

## Teknologi Informasi Efektif Mendeteksi Cyberbullying

Ajeng Ayu Kustianti<sup>1</sup>, Renta Sianturi<sup>2</sup>, Ameliya Sarwani Putri<sup>3</sup>, Anggita Putri Siswadi<sup>4</sup>, Delia Nuralita<sup>5</sup>, Elisa Puspitasari<sup>6</sup>.  
 Jurusan Keperawatan, Sekolah Tinggi Ilmu Kesehatan Mitra Keluarga

### ABSTRACT

Social media users are at risk for mental health disorders. Mental health problems can occur with cyberbullying. Cyberbullying that occurs on social media is in the form of rude comments, threats, insults, slander and even harassment given by netizens. Cyberbullying can shake a person's mental health condition and even have an impact on suicide. Cyberbullying will be very detrimental both mentally and productively. Cyberbullying must be detected early to prevent adverse effects on social media users. With advances in technology, it can be used to detect cyberbullying that occurs on social media. This article uses a literature review method approach, namely narrative literature review of 10 articles on the use of technology for cyberbullying detection in the period 2011 - 2021 with the aim of finding out cyberbullying comments on someone's account/post. Therefore, cyberbullying detection tries to collect global datasets on social media (Facebook, Instagram, Twitter, etc.), by classifying the Machine Learning method. Each algorithm method is evaluated using accuracy, precision, recall, and F1 score to determine the performance of the classification level.

### KEYWORDS

*Cyberbullying, Mental Health Disorders, Social media*

### ABSTRACT

Pengguna media sosial berisiko mengalami gangguan kesehatan mental. Masalah kesehatan mental dapat terjadi dengan cyberbullying. Cyberbullying yang terjadi di media sosial berupa komentar kasar, ancaman, hinaan, fitnah bahkan pelecehan yang diberikan oleh netizen. Cyberbullying dapat menggoyahkan kondisi kesehatan mental seseorang bahkan berdampak pada bunuh diri. Cyberbullying akan sangat merugikan baik secara mental maupun produktif. Cyberbullying harus dideteksi sejak dini untuk mencegah dampak buruk bagi pengguna media sosial. Dengan kemajuan teknologi, dapat digunakan untuk mendeteksi cyberbullying yang terjadi di media sosial. Artikel ini menggunakan pendekatan metode literature review yaitu narrative literature review 10 artikel pemanfaatan teknologi pendeteksi cyberbullying periode 2011 - 2021 dengan tujuan untuk mengetahui komentar cyberbullying pada akun/postingan seseorang. Oleh karena itu, deteksi cyberbullying mencoba mengumpulkan dataset global di media sosial (Facebook, Instagram, Twitter, dll), dengan mengklasifikasikan metode Machine Learning. Setiap metode algoritma dievaluasi menggunakan akurasi, presisi, recall, dan skor F1 untuk menentukan kinerja tingkat klasifikasi.

### KATA KUNCI

*Cyberbullying, Gangguan Kesehatan Jiwa, Sosial Media*

## PENDAHULUAN

Teknologi komunikasi adalah prinsip ilmu komunikasi untuk membuat suatu item material yang efektif dan efisiensi dalam sistem komunikasi (Setiawan, 2018). Perkembangan teknologi informasi dan komunikasi di era globalisasi sangat pesat dan meluas pemakaian di dunia termasuk di Indonesia yang dapat memudahkan masyarakat untuk mengakses internet baik kapan pun dan dimana pun (Ramadhani & Pratama, 2020). Teknologi komunikasi dan informasi merupakan alat untuk memperoleh

informasi atau pesan yang dapat membantu manusia menyelesaikan masalah agar tujuan komunikasi tercapai. Teknologi informasi yang banyak digunakan oleh masyarakat yaitu *handphone* (Setiawan, 2018).

Media sosial seperti seperti Twitter, Facebook, Instagram, dan media sosial lainnya merupakan salah satu produk dari perkembangan teknologi informasi sebagai sarana untuk berkomunikasi secara daring oleh pengguna internet. Menurut data Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) bahwa

pada tahun 2020 jumlah pengguna jejaring internet di Indonesia mencapai 196,7 juta pengguna atau 73,7% dari total penduduk Indonesia (APJII, 2014).

Sementara itu, hasil survei APJII tahun 2017 menyebutkan bahwa aplikasi media sosial yang sering dikunjungi masyarakat Indonesia diantaranya yaitu Facebook (54%), Instagram (15%), Youtube (11%), Google (6%), Twitter (5,5%) dan LinkedIn (0,6%) (Aprilia et al., 2020).

Pemanfaatan Teknologi informasi mempunyai manfaat positif, diantaranya memudahkan seseorang untuk menyelesaikan masalah di kehidupan sehari-hari. Namun demikian teknologi ini mempunyai dampak negatif yaitu adanya fenomena *cyberbullying* pada anak-anak maupun remaja (Rahayu, 2012).

*Cyberbullying* adalah perilaku seperti mengintimidasi untuk melecehkan korban melalui perangkat teknologi yang dilakukan secara sengaja dan terjadi berulang-ulang kepada target atau korban pengguna media sosial (Rifauddin, 2016). *Cyberbullying* atau kekejaman media sosial memberikan dampak yang lebih besar dari pada kekerasan fisik (Rahayu, 2012).

Fenomena *cyberbullying* juga banyak terjadi di Indonesia. Penelitian UNICEF (2016) mengungkapkan sebanyak 50% dari 41 remaja di Indonesia berusia 13-15 tahun telah melakukan tindakan *cyberbullying* antara lain berupa mempublikasikan data pribadi orang lain, *stalker* atau menguntit, balas dendam dengan menyebarkan video atau foto bertujuan adanya

tindakan intimidasi dan pemerasan (Syah & Hermawati, 2018).

*Cyberbullying* di kalangan remaja cukup tinggi (Syah & Hermawati, 2018). Penelitian Patchin dan Hinduja (2006) menyatakan bahwa 11% remaja pernah melakukan *cyberbullying*, 29% menjadi korban *cyberbullying*, dan hampir 50% menjadi saksi *cyberbullying*. Penelitian lainnya oleh Li (2007) didapatkan data bahwa pada siswa Sekolah Menengah Pertama di Kanada dan China yaitu sebanyak 30% merupakan korban *cyberbullying* dengan laki-laki sebanyak 55,6% dan perempuan 54,5%. Selain itu, sebanyak 18% terlibat menjadi pelaku *cyberbullying*, 85% menjadi saksi dari interaksi negatif media sosial, 12% mengatakan bahwa hal tersebut sering terjadi (Syah & Hermawati, 2018).

Penelitian lain menyebutkan bahwa terdapat kasus penghinaan melalui jejaring sosial Facebook. Korban menerima kata-kata hinaan melalui Facebook atas dasar kecemburuan pelaku terhadap kedekatan korban dengan pacar pelaku. Hal tersebut memberikan dampak psikologis bagi korban (Rifauddin, 2016).

Sementara itu penelitian lain pada 20 siswa SMP diperoleh data bahwa perbuatan *cyberbullying* terjadi di jejaring sosial media Facebook berupa penghinaan kata-kata, penghinaan berupa gambar, menyebarkan rahasia, dan membuat akun palsu untuk mengirim pesan ancaman. Dampak yang diterima korban yaitu takut, cemas, sulit bersosialisasi, depresi, frustrasi, dan stress (Pandie & Weismann, 2016).

Penelitian (Rahayu, 2012) dilakukan pada SMP dan SMU usia 12-19 tahun di kota Magelang, Semarang, dan Yogyakarta terjadi cyberbullying pada 55% berada di lingkungan sekolah, 45% terjadi di luar lingkungan sekolah. Cyberbullying cenderung terjadi pada perempuan dibandingkan dengan laki-laki (N. Willard, 2011).

Penelitian (Maya, 2015) dilakukan pada 6 pelajar SMA dan SMK di Kota Malang yang berusia 16-17 tahun diantaranya 2 pelaku cyberbullying, 2 korban cyberbullying, 2 orang pengguna Facebook. Tindakan bullying yang dilakukan pelaku di jejaring sosial Facebook berupa menghina, mengolok, memaki, dan mengancam.

Penelitian Wiryada dan Martiani (2017) pada siswa Sekolah Menengah Atas (SMA) mendapat gambaran bahwa korban cyberbullying berdasarkan jenis kelamin dan rentang usia 15-17 tahun berada pada kategori tinggi.

Penelitian (Syadza, 2017) dilakukan pada remaja SMP X di Kota Perkalangan menunjukkan adanya hubungan remaja antara konformitas dan kematangan emosi terhadap cyberbullying baik positif maupun negatif. Terdapat 23,8% merupakan konformitas, sedangkan 29,5% merupakan kematangan emosi.

Upaya pencegahan yang dapat dilakukan yaitu dengan melibatkan peran orang tua dan guru. Penelitian (James & Yuono, 2020), melaporkan bahwa banyak orangtua yang mengabaikan perilaku

cyberbullying atau sebagian orangtua tidak sadar akan masalah cyberbullying pada anak. Selain itu juga masih adanya anggapan bahwa *cyberbullying* merupakan kenakalan biasa pada anak di media sosial. Pencegahannya dengan orang tua harus bertindak dengan cara positif karena anak lebih cenderung menurut pada orang tua, Dukungan orang tua seperti memberi kasih sayang, memuji, dicintai dan dihargai dengan adanya dukungan tersebut anak-anak akan merasa percaya diri terhadap lingkungan. Program dalam menghadapi cyberbullying dengan melakukan kebijakan, seperti meningkatkan kegiatan-kegiatan kelompok berbasis lingkungan di sekolah, contohnya program anti *cyberbullying* (James & Yuono, 2020).

## METODOLOGI PENELITIAN

Metode penulisan yang digunakan dalam penulisan artikel melalui metode *literature review*. Beberapa database dan website pencarian meliputi Google Scholar, PubMed, Perpustakaan, dan National Center for Biotechnology Information dengan kata kunci "cyberbullying, teknologi, gangguan kesehatan mental pada pengguna media sosial". Kriteria inklusi meliputi artikel tentang penggunaan teknologi, pemantauan kesehatan mental berupa Cyberbullying pada artikel dalam 10 tahun terakhir (2011 - 2021). Sementara kriteria eksklusi yang dipergunakan adalah artikel dengan Bahasa Indonesia dan Bahasa Inggris, artikel tidak spesifik terkait pemantauan gangguan yang dialami pada pengguna media social. Jumlah

artikel yang ditemukan pada database awal ada 30 artikel. Setelah dilakukan skrining, jumlah artikel akhir yang di review adalah 10 artikel. Kemudian diseleksi, disajikan, dianalisis dan diolah secara ringkas dan sistematis. Analisis dilakukan dengan membandingkan atau menemukan kesamaan isi artikel tergantung pada kata kunci dari masalah yang ingin dibahas.

## HASIL DAN PEMBAHASAN

Para peneliti telah bekerja secara intensif pada deteksi cyberbully untuk menemukan cara untuk mengontrol atau mengurangi cyberbully di platform Media Sosial. Cyberbullying meresahkan karena korban tidak dapat mengatasi beban emosional dari pesan kekerasan, intimidasi, merendahkan, dan bermusuhan. Untuk mengurangi dampak cyberbullying perlu dipelajari dalam hal deteksi, pencegahan, dan mitigasi. Berdasarkan isu-isu pendekatan pencegahan di atas, kebutuhan untuk mendeteksi dan menyaring cyberbullying di media sosial sangat dibutuhkan. Dengan demikian, bagian ini bertujuan untuk memeriksa teknik deteksi cyberbully. Sesuai tinjauan literatur, salah satu utama dalam mendeteksi cyberbully menggunakan metode *Machine Learning*.

Banyak penelitian tentang cyberbullying telah dilakukan di dalam dan luar negeri. Penelitian Candra dan Rozana (2020) yang berjudul "Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor" membahas tindakan Bullying yang ada di kolom komentar pada media sosial instagram (Candra & Nanda Rozana, 2020).

Sistem media sosial pada Instagram

menghubungkan satu akun dengan akun lainnya dengan cara menjadi follower atau mengikuti. Hubungan komunikasi antara pengguna Instagram lain dapat dibangun dengan like dan komentar pada foto dan video yang diunggah oleh pengguna lain. Namun penggunaan komunikasi melalui komentar Instagram telah banyak disalahgunakan oleh pengguna yang tidak memahami etika ketika bersosialisasi di dunia maya, dan telah diejek, dan dihina melalui komentar di Instagram. Ini akan menjadi bagian dari tindakan cyberbullying (Candra & Nanda Rozana, 2020).

Kelebihan pada penelitian ini yaitu seseorang dapat mengetahui perbedaan komentar bullying dan yang bukan bullying, Perbandingan menggunakan Riset algoritma naive Bayes classifier untuk mendeteksi komentar negatif di Instagram. Hasil akurasi menggunakan metode Naive Bayesian adalah 76,7%, sedangkan hasil akurasi menggunakan metode K Nearest Neighbor adalah 79,21%. Oleh karena itu, metode K Nearest Neighbor lebih cocok untuk mengklasifikasikan data komentar dalam penelitian ini. Semenatra itu kekurangan dari penelitian ini yaitu diperlukan pencaharian secara menyeluruh dan mengumpulkan data secara manual (Candra & Nanda Rozana, 2020).

Penelitian kedua oleh Munir dan Fati (2020) yang berjudul "A Comparative Analysis of Machine Learning Techniques for Cyberbullying Detection on Twitter" membahas bagaimana cara mendeteksi cyberbullying tanpa melibatkan korban. Cyberbullying berbasis pembelajaran mesin learning adalah arah lain

dari deteksi cyberbullying dan banyak digunakan oleh banyak peneliti. Selain itu, Machine Learning (ML) merupakan turunan dari teknologi kecerdasan buatan yang memungkinkan sistem untuk secara otomatis belajar dan berkembang dari pengalaman tanpa pemrograman khusus. Teknik ini dapat diklasifikasikan sebagai algoritma yang diawasi, semi-diawasi, atau tidak diawasi. Kelebihannya yaitu dapat meningkatkan keamanan pengguna Internet termasuk anak-anak dan dapat digunakan untuk mengekstrak karakteristik pengklasifikasi (Muneer & Fati, 2020).

Pada penelitian selanjutnya membahas tentang cyberbullying pada media sosial facebook dengan menggunakan K-Nearest Neighbor. Untuk mendeteksi cyberbullying dan mengambil keputusan kepada pengguna facebook. Kelebihan deteksi cyberbullying ini yaitu pemanfaatan fitur-fitur dengan mencari kata-kata yang termasuk cyberbullying dan emotikon yang meningkat dalam akurasi klasifikasi (Hasan, 2021).

Penelitian lainnya membahas tentang cyberbullying pada media sosial Twitter dengan menggunakan Random Forest Classifier. Untuk mendeteksi cyberbullying pada isi pesan, pengirim, dan penerima pesan. Kelebihan deteksi cyberbullying dilakukan secara langsung maupun tidak langsung untuk menemukan kata-kata atau kalimat yang mengandung unsur cyberbullying di tweet yang merupakan deteksi akurasi tertinggi yaitu 90%. Kekurangan deteksi ini yaitu mengevaluasi komentar yang membagikan metode random forest classifier

atau pengklasifikasian kalimat bertambah secara yang terus menerus (Natasya Novalita et al., 2019).

Pada penelitian kelima ini berfokus pada aplikasi alat untuk mendeteksi komentar cyberbullying ke postingan media sosial Instagram. Penelitian ini bertujuan untuk menghasilkan aplikasi yang mampu untuk mendeteksi komentar negative yang mengandung unsur cyberbullying. Berdasarkan yang diuraikan maka diciptakan aplikasi pendeteksi cyberbullying pada komentar Postingan Media Sosial Instagram Dengan Metode Naive Bayes Classifier. Kelebihannya yaitu untuk mendapatkan nilai tertinggi yang mengklasifikasinya pada data diuji dan digolongkan paling cepat. Pada penelitian ini untuk data uji yaitu dokumen komentar postingan sosial media Instagram (Hutagalung et al., 2021).

Penelitian keenam ini dengan melakukan beberapa proses. Proses yang pertama untuk mengumpulkan data, preprocessing, klasifikasi, evaluasi, dan deteksi konten. Pada penelitian ini menggunakan metode deep learning dan algoritma yang digunakan adalah LSTM dan BiLSTM. Metode ini sangat penting dilakukan karena belum adanya metode yang dapat mendeteksi tweet (cuitan) yang mengandung cyberbullying. Algoritma ini juga memiliki nilai yang sangat tepat untuk mendeteksi algoritma yang lain. Pengembangan dari penelitian ini dapat membantu orang tua dan melindungi generasi muda dari perbuatan cyberbullying. Kelebihan pada LSTM yaitu dapat merubah dari Recurrent Neural Network (RNN) dengan menambahkan memori cell untuk

menyimpan suatu informasi menggunakan waktu yang sangat panjang, dan dapat menangani sesuatu masalah vanishing gradient yang di RNN (Natasya Novalita et al., 2019).

Pada kelebihan di BILSTM merupakan suaru yang dikembangkan dari model LSTM yang memiliki dua macam untuk masing-masing proses yang saling bertolak belakang, model ini sangatlah baik untuk mengetahui suatu pola pada kalimat karena setiap kata yang terdapat di dokumen dapat diproses secara urutan (Natasya Novalita et al., 2019).

Penelitian selanjutnya berkaitan dengan analisis, desain, dan implementasi perangkat lunak cyberbullying berdasarkan teknologi agen. Perangkat lunak yang dikembangkan dan dapat membantu mencegah dan mengatasi terjadinya cyberbullying. Aplikasi ini dapat melakukan filtering atau penyaringan konten berupa kata-kata atau pesan. Sistem ini dirancang untuk mengenali dan menyimpan kosakata yang digunakan dalam cyberbullying dan melibatkan orang tua untuk memantau penggunaan internet oleh anak (Cohen et al., 2014).

Penelitian lainnya yaitu menggunakan mesin learning untuk mendeteksi cyeberbullying dan untuk mendeteksi perilaku cyberbullying dan tingkat keparahannya di jejaring sosial online. Kelebihan dari sitem ini dapat menghasilkan sebuah konten dengan menggunakan sitem informasi timbal balik yang tepat (Talpur & O'Sullivan, 2020).

Sementara itu terdapat juga model neural network yang mendeteksi konten negatif di media

sosial sebanyak dengan mempertimbangkan konteks kalimat atau frasa serta dua model yang dianalisis dari hasil, yaitu kata dengan ANN dan RNN. Model RNN berkinerja lebih baik dari pada model ANN, dengan pelatihan, validasi, dan akurasi pengujian masing-masing sebesar 94% dan 84%. Metode yang umum termasuk naive bay, mesin vektor pendukung, metode semantik, dan metode obfuscation, yang semuanya memperhitungkan konteks kalimat atau frasa yang tidak dipahami dengan baik oleh komputer (Sahrul et al., 2019). Kelebihannya yitu dapat mendeteksi konten negatif yang biasanya muncul dengan kalimat umpatan, dalam sistem ini dapat mendeteksi tidak hanya kata-per-kata. Kekurangannya yaitu lamanya proses pelatihan. Menggunakan data dalam jumlah besar dapat memakan waktu yang sangat lama (Sahrul et al., 2019).

Penelitian lain yang menggunakan machine learning dapat dilakukan untuk mendeteksi kategori teks; apakah tergolong sebagai ujaran kebencian, serta memprediksi kategori dari ujaran kebencian yang terkandung dalam suatu teks. Kelebihannya yaitu dapat mengklasifikasi pada data dengan *double label* (Perwira et al., 2021).

## SIMPULAN DAN SARAN

Dalam penelitian ini diharapkan dapat membantu untuk mendeteksi adanya *cyberbullying* dengan menggunakan banyak model dalam mendeteksi. Sehingga dalam pendeteksian ini dapat bekerja secara sistematis dan dapat mengidentifikasi tingkat keparahan *cyberbullying* serta

mengklasifikasikannya apakah termasuk *cyberbullying* atau *non-cyberbullying*.

## DAFTAR PUSTAKA

- APJII, P. (2014). Asosiasi Penyelenggara Jasa Internet Indonesia. *Jakarta: APJII*.
- Aprilia, R., Sriati, A., & Hendrawati, S. (2020). Tingkat Kecanduan Media Sosial pada Remaja. *Journal of Nursing Care*, 3(1), 41–53.
- Candra, R. M., & Nanda Rozana, A. (2020). Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor. *IT Journal Research and Development*, 5(1), 45–52. [https://doi.org/10.25299/itjrd.2020.vol5\(1\).4962](https://doi.org/10.25299/itjrd.2020.vol5(1).4962)
- Cohen, R., Lam, D. Y., Agarwal, N., Cormier, M., Jagdev, J., Jin, T., Kukreti, M., Liu, J., Rahim, K., & Rawat, R. (2014). Using computer technology to address the problem of cyberbullying. *Acm Sigcas Computers and Society*, 44(2), 52–61.
- Hasan, N. F. (2021). Deteksi Cyberbullying pada Facebook Menggunakan Algoritma K-Nearest Neighbor. *Journal of Smart System*, 1(1), 35–44. <https://doi.org/10.36728/jss.v1i1.1605>
- Hutagalung, A. S., Negara, A. B. P., & Pratama, E. E. (2021). Aplikasi Pendeteksi Cyberbullying Terhadap Komentar Postingan Media Sosial Instagram dengan Metode Naïve Bayes Classifier Berbasis Website. *JUSTIN (Jurnal Sistem Dan Teknologi Informasi)*, 9(3), 364–371. <https://doi.org/10.26418/justin.v9i3.44843>
- James, B., & Yuono, D. (2020). Pusat Pencegahan Cyberbullying: Pencegahan Cyberbullying Melalui Karya Arsitektur. *Jurnal Sains, Teknologi, Urban, Perancangan, Arsitektur (Stupa)*, 1(2), 1359. <https://doi.org/10.24912/stupa.v1i2.4450>
- Maya, N. (2015). Fenomena Cyberbullying Di Kalangan Pelajar. *JISIP: Jurnal Ilmu Sosial Dan Ilmu Politik*, 4(3), undefined-450.
- Muneer, A., & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on twitter. *Future Internet*, 12(11), 1–21. <https://doi.org/10.3390/fi12110187>
- N. Willard. (2011). Educator's Guide to Cyberbullying and Cyberthreats. *Center for Safe and Responsible*.
- Novalita, N., Herdiani, A., Lukmana, I., & Puspandari, D. (2019). Cyberbullying identification on twitter using random forest classifier. *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012029>
- Novalita, Natasya, Herdiani, A., & Lukmana, I. (2019). *Identifikasi Cyberbullying Pada Media Sosial Twitter Menggunakan Metode Klasifikasi Random Forest*.
- Okik Adishya Banu Wiryada, Nuke Martiarini, T. E. B. (2017). Gambaran Cyberbullying Pada Remaja Pengguna Jejaring Sosial di SMA Negeri 1 dan SMA Negeri 2 Ungaran. *Intuisi Jurnal Psikologi Ilmiah*, 9(1), 26–38.
- Pandie, M. M., & Weismann, I. T. J. (2016). Pengaruh Cyberbullying Di Media Sosial Terhadap Perilaku Reaktif Sebagai Pelaku Maupun Sebagai Korban Cyberbullying Pada Siswa Kristen SMP Nasional Makassar. *Jurnal Jaffray*, 14(1), 43–62. <https://doi.org/10.25278/jj.v14i1.188.43-62>
- Patchin & Hinduja. (2006). Bullies move beyond the schoolyard: a preliminary look at cyberbullying. *Youth Violence Juv. Justice*, 148–169. <https://doi.org/10.1177/1541204006286288>
- Pewira, A., Dwitama, J., & Kunci, K. (2021). *Deteksi Ujaran Kebencian Pada Twitter Bahasa Indonesia Menggunakan Machine Learning : Reviu Literatur*. 1, 31–39.
- Qilla Aulia Suri, A. M. G. (2019). Perancangan Sistem Perangkat Lunak Anti Cyberbullying Berbasis Agen Sebagai Solusi Pencegahan Dan Penanganan Dampak Negatif Penggunaan Teknologi Internet. *Prosiding SNATIF Ke-6 Tahun 2019, 2007*, 96–101.
- Rahayu, F. S. (2012). *CYBERBULLYING SEBAGAI DAMPAK NEGATIF PENGGUNAAN TEKNOLOGI INFORMASI*. 43, 22–31.
- Ramadhani, M. R., & Pratama, & A. R. (2020). Analisis Kesadaran Cyber Security Pada Pengguna Media Sosial Di Indonesia. *Jurnal Aotomata*, 3(2), 1–8.
- Rifauddin, M. (2016). Fenomena Cyberbullying pada Remaja. *Khizanah Al-Hikmah: Jurnal Ilmu Perpustakaan, Informasi, Dan Kearsipan*, 4(1), 35–44. <https://doi.org/10.24252/kah.v4i1a3>
- Sahrul, S., Rahman, A. F., Normansyah, M. D., &

- Irawan, A. (2019). Sistem Pendeteksi Kalimat Umpatan Di Media Sosial Dengan Model Neural Network. *Computatio: Journal of Computer Science and Information Systems*, 3(2), 108–115.
- Setiawan, D. (2018). Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Budaya. *JURNAL SIMBOLIKA: Research and Learning in Communication Study*, 4(1), 62. <https://doi.org/10.31289/simbolika.v4i1.1474>
- Syadza, N. (2017). Ditinjau Dari Konformitas Dan Kematangan Emosi. 12(1), 17–26.
- Syah, R., & Hermawati, I. (2018). The Prevention Efforts on Cyberbullying Case for Indonesian Adolescent Social Media Users. *Jurnal Penelitian Kesejahteraan Sosial*, 17(2), 131–146.
- Talpur, B. A., & O'Sullivan, D. (2020). Cyberbullying severity detection: A machine learning approach. *PLoS ONE*, 15(10 October), 1–19. <https://doi.org/10.1371/journal.pone.0240924>



Tabel Literature Review (Di tulis di akhir)

Nama Penelitian & Tahun	Judul	Responden	Intervensi	Hasil
(Candra & Nanda Rozana, 2020)	Instagram Menggunakan Metode <i>K-Nearest Neighbor</i>	Populasi data Jumlah data yang digunakan sebanyak 1000 data, dimana 500 merupakan bullying dan 500 tidakbully	<i>K-Nearest Neighbor</i>	Hasil pengujian Berdasarkan tahapan-tahapan metodologi penelitian yang telah dilaksanakan, maka dapat ditarik beberapa kesimpulan sebagai berikut: Metode <i>K-Nearest Neighbor</i> dapat diterapkan untuk klasifikasi bully dan tidak bully pada komentar Instagram. Klasifikasi bully dan tidak bully pada komentar Instagram menggunakan <i>K-Nearest Neighbor</i> ini telah berhasil mengklasifikasikan komentar bully dan tidak bully
(Muneer & Fati, 2020)	<i>A Comparative Analysis of Machine Learning Techniques for Cyberbullying Detection on Twitter</i>	Pada penelitian ini menggunakan kumpulan data global 37.373 tweet	Teknik <i>Machine Learning (Machine Learning Techniques)</i>	<p>Hasil eksperimen menunjukkan bahwa <i>Logistic Regression</i> mencapai akurasi dengan skor terbaik, yaitu 90,57% dan 0,9280.</p> <p>Terdapat perbedaan dari performa <i>classifier Logistic Regression (LR)</i>, <i>SGD Classifier (SGD)</i>, dan <i>LGBM Classifier (LGM)</i>, dimana SGD mencapai akurasi 90,6%, namun skor F1 lebih rendah dari LR. Namun, pengklasifikasi LGBM mencapai akurasi 90,55%, dan skor F1 adalah 0,9271. Sedangkan SVM telah mencapai akurasi 67,13% dan presisi 0,6713 terendah dalam riset ini.</p> <p>Ini berarti bahwa LR berkinerja lebih baik daripada pengklasifikasi lainnya.</p>
(Hasan, 2021)	Deteksi Cyberbullying pada Facebook Menggunakan <i>Algoritma K-Nearest Neighbor</i>	Cyberbullying dengan jumlah data Facebook sebanyak 3000, diantaranya 1022 komentar positif cyberbullying, dan 1978 komentar negatif cyberbullying (no cyberbullying)	<i>K-Nearest Neighbor</i>	<p>Hasil akurasi menggunakan K-</p> <p>Hasil akurasi menggunakan K-NN untuk deteksi cyberbullying pada facebook dalam mengenali sentiment positif cyberbullying perolehan akurasi tertinggi saat menggunakan 1-</p>
(N. Novalita et al., 2019)	<i>Cyberbullying Identification On Twitter Using Random Forest Classifier</i>	Penelitian ini pada pengguna Twitter (50 grup Twitter)	<i>Random Forest Classifier</i>	Hasil dari pengujian ini dengan klasifikasi random forest didapatkan bahwa sistem ini berhasil mengidentifikasi tweet cyberbullying dengan f1 score terbaik
(Hutagalung et al., 2021)	Aplikasi Pendeteksi Cyberbullying Terhadap Komentar Postingan Media Sosial Instagram dengan Metode <i>Naïve Bayes Classifier</i> Berbasis <i>Webside</i>	Pada penelitian ini jumlah data terhadap 200 data komentar dari suatu postingan media sosial instagram menunjukkan bahwa terdapat 78 komentar bullying dan 122 komentar bukan bullying	<i>Naïve Bayes Classifier</i>	Aplikasi pendeteksi cyberbullying terhadap komentar postingan media sosial instagram dengan metode <i>naive bayes classifier</i> membuktikan bahwa aplikasi yang telah dibuat dalam penelitian ini dapat mendeteksi dan menghapus ribuan komentar cyberbullying secara massal
(Natasya Novalita	Identifikasi	Pada penelitian ini	LTMS dan	Aplikasi pendeteksi

et al., 2019)	Cyberbullying pada Media Sosial Twitter Menggunakan Metode LSTM dan BiLSTM	jumlah data 6835 tweet (cuitan) seperti yang terlihat pada diantaranya yaitu 3900 tweet (cuitan) non-cyberbullying dan 2935 tweet (cuitan) cyberbullying	BiLSTM	Identifikasi Cyberbullying aplikasi pendeteksi cyberbullying terhadap tweet (cuitan) di media sosial Twitter dengan metode LSTM dan BiLSTM membuktikan bahwa aplikasi yang telah dibuat dalam penelitian ini berhasil mengidentifikasi tweet (cuitan) bermakna cyberbullying pada media sosial Twitter dengan melakukan klasifikasi dua kelas tweet (cuitan) yang tersedia pada dataset
(Qilla Aulia Suri, 2019)	Perancangan Sistem Perangkat Lunak Anti Cyberbullying Berbasis Agen Sebagai Solusi Pencegahan Dan Penanganan Dampak Negatif Penggunaan Teknologi Internet	Remaja dan Pra Remaja sebanyak 4500 korban cyberbullying	Perangkat Lunak Anti Cyberbullying Berbasis Teknologi Agen	Hasil dari perancangan aplikasi ini adalah adanya sistem yang dapat membuat orang tua mengawasi penggunaan Internet oleh anak terutama dalam upaya pencegahan terhadap tindakan cyberbullying melalui media sosial
(Talpur & O'Sullivan, 2020)	<i>Cyberbullying Severity Detection: A Machine Learning Approach</i>	Populasi yang kami dapatkan dari tweet yaitu: dari total 50.000 tweet yang terkumpul, peneliti hanya mengambil sebanyak 24.189 tweet yang dianotasi	Mesin Pendekatan Pembelajaran	Hasil ini menunjukkan bahwa kerangka kerja yang kami usulkan memberikan solusi yang layak untuk mendeteksi perilaku cyberbullying dan tingkat keparahannya di jejaring sosial online
(Sahrul et al., 2019)	Sistem Pendeteksi Kalimat Umpatan Di Media Sosial Dengan Model Neural Network	Data yang berhasil dikumpulkan sebanyak 88,009 data selama 3 hari	<i>Synthetic minority over-sampling technique (SMOTE)</i>	Menunjukkan bahwa model RNN lebih baik dibandingkan dengan model ANN.  Model RNN dapat memprediksi dengan benar 83.8% dari keseluruhan test set yang dilabeli sebagai umpatan dan 84.4% yang bukan umpatan. Sedangkan model ANN hanya dapat memprediksi dengan benar 82.2% yang dilabeli sebagai umpatan dan 80.6% yang bukan umpatan
(Perwira et al., 2021)	Deteksi Ujaran Kebencian Pada Twitter Bahasa Indonesia Menggunakan Machine Learning: Review Literatur	Sebanyak 125 akun media sosial yang terindikasi mengandung unsur Cyber bullying. akun twitter memiliki angka terbesar yakni sebanyak 79 akun	Menggunakan mesin pendekatan pembelajaran	Dalam metode machine learning pada algoritma Neural network memiliki performa sangat baik dalam melakukan klasifikasi artinya sangat akurat dalam mendeteksi rincian dari 79 akun pengguna twitter berhasil memberikan 54.650 teks dengan mentang posting dari jam 11.00 hingga 19.00. Teks-teks tersebut kemudian disaring dengan hasil akhir sebanyak 320 teks yang termasuk ke dalam ujaran kebencian dan 277 lainnya bukan termasuk ke dalam ujaran kebencian