

BAB I PENDAHULUAN

1.1 Latar Belakang

Di era digital, arus informasi berkembang dengan sangat pesat melalui berbagai platform, seperti media sosial dan portal berita daring. Kemudahan dalam mengakses dan menyebarkan informasi memberikan manfaat besar bagi masyarakat, terutama dalam memperoleh berita. Namun, di sisi lain, fenomena ini juga meningkatkan risiko penyebaran informasi yang tidak valid, termasuk berita *hoax*. *Hoax* adalah informasi yang salah atau menyesatkan yang disebarluaskan dengan tujuan tertentu, baik secara sengaja maupun tidak. Penyebaran *hoax*s dapat menimbulkan dampak negatif, seperti membentuk opini yang keliru, menyesatkan masyarakat, hingga memicu konflik sosial. *Hoax* sering kali dikemas agar tampak kredibel, sehingga banyak orang sulit membedakan antara berita yang benar dan yang palsu [1]. Salah satu kategori *hoax* yang paling sering ditemukan adalah *hoax* dalam ranah politik. Berita politik memiliki daya tarik tinggi bagi masyarakat karena berkaitan dengan kepentingan publik, kebijakan pemerintah, dan masa depan demokrasi. Akibatnya, berita politik sering menjadi sasaran disinformasi yang bertujuan untuk membentuk persepsi tertentu terhadap tokoh atau kebijakan politik. Penyebaran *hoax* politik tidak hanya menyesatkan opini publik, tetapi juga dapat memperburuk polarisasi sosial, menurunkan kepercayaan masyarakat terhadap pemerintah, serta mengganggu stabilitas demokrasi [2].

Data dari Kementerian Komunikasi dan Informatika (Kominfo) menunjukkan bahwa hingga Mei 2023, terdapat 11.642 konten *hoax* yang teridentifikasi, dengan 1.373 di antaranya berkategori politik [3]. Selain itu, dalam periode Agustus 2018 hingga Desember 2023, Kominfo mencatat 12.547 kasus *hoax* yang tersebar di berbagai platform digital [4]. Survei Katadata *Insight Center (KIC)* pada 2021 juga menunjukkan bahwa 69,3% responden mengidentifikasi politik sebagai topik yang paling banyak mengandung isu *hoax* [5]. Temuan ini menegaskan bahwa *hoax* politik merupakan ancaman serius

yang dapat memengaruhi stabilitas sosial dan proses demokrasi di Indonesia. Dalam menghadapi maraknya penyebaran *hoax* politik, masyarakat sering kali kesulitan membedakan antara berita valid dan berita palsu, terutama karena banyaknya informasi yang beredar dengan format yang menyerupai berita resmi. Oleh karena itu, diperlukan sistem klasifikasi otomatis yang mampu mengidentifikasi berita *hoax* secara akurat agar masyarakat dapat memperoleh informasi yang lebih kredibel.

Salah satu pendekatan yang dapat digunakan untuk mengembangkan sistem ini adalah dengan memanfaatkan teknologi kecerdasan buatan (*artificial intelligence*), khususnya model *deep learning* berbasis *Transformer*. Salah satu model yang relevan untuk tugas ini adalah IndoBERT, sebuah varian dari BERT (*Bidirectional Encoder Representations from Transformers*) yang telah dilatih secara khusus menggunakan korpus bahasa Indonesia dalam jumlah besar. Dengan kemampuannya dalam memahami konteks kata dalam kalimat, IndoBERT lebih efektif dalam membedakan berita *hoax* dan berita valid dibandingkan dengan metode berbasis fitur statistik atau *machine learning* konvensional. Model ini dapat mengenali pola bahasa yang sering digunakan dalam berita palsu, sehingga meningkatkan akurasi dalam klasifikasi berita.

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi berita *hoax* dalam ranah politik menggunakan IndoBERT. Sistem ini diharapkan dapat membantu masyarakat dalam mendeteksi berita *hoax* dengan lebih cepat dan akurat, sehingga mereka dapat terhindar dari disinformasi dan dapat mengambil keputusan berdasarkan informasi yang valid. Selain itu, sistem ini juga dapat menjadi alat bantu bagi pihak berwenang dalam menangkal penyebaran *hoax* politik di era digital.

1.2 Rumusan Masalah

Dari deskripsi latar belakang permasalahan yang telah disebut diatas, beberapa masalah yang dapat dapat dirumuskan adalah sebagai berikut :

1. Bagaimana mengembangkan sistem klasifikasi otomatis yang mampu mengidentifikasi dan mengelompokkan teks *hoax* secara akurat dengan memanfaatkan model IndoBERT sebagai dasar pemodelan?
2. Bagaimana kinerja model IndoBERT dalam mendeteksi teks *hoax* berdasarkan tingkat akurasi dan efisiensi pemrosesan?

1.3 Batasan Masalah

Batasan masalah yang terdapat pada penelitian ini yaitu :

1. Model dilatih menggunakan dataset yang berisi berita politik tentang Presiden Prabowo, sehingga sistem mungkin tidak akurat dalam mendeteksi *hoax* politik di luar cakupan data latih tersebut.
2. Sistem hanya mengenali teks sebagai input. Gambar, video, atau media lainnya tidak akan diproses dalam deteksi *hoax*.
3. Sistem hanya berfungsi untuk mendeteksi dan mengklasifikasikan teks sebagai *hoax* atau bukan *hoax*. Sistem tidak memberikan referensi mengenai kebenaran suatu berita atau menyediakan sumber yang valid.
4. Sistem hanya dapat memproses teks dalam bahasa Indonesia. Teks dalam bahasa lain tidak akan dikenali dengan baik, sehingga sistem tidak dapat melakukan klasifikasi yang akurat jika terdapat campuran bahasa atau penggunaan istilah asing.
5. Sistem ini hanya dapat memproses kalimat yang disampaikan secara langsung dan eksplisit, serta tidak dirancang untuk mendeteksi makna implisit, sarkasme, atau *hoax* yang tersamar dalam kalimat kompleks.

1.4 Tujuan

Tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mengembangkan sistem klasifikasi otomatis untuk mendeteksi dan mengelompokkan teks *hoax* secara akurat dengan memanfaatkan model IndoBERT sebagai dasar pemodelan.
2. Menganalisis kinerja model IndoBERT dalam mendeteksi teks *hoax* berdasarkan tingkat akurasi dan efisiensi pemrosesan.

1.5 Manfaat

Penelitian tugas akhir ini diharapkan memberikan beberapa manfaat, antara lain :

1.5.1 Bagi Penulis

Penelitian ini akan memberikan manfaat bagi penulis, di antaranya :

- a. Meningkatkan pemahaman dalam penerapan *Natural Language Processing* (NLP), khususnya dalam klasifikasi teks *hoax* menggunakan model IndoBERT.
- b. Menambah wawasan dan keterampilan dalam pengembangan sistem berbasis Python.

1.5.2 Bagi Masyarakat

Penelitian ini akan memberikan manfaat bagi masyarakat atau pengguna, di antaranya :

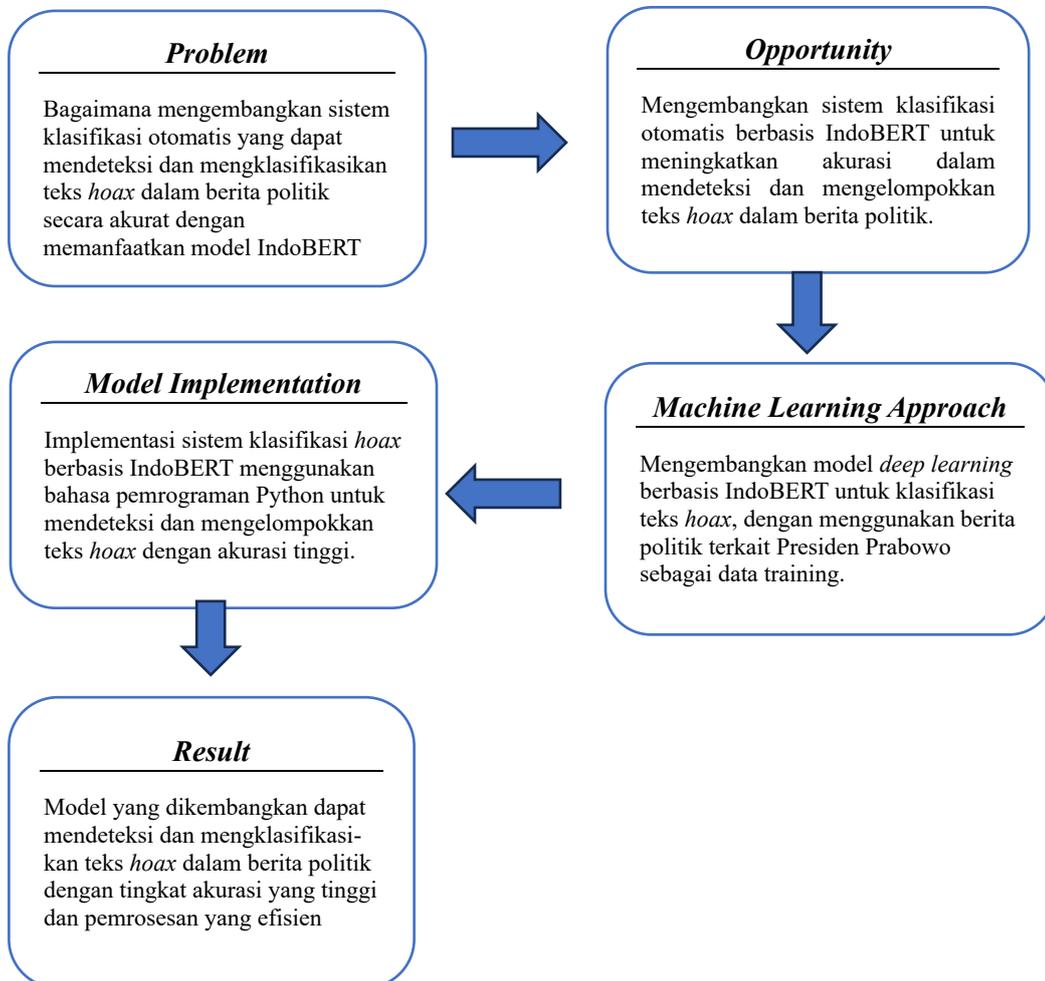
- a. Membantu masyarakat dalam mendeteksi berita *hoax* secara otomatis, sehingga dapat mengurangi penyebaran informasi yang salah.
- b. Meningkatkan kesadaran dan literasi digital masyarakat terhadap penyebaran berita *hoax*, terutama dalam ranah politik.
- c. Memudahkan pengguna dalam membedakan berita asli dan berita palsu, sehingga mereka dapat mengambil keputusan berdasarkan informasi yang lebih kredibel.

1.5.3 Bagi Akademik

Penelitian ini juga memberikan manfaat bagi dunia akademik, di antaranya :

- a. Menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan deteksi berita hoax menggunakan NLP dan model berbasis *deep learning* seperti IndoBERT.
- b. Memberikan wawasan bagi mahasiswa dan akademisi dalam penerapan teknologi AI untuk menyelesaikan permasalahan sosial, khususnya dalam deteksi berita palsu.
- c. Memperkaya sumber pustaka dan penelitian dalam bidang ilmu komputer, khususnya terkait klasifikasi teks dalam bahasa Indonesia.

1.6 Kerangka Pemikiran



Gambar 1. 1 Kerangka Pemikiran

Gambar 1.1 menunjukkan kerangka pemikiran dalam pengembangan sistem klasifikasi teks *hoax* dalam berita politik. Proses dimulai dengan mengidentifikasi masalah, yaitu kebutuhan akan sistem deteksi *hoax* yang cepat dan akurat. Untuk itu, penelitian ini memanfaatkan model IndoBERT sebagai pendekatan utama dalam *machine learning*. Model ini dilatih untuk mengklasifikasikan teks *hoax* dengan menggunakan berita politik terkait Presiden Prabowo sebagai data pelatihan. Setelah pelatihan, model diimplementasikan dalam sistem berbasis Python. Hasilnya, model mampu mendeteksi dan mengklasifikasikan teks *hoax* dengan tingkat akurasi yang tinggi serta pemrosesan yang efisien, menjadikannya solusi yang optimal dalam identifikasi berita *hoax*.