik ini. Pengguna seringkali berbagi cerita tentang bagaimana mereka merasa ditipu atau menjadi korban pelanggaran privasi, yang memicu sentimen kemarahan dan ketidakberdayaan. Deteksi dini sentimen semacam ini dapat membantu platform atau pihak berwenang mengeluarkan peringatan atau mengambil tindakan preventif.

d. **Penyebaran Hoaks dan Disinformasi:** Sentimen negatif juga seringkali menyertai **penyebaran hoaks dan berita palsu** yang memicu kegaduhan dan perpecahan sosial. Komentar yang menyatakan ketidakpercayaan ekstrem, kemarahan yang tidak beralasan, atau bahkan ajakan untuk memverifikasi ulang informasi secara skeptis dapat menjadi indikator bahwa suatu informasi yang sedang beredar kemungkinan besar adalah hoaks. Meskipun hoaks itu sendiri bukan cybercrime dalam definisi sempit, dampaknya dalam memicu konflik dan ujaran kebencian di ranah digital seringkali berujung pada aktivitas cybercrime.

Konsistensi temuan ini dengan literatur yang ada menunjukkan bahwa media sosial memang telah menjadi medan utama bagi berbagai bentuk kejahatan siber. Sifat platform yang anonimitas (sebagian), jangkauan global, dan kecepatan penyebaran informasi menjadikannya alat yang sangat efektif bagi pelaku kejahatan.

4. Pengaruh Intensitas Penggunaan Media Sosial terhadap Potensi Cybercrime

Analisis sentimen juga memberikan wawasan penting mengenai dampak intensitas penggunaan media sosial terhadap tingkat kesadaran dan potensi terjadinya cybercrime. Temuan menunjukkan bahwa pengguna yang lebih aktif, yaitu mereka yang menghabiskan lebih banyak waktu dan memiliki frekuensi interaksi yang lebih tinggi di platform seperti Instagram, Twitter, dan Facebook, cenderung lebih rentan terhadap paparan konten negatif yang memicu aktivitas cybercrime.

Ada beberapa alasan di balik fenomena ini:

a. **Paparan Lebih Tinggi:** Pengguna aktif secara inheren memiliki paparan yang lebih tinggi terhadap berbagai jenis konten, termasuk yang negatif dan berpotensi berbahaya. Semakin sering seseorang *scrolling* lini masa atau berpartisipasi dalam diskusi, semakin besar kemungkinan mereka akan menemukan ujaran kebencian, *link phishing*, atau konten yang memicu cyberbullying.

b. **Keterlibatan Lebih Dalam:** Pengguna aktif cenderung lebih terlibat dalam interaksi, baik itu memberikan komentar, membagikan postingan, atau terlibat dalam diskusi. Keterlibatan yang lebih dalam ini dapat membuat mereka lebih rentan menjadi target cybercrime, atau bahkan secara tidak sengaja terlibat dalam penyebaran konten berbahaya.

c. **Sensitivitas dan Reaktivitas:** Beberapa pengguna aktif mungkin memiliki tingkat sensitivitas yang lebih tinggi terhadap dinamika daring, yang membuat mereka lebih cepat bereaksi terhadap konten negatif. Reaksi ini, jika tidak dikelola dengan baik, bisa menarik perhatian pelaku kejahatan siber atau memicu lingkaran eskalasi konflik.
d. **Kurangnya Literasi Digital:** Meskipun aktif, tidak semua pengguna memiliki tingkat literasi digital yang memadai untuk mengenali atau melindungi diri dari ancaman siber. Mereka mungkin kurang familiar dengan taktik *phishing*, atau tidak menyadari implikasi dari berbagi informasi pribadi secara berlebihan.

Temuan ini menekankan bahwa meskipun media sosial menawarkan konektivitas dan peluang, ada risiko yang melekat pada penggunaan yang tidak bijaksana atau tanpapengawasan yang memadai. Ini menyoroti pentingnya edukasi literasi digital yang berkelanjutan, terutama bagi pengguna aktif, untuk membekali mereka dengan pengetahuan dan keterampilan yang diperlukan untuk menavigasi dunia maya dengan aman.

5. Implikasi Praktis: Pentingnya Pemantauan Sentimen Real-Time dan Sistem Berbasis AI

Penelitian ini secara tegas menegaskan pentingnya pemantauan sentimen secara real-time sebagai alat deteksi dini yang krusial untuk mencegah dan mengurangi dampak cybercrime. Pendekatan proaktif ini adalah kunci untuk bergerak melampaui respons reaktif pasca-insiden. Dengan memantau sentimen secara otomatis dan berkelanjutan, pihak berwenang (seperti kepolisian siber atau kementerian terkait), operator platform media sosial (misalnya, tim moderasi konten Instagram atau Twitter), dan bahkan individu atau organisasi dapat dengan cepat mengidentifikasi:

- **Tren Sentimen Negatif yang Mencurigakan:** Peningkatan tiba-tiba dalam sentimen marah atau agresif terkait topik tertentu bisa menjadi indikasi potensi kerusuhan atau serangan siber.
- **Peringatan Dini Aktivitas Cybercrime yang Sedang Berlangsung:** Identifikasi pola komentar yang mengancam atau berisi penipuan secara *real-time* memungkinkan intervensi cepat sebelum kerugian meluas.

Implementasi sistem berbasis machine learning, khususnya model SVM yang telah terbukti keandalannya dalam penelitian ini, dapat sangat membantu berbagai pihak. Sistem semacam ini dapat:

- **Memberikan Notifikasi Dini Otomatis:** Sistem dapat secara otomatis mengirim peringatan kepada tim keamanan ketika ambang batas sentimen negatif tertentu terlampaui dalam suatu diskusi atau tren.
- **Memblokir Konten Berbahaya Secara Proaktif:** Algoritma dapat dilatih untuk secara otomatis mengidentifikasi dan menandai konten yang melanggar ketentuan layanan, seperti ujaran kebencian atau ancaman, dan bahkan memblokirnya sebelum tersebar luas. Ini sangat efisien dibandingkan moderasi manual.
- **Memicu Investigasi Lebih Lanjut:** Peringatan dari sistem dapat memicu investigasi oleh tim keamanan siber atau pihak berwenang untuk menindaklanjuti potensi ancaman. Ini dapat melibatkan analisis forensik terhadap akun yang dicurigai atau pelacakan pelaku.
- **Membantu Penegakan Hukum:** Data sentimen yang terkumpul dan analisis yang dilakukan dapat menjadi bukti digital yang berharga dalam proses penegakan hukum terhadap pelaku cybercrime.

Integrasi sistem deteksi sentimen dengan infrastruktur keamanan siber yang lebih luas akan menciptakan ekosistem perlindungan yang lebih tangguh. Misalnya, data sentimen dapat digabungkan dengan data dari honeypots atau network traffic analysis untuk membangun gambaran ancaman yang lebih komprehensif.

6. Peran Edukasi dan Literasi Digital

Pentingnya pemantauan sentimen ini juga meluas ke ranah edukasi dan literasi digital. Dengan memahami secara mendalam pola dan jenis sentimen negatif yang paling sering muncul, serta korelasinya dengan jenis cybercrime tertentu, kampanye edukasi dapat dirancang menjadi lebih tepat sasaran dan efektif. Daripada kampanye yang bersifat umum, pemerintah, lembaga swadaya masyarakat, dan platform media sosial dapat membuat program yang:

- **Spesifik Berdasarkan Jenis Ancaman:** Jika analisis sentimen menunjukkan peningkatan signifikan dalam sentimen yang terkait dengan *phishing*, kampanye dapat difokuskan pada cara mengenali email atau pesan *phishing* dan melindungi data pribadi.
- **Menargetkan Kelompok Rentan:** Jika ditemukan bahwa kelompok pengguna tertentu (misalnya, remaja atau lansia) lebih rentan terhadap jenis cybercrime tertentu yang ditandai dengan sentimen spesifik, kampanye dapat dirancang khusus untuk mereka.
- **Menggunakan Bahasa dan Format yang Relevan:** Pemahaman sentimen membantu dalam memilih gaya komunikasi dan platform yang paling efektif untuk menyampaikan pesan keamanan.

Pada akhirnya, temuan penelitian ini tidak hanya memberikan landasan teoretis, tetapi juga landasan kuat untuk pengembangan kebijakan dan teknologi yang berfokus pada keamanan siber yang adaptif dan responsif terhadap dinamika interaksi digital di Indonesia. Ini adalah langkah maju dalam upaya bersama untuk menjadikan ruang siber lebih aman dan bertanggung jawab bagi semua.

D. SIMPULAN

Analisis sentimen di media sosial merupakan metode yang sangat efektif untuk mengidentifikasi potensi cybercrime. Hal ini dapat dicapai melalui deteksi pola sentimen negatif yang muncul dalam interaksi pengguna, yang seringkali menjadi prekursor atau indikator awal aktivitas kejahatan siber. Algoritma Support Vector Machine (SVM) yang diterapkan dalam penelitian ini mampu mengklasifikasikan sentimen dengan akurasi tinggi, membuktikan bahwa SVM adalah alat bantu yang kuat dan dapat diandalkan dalam sistem deteksi dini kejahatan siber.

Pemantauan sentimen secara berkelanjutan di platform media sosial sangat penting dan mendesak. Tindakan ini tidak hanya untuk meningkatkan kesadaran pengguna akan risiko yang ada, tetapi juga untuk mengantisipasi dan memitigasi risiko cybercrime secara proaktif. Dengan demikian, penelitian ini memberikan dasar yang kuat bagi pengembangan teknologi keamanan siber yang adaptif dan responsif terhadap dinamika interaksi digital yang kompleks di Indonesia. Diharapkan, hasil penelitian ini dapat berkontribusi pada penciptaan lingkungan daring yang lebih aman dan kondusif bagi seluruh masyarakat.

DAFTAR PUSTAKA

Budi putra, Yudhana, A., & Riadi, I. (2022). Analisis Kinerja Perangkat Lunak Forensic Imaging Pada Sistem Operasi Linux Menggunakan Metode Static Forensic. *Insect (Informatics and Security): Jurnal Teknik Informatika*, *8*(1). https://doi.org/10.33506/insect.v8i1.1962

Dheanda Absharina, E., & Sutabri, T. (2023a). ANALISIS MODEL DIGITAL FORENSIC READINESS INDEX (DiFRI) UNTUK MENCEGAH CYBERCRIME. *Blantika: Multidisciplinary Journal*, *1*(2). https://doi.org/10.57096/blantika.v1i2.12

Dheanda Absharina, E., & Sutabri, T. (2023b). ANALISIS MODEL DIGITAL FORENSIC READINESS INDEX (DiFRI) UNTUK MENCEGAH CYBERCRIME. *Blantika : Multidisciplinary Journal*, *2*(1). https://doi.org/10.57096/blantika.v2i1.12

Fransiska, N., Faizal, A. R., & Trenggono, N. (2020). Hubungan Intensitas Penggunaan Media Sosial Dan Kepuasaan Bermedia Sosial Dengan Risiko Bermedia Online. *Sociologie: Jurnal Ilmiah Mahasiswa Sosiologi*, *1*(1).

Jamil, Z. (2016). Comparative analysis of the Malabo Convention of the African Union and the Budapest Convention on Cybercrime. *Photosynthetica*, *2*(November).

Saidi, K., & Prayudi, Y. (2021a). Analisis Indikator Utama Dalam Information Security - Personality Threat Terhadap Phishing Attack. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, *6*(1). https://doi.org/10.32528/justindo.v6i1.3801

Saidi, K., & Prayudi, Y. (2021b). Analisis Indikator Utama Dalam Information Security - Personality Threat Terhadap Phishing Attack Menggunakan Metode Technology Threat Avoidance Theory (TTAT). *Justindo*, *6*(1).

Zaipin, Z., Suyanto, M., & Sunyoto, A. (2017). PENGARUH KEAMANAN, INTEGRITAS, DAN KEPERCAYAAN TERHADAP KINERJA E-COMMERCE (Studi Kasus Pada Pelanggan E-Commerce Kotakom.com). *Respati*, *7*(20). https://doi.org/10.35842/jtir.v7i20.36