

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital saat ini, akses mudah ke internet dan media sosial telah membuat masyarakat semakin terhubung dan terinformasi, Namun, dalam lingkungan yang terbuka ini, masalah yang sering muncul adalah maraknya berita palsu atau *hoax* yang dapat menimbulkan kerusakan dan kebingungan. Berita *hoax*, juga dikenal sebagai berita palsu atau berita bohong, merujuk kepada informasi yang disebarkan dengan sengaja untuk menyesatkan atau memanipulasi pembaca [1]. Berita *hoax* sering kali dibuat untuk mempengaruhi opini publik, menyebabkan kebingungan, atau memperoleh keuntungan pribadi atau politik.

Ciri utama dari berita *hoax* adalah ketidakbenaran atau distorsi fakta yang signifikan. Biasanya, berita *hoax* disebarkan melalui platform media sosial, situs web yang tidak terpercaya, atau melalui pesan berantai. Konten berita palsu sering kali dirancang dengan tampilan dan gaya yang mirip dengan sumber berita yang kredibel, sehingga sulit bagi pembaca untuk membedakan antara berita yang benar dan yang palsu [2].

Berita *hoax* dapat mencakup berbagai topik, mulai dari politik, kesehatan, kejahatan, hingga peristiwa berita terkini. Tujuan dibalik penyebaran berita palsu dapat bervariasi, termasuk upaya untuk mempengaruhi pemilihan umum, menciptakan ketakutan dan kepanikan, atau memperoleh keuntungan finansial melalui iklan yang tampil di situs web berita palsu [3].

sebelum mempercayai atau membagikan berita yang diterima, sangat penting untuk mengembangkan kemampuan kritis dan melakukan verifikasi informasi. Mencari sumber informasi yang terpercaya, memeriksa keabsahan sumber berita, dan melihat apakah berita tersebut dilaporkan oleh sumber-sumber yang kredibel adalah langkah-langkah penting untuk menghindari penyebaran berita *hoax* dan menjaga informasi yang akurat. Terutama mewaspadaai berita *hoax* berbahasa Inggris yang dimana dalam penelitian yang dipaparkan oleh Wahyuni A dkk [4], mengungkapkan bahwa Kemampuan bahasa Inggris masyarakat Indonesia masih masuk dalam kategori rendah sehingga kesulitan masyarakat dalam mengetahui kebenaran suatu berita menjadi semakin rumit. Selain itu Berita palsu yang tidak

terverifikasi dapat merusak reputasi seseorang atau organisasi, mempengaruhi opini publik, dan memicu konflik di masyarakat. Oleh karena itu, deteksi dan klasifikasi berita palsu menjadi sangat penting [2].

Beberapa metode telah dikembangkan untuk mengatasi masalah ini, termasuk teknik-teknik *Natural Language Processing* (NLP). Salah satu teknik yang dapat diterapkan dalam deteksi dan klasifikasi berita adalah TF-IDF (*Term Frequency-Inverse Document Frequency*) dan *Long Short-Term Memory* (LSTM).

TF-IDF adalah teknik pengukuran kepentingan kata dalam dokumen [1]. Dalam teknik ini, frekuensi kemunculan kata dalam dokumen dihitung dan dibagi dengan jumlah kata dalam dokumen tersebut. Kemudian, dibandingkan dengan frekuensi kemunculan kata yang sama dalam seluruh korpus dokumen untuk mendapatkan nilai TF-IDF yang lebih akurat. Metode ini dapat membantu mengidentifikasi kata-kata yang paling relevan dan penting dalam dokumen. Sementara itu, LSTM adalah salah satu jenis metode jaringan saraf yang telah digunakan dalam banyak aplikasi NLP. LSTM dapat memperhitungkan konteks dan mengatasi masalah yang terkait dengan pemrosesan teks panjang. Dalam konteks deteksi berita palsu, LSTM dapat membantu memprediksi apakah sebuah artikel berita termasuk kategori *hoax* atau *non-hoax* [5].

Pada Penelitian sebelumnya yang berjudul *Automating fake news detection system using multi-level voting model* yang dipaparkan oleh Kaur, S., Kumar, P. & Kumaraguru, P [6]. hasil akurasi yang dicapai untuk mendeteksi berita palsu menggunakan menggunakan TF-IDF memiliki tingkat akurasi lebih baik daripada menggunakan metode lain seperti *Count vectorized* dan *Hashing Vectorized* pada tiap-tiap metode klasifikasi yang digunakan.

Kemudian pada penelitian lain yang berjudul “*Fake News Classification Using Different Methods*” yang dipaparkan oleh menggunakan 3 metode klasifikasi yakni SVM (*Support Vector Machine*), *Neural Network with Tensorflow*, dan LSTM (*Long Short-Tem Memory*). Dari ketiga metode tersebut, LSTM memiliki hasil akurasi tertinggi yakni 94% [7].

Oleh karena itu, dalam penelitian ini, Metode TF-IDF dan LSTM akan digunakan dalam pembuatan sistem deteksi terhadap berita *hoax*. Metode ini akan diterapkan pada dataset dan diuji untuk mengukur akurasi dan kinerja model. Selain

itu penerapan berbagai parameter yang terdapat di LSTM juga patut untuk diuji mengingat dalam pembuatan sistem deteksi berita hoax, kinerja suatu model sangat dibutuhkan agar sistem dapat mendeteksi berita *hoax* dengan sangat baik. Diharapkan hasil dari penelitian ini dapat membantu meningkatkan deteksi berita *hoax* secara efektif, sehingga masyarakat dapat memperoleh informasi yang akurat dan terpercaya.

1.2 Rumusan Masalah

Rumusan masalah yang dapat diperoleh berdasarkan latar belakang adalah sebagai berikut:

1.2.1 Permasalahan

Permasalahan yang ada adalah maraknya penyebaran berita hoax berbahasa Inggris yang terdapat di Internet.

1.2.2 Solusi

Solusi dari permasalahan adalah dengan membuat sistem deteksi berita *hoax* menggunakan pembobotan TF-IDF dan dimodelkan dengan metode LSTM sehingga mendapatkan hasil. Hasil tersebut tentunya akan menjadi tambahan informasi kepada masyarakat mengenai pengkategorian teks berita *hoax* atau non *hoax*.

1.2.3 Pertanyaan Penelitian

Mengenai pertanyaan yang dapat dikutip dari keadaan pada latar belakang dan permasalahan serta solusi dari penelitian ini, beberapa pertanyaan dapat ditarik guna menjadi tujuan dari disusunnya penelitian ini :

1. Bagaimana pengaruh parameter LSTM terhadap nilai keakuratan suatu model?
2. Bagaimana pengaruh *preprocessing* pada nilai keakuratan suatu model?

1.3 Tujuan

Berdasarkan rumusan masalah diatas, didapatkan Tujuan dari penelitian sebagai berikut :

1. Mengetahui pengaruh parameter LSTM terhadap nilai keakuratan suatu model.
2. Mengetahui pengaruh *preprocessing* pada nilai keakuratan suatu model.

1.4 Manfaat

Berdasarkan rumusan masalah diatas, didapatkan manfaat dari penelitian sebagai berikut :

1. Menghasilkan sistem yang efisien untuk mendeteksi berita hoax secara otomatis, yang bisa digunakan oleh media, platform sosial, dan pengguna individu untuk mengidentifikasi dan mengurangi penyebaran informasi palsu.
2. Sebagai pembelajaran penulis mempelajari metode LSTM dengan pembobotan kata TF-IDF dalam pembuatan sistem deteksi berita hoax.

1.5 Batasan Masalah

Dalam penelitian ini penulis membatasi permasalahan agar pemahaman terhadap penelitian tidak meluas dan menyimpang, diantaranya :

1. *Dataset* yang digunakan dalam penelitian ini adalah *Fake News Dataset* yang diperoleh dari kaggle yang berisikan teks berita bahasa inggris.
2. Metode yang digunakan dalam mencari *vector space model* adalah TF-IDF
3. Metode yang digunakan dalam pemodelan adalah *Long Short Term Memory*.
4. Dalam melakukan pengujian model hanya berfokus pada mencari nilai akurasi

1.6 Metodologi Penulisan Skripsi

Agar penulisan dalam penelitian sesuai dengan standar yang telah ditetapkan, maka terdapat 5 topik yang berisikan ringkasan bab, yaitu :

BAB I PENDAHULUAN

Pendahuluan berisi tentang garis besar persoalan yang dihadapi dalam penelitian ini, yang terdapat Latar belakang penelitian ini, rumusan masalah dari penjabaran latar belakang, batasan masalah , dan tujuan pada penelitian ini, serta manfaat pada penelitian

BAB II KAJIAN PUSTAKA

Tinjauan pustaka berisi tentang teori-teori penunjang yang menjadi tumpuan di tugas akhir ini.

BAB III METODOLOGI PENELITIAN

Bagian yang menjabarkan tentang metode Penelitian dan alur pada sistem yang diinginkan.

BAB IV HASIL PEMBAHASAN

Pada bab ini menjelaskan tentang hasil uji coba dan pengimplementasian sistem dalam bentuk website. Selain itu, akan dijelaskan pula ruang lingkup kerja, tahapan pembangunan sistem, tahap uji coba, dan penjelasan sistem.

BAB V HASIL PEMBAHASAN

Pada bab ini berisi kesimpulan dari hasil uji coba klasifikasi berita dan saran untuk pengembangan skripsi.