

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Menjelajah *platform* media sosial menjadi hal yang normal untuk masyarakat saat ini dikarenakan persebaran internet yang semakin merata [1]. Tercatat pada April 2024 terdapat lebih dari 5 milyar pengguna media sosial di seluruh dunia[2]. Kemunculan media sosial telah membantu masyarakat untuk berkomunikasi dengan siapapun, kapanpun dan dimanapun[3]. Namun, bersamaan dengan itu media sosial juga menjadi alasan munculnya masalah-masalah baru yang diantaranya adalah *cyberbullying*[4].

Cyberbullying merupakan sebuah aksi untuk mengancam atau melecehkan seseorang menggunakan media komunikasi modern seperti media sosial. *Cyberbullying* dapat meliputi berbagai aksi seperti rasisme, pelecehan seksual, penghinaan fisik, pengancaman, penyebaran rumor yang tidak benar, penyebaran pesan atau gambar yang menyakiti seseorang, dan beberapa tindakan lainnya[5], [6], [7]. *Cyberbullying* dianggap sebagai salah satu kejahatan siber yang serius karena dapat menyebabkan luka emosional yang berat terhadap korbannya[1]. *Cyberbullying* dapat berdampak pada turunnya kesehatan mental korban mulai dari emosi, kepercayaan diri, dan rasa takut untuk bertemu orang lain[5]. Selain itu, *cyberbullying* juga dapat berdampak pada kesehatan fisik karena dapat memunculkan upaya menyakiti diri sendiri atau bahkan dalam kasus yang berat dapat berujung pada tindakan bunuh diri[1][8]. *Cyberbullying* kerap kali ditemukan di *platform* media sosial dengan banyak pengguna seperti media sosial “X”[4].

Twitter atau yang sekarang lebih dikenal sebagai “X” merupakan salah satu *platform* media sosial populer dengan pengguna yang beragam mulai dari pejabat, selebritis, hingga masyarakat biasa. Setidaknya hingga April 2024 “X” memiliki jumlah pengguna sebanyak 611 juta dan berada di urutan ke-12 media sosial yang paling banyak digunakan di seluruh dunia (dengan memasukkan media sosial lokal China)[2]. “X” merupakan media sosial yang

berbentuk *microblogging* atau blog singkat dengan panjang tulisan tidak lebih dari 280 huruf yang disebut sebagai *tweet* [9]. Dari berbagai media sosial yang tersedia, “X” merupakan sumber data yang paling banyak digunakan oleh para peneliti dalam deteksi *cyberbullying* [1], [10]. Hal tersebut dikarenakan *platform* “X” masuk ke dalam daftar *platform* media sosial dengan presentase kejadian *cyberbullying* tertinggi bersama dengan *platform* lain seperti Facebook, Youtube, dan Instagram [11].

Kejadian *Cyberbullying* menasar ke semua bidang kehidupan termasuk dunia olahraga khususnya sepak bola dimana rasisme dan pengancaman merupakan bentuk *cyberbullying* yang paling sering ditemukan di sini[12]. Fans sepak bola kerap kali mengekspresikan kekecewaan mereka pada para pemain di media sosial ketika tim yang mereka dukung mengalami kekalahan. [13]. Hasil analisis yang diterbitkan oleh Ofcom pada tanggal 2 Agustus 2022 menemukan jika dari 2,4 juta *tweet* yang didapatkan pada paruh pertama kompetisi *Premier League* musim 2021/2022, terdapat sekitar 60.000 postingan abusif di *platform* “X” yang menasar 70% pemain *Premier League*[14]. Contoh kasus lain di luar data yang telah disebutkan adalah ketika Marcus Rashford, Jadon Sancho, dan Bukayo Saka gagal dalam mengeksekusi penalti di Final Euro 2020 antara Inggris vs Italia. Setelah pertandingan selesai dengan kekalahan Tim Nasional Inggris, muncul hinaan-hinaan rasis terhadap ketiga pemain Inggris berkulit gelap tersebut di media sosial. Meskipun pihak “X” mengklaim telah menghapus lebih dari 1.000 *tweet* dengan bantuan *Artificial Intelligence* dan moderator manusia, fakta bahwa kejadian serupa terus berulang menunjukkan bahwa masalah *cyberbullying* dalam sepak bola masih belum terselesaikan dan metode yang ada sekarang belum mampu mendeteksi *cyberbullying* secara spesifik terhadap pemain sepak bola [12][15].

Deteksi *cyberbullying* merupakan proses mendeteksi atau mengklasifikasi konten *cyberbullying* berdasarkan fitur tertentu seperti fitur semantik, sintaktik, sentimen, dan fitur lainnya [10]. Hal tersebut dapat

dilakukan melalui implementasi *Natural Language Processing (NLP)* dan *Machine learning*[1]. Pendekatan *NLP* cocok untuk digunakan pada masalah ini dikarenakan area penelitian yang berfokus pada deteksi *cyberbullying* seringkali fokus untuk menganalisis konten tekstual dibanding konten dalam bentuk gambar atau video[1], [5]. Model yang dibangun akan bekerja dengan cara mendeteksi apakah suatu kalimat termasuk ke dalam kategori *cyberbullying* atau *non-cyberbullying* sesuai dengan metode yang dipilih. Metode yang banyak digunakan dalam penelitian yang berkaitan dengan deteksi *cyberbullying* lainnya adalah metode yang berasal dari model *machine learning* dan *deep learning* [16].

Banyak model *machine learning* tradisional telah diimplementasikan dalam kasus deteksi *cyberbullying* di media sosial seperti *Naïve Bayes*, *Support Vector Machine (SVM)*, *K-Nearest Neighbor (KNN)* dan *Decision Tree* [17], [18]. Dari beberapa model tersebut *Naïve Bayes* dan *Support Vector Machine (SVM)* menjadi yang paling sering digunakan [19]. Selain teknik *machine learning* tradisional, belakangan digunakan pula model *deep learning* karena hasilnya yang cukup impresif [18]. Model *deep learning* seperti *Recurrent Neural network (RNN)*, *Long Short-Term Memory (LSTM)* dan *Dynamic CNN* berhasil mengungguli *accuracy* yang didapatkan oleh model *machine learning* tradisional[20].

Dalam pemrosesan teks, *RNN* lebih sering digunakan dibanding dengan model *deep learning* lain karena *RNN* didesain untuk memproses data dalam bentuk sekuensial yang sesuai dengan bentuk teks yang terdiri atas kata-kata yang saling terhubung dan memiliki korelasi antar tiap katanya [21]. Namun, arsitektur *RNN* tradisional tidak mampu memproses *Long-Term dependencies* yang menyebabkan hubungan antar kata menjadi semakin kecil ketika jaraknya semakin jauh atau sering disebut sebagai *Vanishing Gradient Problem* [20], [22], [23]. *LSTM* diciptakan untuk mengatasi masalah ini, *LSTM* merupakan salah satu arsitektur dari *RNN* yang mampu memproses *long-term dependencies*. Artinya, *LSTM* mampu mengingat kata yang

jaraknya jauh sekalipun dengan baik [20], [21], [24]. Hal tersebut membuat *LSTM* cocok digunakan pada *text classification* karena mampu mengetahui hubungan antar kata bahkan dengan jarak yang jauh. Beberapa studi sebelumnya menunjukkan jika *LSTM* mampu mengalahkan model *machine learning* tradisional bahkan sesama model *deep learning* dalam berbagai kasus *text classification* [20], [24], [25], [26], [27].

Berdasarkan uraian-uraian yang telah dijelaskan, penelitian ini mencoba mengimplementasikan metode *LSTM* dalam mendeteksi *cyberbullying* terhadap pemain sepak bola dengan mengangkat judul “Deteksi *Cyberbullying* pada Pemain Sepak Bola di Platform Media Sosial ‘X’ Menggunakan Metode Long Short-Term Memory (*LSTM*)”.

1.2 Perumusan Masalah

Kejadian *cyberbullying* telah menjadi masalah serius dalam dunia sepak bola[12]. Fans sepak bola kerap kali mengekspresikan kekecewaan mereka terhadap pemain ketika tim yang mereka dukung mengalami kekalahan[13]. Metode deteksi *cyberbullying* yang digunakan media sosial saat ini memiliki kelemahan dalam memahami konteks, sarkasme, dan perbedaan bahasa. Hal ini menjadi tantangan besar terutama dalam konteks sepak bola, dimana *cyberbullying* sering kali disamarkan dengan penggunaan istilah-istilah khusus. Akibatnya, metode yang ada masih belum efektif dalam mendeteksi *cyberbullying* secara akurat, khususnya dalam konteks sepak bola. Untuk mengatasi keterbatasan ini, penelitian ini mengusulkan penggunaan Long Short-Term Memory (*LSTM*) dalam mendeteksi *cyberbullying*. *LSTM* yang memiliki kemampuan untuk memahami hubungan kontekstual dalam teks sekuensial diharapkan dapat meningkatkan akurasi dalam pendeteksian *cyberbullying* dalam dunia sepak bola.

1.3 Pertanyaan Penelitian

1. Bagaimana arsitektur model *LSTM* terbaik dalam mendeteksi *cyberbullying* pada pemain sepak bola di media sosial dan berapa hasil *accuracy*, *precision*, *recall*, *f1-score*, serta *loss*-nya?

1.4 Batasan Masalah

1. Penelitian membahas tentang *cyberbullying* terhadap pemain sepak bola.
2. Data yang digunakan merupakan *tweet* yang berasal dari *platform* “X” dalam kurun waktu 01 Maret - 17 November 2024.
3. Data yang diambil merupakan *tweet* dalam Bahasa Indonesia.
4. Data yang digunakan hanyalah data teks dan tidak melibatkan data dalam bentuk lain seperti gambar atau video.

1.4 Tujuan Penelitian

1. Mendeteksi *cyberbullying* pada postingan di media sosial “X” menggunakan metode *Long Short-Term Memory (LSTM)*.
2. Mendapatkan model dengan hasil *accuracy*, *precision*, *recall*, *f1-score*, dan *loss* terbaik dari model *LSTM* yang dibangun.

1.5 Manfaat Penelitian

1. Mencegah terjadinya *cyberbullying* terhadap pemain sepak bola di media sosial.
2. Menaikkan performa pemain sepak bola baik dalam level klub maupun Tim Nasional.
3. Mencegah dampak negatif *cyberbullying* baik dampak emosional maupun fisik.
4. Suporter sepak bola menjadi teredukasi tentang bahaya *cyberbullying*.