

PEMANFAATAN METODE ANALISIS SENTIMEN SEBAGAI PENDEKATAN DETEKSI DINI BULLYING

Arfian Alhafidz¹, Khanun Roisatul Ummah²

^{1,2}Universitas Muhammadiyah Surakarta, Jl. A. Yani, Pabelan, Kartasura, Sukoharjo,
Jawa Tengah, Indonesia

 Email korespondensi: arfian.alhafidz@gmail.com

Abstrak. *Bullying* adalah sebuah bentuk tindakan kekerasan anak yang dilakukan teman sebaya maupun yang lebih tua kepada seseorang yang lebih 'rendah' atau lebih lemah untuk mendapat keuntungan atau kepuasan tertentu. *bullying* masih menjadi teror bagi anak-anak di lingkungan sekolah maupun di luar sekolah. Berdasarkan masalah tersebut, maka dibutuhkan suatu sistem yang memungkinkan mendeteksi potensi tindakan *bullying* secara dini dan memberikan informasi yang relevan untuk membantu pihak terkait dalam menindak lanjuti kejadian *bullying* tersebut. Sistem ini memanfaatkan sebuah model analisis sentimen berbasis NLP (*Natural Language Processing*), dengan model analisis sentimen yang diterapkan pada platform *mobile* memanfaatkan teknologi ML (*Machine Learning*) untuk memproses dan menganalisis data teks secara otomatis. Metode yang digunakan dimulai dengan pengumpulan data teks, yang kemudian diproses menggunakan teknik NLP untuk ekstraksi fitur. Penelitian ini mengembangkan sistem deteksi *bullying* melalui analisis sentimen dengan membandingkan tiga model *machine learning* untuk menemukan kinerja terbaik: BiLSTM (*Bidirectional Long Short-Term Memory*) sebagai model utama berbasis deep learning, Naïve Bayes sebagai baseline statistik, dan IndoBERT sebagai representasi model state-of-the-art. Hasil penelitian menunjukkan bahwa model BiLSTM mencapai kinerja paling unggul dengan akurasi keseluruhan tertinggi sebesar 90,00%. Berdasarkan hasil tersebut, model BiLSTM dipilih sebagai model final, model ini kemudian berhasil diimplementasikan ke dalam aplikasi mobile, fungsional yang mampu menyediakan deteksi teks perundungan secara real-time.

Kata kunci: analisis sentimen; BiLSTM; *bullying*; *machine learning*; NLP



PENDAHULUAN

Perundungan dan kekerasan terhadap anak dan perempuan merupakan isu serius di berbagai lingkungan, termasuk lembaga pendidikan, lingkungan keagamaan, dan keluarga. Menurut data FSGL, terdapat 23 kasus perundungan di lembaga pendidikan dari Januari hingga September 2023, dengan mayoritas terjadi di jenjang SMP (50%), diikuti SD (23%), SMA, dan SMK (masing-masing 13,5%). Selain itu, data dari Kementerian Pemberdayaan Perempuan dan Perlindungan Anak menunjukkan 22.762 kasus kekerasan terhadap anak dan perempuan dari Januari hingga Oktober 2023, dengan kekerasan seksual (10.104 kasus), fisik (7.732 kasus), dan psikis (7.119 kasus) sebagai jenis kekerasan yang paling banyak terjadi. Angka-angka ini menekankan pentingnya upaya pencegahan dan penanganan yang efektif terhadap *bullying* dan kekerasan.

Dengan perkembangan teknologi dan meningkatnya penggunaan media sosial, mengakibatkan tindakan *bullying* juga dapat terjadi di dunia maya, yang dikenal sebagai *cyberbullying* [1]. *Cyberbullying* memiliki karakteristik yang berbeda dari *bullying* konvensional karena dapat terjadi kapan saja dan di mana saja, serta dapat menyebar dengan cepat melalui platform digital [2]. Korban *cyberbullying* sering kali merasa lebih terisolasi karena serangan dapat datang dari anonim atau seseorang yang tidak dikenal. Selain itu, jejak digital dari *cyberbullying* dapat bertahan lama dan sulit dihapus, sehingga memperpanjang penderitaan korban [3]. Sampai saat ini, banyak korban *bullying* yang tidak melaporkan kejadian yang mereka alami, baik karena takut maupun tidak tahu harus melapor ke mana. Hal ini menyebabkan banyak kasus *bullying* yang tidak terdeteksi dan tidak tertangani dengan baik.

Oleh karena itu, diperlukan solusi teknologi yang efektif untuk mendeteksi dan mencegah *bullying*, baik di dunia nyata maupun dunia maya. Dengan suatu sistem yang dapat memfasilitasi pelaporan, menawarkan dukungan yang diperlukan, dan bertujuan untuk mengurangi dampak jangka panjang dari pengalaman traumatis dari kejadian *bullying* yang dialami korban. Salah satu pendekatan yang mungkin dilakukan adalah penggunaan ML (*machine learning*) dan NLP (*Natural Language Processing*) untuk menganalisis teks dan mendeteksi tanda-tanda *bullying*. Dengan mengembangkan aplikasi pendeteksi tindakan *bullying* yang menggunakan model Analisis Sentimen berbasis *machine learning* dengan metode NLP [4]. NLP sendiri adalah bagian dari kecerdasan buatan yang berhubungan dengan pemahaman dan memperoleh wawasan dari bahasa manusia seperti teks dan ucapan [5]. Beberapa aplikasi umum NLP adalah analisis sentimen dengan menggunakan metode *Deep Learning* yaitu LSTM (*Long Short-Term Memory*) dapat meningkatkan proses analisis sentimen agar mendapat hasil yang lebih menjanjikan [6]. LSTM sangat efektif untuk menangani data urutan seperti teks,



yang merupakan inti dari banyak aplikasi NLP, termasuk analisis sentimen, penerjemahan bahasa, dan pengenalan suara [7]. Namun, arsitektur LSTM standar hanya memproses informasi dalam satu arah (dari depan ke belakang). Keterbatasan ini membuat pemahaman konteks sebuah kata hanya didasarkan pada kata-kata yang telah muncul sebelumnya. Untuk mendapatkan pemahaman konteks yang lebih kaya, penelitian ini mengimplementasikan BiLSTM (*Bidirectional Long Short-Term Memory*). BiLSTM merupakan pengembangan dari LSTM yang memproses teks dari dua arah secara bersamaan: satu lapisan memproses dari depan ke belakang (konteks masa lalu), dan lapisan kedua memproses dari belakang ke depan (konteks masa depan). Dengan menggabungkan informasi dari kedua arah, BiLSTM mampu menangkap makna kata secara lebih utuh dan akurat, yang sangat krusial untuk tugas deteksi *bullying* yang sensitif terhadap konteks kalimat. Dikarenakan sistem ini menggunakan sebuah teks atau narasi sebagai dasar penilaian, sehingga penggunaan analisis sentimen sangat cocok karena sistem ini mengklasifikasikan sebuah teks menjadi 2 label yaitu “positif” atau “negatif”, sehingga sangat cocok untuk melakukan pendeteksian *bullying* secara otomatis [8].

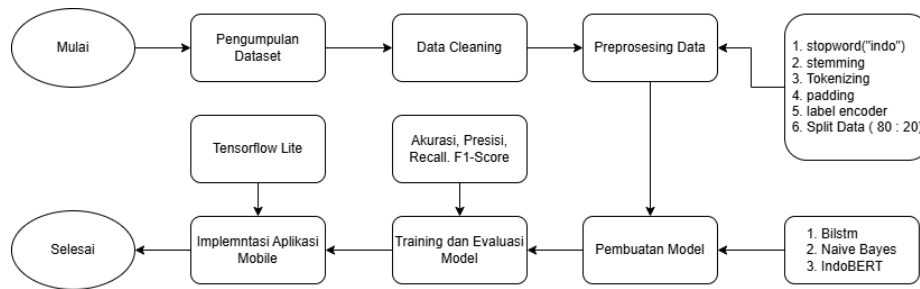
Kebaruan dari penelitian ini terletak pada integrasi teknologi analisis sentimen berbasis BiLSTM (*bidirectional Long Short-Term Memory*) dengan implementasi pada aplikasi *mobile* yang memungkinkan deteksi secara *real-time*. Berdasarkan penelitian sebelumnya, aplikasi anti-*bullying* berfokus pada sistem konsultasi dengan pakar psikiater [9]. Penelitian ini berbeda karena dikembangkan dengan bantuan *machine learning*, serta *TensorFlow Lite* untuk memberikan solusi yang efisien, portabel, dan memungkinkan mendeteksi teks atau narasi secara otomatis. Sistem ini diharapkan dapat membantu mengidentifikasi kasus *bullying* secara otomatis, memberikan intervensi yang tepat waktu, serta memberikan rekomendasi tindakan yang dapat diambil untuk mengatasi masalah tersebut secara efektif.

METODE

Alur penelitian ini menggambarkan langkah-langkah dalam pengembangan aplikasi *mobile* untuk deteksi *bullying* menggunakan analisis sentimen dengan model *machine learning*. Proses dimulai dengan pengumpulan data set dari sumber relevan, diikuti oleh *preprocessing* data, yang mencakup pembersihan data untuk meningkatkan kualitas analisis. Setelah data siap, model *machine learning*, khususnya BiLSTM (*Bidirectional Long Short-Term Memory*), dibuat dan dilatih menggunakan data set yang telah diproses. Evaluasi model dilakukan untuk mengukur kinerjanya dengan metrik seperti akurasi dan presisi. Terakhir, model yang telah dievaluasi diimplementasikan ke dalam aplikasi



mobile, yang dirancang untuk menganalisis teks secara *real-time* dan mendeteksi potensi *bullying*. Metode yang dilakukan pada penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

Dataset

Data yang digunakan berasal dari *dataset* publik di Kaggle dan Github, berisi teks berbahasa Indonesia dengan label *bullying* dan *non-bullying* [10]. *Dataset* ini dikumpulkan dari media sosial seperti Twitter dan Instagram, yang mayoritas penggunaannya adalah remaja. Awalnya, dataset terdiri dari 14.571 data dengan distribusi 55,07% berlabel *Bullying* dan 44,93% berlabel *Tidak-Bullying*. Untuk mengatasi ketidakseimbangan, dilakukan *oversampling* pada kelas minoritas sehingga kedua kelas memiliki jumlah data yang sama, yaitu 8.024 sampel.

Data Pre-processing

Tahap pra-pemrosesan data bertujuan untuk mengubah data mentah menjadi format yang lebih terstruktur dan siap untuk dianalisis [11]. Proses ini diawali dengan *data cleaning* untuk memastikan kualitas data dengan menghilangkan bias seperti salah tulis, perbedaan huruf besar-kecil, dan singkatan. Setelah data bersih, dilakukan serangkaian langkah pemrosesan yang lebih luas, dimulai dengan tokenisasi, yaitu memecah teks menjadi unit-unit kecil seperti kata atau frasa. Selanjutnya, dilakukan normalisasi dengan mengubah semua huruf menjadi kecil, menghapus tanda baca yang tidak diperlukan, serta menghilangkan kata-kata umum yang tidak signifikan (*stop words*), langkah ini diikuti dengan *padding* untuk memastikan semua urutan teks memiliki panjang yang seragam (misalnya 100 token), yang merupakan syarat input untuk model LSTM. Proses ini dilanjutkan dengan *stemming* dan *lemmatization* untuk mengurangi setiap kata ke bentuk dasarnya [12]. Setelah teks diproses, label kategori seperti "*bullying*" dan "*non-bullying*" diubah menjadi representasi numerik melalui proses *encoding* [13]. Terakhir, *dataset* yang telah siap dibagi menjadi 80% data latih dan 20% data uji untuk melatih serta mengevaluasi model secara objektif.

Pembuatan dan Evaluasi Model



Tiga model dievaluasi untuk perbandingan:

1. **BiLSTM:** Sebagai model utama, BiLSTM diimplementasikan karena kemampuannya menangkap informasi kontekstual dari dua arah (maju dan mundur)[14], yang krusial untuk memahami konteks kalimat secara mendalam. Arsitekturnya terdiri dari lapisan *Embedding*, LSTM, BiLSTM, *Dropout*, dan *Dense*.
2. **Naïve Bayes:** Model pembandingan ini adalah algoritma klasifikasi statistik yang menggunakan Teorema Bayes [15], [16]. Teks diubah menjadi representasi numerik menggunakan TF-IDF (*Term Frequency-Inverse Document Frequency*) sebelum diklasifikasikan.
3. **IndoBERT:** Model ini merepresentasikan pendekatan *state-of-the-art* untuk Bahasa Indonesia. Penelitian ini menggunakan pendekatan *fine-tuning* pada model *TFBertForSequenceClassification* yang telah dilatih sebelumnya.

Evaluasi model dilakukan menggunakan metrik seperti akurasi, presisi, *recall*, F1-Score, dan AUC-ROC berdasarkan *confusion matrix*.

Implementasi pada Aplikasi Mobile

Model BiLSTM yang terpilih dikonversi ke format TensorFlow Lite (.tflite) untuk efisiensi pada perangkat *mobile*. Model ini kemudian diintegrasikan ke dalam aplikasi Android *native* yang dikembangkan menggunakan Java dan XML. Aplikasi ini memiliki alur kerja yang mencakup pra-pemrosesan teks masukan dari pengguna, inferensi dengan model TFLite, dan penampilan hasil deteksi kepada pengguna.

HASIL

Hasil Kinerja Model Setiap model diuji melalui beberapa iterasi untuk menemukan performa optimal.

1. **Model BiLSTM:** Skenario *baseline* (Iterasi 1) mencapai akurasi validasi 90,09% namun menunjukkan potensi *overfitting*. Upaya regularisasi (Iterasi 2) sedikit menurunkan akurasi menjadi 89,35%, dan penggunaan *embedding GloVe* (Iterasi 3) menurunkan performa secara signifikan ke 80,09%. Model *baseline* berhasil mengklasifikasikan 1.428 kasus *bullying* (TP) dan 1.483 kasus *non-bullying* (TN) dengan benar.
2. **Model Naïve Bayes:** Performa terbaik dicapai pada iterasi ketiga dengan akurasi 86,92%, menggunakan fitur TF-IDF dan parameter *smoothing Alpha* 0.5. Model ini



mampu mengklasifikasikan 1.358 kasus *bullying* (TP) dan 1.432 kasus *non-bullying* (TN) dengan benar.

- 3. Model IndoBERT:** Konfigurasi *baseline* melalui *fine-tuning* menghasilkan akurasi 89,53% dan skor AUC-ROC tertinggi dari semua eksperimen (96,03%). Model ini unggul dalam jumlah *False Negative* yang paling rendah (111 kasus).

Tabel 1. Perbandingan Model Terbaik

Model	konfigurasi	akurasi	Presisi	Recall	F1-Score	AUC-ROC
BiLSTM	Baseline (Iterasi 1)	90%	89%	91%	90%	94,71%
Naive Bayes	TF-IDF, Alpha 0.5 (Iterasi 3)	86%	85%	90%	87%	94,60%
IndoBERT	aseline Fine-tuning (Iterasi 1)	89,53%	93%	88%	90%	96,03%

PEMBAHASAN

Berdasarkan hasil perbandingan, model BiLSTM mencapai akurasi keseluruhan tertinggi sebesar 90,00%. Meskipun IndoBERT menunjukkan presisi (93%) dan AUC-ROC (96,03%) yang lebih unggul, menandakan keandalannya saat memprediksi "Bullying", akurasi total menjadi metrik yang lebih penting untuk aplikasi deteksi dini. Kemampuan BiLSTM memproses teks dari dua arah terbukti sangat efektif untuk menangkap dependensi kata dalam kalimat yang krusial untuk deteksi *bullying*.

Analisis kesalahan pada model BiLSTM menunjukkan beberapa tantangan, seperti kesulitan mendeteksi hinaan implisit tanpa kata makian yang jelas atau memahami konteks kalimat yang tidak utuh. Investigasi terhadap sampel *False Positive* (salah deteksi sebagai *bullying*) dan *False Negative* (*bullying* yang terlewat) dari data uji mengidentifikasi beberapa pola kesalahan yang menjadi tantangan utama, seperti yang dirangkum pada Tabel 2.



Tabel 2. Contoh Analisis Kesalahan pada Model BiLSTM

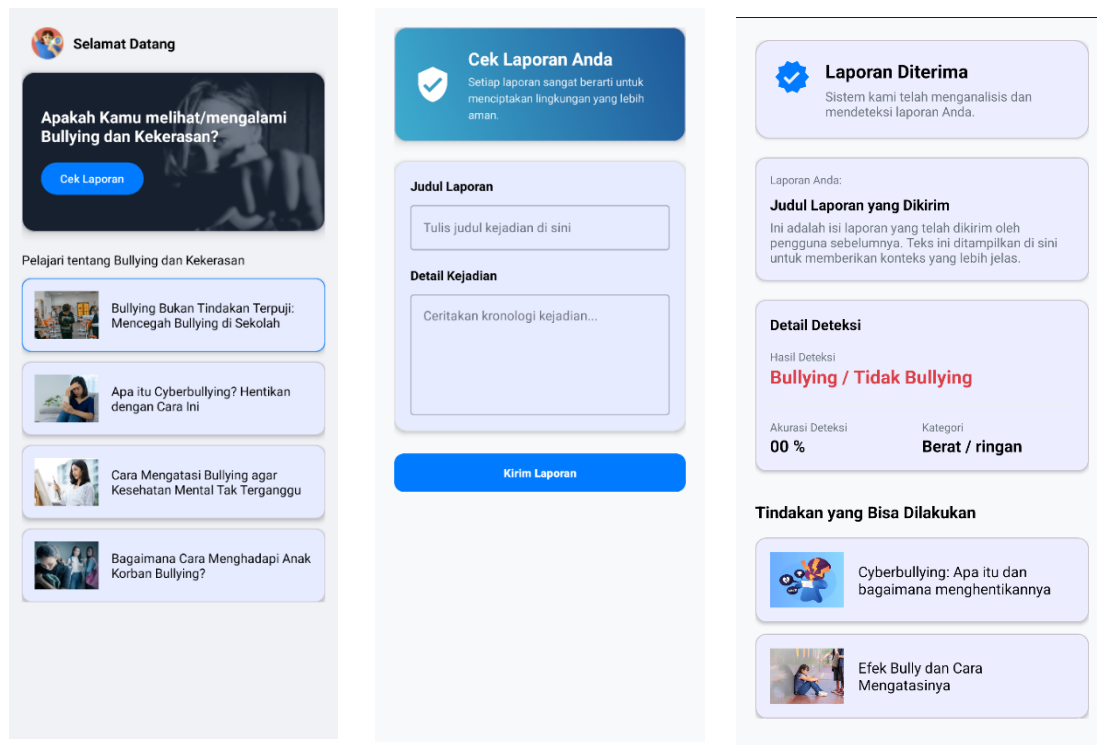
Tipe Kesalahan	Contoh Teks dari Data Uji	Pola Kesalahan (Analisis Awal)	Penjelasan Detail
False Negative (Bullying terlewat)	"Muka sama kelakuan sama sama buruk"	Hinaan Implisit	Model tidak menangkap hinaan deskriptif karena tidak ada kata makian yang jelas sebagai pemicu.
False Negative (Bullying terlewat)	"otak nya ga di pake buat mikir sih jadi nya tolol"	Konteks Kalimat Tidak Utuh	Gagal memahami frasa "otak ga dipake" sebagai penguat hinaan, meskipun terdapat kata kasar.
False Positive (Salah deteksi)	"ih kakak kelas sok cantik deh"	Sentimen Negatif Kuat	Ekspresi sentimen negatif yang kuat keliru diinterpretasikan sebagai perundungan, bukan sekadar cibiran biasa.

Dari tabel di atas, terlihat jelas bahwa keterbatasan utama model terletak pada pemahaman nuansa bahasa yang kompleks. Pada kasus *False Negative*, model menunjukkan ketergantungan pada kata kunci eksplisit (makian) dan kesulitan mengenali perundungan yang bersifat implisit atau sarkastik. Sebaliknya, pada kasus *False Positive*, model cenderung terlalu sensitif terhadap sentimen negatif yang kuat, sehingga dapat salah mengklasifikasikan kritik atau cibiran biasa sebagai tindakan perundungan. Keterbatasan ini menyoroti betapa krusialnya pengembangan dataset di masa depan yang divalidasi oleh ahli psikologi dan bahasa untuk menangani konteks kalimat yang lebih rumit, sehingga dapat meningkatkan keandalan deteksi secara signifikan.

Pemilihan model akhir untuk aplikasi *mobile* harus menyeimbangkan akurasi, keandalan, dan efisiensi komputasi. Model BiLSTM dipilih karena keunggulannya dalam akurasi total (90,00%) dan arsitekturnya yang secara inheren lebih ringan dibandingkan model *Transformer* besar seperti IndoBERT. Efisiensi ini krusial untuk memastikan aplikasi tetap responsif dan tidak menguras daya baterai pada perangkat seluler. Dengan demikian, model BiLSTM *baseline* dipilih sebagai model final untuk implementasi.



Model ini kemudian berhasil diimplementasikan ke dalam aplikasi *mobile* fungsional. Prosesnya meliputi konversi model ke format TensorFlow Lite, merancang arsitektur aplikasi *native* Android, dan mengembangkan antarmuka pengguna yang intuitif. Pengguna dapat memasukkan teks, yang kemudian diproses oleh aplikasi untuk memberikan hasil deteksi secara *real-time* beserta rekomendasi tindakan yang bisa dilakukan.



Gambar 2. Tampilan Antarmuka Aplikasi Mobile (A) Halaman Awal, (B) Halaman Lapor, (C) Halaman Hasil

KESIMPULAN

Berdasarkan hasil penelitian, kesimpulan utama yang dapat ditarik adalah bahwa model BiLSTM (*Bidirectional Long Short-Term Memory*) terbukti menjadi arsitektur paling efektif untuk mendeteksi teks *bullying* berbahasa Indonesia, dengan mencapai akurasi tertinggi sebesar 90,00%. Keunggulan ini berasal dari kemampuan BiLSTM untuk memproses informasi kontekstual dari dua arah, yang sangat krusial untuk memahami nuansa dalam kalimat. Meskipun model IndoBERT menunjukkan presisi yang tinggi, BiLSTM dipilih sebagai model final karena menawarkan keseimbangan terbaik antara akurasi, efisiensi



komputasi, dan kesesuaian untuk implementasi pada platform *mobile* yang memiliki sumber daya terbatas.

Puncak dari penelitian ini adalah keberhasilan implementasi model tersebut ke dalam aplikasi *mobile* fungsional yang mampu melakukan deteksi secara *real-time*, sehingga mengubah hasil studi teoritis menjadi sebuah alat intervensi dini yang praktis. Untuk pengembangan selanjutnya, disarankan agar fokus diarahkan pada peningkatan kapabilitas model dan fungsionalitas aplikasi. Kemampuan model dapat ditingkatkan dengan memperkaya *dataset* menggunakan validasi dari ahli psikologi dan bahasa untuk menangani kasus yang lebih kompleks seperti hinaan implisit atau sarkasme. Dari sisi aplikasi, penambahan fitur yang berorientasi pada pengguna seperti tombol pelaporan anonim atau integrasi dengan sumber daya bantuan profesional dapat secara signifikan meningkatkan dampak dan manfaatnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang tulus kepada Ibu Khanun Roisatul Ummah, S.Tr.T., M.Tr.Kom., selaku dosen pembimbing, atas segala bimbingan, arahan, dan dukungan yang diberikan selama proses penelitian hingga penyusunan naskah ilmiah ini. Ucapan terima kasih juga ditujukan kepada seluruh dewan penguji atas masukan dan saran yang membangun. Terakhir, penulis berterima kasih kepada pihak Program Studi Teknik Informatika, Fakultas Komunikasi dan Informatika, Universitas Muhammadiyah Surakarta, yang telah menyediakan fasilitas dan lingkungan akademis yang mendukung terselesainya penelitian ini.



DAFTAR PUSTAKA

- [1] A. Siwi, F. Utami, and N. Baiti, "Pengaruh Media Sosial Terhadap Perilaku Cyber Bullying Pada Kalangan RSiwi, A., Utami, F., & Baiti, N. (2018). Pengaruh Media Sosial Terhadap Perilaku Cyber Bullying Pada Kalangan Remaja. 18(2), 257–262.emaja," vol. 18, no. 2, pp. 257–262, 2018, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/cakrawala%0APengaruh>
- [2] A. Efaningrum, S. I. A. Dwiningrum, and R. Nurhayati, "Cyberbullying pelajar SMA di media sosial: Prevalnsi dan rekomendasi," *J. Pembang. Pendidik. Fondasi dan Apl.*, vol. 8, no. 1, pp. 144–153, 2021, doi: 10.21831/jppfa.v8i2.38300.
- [3] B. Rusyidi, "Memahami Cyberbullying Di Kalangan Remaja," *J. Kolaborasi Resolusi Konflik*, vol. 2, no. 2, p. 100, 2020, doi: 10.24198/jkrk.v2i2.29118.
- [4] W. Widayat, "Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 1018, 2021, doi: 10.30865/mib.v5i3.3111.
- [5] A. Jadhav Patil, M. Nakhate, R. Bulani, R. Verma, A. Jain, and A. Professor, "Realtime Cyberbullying Detection Using MI and Nlp," *Int. J. Creat. Res. Thoughts*, vol. 11, no. 5, pp. 2320–2882, 2023, [Online]. Available: www.ijcrt.org
- [6] M. A. Amrustian, W. Widayat, and A. M. Wirawan, "Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 535, 2022, doi: 10.30865/mib.v6i1.3527.
- [7] A. Dewani, M. A. Memon, and S. Bhatti, "Cyberbullying detection: advanced preprocessing techniques & deep learning architecture for Roman Urdu data," *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00550-7.
- [8] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. S. T. Ayu, and W. F. Dicka, "Dataset Indonesia untuk Analisis Sentimen," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, p. 334, 2019, doi: 10.22146/jnteti.v8i4.533.
- [9] A. A. Pramudji, B. O. Oktavianus, R. Dwi, and H. L. H. S. Warnars, "Aplikasi Smartphone Untuk Anti Bully," *Petir*, vol. 12, no. 2, pp. 271–280, 2019, doi: 10.33322/petir.v12i2.601.
- [10] W. Athira Luqyana, I. Cholissodin, and R. S. Perdana, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4704–4713, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] H. Imaduddin, F. Y. A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach," *Int. J. Adv. Comput. Sci.*



- Appl.*, vol. 14, no. 8, pp. 113–117, 2023, doi: 10.14569/IJACSA.2023.0140813.
- [12] M. A. Palomino and F. Aider, “Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis,” *Appl. Sci.*, vol. 12, no. 17, 2022, doi: 10.3390/app12178765.
- [13] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, “Big data preprocessing: methods and prospects,” *Big Data Anal.*, vol. 1, no. 1, pp. 1–22, 2016, doi: 10.1186/s41044-016-0014-0.
- [14] Merinda Lestandy and Abdurrahim, “Effect of Word2Vec Weighting with CNN-BiLSTM Model on Emotion Classification,” *J. Nas. Pendidik. Tek. Inform.*, vol. 12, no. 1, pp. 99–107, 2023, doi: 10.23887/janapati.v12i1.58571.
- [15] W. Darmawan, A. P. Wibowo, and B. Ismanto, “Klasifikasi Teks Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Forward Selection,” *IC-Tech J. Inform. Comput. Technol.*, vol. 16, no. 1, pp. 1–7, 2021, [Online]. Available: <https://ejournal.stmik-wp.ac.id/index.php/ictech/article/view/151>
- [16] G. Noer, “Implementasi Algoritma Naïve Bayes dan TF-IDF Dalam Analisis Sentimen Data Ulasan (Studi Kasus: Ulasan Review Aplikasi E-commerce Shopee di Situs Google ...,” *Repository.Uinjkt.Ac.Id*, 2023, [Online]. Available: [https://repository.uinjkt.ac.id/dspace/handle/123456789/68747%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/68747/1/GERALD HALIM AL RASYID NOER-FST.pdf](https://repository.uinjkt.ac.id/dspace/handle/123456789/68747%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/68747/1/GERALD%20HALIM%20AL%20RASYID%20NOER-FST.pdf)

