

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pentingnya pengaruh teknologi pada masa modern ini tak dapat dipungkiri karena teknologi menyediakan segala kemudahan untuk mendukung aktivitas manusia. Fenomena ini memicu perkembangan metode dan penelitian di berbagai bidang teknologi, termasuk metode komputasi. Salah satu metode yang menonjol adalah kecerdasan buatan, yang bersifat sangat fleksibel dan merangkum berbagai disiplin ilmu. Kecerdasan buatan sebagai bidang teknologi dan ilmu memfokuskan diri pada pemahaman cara membuat komputer mampu melakukan tugas-tugas serupa dengan manusia. Dengan kata lain, studi kecerdasan buatan bertujuan untuk mengajarkan komputer melakukan hal-hal yang saat ini hanya mampu dilakukan oleh manusia [1].

Di beberapa daerah, masyarakat memilih profesi sebagai petani secara turun-temurun atau karena terbatasnya peluang pekerjaan di bidang lain, seringkali disebabkan oleh keterbatasan pendidikan. Para petani di daerah umumnya mengelola pertanian berdasarkan pengetahuan yang diwarisi dari generasi sebelumnya atau didasarkan pada pengalaman yang dimiliki oleh para petani itu sendiri [3].

Saat ini, banyak masyarakat belum sepenuhnya mengembangkan dan memanfaatkan sarana secara optimal, terutama karena keterbatasan sarana dan minimnya akses pengetahuan. Teknologi di sektor pertanian belum sepenuhnya efektif dalam pengelolaan lahan, sehingga petani masih menghadapi berbagai kendala. Banyak petani mengandalkan metode tradisional, meskipun beberapa milenial mulai mengadopsi teknologi modern [2]. Pengetahuan petani tentang hama dan penyakit tanaman terbatas dan sering hanya berasal dari sesama petani. Jika muncul jenis hama atau penyakit baru, mereka sering tidak paham. Ada pakar yang memiliki pengetahuan luas, namun jumlahnya tidak sebanding dengan jumlah petani [4]. Kesenjangan ini membuat informasi tidak

tersampaikan secara efisien. Oleh karena itu, diperlukan peningkatan komunikasi dan pertukaran informasi antara pakar dan petani agar pengetahuan tentang hama dan penyakit tanaman dapat disebarkan lebih efektif.

Solusi yang dihadirkan dalam penelitian ini adalah pengembangan dan menganalisis Chatbot berbasis Bahasa Indonesia dengan model IndoBERT dirancang khusus untuk memberikan bantuan kepada petani dalam menangani masalah penyakit tanaman. Chatbot ini dapat mendeteksi gejala, mengidentifikasi jenis penyakit, dan memberikan solusi serta rekomendasi perawatan yang sesuai. Sebagai program komputer berbasis, chatbot ini memanfaatkan pemrosesan Bahasa alami untuk memberikan pengalaman interaktif kepada pengguna.

Penelitian yang dilakukan oleh Erwin Apriliyanto, Kusriani, dan Rudyanto Arie terkait *Chatbot* untuk Pengendalian Hama Tanaman Padi dengan Metode *Artificial Intelligence Markup Language* dan Normalisasi menghasilkan akurasi sebesar 90.9%, dengan waktu respon kurang dari 5 kata adalah 0.01 detik dan lebih dari 5 kata adalah 0.02 detik menggunakan data set 1000 baris [5].

Transformers, sebagai model transmisi pionir, sepenuhnya mengandalkan *Self-attention* untuk menghasilkan representasi input dan outputnya. *Self-attention* adalah mekanisme yang memungkinkan model memahami hubungan kontekstual antar kata dalam suatu urutan data, terutama dalam konteks pemrosesan bahasa alami dan *computer vision*. Model ini khususnya diterapkan dalam tugas-tugas seperti terjemahan bahasa dan peringkasan teks, menunjukkan keberhasilan yang luar biasa dalam pemrosesan data urutan, seperti bahasa alami. *Transformasi* ini telah menjadi standar dalam bidang kecerdasan buatan, menawarkan keandalan dan efisiensi yang tinggi dalam berbagai aplikasi. *Transformers* terdiri dari dua buah tumpukan.

BERT menerapkan pembelajaran dua arah, yang memperhatikan token sebelum dan setelah dalam suatu urutan atau sekuensial. Dalam struktur pembobotan atau *embedding* BERT, terdapat tiga lapisan, yaitu *vocab*

embedding, *segment embedding*, dan *position embedding*, yang bekerja secara bersamaan [10]. Mekanisme ini dirancang untuk menangkap konteks kata berdasarkan posisi dan kedudukannya dalam kalimat.

Selain itu, BERT mendukung skema *fine-tuning* yang memungkinkan pengguna untuk memanfaatkan model dari pelatihan sebelumnya. Proses *fine-tuning* melibatkan pelatihan ulang model yang sudah di-*pre-trained* pada dataset lain, dengan penyesuaian konfigurasi dan parameter sesuai dengan kebutuhan spesifik tugas yang dihadapi [11]. Pendekatan *fine-tuning* ini memberikan fleksibilitas dan adaptabilitas pada pengguna untuk menggunakan BERT dalam berbagai konteks dan aplikasi.

IndoBERT merupakan varian bahasa Indonesia dari model *pre-trained* yang dikembangkan oleh sejumlah peneliti dari *Google AI Language*, yaitu BERT. BERT sendiri dapat membaca teks sekaligus sehingga dapat digunakan untuk data sekuensial. IndoBERT telah melalui proses pelatihan dengan lebih dari 220 juta kata yang diperoleh dari beberapa sumber utama [1].

Dalam kerangka kerja pelatihan IndoBERT terdapat dua tahap, yaitu *pretraining* dan *fine tuning*. Pada tahap *pretraining*, IndoBERT dilatih dengan dataset yang tidak memiliki label. Pada tahap ini, IndoBERT dilatih untuk menggunakan dua tugas yang diawasi sendiri, yaitu *Masked LM (Masked Language Model)* dan *Prediksi Kalimat Berikutnya*. IndoBERT belajar untuk memahami bahasa dan konteksnya. Untuk IndoBERT, pada tahap praplatihan, dilatih menggunakan dataset Indo4B. Kemudian pada tahap *fine tuning*, model yang telah dilatih pada tahap praplatihan disesuaikan dan kemudian dilatih dengan dataset tertentu yang memiliki label [2].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan diatas, permasalahan yang dapat diidentifikasi yaitu:

1. Bagaimana mengukur efektivitas *chatbot* dalam mendukung petani dalam mengelola hama dan penyakit pertanian serta meningkatkan hasil panen.

2. Bagaimana meningkatkan aksesibilitas dan akurasi model untuk *chatbot* hama dan penyakit tanaman.

1.3 Pertanyaan Penelitian

Berdasarkan rumusan masalah pada latar belakang diatas, maka pertanyaan-pertanyaan dalam penelitian ini yaitu:

1. Apakah penerapan model IndoBERT dalam Chatbot Penyakit Tanaman meningkatkan pemahaman bahasa alami dan responsivitas terhadap pertanyaan?
2. Bagaimana hasil evaluasi model pada data validasi dapat memberikan wawasan terhadap kualitas dan akurasi Chatbot Penyakit Tanaman?

1.4 Batasan Masalah

Dalam usaha mencapai pemahaman yang komprehensif terhadap masalah yang sedang dihadapi, perlu ditetapkan batasan masalah sebagai panduan dalam penelitian. Berdasarkan evaluasi terhadap rumusan masalah dan tujuan penelitian, berikut adalah sejumlah batasan yang diberlakukan:

1. Penelitian ini terbatas pada penggunaan Bahasa Indonesia sebagai Bahasa utama interaksi dengan Chatbot Penyakit Tanaman.
2. Penelitian ini berfokus pada analisis model IndoBERT sebagai basis untuk Chatbot Penyakit Tanaman.
3. Dataset yang dipakai merupakan dari buku berjudul “Hama dan Penyakit Tanaman” yang diterbitkan oleh Yayasan Kita Menulis yang diterbitkan tahun 2021.
4. Fokus klasifikasi dan identifikasi penyakit tanaman dibatasi pada jenis penyakit tertentu yang tersedia pada buku berjudul berjudul “Hama dan Penyakit Tanaman” yang diterbitkan oleh Yayasan Kita Menulis yang diterbitkan tahun 2021.

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah dipaparkan, dapat dijabarkan tujuan penelitian sebagai berikut:

1. Mengukur performa yang dihasilkan dari model IndoBERT dalam membangun *chatbot* penyakit tanaman.
2. Menganalisis performa dari berbagai *hyper parameter* yang disesuaikan dengan model IndoBERT.

1.6 Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah dipaparkan, maka dapat diketahui manfaat dari penelitian ini bahwa:

1. Manfaat bagi peneliti, menambah pengetahuan tentang pemanfaatan teknologi kecerdasan buatan, khususnya dalam pembuatan chatbot menggunakan IndoBERT.
2. Manfaat teoritis, hasil penelitian dapat digunakan sebagai referensi untuk penelitian selanjutnya dengan metode yang sejenis.