Repeater: Publikasi Teknik Informatika dan Jaringan
Volume. 3 Nomor. 3 Juli 2025
e-ISSN : 3046-7276, p-ISSN : 3046-7284, Hal 105-121
DOI: https://doi.org/10.62951/repeater.v3i1.382
Available online at: https://journal.aptii.or.id/index.php/Repeater

Deteksi *Cyberbullying* pada Pemain Sepak Bola di Platform Media Sosial "X" Menggunakan Metode *Long Short-Term Memory* (*LSTM*)

Pawit Widiyantoro[1], Paradise[2*], Yogo Dwi Prasetyo[2]
[1-3]Universitas Telkom Purwokerto, Indonesia

Alamat: Jalan D. I Panjaitan No. 128 Purwokerto Jawa Tengah Indonesia
Korespondensi penulis: paradise@ittelkom-pwt.ac.id *

Abstract. *Social media has become a crucial part of modern life around the globe, providing users with various conveniences. However, its widespread use has also brought about new challenges, one of which is cyberbullying. This harmful issue can have serious emotional and physical effects on those targeted. Cyberbullying occurs in many areas, including sports, and soccer—a sport loved by millions—is no exception. Soccer players often face severe criticism, hate speech, and harassment on social media platforms. To tackle this problem, this study aims to create a strong model for detecting cyberbullying on the social media platform "X" using the Long Short-Term Memory (LSTM) method. By utilizing advanced machine learning techniques, the proposed model intends to identify and reduce instances of cyberbullying, helping to create a safer online space for athletes and the wider community.*

Keywords: *Cyberbulliying Detecttion, Football Player, Social Media, Long Short-Term Memory, LSTM.*

Abstrak. Media sosial telah menjadi bagian yang sangat penting dalam kehidupan modern di seluruh dunia, memberikan berbagai kemudahan bagi penggunanya. Namun, penggunaan yang meluas ini juga membawa tantangan baru, salah satunya adalah *cyberbullying*. Masalah serius ini dapat memberikan dampak yang signifikan, baik secara emosional maupun fisik, bagi mereka yang menjadi korban. *Cyberbullying* dapat terjadi di berbagai bidang, termasuk olahraga, dan sepak bola yang merupakan olahraga yang dicintai oleh jutaan orang. Pemain sepak bola sering kali menghadapi kritik tajam, ujaran kebencian, dan pelecehan di platform media sosial. Untuk mengatasi masalah ini, penelitian ini bertujuan untuk mengembangkan model yang efektif dalam mendeteksi *cyberbullying* di platform media sosial "X" dengan menggunakan metode *Long Short-Term Memory (LSTM)*. Dengan memanfaatkan teknik pembelajaran mesin yang canggih, model yang diusulkan ini bertujuan untuk mengidentifikasi dan mengurangi kasus *cyberbullying*, sehingga membantu menciptakan ruang daring yang lebih aman bagi atlet dan masyarakat secara umum.

Kata kunci: *Deteksi Cyberbullying*, Pemain Sepak Bola, Media Sosial, *Long Short-Term Memory*, LSTM.

1. LATAR BELAKANG

Di era digital saat ini, media sosial telah menjadi bagian integral dari kehidupan sehari-hari masyarakat global. Pada April 2024, tercatat lebih dari 5 miliar orang di seluruh dunia menggunakan media sosial (Datareportal, 2024). Platform ini memungkinkan pengguna untuk berkomunikasi secara bebas tanpa batasan ruang dan waktu, menciptakan koneksi yang lebih cepat dan efisien (Akhter et al., 2019). Namun, di balik kemudahan tersebut, media sosial juga membawa tantangan baru, salah satunya adalah *cyberbullying* (K.D. Gorro et al., 2018).

Cyberbullying merujuk pada tindakan mengancam atau melecehkan seseorang melalui media komunikasi modern seperti media sosial. Bentuknya bisa bervariasi, mulai dari rasisme, pelecehan seksual, penghinaan fisik, penyebaran rumor palsu, hingga pesan atau gambar yang menyakiti korban (Nikmah et al., 2022; Perera & Fernando, 2021; R. R. Dalvi et al., 2020).

Received: November 13, 2024; Revised: Desember 15, 2024; Accepted: Januari 18, 2025; Published: Januari 31, 2025

Dampaknya tidak hanya terbatas pada aspek emosional, tetapi juga dapat memengaruhi kesehatan fisik korban. Banyak korban mengalami penurunan kepercayaan diri, ketakutan berlebih, bahkan trauma mendalam. Dalam kasus ekstrem, *cyberbullying* dapat berujung pada tindakan bunuh diri (Alsaed & Eleyan, 2021). Salah satu platform yang sering menjadi tempat terjadinya *cyberbullying* adalah "X" (sebelumnya dikenal sebagai Twitter). Hingga April 2024, platform ini memiliki sekitar 611 juta pengguna aktif dan menempati peringkat ke-12 sebagai media sosial paling banyak digunakan di dunia (Datareportal, 2024). "X" merupakan platform *microblogging* dengan batas karakter *tweet* maksimal 280 huruf [9]. Sayangnya, platform ini juga masuk dalam daftar media sosial dengan persentase tertinggi kasus *cyberbullying*, bersama dengan Facebook, YouTube, dan Instagram (GilPress, 2023).

Tidak hanya sebagai individu biasa, dalam dunia olahraga termasuk sepak bola juga tidak luput dari ancaman *cyberbullying*. Pemain sepak bola sering kali menjadi sasaran kritik pedas, ujaran rasis, dan pelecehan, terutama setelah tim mereka mengalami kekalahan (Arimoro & Elgujja, 2019). Sebagai contoh, analisis Ofcom pada Agustus 2022 menunjukkan bahwa dari 2,4 juta tweet selama paruh pertama musim Premier League 2021/2022, sekitar 60.000 di antaranya bersifat abusif dan menyasar pemain. Bahkan, 12 pemain menerima separuh dari total *tweet* abusif tersebut, dengan rata-rata 15 *tweet* abusif setiap hari (Ofcom & The Alan Turing Institute, 2022). Kasus lainnya terjadi saat Marcus Rashford, Jadon Sancho, dan Bukayo Saka gagal mengeksekusi penalti di final Euro 2020. Ketiga pemain kulit hitam itu menerima hinaan rasis di media sosial setelah kekalahan Inggris. Meskipun "X" berhasil menghapus lebih dari 1.000 tweet menggunakan *Artificial Intelligence (AI)* dan moderator manusia, fenomena ini menunjukkan bahwa masalah *cyberbullying* belum sepenuhnya teratasi (de Souza Dias & Thapa, 2021). Sehingga untuk mengatasi masalah ini, deteksi *cyberbullying* menjadi langkah penting. Deteksi ini dilakukan dengan mengidentifikasi konten abusif berdasarkan fitur tertentu seperti semantik, sintaksis, dan sentimen (Muneer & Fati, 2020). Proses ini umumnya menggunakan pendekatan *Natural Language Processing (NLP)* dan *Machine Learning* (Alsaed & Eleyan, 2021). NLP sangat cocok untuk masalah ini karena fokusnya pada analisis teks dibandingkan konten visual (Alsaed & Eleyan, 2021; R. R. Dalvi et al., 2020). Model yang dikembangkan bertujuan untuk mengklasifikasikan apakah suatu kalimat termasuk *cyberbullying* atau *non-cyberbullying.* Metode yang sering digunakan dalam penelitian mengenai *cyberbullying* ini meliputi model machine learning tradisional seperti *Naïve Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN)*, dan *Decision Tree* (Nandakumar, 2018; Setiawan et al., 2022). Namun, belakangan ini, model deep learning

seperti *Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM)*, dan *Dynamic CNN* menunjukkan hasil yang lebih baik (N Murthy et al., 2020).

Di antara model *deep learning* lainnya, LSTM lebih populer karena kemampuannya dalam memproses data teks secara sekuensial dan mengatasi masalah *long-term dependencies* yang sering dialami oleh RNN tradisional (Huang et al., 2022; N Murthy et al., 2020; Wang et al., 2018). LSTM dirancang untuk mengingat hubungan antar kata meskipun jaraknya cukup jauh, sehingga sangat cocok untuk tugas klasifikasi teks (N Murthy et al., 2020; Rehman et al., 2019; Srinivas et al., 2021). Beberapa studi menunjukkan bahwa LSTM mampu mengungguli model *machine learning* tradisional dan model *deep learning* lainnya dalam berbagai aplikasi text classification (Monika et al., 2019; Muhammad et al., 2021; N Murthy et al., 2020; Srinivas et al., 2021; Zhou et al., 2019). Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan metode LSTM dalam mendeteksi *cyberbullying* terhadap pemain sepak bola di platform media sosial "X". Penelitian ini berjudul "Deteksi *Cyberbullying* pada Pemain Sepak Bola di Platform Media Sosial 'X' Menggunakan Metode *Long Short-Term Memory (LSTM)*".

2. KAJIAN TEORITIS

Cyberbullying di media sosial telah menjadi topik penelitian yang semakin mendapat perhatian dalam beberapa tahun terakhir. Sejumlah studi telah dilakukan untuk mengembangkan metode deteksi yang lebih akurat dengan berbagai pendekatan, termasuk *machine learning* dan *deep learning*. Kajian terhadap penelitian terdahulu sangat penting untuk memahami tren, metodologi, serta tantangan dalam mendeteksi *cyberbullying* secara efektif. Pada bagian ini, akan diulas beberapa penelitian yang berkontribusi dalam bidang ini, dengan fokus pada penggunaan model *machine learning* dan *deep learning* dalam deteksi *cyberbullying* di media sosial, khususnya platform "X".

Penelitian Z. Alsaed dan D. Eleyan (Alsaed & Eleyan, 2021) melakukan survei terhadap berbagai penelitian yang membahas deteksi *cyberbullying* di media sosial. Studi ini menemukan bahwa beberapa platform digunakan sebagai sumber data, di antaranya Twitter ("X") dengan lima penelitian, Facebook dengan tiga penelitian, Instagram dengan dua penelitian, serta satu penelitian yang menggunakan dataset dari Kaggle. Model machine learning yang sering digunakan dalam penelitian ini meliputi SVM (Support Vector Machine) yang digunakan dalam sembilan penelitian dan Naïve Bayes dalam lima penelitian lainnya. Evaluasi model dilakukan menggunakan metrik seperti accuracy, precision, recall, dan F1-

score. Studi ini memberikan gambaran komprehensif tentang tren dan metode dalam deteksi *cyberbullying* di berbagai platform media sosial.

Akhter et al. (Akhter et al., 2019) mengusulkan metode deteksi dan klasifikasi konten *cyberbullying* menggunakan Multinomial Naïve Bayes dan Logika Fuzzy. Cyberbullying dikategorikan menjadi beberapa jenis, seperti pelecehan dan rasisme. Dataset yang digunakan berasal dari Kaggle dan berisi data dari Facebook. Model ini menggunakan tiga fitur utama: umur pengguna, jumlah kata kasar, dan tingkat intensitas bully berdasarkan aturan fuzzy. Model yang diusulkan berhasil mencapai akurasi 88,76%, lebih baik dibandingkan metode sebelumnya menggunakan SVM dengan akurasi 76,38%. Penelitian R.R. Dalvi et al. [5] mengembangkan software untuk mendeteksi postingan *cyberbullying* di Twitter dengan menggunakan teknik feature extraction TF-IDF. Proses modeling dilakukan menggunakan SVM dan Naïve Bayes. Jika probabilitas postingan di atas 0,5, *tweet* akan dimasukkan ke dalam database, dan dilakukan crawling tambahan terhadap sepuluh *tweet* dari pengguna yang sama untuk memastikan pola cyberbullying. Hasil penelitian menunjukkan bahwa SVM mendapatkan akurasi lebih tinggi (71,25%) dibandingkan Naïve Bayes (52,70%).

Seiring berkembangnya teknik *deep learning*, model berbasis neural *network* mulai digunakan dalam penelitian deteksi *cyberbullying*. Desai et al. (Desai et al., 2021) menggunakan model pre-trained BERT dengan lima fitur utama: sentimen, sarkasme, sintaksis, semantik, dan sosial. Dataset penelitian ini berasal dari kumpulan *tweet* yang diklasifikasikan sebagai bullying dan non-bullying. Model BERT menunjukkan akurasi 91,90%, mengungguli SVM (71,25%) dan Naïve Bayes (52,70%), yang menunjukkan keunggulan model deep learning dibandingkan machine learning tradisional.

Maslej-Kresnakova et al. (Maslej-Krešňáková et al., 2020) membandingkan beberapa teknik *preprocessing* seperti TF-IDF, Word2Vec, GloVe, FastText, dan BERT dalam deteksi cyberbullying. Model BiLSTM+CNN dipilih setelah hasil evaluasi menunjukkan performa terbaik. Hasil penelitian menunjukkan bahwa model BiLSTM+CNN dengan word embedding GloVe tanpa *preprocessing* standar mencapai akurasi rata-rata 0,9725, hampir setara dengan model transformer BERT. Studi ini menekankan pentingnya pemilihan teknik text representation dalam meningkatkan akurasi deteksi *cyberbullying*.

Penelitian Wang et al. (Wang et al., 2018) meneliti efektivitas LSTM dengan word embedding Word2Vec dalam analisis sentimen. Studi ini menggunakan dataset review film dari IMDB, Douban, dan komentar dari media sosial Taiwan, PTT. Dibandingkan dengan Naïve Bayes dan Extreme Machine Learning (ELM), model LSTM menunjukkan hasil yang lebih baik dalam beberapa dataset, terutama saat jumlah data meningkat. Hasil ini

menunjukkan bahwa model deep learning seperti LSTM dapat meningkatkan efektivitas analisis sentimen dibandingkan dengan metode tradisional.

Monika et al. (Monika et al., 2019) mengklasifikasikan tweet tentang maskapai penerbangan di Amerika Serikat berdasarkan sentimen layanan penerbangan. Dua model word embedding, Word2Vec dan GloVe, digunakan dalam kombinasi dengan RNN dan LSTM. Hasil eksperimen menunjukkan peningkatan akurasi dengan data latih 80% dan data uji 20%, serta penurunan nilai loss setelah 14 epoch pelatihan. Murthy et al. (N Murthy et al., 2020) juga melakukan analisis sentimen menggunakan LSTM dengan dataset review film dari IMDB. Hasil penelitian menunjukkan peningkatan akurasi selama 10 epoch pelatihan, menandakan model mampu belajar pola data dengan baik. Srinivas et al. (Srinivas et al., 2021) membandingkan tiga model deep learning—simple neural network, CNN, dan LSTM—untuk analisis sentimen data dari Twitter. Dataset dibagi dalam rasio 80:20 untuk pelatihan dan pengujian. Parameter yang digunakan termasuk batch size = 18, epoch = 128, ReLU sebagai activation function, sigmoid untuk output layer, Adam optimizer, dan dropout = 5. Model LSTM berhasil mengungguli CNN dan simple neural network dengan akurasi 87%.

Muhammad et al. (Muhammad et al., 2021) melakukan analisis sentimen terhadap review hotel dalam bahasa Indonesia menggunakan LSTM dan Word2Vec (skip-gram). Dataset yang diperoleh melalui web scraping dari Traveloka berjumlah 2.500 data. Model LSTM dengan dropout 0,2, average pooling, dan learning rate 0,001 menghasilkan akurasi 85,96%, menunjukkan efektivitas pendekatan deep learning dalam analisis sentimen bahasa Indonesia.

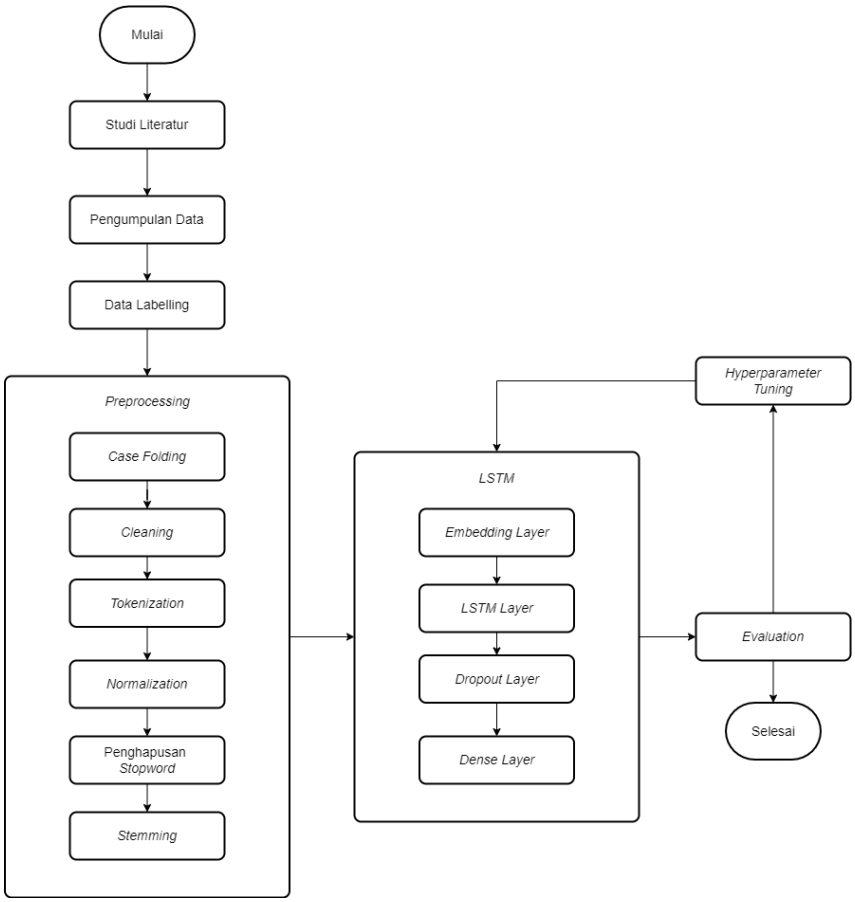
Cyberbullying dalam dunia sepak bola juga menjadi perhatian akademik. Penelitian Arimoro A. dan Dr. Elgujja A. (Arimoro & Elgujja, 2019) membahas penyalahgunaan media sosial oleh fans sepak bola untuk menyerang pemain secara verbal. Studi ini mengungkapkan bahwa platform media sosial tidak secara otomatis mengklasifikasikan ujaran kebencian sebagai *cyberbullying* kecuali terkait dengan agama atau orientasi seksual. Oleh karena itu, mereka merekomendasikan regulasi lebih ketat untuk menangani serangan verbal terhadap pemain sepak bola di media sosial.

Berdasarkan penelitian terdahulu, dapat disimpulkan bahwa pendekatan deep learning, khususnya LSTM, menunjukkan hasil yang menjanjikan dalam deteksi cyberbullying dibandingkan metode machine learning tradisional. Penggunaan word embedding seperti Word2Vec, GloVe, dan BERT juga memberikan dampak signifikan terhadap performa model. Studi yang berfokus pada cyberbullying terhadap pemain sepak bola masih terbatas, sehingga penelitian ini bertujuan untuk mengembangkan sistem deteksi cyberbullying yang lebih akurat

dengan menggunakan metode LSTM pada platform "X". Oleh karena itu, penelitian ini berkontribusi dalam memperkaya literatur tentang deteksi *cyberbullying*, khususnya dalam konteks olahraga, serta mengusulkan pendekatan berbasis deep learning untuk meningkatkan efektivitas deteksi ujaran kebencian terhadap pemain sepak bola.

3. METODE PENELITIAN

Penelitian ini bertujuan untuk mendeteksi *cyberbullying* terhadap pemain sepak bola di platform media sosial "X" menggunakan metode *Long Short-Term Memory (LSTM).* Subjek penelitian adalah para pemain sepak bola yang menjadi bahan perbincangan di kalangan pengguna platform "X", dengan fokus pada interaksi berupa komentar yang ditujukan kepada mereka. Selain itu, objek penelitian adalah dataset komentar yang diperoleh melalui proses *crawling* dari media sosial "X". Dataset ini mencakup berbagai jenis komentar, baik yang mengandung unsur *cyberbullying* maupun yang tidak, sehingga dapat digunakan untuk melatih dan menguji model LSTM. Pendekatan ini diharapkan mampu menghasilkan model yang efektif dalam mengklasifikasikan konten abusif secara akurat. Langkah-langkah metodologi yang digunakan dalam penelitian ini, mulai dari pengumpulan data hingga evaluasi model disajikan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian Deteksi *Cyberbullying* Pada Pemain Bola dengan LSTM

Studi Literatur

Langkah pertama yang dilakukan adalah pencarian dan pengkajian literatur penelitian-penelitian terdahulu yang sesuai dengan topik penelitian yang diambil. Dikumpulkan berbagai literatur dalam bentuk artikel, jurnal dan prosiding baik nasional maupun internasional yang didapatkan melalui situs Google Scholar. Dari semua literatur yang berhasil dikumpulkan, beberapa dimasukkan ke dalam tabel kajian pustaka sebagai referensi utama dari penelitian ini.

Pengumpulan Data

Pengumpulan data dilakukan dengan cara *crawling* data dari media sosial "X". Data dikumpulkan menggunakan *tool Tweet-Harvest* dengan memanfaatkan token dari akun media sosial "X". Data berisi *tweet cyberbullying* terhadap pemain sepak bola nasional dan internasional dalam berbagai bahasa. Digunakan *keyword* pemain yang sedang terlibat kontroversi sebagai data untuk kelas *cyberbullying* dan pemain yang sedang menuai pujian untuk kelas *non-cyberbullying*. Selain menggunakan bantuan *keyword*, proses *crawling* juga dilakukan dengan mengambil komentar dari beberapa postingan tentang pemain yang menjadi *keyword* dari akun media sosial "X" yang membahas tentang sepak bola. Data hasil *crawling* kemudian di-*filter* agar hanya data dalam Bahasa Indonesia yang akan disertakan.

Data *Labelling*

Proses *labelling* dilakukan secara manual untuk menentukan *tweet* yang termasuk dalam kelas *cyberbullying* dan *non-cyberbullying*. Kata-kata kasar yang merujuk pada perilaku *cyberbullying* didapatkan dari kamus terminologi kasar yang berasal dari beberapa sumber seperti Kaggle, Github, dan *website* KBBI. Beberapa contoh kata yang dianggap kasar adalah "tolol", "anjing", "babi", "bego", dan lain sebagainya. Tweet yang berisi *cyberbullying* diberi label 1 sementara tweet dengan kelas non *cyberbullying* diberi label 0.

Data *Preprocessing*

Data yang telah memiliki label kemudian akan memasuki tahap *data pre-processing*. *Pre-processing* akan sangat berpengaruh pada hasil pemrosesan karena dapat mengoptimalkan hasil pemrosesan (Candra & Rozana, 2020). Pada tahap ini, dilakukan beberapa langkah *pre-processing* untuk menyiapkan data teks agar lebih rapi dan siap digunakan dalam melatih model. Langkah pertama adalah *Case Folding*, yaitu mengubah semua huruf menjadi huruf kecil agar format teks seragam. Selanjutnya, proses *Cleaning* dilakukan untuk membersihkan teks dari elemen-elemen yang tidak diperlukan, seperti tanda baca, angka, atau simbol-simbol lain yang tidak relevan. Setelah itu, teks dipecah menjadi bagian-bagian kecil atau kata-kata melalui proses *Tokenization* .

Untuk menyamakan bentuk kata-kata yang tidak baku, seperti singkatan atau kesalahan penulisan, dilakukan proses *Normalization* . Kemudian, kata-kata umum yang tidak memberikan makna penting dalam analisis, seperti "dan", "atau", dan "yang", dihapus melalui tahap Penghapusan *Stopwords* . Terakhir, *Stemming* digunakan untuk mengubah kata-kata menjadi bentuk dasarnya sehingga variasi kata yang mirip dapat dikelompokkan dengan lebih efisien. Semua langkah ini bertujuan untuk memastikan data yang digunakan lebih berkualitas, sehingga model dapat bekerja lebih optimal dalam mendeteksi *cyberbullying*.

Data *Splitting*

Proporsi 8:1:1 digunakan sebagai konfigurasi default dalam membagi data, karena banyak penelitian sebelumnya (Gill & Khehra, 2021; Monika et al., 2019; Nikmah et al., 2022; Srinivas et al., 2021) menunjukkan bahwa skema pembagian 8:2 (80% untuk training dan 20% untuk testing) menghasilkan akurasi yang baik. Namun, dalam penelitian ini, data dibagi menjadi tiga kelompok: training, validation, dan testing. Untuk memenuhi kebutuhan tersebut, proporsi data testing yang sebelumnya 20% dibagi menjadi dua bagian yang sama besar, menghasilkan 10% untuk data validation dan 10% untuk data testing. Pendekatan ini memungkinkan model untuk dilatih menggunakan sebagian besar data (80%), sementara 10% data digunakan untuk memvalidasi performa model selama pelatihan, sehingga membantu mencegah overfitting. Sisanya, 10%, digunakan sebagai data testing untuk mengevaluasi kemampuan model pada data yang belum pernah dilihat sebelumnya, memastikan bahwa performa model dapat digeneralisasikan dengan baik. Skema pembagian ini dirancang untuk memberikan keseimbangan yang baik antara pelatihan, validasi, dan evaluasi model.

Modelling

Data kemudian masuk ke dalam model *LSTM* yang terdiri atas empat *layer* yaitu *embedding layer*, *LSTM layer*, *dropout layer,* dan *dense layer*. *Layer embedding* berfungsi untuk mengkonversi teks menjadi vektor. Kemudian, *layer LSTM* berfungsi untuk mencari hubungan antar urutan kata dalam kalimat menggunakan bobot dan bias. *Layer dropout* kemudian ditambahkan untuk mengurangi risiko *overfitting*. Dan terakhir, *layer dense* akan memberikan output menggunakan *activation function* yaitu sigmoid, sigmoid digunakan pada karena penelitian mengambil kasus *binary classification*. Terdapat beberapa parameter yang digunakan seperti jumlah layer dan unit LSTM, jumlah *epoch*, *batch size*, jenis *optimizer*, *learning rate*, *dropout rate*, dan lain lain. Terdapat pula teknik *early stopping* yang sering digunakan untuk meminimalisir *overfitting* sehingga *training* akan berhenti ketika kurva peningkatan *accuracy* mulai turun (Nikmah et al., 2022).

Evaluation

Pada tahap ini model akan dievaluasi menggunakan nilai *accuracy, precision, recall, f-1 score,* dan juga *loss* nya. Jika nilai *accuracy, precision, recall, dan f1-score* setiap *epoch* semakin naik dan *loss* semakin turun maka model mampu mempelajari data dengan baik. Sebaliknya jika nilai *accuracy, precision, recall, f1-score,* dan *loss* konstan maka model tidak terlalu baik dalam mempelajari data.

Hyperparameter Tuning

Hasil evaluasi dalam bentuk *accuracy, precision, recall, f1-score,* dan *loss* tadi kemudian akan ditinjau ulang. Apabila hasil *accuracy, precision, recall, f1-score* masih rendah dan *loss* tidak turun secara signifikan, maka model akan di-*training* ulang dengan perubahan di beberapa parameter. Terdapat beberapa parameter yang dapat diubah mulai dari jumlah *epoch*, *batch size*, jumlah *LSTM unit*, jenis *optimizer*, *learning rate*, *dropout rate*, dan lain lain. *Training* model akan dihentikan jika nilai *accuracy, precision, recall, f1-score* telah mencapai nilai yang tinggi dan *loss* semakin rendah (Nikmah et al., 2022).

4. HASIL DAN PEMBAHASAN

Data yang berhasil dikumpulkan pada penelitian ini berdasarkan rentang waktu yang telah ditetapkan pada batasan masalah. Proses ini berhasil mengumpulkan 9.060 data dalam format *Comma Separated Value (CSV)*. Data hasil proses pengumpulan data dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pengumpulan Data

Text
Next match gak usah lg mainin Rashford & Martial. Drpd kena bully lg...
Martial buang aja lah cuk. Mainin pas lawan Everton aja.

Data pada Tabel 1 kemudian diimplementasikan ke beberapa teknik *pre-processing,* antara lain: *Case Folding, Cleaning, Tokenization, Normalization, Penghapusan Stopwords, dan Stemming* yang hasilnya disajikan pada Tabel 2.

Tabel 2. Hasil *Pre-processing*

Raw text	***Case Folding***	***Cleaning***
@makanbolaID Circelnya asnawi dan temen istrinya arhan itu virus negatif bgt buat perkembangan arhan asnawi di dunia sepakbola"	@makanbolaid circelnya asnawi dan temen istrinya arhan itu virus negatif bgt buat perkembangan arhan asnawi di dunia sepakbola"	circelnya asnawi dan temen istrinya arhan itu virus negatif bgt buat perkembangan arhan asnawi di dunia sepakbola
Tokenization	***Normalization***	***Stemming***
['circelnya', 'asnawi', 'dan', 'temen', 'istrinya', 'arhan', 'itu', 'virus', 'negatif', 'bgt', 'buat', 'perkembangan', 'arhan', 'asnawi', 'di', 'dunia', 'sepakbola']	['circelnya', 'asnawi', 'dan', 'teman', 'istrinya', 'arhan', 'itu', 'virus', 'negatif', 'banget', 'buat', 'perkembangan', 'arhan', 'asnawi', 'di', 'dunia', 'sepakbola']	circelnya asnawi teman istri arhan virus negatif banget kembang arhan asnawi dunia sepakbola

Hasil *pre-processing* data kemudian dilakukan proses *splitting* sesuai dengan proporsi yang ditetapkan untuk selajutnya akan diproses ke tahap selanjutnya yaitu *modelling*. Pada eksperimen awal, model *LSTM* dasar dengan konfigurasi parameter standar diuji untuk memberikan *baseline* bagi eksperimen lanjutan yang melibatkan model dengan parameter yang lebih kompleks. Model ini dibangun menggunakan bahasa pemrograman *python* dengan *framework* tensorflow. Langkah pertama yang dilakukan adalah proses *word embedding* menggunakan GloVe. GloVe menyediakan beberapa ukuran dimensi vektor, yaitu 50, 100, 200, dan 300. Dengan mempertimbangkan efisiensi komputasi, penulis memilih menggunakan 100 dimensi untuk proses *word embedding*, yang dianggap cukup representatif untuk menangkap makna kata tanpa membebani komputasi.

Pemilihan satu *layer* LSTM bertujuan untuk menyederhanakan model agar lebih mudah untuk dievaluasi pada eksperimen awal. Nilai LSTM unit sebesar 64 dipilih setelah mempertimbangkan variasi nilai yang digunakan dalam penelitian lain, yaitu 32, 64, dan 128 (Zhou et al., 2019). Nilai 64 dipilih karena berada pada titik tengah, yang tidak terlalu tinggi maupun terlalu rendah, sehingga memberikan keseimbangan yang baik antara kemampuan model untuk mengekstrak fitur dan kompleksitas model. *Batch size* sebesar 64 dipilih (comparison of lstm, svm, nb), untuk memberikan model *batch size* yang cukup besar tanpa membebani komputasi. Adam *optimizer* dipilih sebagai *optimizer* dasar berdasarkan hasil penelitian terdahulu (Ni & Cao, 2020; Zhou et al., 2019) yang menunjukkan bahwa adam menghasilkan evaluasi yang lebih baik dibanding *optimizer* lainnya. *Learning rate* sebesar 0,001 (Ni & Cao, 2020) yang merupakan nilai *learning rate default* dari *optimizer* adam digunakan karena telah terbukti memberikan hasil yang optimal pada penelitian terdahulu. *Dropout rate* sebesar 0,5 dipilih untuk *dropout layer* setelah mempertimbangkan penelitian terkait yang menunjukkan efektivitas nilai tersebut dalam mengurangi *overfitting* (Indonesian hotel review)(Muhammad et al., 2021). Untuk efisiensi waktu komputasi, *epoch* ditetapkan sebesar 10 agar proses pelatihan dapat diselesaikan lebih cepat, memberikan gambaran awal mengenai performa model tanpa memakan banyak waktu. Dengan konfigurasi parameter ini, model dasar diharapkan memberikan hasil yang cukup baik sebagai titik acuan, sebelum melanjutkan dengan eksperimen yang lebih kompleks untuk meningkatkan performa model. Hasil evaluasi model *LSTM* dasar dapat dilihat pada Tabel 3.

Tabel 3. Hasil evaluasi model dasar

Precision	*Recall*	*F1-Score*	*Loss*	*Accuracy*
36	60	45	0.6428975462913513	59,7%

Hasil evaluasi terhadap model dasar menunjukkan bahwa model masih memiliki nilai *loss* yang relatif tinggi dan akurasi yang rendah. Hal ini mengindikasikan bahwa model belum mampu mempelajari pola data secara optimal. Kemungkinan penyebabnya adalah kompleksitas model yang belum cukup untuk menangkap pola-pola dalam data, terutama jika data memiliki struktur yang kompleks. Selain itu, hasil ini juga dapat disebabkan oleh pemilihan parameter yang kurang sesuai, seperti *learning rate*, jumlah unit LSTM, atau *dropout rate* yang belum optimal untuk dataset yang digunakan.

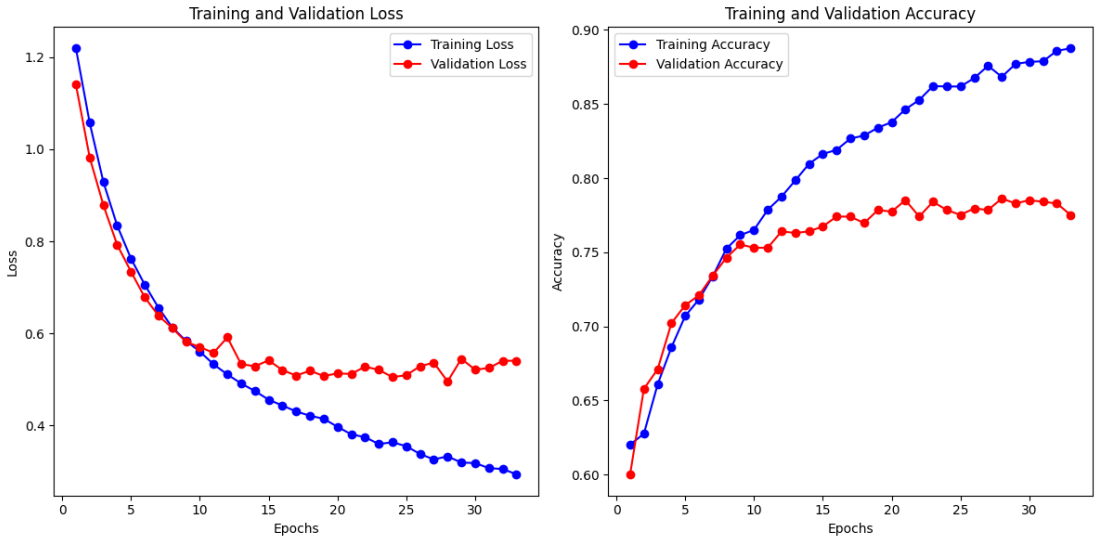
Kondisi ini menunjukkan perlunya dilakukan eksperimen lanjutan untuk meningkatkan performa model, misalnya dengan menyesuaikan arsitektur, menambahkan jumlah *layer*, mengatur ulang paremter seperti *learning rate* dan *batch size*, atau menerapkan teknik regulasi tambahan seperti *L2 regularization* atau dropout untuk membantu model menangkap pola data dengan lebih baik tanpa *overfitting*. Eksperimen tersebut bertujuan untuk menemukan konfigurasi yang lebih sesuai dengan kebutuhan dataset dan kompleksitas pola yang ingin dipelajari.

Karena hasil evaluasi model dasar belum optimal, dilakukan proses *hyperparameter tuning* untuk mengoptimalkan performa model. Pada tahap ini, digunakan metode *trial and error*, yaitu metode pencarian parameter terbaik secara manual melalui berbagai eksperimen. Metode ini dipilih karena memiliki keunggulan dalam hal biaya komputasi yang rendah, meskipun memerlukan waktu yang lebih lama untuk menguji setiap kombinasi parameter secara menyeluruh. Variasi parameter yang diuji mencakup jumlah *layer LSTM*, jumlah unit *LSTM*, *batch size*, jenis *optimizer*, *learning rate*, *dropout rate*, dan jumlah *epoch*

Hasil evaluasi model dapat lebih mudah dipahami dengan menggunakan kurva yang menampilkan perubahan nilai akurasi dan *loss* selama proses *modelling*. Kurva ini memberikan visualisasi yang jelas mengenai performa model pada data *training* dan data *validation*. Dengan menganalisis kurva tersebut, kita dapat mengamati bagaimana model belajar dari data selama proses pelatihan dan mengevaluasi apakah model mengalami *overfitting* atau *underfitting*. Berikut adalah kurva hasil evaluasi yang menunjukkan tren akurasi dan *loss* pada data *training* dan data *validation* untuk setiap *epoch* pelatihan yang dapat dilihat pada Gambar 2.

Dari hasil kurva evaluasi pada Gambar 2, dapat diamati bahwa meskipun model dirancang untuk berjalan selama 100 *epoch*, proses *training* sebenarnya berhenti lebih awal pada *epoch* ke-30. Hal ini disebabkan oleh mekanisme *early stopping,* yang mendeteksi bahwa *validation loss* telah mencapai kondisi stabil yang ditandai dengan melandainya kurva dan

tidak menunjukkan potensi untuk menurun lebih jauh. Oleh karena itu, *training* dihentikan lebih cepat untuk mencegah pemborosan sumber daya dan potensi *overfitting*.



Gambar 2. Hasil evaluasi model akhir

Dari tren kurva pada Gambar 2, terlihat bahwa *loss* pada data *training* dan *validation* terus menurun, menandakan bahwa model dapat belajar dengan baik. Namun, *validation loss* mulai stabil di sekitar *epoch* ke-15, sementara *training loss* terus menurun hingga *training* dihentikan. Hal ini mengindikasikan bahwa model mulai menunjukkan tanda-tanda *overfitting*, di mana model semakin menyesuaikan diri dengan data *training* tetapi tidak meningkatkan kinerja pada data *validation*. Sementara itu, tren akurasi menunjukkan peningkatan pada data *training* maupun data *validation*. Namun, akurasi pada data *validation* mulai melandai di sekitar *epoch* ke-20, sedangkan akurasi data *training* terus meningkat. Ini kembali memperkuat indikasi awal adanya *overfitting*. Berkat mekanisme *early stopping*, model dihentikan sebelum *overfitting* menjadi lebih parah, sehingga performa evaluasi pada data *testing* tetap relatif tinggi. Setelah proses training selesai, performa model dievaluasi menggunakan data testing untuk mengukur kemampuan model dalam menggeneralisasi data yang belum pernah dilihat sebelumbya. Hasil Evaluasi dapat dilihat pada tabel 4.

Tabel 4 Hasil evaluasi dengan 100 *epoch*

Epoch	*Precision*	*Recall*	*F1-Score*	*Loss*	*Accuracy*
100	74	73	72	0.5904784202575684	73,3%

Nilai *precision*, *recall*, *F1-score*, *loss*, dan *accuracy* menunjukkan bahwa model memiliki performa yang cukup baik. Meskipun nilai akurasi pada data testing tidak setinggi pada data *training*, hal ini mengindikasikan bahwa model tidak mengalami *overfitting* yang signifikan, berkat implementasi *early stopping* dan pemilihan parameter yang tepat. Selain itu, hasil evaluasi pada data testing ini menegaskan bahwa model mampu menangkap pola pada

data dengan cukup baik meskipun terdapat tantangan dalam menjaga keseimbangan antara generalisasi dan pembelajaran mendalam.

Selanjutnya, dilakukan proses *undersampling* untuk mengubah dataset yang sebelumnya tidak seimbang (*imbalance*) menjadi lebih seimbang (*balance*). Proses ini dipilih sebagai metode penyesuaian karena *undersampling* memungkinkan penggunaan data asli yang sepenuhnya berasal dari proses pengumpulan data tanpa modifikasi buatan. Jika menggunakan metode *oversampling*, data pada kelas minoritas akan digandakan yang dapat menambahkan elemen data sintetis dan mengurangi kemurnian dataset asli. Dalam proses *undersampling* ini, metode manual digunakan untuk memilih data dari kelas mayoritas, dibandingkan dengan sekadar menghapus data secara acak menggunakan *python*. Pendekatan manual memberikan kontrol yang lebih baik terhadap data yang dipilih, sehingga tetap dapat menjaga representasi distribusi fitur yang relevan dalam kelas mayoritas. Dataset yang awalnya memiliki perbandingan 2:1 antara kelas *non-cyberbullying* (0) dan kelas *cyberbullying* (1), diubah menjadi 1:1 setelah melalui proses *undersampling*. Hasil evaluasi model menggunakan dataset yang seimbang yang menunjukkan pola grafik yang relatif serupa dengan dataset yang tidak seimbang. Namun, terdapat perbedaan penting, yaitu selisih antara performa data *training* dan data *validation* menjadi lebih kecil. Sama seperti pada proses *training* sebelumnya, pelatihan model dihentikan pada *epoch* ke-30 karena mekanisme *early stopping*, yang mendeteksi bahwa *validation loss* telah mencapai kondisi stabil.

Pada data *training*, *loss* terus menurun secara konsisten, sedangkan pada data *validation*, *loss* mulai melandai pada sekitar *epoch* ke-20 dengan nilai berkisar antara 0,5 hingga 0,6. Hal ini menunjukkan bahwa model berhasil belajar dengan baik pada data *training* tanpa kehilangan generalisasi terhadap data *validation*. Dari segi akurasi, grafik menunjukkan peningkatan pada data *training* dan data *validation*. Namun, peningkatan pada data *validation* mulai melandai di sekitar nilai 0,80 sebelum akhirnya proses *modelling* dihentikan pada *epoch* ke-30. Selanjutnya, hasil evaluasi model pada data *testing* dapat dilihat pada tabel 4.

Tabel 5. Hasil evaluasi dengan 100 *epoch* setelah *undersampling*

Dataset	*Precision*	*Recall*	*F1-Score*	*Loss*	*ccuracy*
2:1	74	73	72	0.5904784202575684	73,3%
1:1	**78**	**78**	**78**	**0.5107450485229492**	**78%**

Hasil evaluasi pada data *testing* yang dapat dilihat pada tabel 4 menunjukkan hasil yang lebih baik dibandingkan dengan hasil evaluasi menggunakan dataset yang tidak seimbang. Hal ini mengindikasikan bahwa perbandingan antar kelas dalam dataset memiliki pengaruh signifikan terhadap performa model. Dataset dengan kelas yang seimbang memungkinkan

model untuk bekerja lebih efektif, menghasilkan evaluasi yang lebih akurat, dan meminimalkan bias terhadap salah satu kelas.

Penelitian ini berhasil menunjukkan potensi metode Long Short-Term Memory (LSTM) dalam mendeteksi cyberbullying terhadap pemain sepak bola di platform media sosial "X". Dengan performa akurasi, precision , recall , dan F1-score sebesar 78%, model yang dikembangkan mampu memberikan hasil yang cukup baik dalam mengklasifikasikan konten cyberbullying dan non-cyberbullying. Meskipun demikian, hasil ini juga menunjukkan bahwa masih ada ruang untuk perbaikan, terutama dalam menghadapi variasi bahasa, konteks, serta tantangan lain seperti data yang tidak seimbang. Oleh karena itu, pengembangan lebih lanjut dengan pendekatan teknik preprocessing yang lebih mendalam, representasi teks yang lebih optimal, atau eksplorasi arsitektur model yang lebih canggih dapat menjadi langkah strategis untuk meningkatkan kinerja deteksi. Penelitian ini diharapkan dapat menjadi dasar bagi penelitian mendatang dan berkontribusi dalam menciptakan lingkungan media sosial yang lebih aman bagi para atlet maupun masyarakat umum.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan model berbasis Long Short-Term Memory (LSTM) untuk mendeteksi *cyberbullying* terhadap pemain sepak bola di platform media sosial "X". Model LSTM yang dirancang menggunakan tiga layer dengan jumlah unit masing-masing 128, 128, dan 64, batch size 32, learning rate 0.0001, serta *dropout rate* sebesar 0.2 mampu mencapai hasil evaluasi terbaik. Hasil pengujian menunjukkan nilai akurasi, *precision*, *recall*, dan *F1-score* sebesar 78%, dengan nilai *loss* sebesar 0,5107450485229492. Secara keseluruhan, model ini berhasil memprediksi 2.904 data *non-cyberbullying* dan 2.313 data *cyberbullying* pada data pelatihan, 366 data non-*cyberbullying* dan 286 data *cyberbullying* pada data validasi, serta 349 data non-*cyberbullying* dan 286 data *cyberbullying* pada data pengujian. Dari total data yang dianalisis, terdapat 3.619 data non-*cyberbullying* dan 2.903 data *cyberbullying*, dengan distribusi persentase sebesar 55,48% untuk *non-cyberbullying* dan 44,52% untuk *cyberbullying*. Meskipun model telah menunjukkan performa yang baik, hasil ini masih dapat dioptimalkan lebih lanjut untuk meningkatkan akurasi dan efisiensi deteksi. Untuk penelitian mendatang, ada beberapa saran yang bisa diterapkan untuk meningkatkan model deteksi *cyberbullying*. Pertama, proses *preprocessing* data dapat diperluas dengan normalisasi teks yang lebih teliti, menghapus data yang tidak relevan, atau mengelompokkan data berdasarkan konteks. Kedua, pendekatan representasi teks yang lebih sesuai untuk domain ini dapat membantu model dalam memahami makna semantik dengan lebih baik. Ketiga,

penggunaan arsitektur model yang lebih kompleks atau modern dapat meningkatkan kemampuan model dalam menangkap hubungan dalam teks. Selain itu, menambah dataset yang lebih besar dan bervariasi akan membantu model berfungsi lebih baik dalam menghadapi berbagai konteks bahasa. Dengan langkah-langkah ini, diharapkan model yang dikembangkan di masa depan akan lebih optimal dan bermanfaat bagi masyarakat.

UCAPAN TERIMA KASIH

Jurnal ini merupakan hasil dari skripsi dengan judul skirpsi Deteksi Cyberbullying pada Pemain Sepak Bola di Platform Media Sosial "X" Menggunakan Metode Long Short-Term Memory (LSTM) di Universitas Telkom Purwokerto.

DAFTAR REFERENSI

Akhter, A., Uzzal, K. A., & Polash, M. M. A. (2019). Cyber bullying detection and classification using multinomial naïve Bayes and fuzzy logic. *International Journal of Mathematical Sciences and Computing, 5*(4), 1–12.

Alsaed, Z., & Eleyan, D. (2021). Approaches to cyberbullying detection on social networks: A survey. *Journal of Theoretical and Applied Information Technology, 15*(13).

Arimoro, A., & Elgujja, A. (2019). When dissent by football fans on social media turns to hate: Call for stricter measures. *University of Maiduguri Journal of Public Law, 6*(1).

Candra, R. M., & Rozana, A. N. (2020). Klasifikasi komentar bullying pada Instagram menggunakan metode K-nearest neighbor. *IT Journal Research and Development, 5*(1), 45–52.

Dalvi, R. R., Chavan, S. B., & Halbe, A. (2020). Detecting a Twitter cyberbullying using machine learning. *International Conference on Intelligent Computing and Control System (ICICCS 2020)*.

Datareportal. (2024). Global social media statistics. *https://datareportal.com/*.

de Souza Dias, T., & Thapa, S. (2021). Tackling football-related online hate speech: The role of international human rights law: Part I. *https://www.ejiltalk.org/*.

Desai, A., Kalaskar, S., Kumbhar, O., & Dhumal, R. (2021). Cyber bullying detection on social media using machine learning. *ITM Web of Conferences, 40*, 1–5.

Gill, H. S., & Khehra, B. S. (2021). Hybrid classifier model for fruit classification. *Multimedia Tools and Applications, 80*(18), 27495–27530. https://doi.org/10.1007/s11042-021-10772-9

GilPress. (2023). Cyberbullying facts & statistics (2024). *https://whatsthebigdata.com/cyberbullying-statistics-facts*.

Gorro, K. D., Sabellano, M. J. G., Gorro, K., Maderazo, C., & Capao, K. (2018). Classification of cyberbullying in Facebook using Selenium and SVM. *2018 3rd International Conference on Computer and Communication System (ICCCS)*, 183–186.

Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J., & Qiao, S. (2022). Attention-emotion-enhanced convolutional LSTM for sentiment analysis. *IEEE Transactions on Neural Networks and Learning Systems, 33*(9), 4332–4345.

Maslej-Krešňáková, V., Sarnovský, M., Butka, P., & Machová, K. (2020). Comparison of deep learning models and various text pre-processing techniques for the toxic comments classification. *Applied Sciences (Switzerland), 10*(23), 1–26.

Monika, R., Deivalakshmi, S., & Janet, B. (2019). Sentiment analysis of US airlines tweets using LSTM/RNN. *IEEE 9th International Conference on Advanced Computing (IACC)*, 92–95.

Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment analysis using Word2vec and long short-term memory (LSTM) for Indonesian hotel reviews. *Procedia Computer Science, 179*, 728–735.

Muneer, A., & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. *Future Internet, 12*(11), 1–21.

Murthy, G. S. N., Rao Allu, S., Andhavarapu, B., Bagadi, M., & Belusonti, M. (2020). Text-based sentiment analysis using LSTM. *International Journal of Engineering Research & Technology (IJERT), 9*(5), 299–303.

Nandakumar, V. (2018). Cyberbullying revelation in Twitter data using naïve Bayes classifier algorithm. *International Journal of Advanced Research in Computer Science, 9*(1), 510–513.

Ni, R., & Cao, H. (2020). Sentiment analysis based on GloVe and LSTM-GRU. *39th Chinese Control Conference*, 7492–7497.

Nikmah, T. L., Ammar, M. Z., Allatif, Y. R., Husna, R. M. P., Kurniasari, P. A., & Bahri, A. S. (2022). Comparison of LSTM, SVM, and Naïve Bayes for classifying sexual harassment tweets. *Journal of Soft Computing Exploration, 3*(2), 131–137.

Ofcom, & The Alan Turing Institute. (2022). Crossing the line: Seven in ten Premier League footballers face Twitter abuse. *https://fcbusiness.co.uk/news/crossing-the-line-seven-in-ten-premier-league-footballers-face-twitter-abuse/*.

Perera, A., & Fernando, P. (2021). Accurate cyberbullying detection and prevention on social media. *Procedia Computer Science, 181*, 605–611.

Rehman, A. U., Malik, A. K., Raza, B., & Ali, W. (2019). A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. *Multimedia Tools and Applications, 78*(18), 26597–26613.

Setiawan, Y., Ulva Maulidevi, N., Surendro, K., & Korespondensi, P. (2022). Deteksi cyberbullying dengan mesin pembelajaran klasifikasi (supervised learning): Peluang dan tantangan. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), 9*(7), 1577–1582.

Srinivas, A. C. M. V., Satyanarayana, Ch., Divakar, Ch., & Sirisha, K. P. (2021). Sentiment analysis using neural network and LSTM. *IOP Conference Series: Materials Science and Engineering, 1074*(1), 012007.

Wang, J.-H., Liu, T.-W., Luo, X., & Wang, L. (2018). An LSTM approach to short text sentiment classification with word embeddings. *Conference on Computational Linguistics and Speech Processing*, 214–223.

Zhou, J., Lu, Y., Dai, H. N., Wang, H., & Xiao, H. (2019). Sentiment analysis of Chinese microblog based on stacked bidirectional LSTM. *IEEE Access, 7*, 38856–38866.