

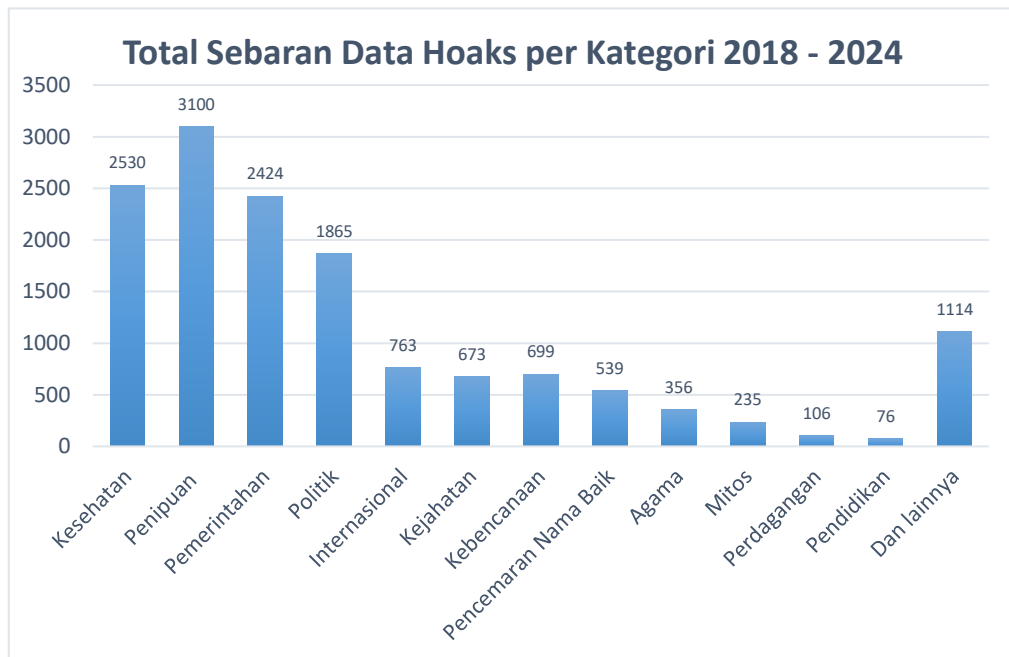
BAB I

PENDAHULUAN

1.1. Latar Belakang

Pesatnya perkembangan teknologi informasi telah mempermudah akses dan distribusi informasi melalui berbagai platform digital, terutama media sosial. Kondisi ini memungkinkan masyarakat kini dapat menerima dan membagikan informasi dalam hitungan detik tanpa batas geografis. Namun, di balik kemudahan tersebut, muncul tantangan serius berupa penyebaran berita hoaks, yaitu informasi bohong yang sengaja direkayasa untuk menyesatkan pembaca dan seringkali disebarluaskan melalui berbagai platform daring [1], [2]. Fenomena ini telah menjadi ancaman serius di Indonesia karena penyebarannya yang masif di ruang digital. Berdasarkan penelusuran Tim Anyar, Inspiratif, dan Solutif (AIS) Kementerian Komunikasi dan Digital (Komdigi), tercatat total 12.547 konten berita hoaks tersebar di situs web Indonesia dalam kurun waktu lima tahun terakhir (2018–2023) [3], [4]. Konten-konten palsu ini mencakup beragam isu seperti kesehatan, penipuan, pemerintahan, politik, hingga bencana alam, menunjukkan kompleksitas masalah disinformasi.

Tantangan penyebaran hoaks terus berlanjut hingga saat ini. Kominfo mencatat 1.923 berita hoaks teridentifikasi sepanjang tahun 2024, terhitung mulai dari 1 Januari hingga 31 Desember [5] seperti pada gambar 1.1. Angka-angka ini mengindikasikan bahwa volume berita hoaks tetap tinggi dan menjadi ancaman berkelanjutan terhadap validitas informasi di masyarakat. Sebagai contoh dampak nyata dari penyebaran hoaks adalah kasus konflik SARA di Tanjung Balai dan di Depok, yang menunjukkan bagaimana paparan hoaks SARA dapat memengaruhi perilaku individu dan memicu tindakan destruktif [6].



Gambar 1.1 Total Sebaran Data Hoaks per Kategori 2018 - 2024

Dampak merusak dari hoaks diperparah oleh rendahnya literasi di Indonesia [7], yang tercermin dari survei yang menunjukkan bahwa sebagian besar responden masih kesulitan mengidentifikasi hoaks bahkan untuk isu-isu sensitif seperti SARA[6], mengingat hoaks seringkali dirancang menggunakan bahasa yang menarik dengan unsur profokatif [8]. Fenomena penyebaran hoaks ini tidak hanya menjadi tantangan di negara Indonesia, tapi juga untuk berbagai negara lain yaitu Malaysia, berita hoaks juga menjadi permasalahan yang sulit untuk ditangani. Berdasarkan data yang dikeluarkan oleh Malaysian Communications and Multimedia (MCMC) terdapat 3,285 laporan berita berpotensi menjadi hoaks dengan rentang 2020 – 2022 [9]. Sedangkan MCMC juga berhasil menanggulangi penyebaran berita hoaks dengan total 6,618 unggahan mulai dari 1 Januari 2023 – 15 Februari 2024 [10].

Keterbatasan kognitif manusia dalam memproses dan memverifikasi volume informasi yang terus membanjir secara cepat menjadikan deteksi manual tidak lagi memadai dan rentan terhadap kesalahan[11]. Oleh karena itu, diperlukan pendekatan inovatif yang dapat mengatasi volume dan kompleksitas hoaks secara

otomatis dan akurat guna memperkuat ketahanan siber [12] dan integritas informasi digital [13].

Berbagai pendekatan manual telah digunakan untuk menanggulangi penyebaran berita palsu, seperti pelabelan konten atau pemeriksaan fakta oleh lembaga tertentu. Namun, metode ini terbukti tidak cukup efisien mengingat volume informasi yang sangat besar dan kecepatan penyebarannya. Oleh karena itu, diperlukan solusi otomatis berbasis teknologi yang mampu melakukan deteksi hoaks secara cepat dan akurat[14].

Dalam konteks penelitian, sejumlah studi telah dikembangkan untuk mendeteksi berita palsu secara otomatis menggunakan pendekatan machine learning dan deep learning. Salah satu studi oleh melakukan tinjauan menyeluruh terhadap metode deteksi hoaks berbasis teks, gambar, dan video. Penelitian tersebut mengevaluasi berbagai teknik seperti Support Vector Machine (SVM), Naïve Bayes, Random Forest, serta pendekatan berbasis deep learning seperti Convolutional Neural Network (CNN) dan BiLSTM. Hasil studi tersebut menunjukkan bahwa pendekatan hibrida serta model CNN- Long Short Term Memory (LSTM) mampu mencapai akurasi yang cukup tinggi dalam mendeteksi hoaks di media sosial dengan nilai akurasi mencapai 99%[15]. Penelitian lain menggunakan LSTM dengan hasil presisi 81%, recall 80% dan f1-measure 80% menggunakan *hierarchical softmax* yang lebih baik dari *negative sampling* [16]

Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada data berbahasa Inggris dan cenderung bekerja secara efektif hanya setelah berita menyebar secara luas [17]. Selain itu, metode yang digunakan belum secara optimal menangkap konteks dua arah dari teks berita yang kompleks. Ini menjadi hambatan serius dalam mendeteksi hoaks sejak dini, terutama dalam konteks bahasa Indonesia yang memiliki struktur kalimat yang kaya dan fleksibel [18].

Kesenjangan inilah yang mendorong perlunya pengembangan sistem deteksi berita hoaks berbasis bahasa Indonesia yang mampu memahami konteks linguistik secara lebih mendalam. Salah satu pendekatan yang potensial untuk menjawab

tantangan ini adalah penggunaan model BiLSTM[19]. Berbeda dengan LSTM konvensional, BiLSTM mampu memproses informasi secara simultan dari dua arah—baik dari awal maupun akhir teks—sehingga dapat menangkap hubungan antar kata dalam konteks yang lebih utuh.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun deteksi berita hoaks berbasis BiLSTM dengan fokus pada berita berbahasa Indonesia dan Malaysia. Pendekatan ini diharapkan tidak hanya berkontribusi dalam pengembangan teknologi keamanan siber, tetapi juga mendukung upaya literasi digital nasional dan pencegahan penyebaran informasi palsu di ruang digital.

1.2. Rumusan Masalah

1. Bagaimana penerapan algoritma BiLSTM dalam mendeteksi berita hoaks berbasis teks berbahasa Indonesia dan Malaysia?
2. Sejauh mana tingkat akurasi model BiLSTM dalam mendeteksi berita hoaks berbahasa Indonesia dan Malaysia?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengembangkan solusi berbasis kecerdasan buatan dalam upaya menangani penyebaran berita hoaks di Indonesia dan Malaysia. Secara khusus, penelitian ini bertujuan untuk:

1. Menerapkan algoritma Bidirectional Long Short-Term Memory (BiLSTM) dalam membangun model deteksi berita hoaks yang efektif terhadap teks berbahasa Indonesia dan Malaysia.
2. Mengevaluasi kinerja model BiLSTM dalam hal akurasi, presisi, recall, dan f1-score untuk mengetahui efektivitasnya dalam mengklasifikasikan berita hoaks bahasa Indonesia dan Malaysia.

1.4. Urgensi Penelitian

Penelitian ini memiliki urgensi yang tinggi dalam konteks keamanan siber dan *Natural Language Processing* (NLP) di Indonesia. Meskipun banyak upaya deteksi hoaks telah dilakukan secara global, sebagian besar model yang

ada masih terfokus pada bahasa Inggris dan belum sepenuhnya mengakomodasi karakteristik serta kompleksitas linguistik Bahasa Indonesia dan Malaysia. Hal ini menciptakan celah signifikan dalam upaya perlindungan informasi di ruang siber nasional. Oleh karena itu, penelitian ini penting untuk mengembangkan solusi berbasis teknologi yang dapat mengotomatisasi klasifikasi berita hoaks.

Pengembangan sistem deteksi hoaks yang secara spesifik berbasis algoritma BiLSTM dan dirancang untuk teks berbahasa Indonesia menjadi sangat vital. Dengan membangun model yang mampu memahami konteks dan nuansa Bahasa Indonesia, penelitian ini tidak hanya berkontribusi langsung pada penguatan keamanan digital nasional dan mendukung kebijakan pemerintah dalam memberantas hoaks. Diharapkan deteksi berita hoaks dapat dilakukan lebih akurat dan efisien, mengurangi ketergantungan pada verifikasi manual, serta memberikan kontribusi signifikan dalam menangani penyebaran disinformasi di Indonesia. Lebih jauh, riset ini membuka jalan bagi pengembangan inovasi lanjutan di bidang kecerdasan buatan dan NLP, serta meningkatkan kapasitas bangsa dalam menghadapi tantangan disinformasi di masa mendatang.

1.5. Luaran Penelitian

Penelitian ini diharapkan menghasilkan sebuah sistem klasifikasi hoaks berbasis teks Bahasa Indonesia yang andal dan akurat, dibangun melalui penerapan algoritma BiLSTM. Model BiLSTM terlatih yang diimplementasikan menggunakan Python dan *framework* TensorFlow/Keras. Model ini dirancang khusus untuk mengklasifikasikan berita hoaks berdasarkan teks berbahasa Indonesia dan Malaysia. Keberhasilan implementasi model ini telah menunjukkan efektivitasnya dalam penanganan kompleksitas teks berbahasa campuran. Seluruh kode sumber, dataset yang digunakan, serta model terlatih telah didokumentasikan dengan baik dan tersedia di repositori internal penelitian sebagai bukti implementasi dan dapat direplikasi untuk pengembangan selanjutnya. Model ini memiliki potensi besar untuk

dikembangkan lebih lanjut sebagai bagian dari sistem deteksi hoaks yang lebih luas untuk mendukung keamanan siber

1.6. Manfaat Penelitian

Penelitian ini tidak hanya berfokus pada pengembangan model deteksi berita hoaks, tetapi juga memiliki manfaat yang lebih luas baik di bidang akademik mencakup berbagai aspek berikut:

1. Meningkatkan Kemampuan Deteksi Hoaks Berbahasa Indonesia dan Malaysia. Penelitian ini akan menghasilkan model deteksi hoaks yang dirancang khusus untuk Bahasa Indonesia dan Malaysia, memberikan solusi yang lebih akurat dan efisien dibandingkan metode manual atau model generik yang kurang adaptif terhadap kompleksitas linguistik lokal.
2. Memberikan kontribusi ilmiah dalam bidang keamanan siber, dengan mengusulkan pendekatan deteksi berita hoaks yang relevan dan kontekstual terhadap bahasa dan kondisi di Indonesia dan Malaysia.
3. Mendukung program Merdeka Belajar Kampus Merdeka (MBKM) dengan menghasilkan luaran penelitian berupa model BiLSTM untuk akademik.

