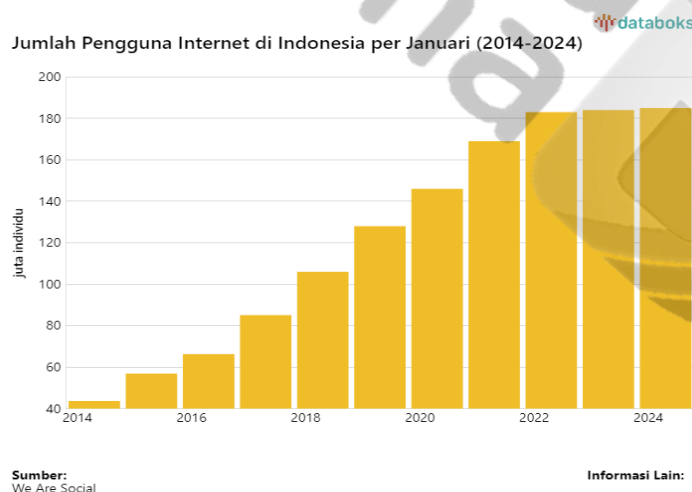


BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi saat ini telah mengalami perkembangan yang pesat dan canggih sehingga dapat diakses dengan mudah serta cepat. Era digitalisasi merupakan kondisi kemajuan zaman yang serba digital, sehingga komunikasi dan mencari informasi saat ini dapat diperoleh di mana saja dan kapanpun pada media daring yang terhubung dengan jaringan internet. Internet memang menjadi bagian yang penting dalam kehidupan saat ini. Banyak manfaat yang diperoleh dengan perkembangan internet, salah satunya adalah penyampaian berita yang dipublikasikan di media daring.



(Sumber : <https://datareportal.com/reports/digital-2024-indonesia>)

Berdasarkan gambar 1.1, setiap tahun perkembangan pengguna internet yang aktif di Indonesia terus meningkat. Internet merupakan sebagai media yang tepat bagi pengguna sebagai komunikasi dan penyebaran informasi berita dengan mudah. Namun, tidak semua informasi yang disebar di media internet berdasarkan fakta. Ada banyak berita yang beredar di media daring adalah berita yang tidak benar atau disebut berita palsu (*hoax*s).

Media sosial juga sering kali disalahgunakan sebagai target orang-orang yang tidak bertanggung jawab menyebarkan berita palsu secara bebas. Rendahnya tingkat pendidikan dan tingkat literasi masyarakat membuat masyarakat mudah untuk menerima berita palsu. Jika hal ini dibiarkan dapat menjadi penyebab terpecahnya persepsi dan keyakinan masyarakat dalam menangkap informasi yang diterima. Penting untuk masyarakat dalam meningkatkan kesadaran serta kewaspadaan terhadap berita palsu melalui media daring seperti website dan media online lainnya dengan selalu mempertimbangkan atau memastikan kebenaran informasi berita yang tersebar.

Salah satu tantangan utama dalam mendeteksi berita palsu dengan menggunakan teknik pembelajaran mesin (*Machine Learning*, ML) adalah kendala bahasa, khususnya dengan bahasa seperti Bahasa Indonesia. Banyak alat pemrosesan bahasa alami (*Natural Language Processing*, NLP) tingkat lanjut, seperti TextBlob, yang sangat penting untuk melakukan analisis sentimen, tidak mendukung Bahasa Indonesia secara langsung (Molenaar et al., 2024). Hal ini mengharuskan penerjemahan teks ke dalam bahasa Inggris sebelum dianalisis, yang dapat menyebabkan kesalahan besar dan salah tafsir budaya.

Memahami nilai polaritas dan subjektivitas teks sangat penting dalam mendeteksi berita palsu. Polaritas yang menunjukkan apakah sentimen yang disampaikan dalam teks itu positif atau negatif, dapat mengungkap maksud emosional di balik sebuah berita (Alonso et al., 2021). Berita palsu sering kali berupaya membangkitkan respons emosional yang kuat berdasarkan polaritas ini (Nanath et al., 2022). Subyektivitas, sebaliknya, menilai apakah teks tersebut berdasarkan opini atau faktual. Berita palsu sering kali mengaburkan batasan ini dan menampilkan opini yang bias sebagai fakta.

Mendeteksi berita palsu memerlukan lebih dari sekedar pemrosesan dasar data tekstual; hal ini memerlukan pemahaman nuansa dan konteks yang menentukan konten asli versus konten berbahaya (Tajrian et al., 2023). Model ML, khususnya yang menggunakan teknik seperti jaringan Long Short-Term Memory

(Long Short Term Memory, LSTM), mahir dalam menguraikan kompleksitas ini. Model-model ini dapat secara efektif mengidentifikasi dan menganalisis pola dan ketergantungan dalam teks yang mengindikasikan berita palsu.

Studi terbaru menggaris bawahi pentingnya mengintegrasikan alat analisis sentimen seperti TextBlob dan VADER dengan model ML (Anderson et al., 2024). Alat-alat ini menganalisis nada emosional dan subjektivitas konten, memberikan wawasan tentang bahasa manipulatif yang sering kali menjadi ciri berita palsu. Dengan mengkaji bagaimana sentimen terbentuk dalam teks-teks tersebut, para peneliti dapat lebih memahami bagaimana berita palsu berupaya memanipulasi persepsi publik.

Jaringan LSTM terkenal karena kemampuannya menangani urutan data, sehingga ideal untuk analisis teks (Ait Hammou et al., 2020). Dalam konteks berita palsu, LSTM dapat memproses artikel yang panjang atau aliran data yang berkelanjutan, mengenali pola yang mungkin menghindari model statis yang lebih sederhana (Rajalaxmi et al., 2022). Kemampuan ini menjadikan mereka alat yang ampuh dalam mengidentifikasi berita palsu, karena konteks dan aliran informasi sangat penting untuk membuat penentuan yang akurat.

Meskipun alat seperti TextBlob dan VADER menawarkan keuntungan yang signifikan, alat tersebut juga menghadirkan tantangan, terutama dengan teks non-Inggris yang memerlukan terjemahan (Hussein & Al-Naymat, 2023). Kebutuhan ini dapat melemahkan nuansa bahasa aslinya, sehingga berpotensi menyebabkan analisis sentimen menjadi kurang akurat. Terlepas dari tantangan-tantangan ini, alat-alat ini tetap sangat diperlukan dalam mendeteksi berita palsu, terutama jika ditingkatkan dengan model ML yang canggih (Özcan et al., 2024).

Studi ini mengusulkan analisis komparatif model ML tradisional yang menggunakan alat analisis sentimen yang memerlukan terjemahan terhadap model yang tidak. Dengan menguji keefektifan model-model ini di kedua situasi, kami bertujuan untuk menyoroti kekuatan dan keterbatasan yang melekat pada masing-

masing pendekatan dan menentukan pendekatan mana yang paling menjanjikan dalam pendeteksian berita palsu yang kuat.

Pendekatan kami membandingkan algoritma jaringan LSTM dengan pendekatan TextBlob dan VADER. Meskipun terdapat tantangan dalam penerjemahan, kombinasi ini diuji terhadap kumpulan data artikel berita yang diterjemahkan ke dalam bahasa Inggris dari Bahasa Indonesia. Kami menilai seberapa baik kinerja model ini dalam mendeteksi berita palsu, terutama berfokus pada kemampuan model tersebut dalam menafsirkan polaritas dan subjektivitas secara akurat.

Performa model ML, TextBlob dan VADER ini akan dievaluasi menggunakan metrik standar seperti akurasi, presisi, perolehan, dan skor F1 (Naidu et al., 2023). Metrik ini akan membantu kita memahami seberapa efektif setiap model memanfaatkan fitur linguistik untuk mendeteksi berita palsu. Evaluasi ini akan melibatkan kumpulan data terkontrol dari artikel berita beranotasi dengan label fakta yang diketahui, sehingga memungkinkan dilakukannya perbandingan komprehensif mengenai kemampuan masing-masing model.

Berdasarkan latar belakang masalah di atas, kami mengusulkan penelitian dengan judul “Deteksi berita palsu pada media daring menggunakan metode pembelajaran mesin.”

1.2 Rumusan Masalah

Dari latar belakang di atas maka rumusan masalah pada penelitian ini yaitu:

1. Bagaimana melakukan deteksi berita palsu pada media daring menggunakan model pembelajaran mesin.
2. Bagaimana perbandingan performa masing-masing model pembelajaran mesin dalam mendeteksi berita palsu pada media daring.

1.3 Batasan Masalah

Penelitian ini hanya menggunakan data artikel berita indonesia dari portal kompas.id untuk berita yang dikonfirmasi benar dan portal kominfo.go.id yang telah mengkonfirmasi tentang berita palsu pada periode tahun 2023 sampai 2024.

Model yang akan dibandingkan adalah LSTM, TextBlob dan gabungan antara TextBlob dan VADER. Pada bagian kajian Pustaka akan dijelaskan alasan mengapa dilakukan perbandingan performa antara 3 (tiga) pendekatan tersebut.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk mendeteksi berita palsu dengan mengklasifikasikan berita menggunakan pendekatan pembelajaran mesin.

1.5 Manfaat Penelitian

Melalui penelitian ini diharapkan dapat memberikan manfaat bagi pihak-pihak yang membutuhkan, khususnya pengambil kebijakan yang berakibat terhadap dampak maraknya berita palsu di masyarakat. Selain itu, meningkatkan ilmu pengetahuan sebagai referensi mengenai klasifikasi berita palsu menggunakan pendekatan pembelajaran mesin serta dapat digunakan untuk pengembangan model suatu sistem yang dapat mendeteksi berita palsu.

1.6 Ruang Lingkup Penelitian

Untuk memastikan model yang dikembangkan benar-benar mencerminkan situasi dunia nyata, penelitian ini menggunakan link berita yang ada di media daring, bersumber dari portal kompas.com dan portal kominfo.go.id. Beberapa pemberitaan di bidang politik dan isu sosial digunakan sebagai studi kasus. Kategori pengambilan sampel untuk bidang-bidang tersebut (politik dan sosial) dipilih berdasarkan tema berita di tingkat nasional.

1.7 Sistematika Penulisan

Sistematika penulisan proposal tesis yang digunakan, dapat diuraikan sebagai berikut:

BAB I PENDAHULUAN

Dalam BAB ini, berisi penjelasan mengenai latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian serta sistematika penulisan.

BAB II KAJIAN PUSTAKA

Berisi penjelasan terkait uraian mengenai teori-teori yang digunakan sebagai landasan atau acuan dalam melaksanakan penelitian.

BAB III METODE PENELITIAN

Bab ini menjelaskan tentang metode yang digunakan untuk melaksanakan penelitian.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini, akan disajikan hasil dari penelitian ini dan juga diberikan penjelasan dan ulasan tentang hasil. Pada bagian ini juga akan dibandingkan hasil dari 3 (tiga) pendekatan yang dipilih dan juga pembahasan kelebihan dan kekurangan dari masing-masing model.

BAB V KESIMPULAN

Pada bab ini, akan dilakukan penarikan kesimpulan dan pemberian saran dan masukan untuk peningkatan dari penelitian ini.