

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Dalam penelitian terkini, perkembangan *chatbot*, terutama yang menggunakan model BERT dan IndoBERT, telah menunjukkan manfaat signifikan dalam penyediaan informasi dan layanan. Dari literatur terdahulu, jelas bahwa pengembangan *chatbot* dapat meningkatkan efisiensi dan kenyamanan pengguna. Dengan fokus pada konteks penyakit tanaman, implementasi chatbot berbasis model bahasa alami menjanjikan kontribusi positif dalam mendukung petani dengan solusi informasi yang cepat dan tepat.

Penelitian yang dilakukan oleh Adriansyah Dwi Rendragraha, Moch. Arif Bijaksana, dan Ade Romadhony pada tahun 2021 mengenai deteksi Bahasa kasar dalam komentar berita online Indonesia menggunakan model BERT, bertujuan untuk mengimplementasikan dan mengevaluasi metode *transformer* BERT dalam mendeteksi Bahasa kasar pada teks bahasa Indonesia. Penelitian ini mencakup evaluasi sistem untuk memastikan keakuratannya dalam melakukan klasifikasi pada kalimat uji. Metode evaluasi menggunakan nilai *Macro Average F1-Score* dan nilai *F1-score* untuk masing-masing kelas label. Hasil penelitian menunjukkan bahwa model BERT *Multilingual* memiliki *Macro Average F1 Score* sebesar 54%, sedangkan model yang dibangun dari awal (*scratch model*) dengan dataset bahasa Indonesia mendapatkan nilai *Macro Average F1 Score* sebesar 50% [14].

Penelitian yang dilakukan oleh Erwin Apriliyanto, Kusriani, dan Rudyanto Arief pada tahun 2020 membahas pembuatan *chatbot* untuk pengendalian Hama tanaman Padi dengan Metode *Artificial Intelligence Markup Language* dan Normalisasi. Penelitian ini bertujuan untuk menguji akurasi nilai jawaban pada *chatbot*. Metode penelitian mencakup penggunaan data pertanyaan dengan jumlah kata di bawah 5 dan di atas 5, serta penggunaan data pertanyaan sesuai dan di luar kata kunci pada *chatbot* ini. Sebanyak 50 data pertanyaan diuji empat kali, kemudian diambil rata-rata hasilnya. Hasil

penelitian menunjukkan akurasi sebesar 90.9%, dengan waktu respon untuk menjawab pertanyaan kurang dari 5 kata adalah 0.01 detik, dan untuk lebih dari 5 kata adalah 0.02 detik, menggunakan dataset sebanyak 1000 baris [5].

Ketiga, penelitian yang dikerjakan oleh Arnetta Rahmawati, Andry Alamsyah, dan Ade Romadhony tentang analisis deteksi berita *hoax* menggunakan IndoBERT yang ditulis pada tahun 2022. Penelitian ini bertujuan untuk mendeteksi berita hoaks menggunakan IndoBERT, *Support Vector Machine*, dan *Naïve Bayes* dengan mengklasifikasikan berita dalam bahasa Indonesia dan menemukan model terbaik. Data yang dikumpulkan mencakup berbagai topik di situs berita seperti Detik, Liputan 6, Kompas, Cek Fakta, dan Turnbackhoax. Sebanyak 2000 berita terkumpul yang berisi 1000 data berita hoaks dan 1000 data berita non-hoaks. Akurasi terbaik diperoleh oleh algoritma IndoBERT, dengan nilai 90%. Dari hasil ini diketahui bahwa IndoBERT memiliki kinerja terbaik dalam mendeteksi berita hoaks dengan mendeteksi aspek kontekstual yang terkait dengan pola yang lebih kompleks [15].

Keempat, penelitian yang dikerjakan oleh Ilham Rizky Hidayat, dan Warih Maharani pada tahun 2022 mengenai analisis deteksi depresi umum dengan IndoBERT. Dalam penelitian ini, pengguna yang dianalisis adalah pengguna yang memiliki lebih dari 1000 tweet dalam bahasa Indonesia. Kemudian, dilakukan *crawling*/pengambilan data tweet pengguna. Setelah itu, dilakukan pra-pemrosesan data. Setelah itu, klasifikasi data yang diperoleh menggunakan metode IndoBERT, dan model kemudian memberikan nilai akurasi analisis deteksi ini menggunakan metode IndoBERT dengan nilai akurasi sebesar 51% dan F1-Score sebesar 31% [12].

Kelima, penelitian yang dikerjakan oleh Ni Made Dwipani Puspitarini, Yuliant Sibaroni, dan Sri Suryani Prasetyowati untuk mendeteksi kepribadian *Big Five* menggunakan IndoBERT dan *Gaussian Naïve Bayes* pada tahun 2023. Tujuan dari penelitian ini adalah untuk mendeteksi kepribadian *Big Five* berdasarkan cuitan Bahasa Indonesia dan meningkatkan performanya dengan menggabungkan pembelajaran mesin dengan deep learning, yaitu *Gaussian*

Naive Bayes dan model IndoBERT. Model gabungan yang diusulkan dalam penelitian ini adalah menjumlahkan vektor probabilitas log pada setiap model. Terkumpul 3.342 cuitan dari 111 akun Twitter yang digunakan sebagai dataset. Penelitian ini juga menerapkan normalisasi *min-max* untuk mereskalasi data. Hasil penelitian menunjukkan bahwa untuk seluruh dataset, model gabungan memiliki skor akurasi lebih tinggi daripada *Gaussian Naive Bayes* sebesar 5,42% dan IndoBERT sebesar hampir 2%, yang menunjukkan bahwa model gabungan lebih baik daripada model *Gaussian Naive Bayes* dan IndoBERT [16].

<i>Year</i>	<i>Author</i>	<i>Country of the Author</i>	<i>Objective</i>	<i>Contribution</i>	<i>Data</i>	<i>Methodology</i>	<i>Result/Conclusion</i>
2021	Adriansyah Dwi Rendragraha, Moch. Arif Bijaksana, Ade Romadhony	Indonesia	Mengimple mentasikan dan mengevalua si BERT dalam deteksi bahasa kasar pada teks bahasa Indonesia	Penerapan dan Pengenmbangan model BERT untuk mendeteksi Bahasa kasar	<i>Crawling</i> data pada twitter, berita kompas, tempo dan kalimat mix Wikipedia dengan total ukuran 95, 876KB	Dalam penelitian ini, model yang digunakan berupa model BERT dan model pre-train BERT Multilingual untuk menjadi baseline	Hasil dari Scratch model yang dilatih dengan dataset bahasa Indonesia mendapat Macro Average F1 Score sebesar 50% dibandingkan dengan BERT Multilingual sebesar 54%.
2020	Erwin Apriliyanto, Kusriani,	Indonesia	Menguji akurasi hasil jawaban	Penerapan dan pengembang	Data dibuat dengan data pertanyaan	Dalam penelitian ini menggunakan Metode Artificial Intelligence	Hasil penelitian n ini adalah mendapatkan akurasi 90.9 %, sedangkan waktu responsetime

	Rudyanto Arief		pada chatbot	n chatbot untuk pengendalian hama dan tanaman padi	dengan kata dibawah 5 kata dan diatas 5 kata, serta menggunakan data pertanyaan sesuai dengan kata kunci dan diluar kata kunci pada chatbot	Markup Language dan Normalisasi	untuk menjawab pertanyaan kurang dari 5 kata adalah 0.01 detik, dan untuk lebih dari 5 kata adalah 0.02 detik dengan data set 1000 baris.
2022	Arnetta Rahmawati, Andry Alamsyah,	Indonesia	Mendeteksi berita palsu menggunakan IndoBERT,	Membandingkan ketiga model untuk mencari hasil terbaik	Dataset tersebut terdiri dari total 2000 judul berita.	Penelitian ini menggunakan 3 model yaitu IndoBERT, Naïve Bayes, dan Support Vector Machine.	Hasil analisis menemukan bahwa IndoBERT memiliki kinerja terbaik dalam mendeteksi berita palsu dengan mendeteksi aspek

	Ade Romadhony		Support Vector Machine, dan Naïve Bayes		Dataset non-palsu berasal dari situs berita resmi di Indonesia, termasuk Liputan 6, Detik, dan Kompas, semua dari tahun 2021		kontekstual yang terkait dengan pola yang lebih kompleks
2022	Ilham Rizky Hidayat, Warih Maharani	Indonesia	Analisis deteksi depresi umum dengan IndoBERT	Penerapan model IndoBERT untuk menganalisa deteksi depresi	Dalam penelitian ini, dataset berasal dari crawling tweet pengguna di	Penelitian ini menggunakan model IndoBERT yang sebelumnya telah dilakukan proses pengambilan data dan <i>pre-processing</i> pada	Hasil penelitian memberikan nilai akurasi dari analisis deteksi menggunakan metode IndoBERT dengan nilai akurasi sebesar 51% dan F1-Score sebesar 31%.

				dengan IndoBERT	Twitter	dataset.	
2023	Ni Made Dwipani Puspitarini, Yuliant Sibaroni, Sri Suryani Prasetyowati	Indonesia	Mendeteksi kepribadian Big Five menggunakan an IndoBERT dan Gaussian Naïve Bayes	Penerapan pengembang dan perbandingan dari model IndoBERT dan Gaussian Naive Bayes untuk mendeteksi kepribadian Big Five	Dataset yang berhasil dikumpulkan dalam penelitian ini adalah 90.260 tweet dari 111 akun Twitter. Sebanyak 22.208 tweet dalam bahasa Indonesia dipilih dari 90.260 data	Penelitian ini menggunakan model IndoBERT dan Gaussian Naive Bayes kemudian penjumlahan vektor probabilitas log pada setiap model. Dikumpulkan 3.342 tweet dari 111 akun Twitter yang digunakan sebagai dataset. Penelitian ini juga menerapkan normalisasi min–maks untuk mereskalasi data.	Hasil menunjukkan bahwa untuk seluruh dataset, model gabungan memiliki skor akurasi lebih tinggi dibandingkan dengan Gaussian Naive Bayes sebesar 5,42% dan IndoBERT hampir 2%.

Tabel 2. 1 Ringkasan Penelitian Sebelumnya

Penelitian yang sudah dilakukan menggunakan IndoBERT mendapatkan performa yang baik dalam membangun dan menggunakan IndoBERT untuk Chatbot. Penelitian [1], [3], dan [4] dijadikan rujukan utama dalam mengerjakan penelitian ini. Penelitian ini menggunakan metode-metode pada penelitian tersebut untuk dijadikan referensi dalam mengembangkan penelitian ini, terutama dalam mengoptimalkan model yang digunakan dengan melakukan skenario riset untuk mendapatkan model yang optimal dalam pembuatan *chatbot*.

2.2 Landasan Teori

2.2.1 *Natural Language Processing*

Pemrosesan bahasa alami (*Natural Language Processing* - NLP) merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memberikan kemampuan pada komputer untuk memahami teks dan kata-kata yang diucapkan dengan cara yang sama seperti yang dapat dilakukan manusia. Klasifikasi teks merupakan salah satu tugas NLP yang dapat memberikan kategori terhadap teks secara otomatis berdasarkan konteks dari teks tersebut dengan bantuan metode *machine learning* maupun *deep learning* [17].

2.2.2 Chatbot

Chatbot adalah suatu program atau perangkat lunak yang menggunakan bahasa pemrosesan alami (*Natural Language Processing*) dalam pertanyaan dan sistem jawab (*system QA*). Merupakan suatu aplikasi komputer yang dapat melakukan percakapan, baik melalui metode auditori maupun tekstual. Umumnya, *chatbot* dirancang untuk meniru atau mensimulasikan perilaku manusia dalam percakapan, mampu menginterpretasikan dan memberikan respons terhadap berbagai jenis input manusia. Dalam interaksinya, *chatbot* dapat memindai kata kunci dalam input yang diterima dan memberikan respons dengan kata kunci yang paling relevan atau dengan pola kata yang serupa dari basis data tekstual [17].

2.2.3 Hama dan Penyakit Tanaman

Dalam konteks budidaya tanaman, pengertian hama secara sempit merujuk pada semua jenis hewan yang dapat merusak tanaman atau hasil pertanian dengan aktivitas hidupnya, yang pada akhirnya dapat menyebabkan kerugian ekonomis. Hama-hama ini dapat mencakup berbagai jenis organisme seperti serangga, burung, dan hewan lainnya yang berpotensi merugikan tanaman.

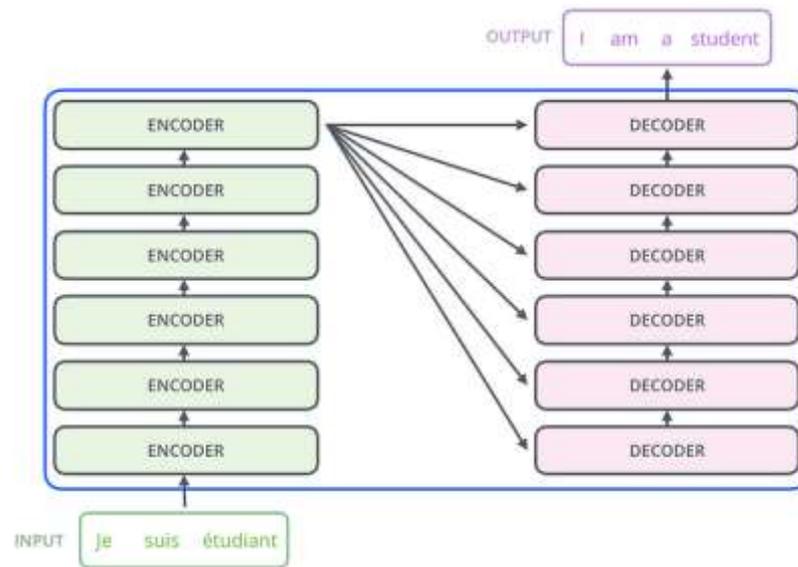
Sementara itu, penyakit tumbuhan didefinisikan sebagai gangguan fisiologi pada tanaman yang disebabkan oleh faktor primer, baik yang bersifat biotik (berkaitan dengan organisme hidup) maupun abiotik (berkaitan dengan faktor lingkungan non-hidup). Penyakit tumbuhan dapat memengaruhi sel atau jaringan tanaman, menyebabkan tanaman mengalami ketidaknormalan, dan gangguan ini dapat bersifat terus menerus dan merugikan bagi pertumbuhan tanaman [18].

2.2.4 *Transformers*

Transformers telah merevolusi bidang pemrosesan bahasa alamiah. Berbeda dengan RNN dan jaringan *Long Short Term Memory* (LSTM), di mana pelatihan dilakukan secara berurutan, desain arsitektur *transformer* memungkinkan pemrosesan paralel dan memungkinkan pembuatan penyandian laten yang kaya (Vaswani et al., 2017). Representasi kontekstual laten dari ucapan melalui mekanisme *self-attention* membuat *transformers* menjadi alat yang kuat untuk berbagai aplikasi turunan seperti jawaban pertanyaan dan ringkasan teks (Devlin et al., 2018) [5].

Dalam jaringan *neural*, lapisan *encoder* memiliki tugas memproses nilai input dan menghasilkan vektor representasi yang menyimpan informasi dan fitur-nilai dari input tersebut. Dalam model *transformer*, setiap lapisan *encoder* terdiri dari dua komponen utama, yaitu *self-attention* dan lapisan *feed forward*. Perbedaan mendasar dengan *Convolutional Neural Network* (CNN) terletak pada pendekatan *transformer* terhadap panjang *input* dan *output*. *Transformer* tidak membatasi panjang *input* atau *output*, sehingga mampu menangani urutan

data dengan panjang yang bervariasi, menjadi lebih fleksibel untuk tugas-tugas seperti pemrosesan bahasa alami, berbeda dengan CNN yang lebih efektif untuk data matriks seperti citra.

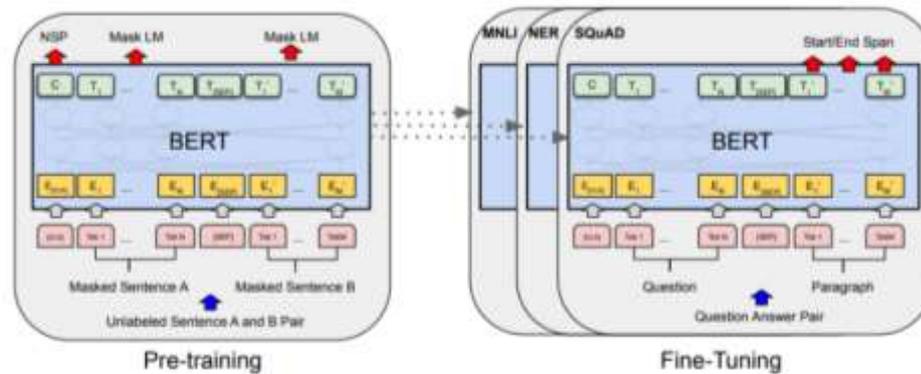


Gambar 2.1 Lapisan Encoder Decoder dalam Arsitektur Transformer

Pemanfaatan mekanisme *self-attention* memungkinkan setiap kata dalam suatu kalimat untuk berinteraksi satu sama lain (*self*) dan menentukan sejauh mana tingkat perhatian (*attention*) diberikan pada setiap kata.

2.2.5 IndoBERT

IndoBERT merupakan hasil modifikasi dari BERT Base yang mengikuti konfigurasi BERT-Base (*uncased*). Dalam implementasinya, IndoBERT menggunakan mekanisme *transformer*, suatu mekanisme yang mempelajari interaksi antar kata dalam teks atau kalimat. Secara formal, *transformer* terdiri dari dua bagian, yaitu *encoder* yang bertugas membaca teks masukan, dan *decoder* yang bertanggung jawab untuk menghasilkan prediksi. Berbeda dengan model sekuensial lainnya yang membaca teks secara berurutan dari kiri ke kanan atau sebaliknya, *encoder transformer* dapat membaca seluruh kata secara simultan. Keunikan ini memungkinkan model untuk memahami konteks suatu kata berdasarkan lingkungannya [12].



Gambar 2.2 Kerangka IndoBERT

Dalam kerangka pelatihan BERT, terdapat dua tahap, yaitu pretraining dan fine tuning. Pada tahap pretraining, BERT dilatih dengan dataset yang tidak memiliki label. Pada tahap ini, BERT dilatih untuk menyelesaikan dua tugas yang bersifat self-supervised, yaitu Masked LM (Masked Language Model) dan Next Sentence Prediction. BERT belajar untuk memahami bahasa dan konteksnya. Untuk IndoBERT, pada tahap pretraining, model dilatih menggunakan dataset Indo4B. Kemudian pada tahap fine tuning, model yang telah dilatih pada tahap pretraining disesuaikan dan dilatih kembali dengan dataset khusus yang memiliki label [2].

1. *Encoder*

Encoder terdiri dari dua lapisan utama, yaitu lapisan self-attention dan jaringan saraf pengumpan balik (feedforward). Mekanisme self-attention memungkinkan simpul (node) saat ini untuk tidak hanya fokus pada kata yang sedang diproses, melainkan juga memperoleh makna semantik dari konteks keseluruhan. Sementara itu, decoder juga terdiri dari dua lapisan seperti yang dijelaskan pada encoder, namun di antara kedua lapisan tersebut terdapat lapisan perhatian (attention) yang membantu simpul saat ini untuk mendapatkan informasi kunci yang perlu diberikan perhatian khusus [6].

2. *Decoder*

Bagian Decoder mirip dengan bagian Encoder, namun memiliki masked multi-head attention di bagian bawahnya. Mask di sini merujuk pada suatu masker yang digunakan untuk menyembunyikan nilai tertentu, sehingga nilai-nilai tersebut tidak mempengaruhi proses pembaharuan parameter. Terdapat dua jenis masker dalam model *Transformer* — masker padding dan masker urutan. Masker padding digunakan dalam semua operasi perhatian dot-produk berskala, sementara masker urutan hanya digunakan dalam self-attention decoder. Masker tersebut berperan penting dalam memastikan bahwa proses pembelajaran tidak terpengaruh oleh nilai-nilai yang seharusnya diabaikan, seperti nilai-nilai padding dalam urutan [7].

3. *Self Attention*

Self-attention Self-attention merupakan metode yang digunakan oleh *Transformer* untuk mengubah "pemahaman" kata-kata terkait menjadi pemahaman terhadap kata yang sedang diproses. Proses self-attention dimulai dengan perhitungan tiga vektor baru. Dalam penelitian ini, dimensi vektor tersebut adalah 512. Ketiga vektor ini dikenal sebagai Query, Key, dan Value. Proses pembuatan vektor-vektor ini melibatkan perkalian vektor embedding kata dengan matriks yang diinisialisasi secara acak (dengan dimensi (64, 512) dalam penelitian ini) yang nilainya diperbarui selama proses back-propagation [8].

4. *Output Layer*

Setelah lapisan decoder dieksekusi, ditambahkan lapisan fully connected dan softmax pada akhirnya untuk memetakan vektor hasil ke kata-kata dalam kosakata. Lapisan Linear adalah jaringan saraf penghubung penuh yang memproyeksikan vektor hasil decoder ke vektor logits dengan lebar sesuai dengan kosakata model. Lapisan softmax mengubah skor-skor tersebut menjadi probabilitas, dan kata dengan probabilitas tertinggi dipilih sebagai output [9].

2.2.6 Evaluasi F1

F1-score adalah ukuran yang menggabungkan presisi dan recall menjadi satu nilai tunggal yang menggambarkan kinerja keseluruhan model. F1-score didefinisikan sebagai harmonic mean dari presisi dan recall. Rumusnya adalah:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

2.2.7 Evaluasi *Precision* dan *Recall*

Precision mengukur akurasi model dalam mengidentifikasi contoh positif dari semua contoh yang diprediksi sebagai positif. Ini didefinisikan sebagai rasio antara jumlah prediksi positif yang benar (true positive) dibagi dengan total jumlah prediksi positif yang dilakukan oleh model.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Precision memberikan informasi tentang seberapa banyak prediksi positif yang relevan dibandingkan dengan total prediksi positif yang dilakukan oleh model. Precision yang tinggi menunjukkan bahwa model memiliki sedikit kesalahan dalam memprediksi contoh positif.

Sebaliknya, recall (atau sensitivity) mengukur kemampuan model dalam menemukan semua contoh positif yang seharusnya diprediksi sebagai positif.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Recall memberikan informasi tentang seberapa banyak dari semua contoh positif yang berhasil ditemukan oleh model. Recall yang tinggi menunjukkan bahwa model mampu mengidentifikasi sebagian besar contoh positif yang seharusnya diprediksi sebagai positif.