SKRIPSI

IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA

DISUSUN OLEH

ARNIYANTIKA PUTRI MEDIANI
2109020021

PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025

IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) dalam Program Studi Teknologi Informasi pada Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara

ARNIYANTIKA PUTRI MEDIANI
2109020021

PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025

LEMBAR PENGESAHAN

Judul Skripsi	:	IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA
Nama Mahasiswa	:	ARNIYANTIKA PUTRI MEDIANI
NPM	:	2109020021
Program Studi	:	TEKNOLOGI INFORMASI

Menyetujui

Komisi Pembimbing

(Farid Akbar Siregar, S.Kom., M.Kom)
NIDN. 0104049401

Ketua Program Studi

(Fatma Sari Hutagalung, S.Kom., M.Kom)
NIDN. 0117019301

Dekan

(Dr. Al-Khowarizmi, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, Oktober 2025

Yang membuat pernyataan

ARNIYANTIKA PUTRI MEDIANI
NPM. 2109020021

PERNYATAAN PERSETUJUAN PUBLIKASI

KARYA ILMIAH UNTUK KEPENTINGAN

AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama	: ARNIYANTIKA PUTRI MEDIANI
NPM	: 2109020021
Program Studi	: TEKNOLOGI INFORMASI
Karya Ilmiah	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif *(Non-Exclusive Royalty free Right)* atas penelitian skripsi saya yang berjudul:

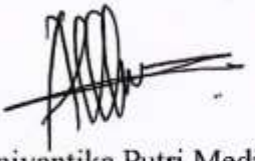
IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, Oktober 2025

Yang membuat pernyataan



Arniyantika Putri Mediani

NPM. 2109020021

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap	: Arniyantika Putri Mediani
Tempat dan Tanggal Lahir	: Bdr. Klipa, 21 April 2003
Alamat Rumah	: Dusun XVII Jl. Pusaka Gg. Bala
Telepon/Faks/HP	: 085362095646
E-mail	: arniyantika.putriii@gmail.com
Instansi Tempat Kerja	: -
Alamat Kantor	: -

DATA PENDIDIKAN

SD SWASTA AL MU'MIN	TAMAT: 2015
SMP CERDAS MURNI	TAMAT: 2018
SMK SWASTA TELADAN MEDAN	TAMAT: 2021

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh,

Puji syukur kehadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan/proposal ini dengan baik. Laporan ini disusun sebagai bagian dari tugas akhir dalam menyelesaikan studi pada Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara (UMSU). Melalui laporan ini, penulis berupaya mengangkat isu dan solusi berbasis teknologi yang relevan dengan kebutuhan masyarakat modern, dengan harapan dapat memberikan kontribusi positif dalam pengembangan ilmu pengetahuan dan pemanfaatan teknologi secara berkelanjutan. Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Agussani, MAP, selaku Rektor Universitas Muhammadiyah Sumatera Utara.
2. Bapak Prof. Dr. Muhammad Arifin, S.H., M.Hum, Bapak Prof. Dr. Akrim, M.Pd, dan Bapak Assoc. Prof. Dr. Rudianto, S.Sos., M.Si, selaku Wakil Rektor I, II, dan III Universitas Muhammadiyah Sumatera Utara.
3. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom, selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi serta Bapak Halim Maulana, S.T., M.Kom, dan Bapak Dr. Lutfi Basit, S.Sos., M.I.Kom., selaku Wakil Dekan I dan II Fakultas Ilmu Komputer dan Teknologi Informasi.
4. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom. Ketua Program Studi Teknologi Informasi serta Bapak Mhd Basri, S.Si, M.Kom., Sekretaris Program Studi Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.

5. Bapak Farid Akbar Siregar, S.Kom., M. Kom selaku dosen pembimbing yang telah membantu dan membimbing penulis serta memberikan arahan penulis dalam mengerjakan tugasnya.
6. Terima kasih yang tulus penulis sampaikan kepada Bapak dan Ibu tercinta selaku orangtua penulis. Berkat doa, kasih sayang, dan pengorbanan kalian menjadi kekuatan terbesar dalam setiap langkah penulis hingga mampu melewati proses ini. Semoga pencapaian ini menjadi bagian dari kebahagiaan dan kebanggaan kalian, sebagaimana kalian yang tidak pernah menyerah untuk pencapaian ini.
7. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini. Penulis juga berharap, laporan ini dapat memberikan manfaat nyata bagi pembaca serta menjadi kontribusi kecil dalam pengembangan ilmu pengetahuan, khususnya di bidang teknologi informasi. Semoga segala bantuan dan dukungan yang telah diberikan mendapatkan balasan kebaikan yang berlipat ganda dari Allah SWT.

Akhir kata, penulis berharap laporan ini dapat memberikan manfaat bagi pembaca, serta menjadi referensi bagi pengembangan penelitian dan inovasi di bidang teknologi informasi, khususnya di lingkungan akademik UMSU.

Wassalamu'alaikum warahmatullahi wabarakatuh.

Medan, 8 Agustus 2025
Penulis

Arniyantika Putri Mediani

IMPLEMENTASI TRANSFER LEARNING DENGAN INDOBERT PADA QUESTION ANSWERING BAHASA INDONESIA

ABSTRAK

Penelitian ini membahas pengembangan sistem Question Answering (QA) berbasis teks yang mampu menjawab pertanyaan secara otomatis dengan mengandalkan konteks yang diberikan oleh pengguna. Sistem dirancang menggunakan model pretrained BERT (Bidirectional Encoder Representations from Transformers), yang mampu memahami hubungan antara pertanyaan dan konteks untuk menentukan jawaban secara tepat. Proses kerja sistem meliputi tokenisasi, penerapan multi-head self-attention melalui arsitektur Transformer, serta perhitungan nilai start logit dan end logit untuk menentukan rentang teks yang paling mungkin sebagai jawaban.

Sistem tidak hanya menampilkan satu jawaban terbaik, tetapi juga menyajikan sejumlah kemungkinan jawaban lain yang disusun berdasarkan skor akurasi. Evaluasi dilakukan menggunakan metrik Exact Match (EM) dan F1 Score untuk mengukur kesesuaian antara jawaban sistem dan jawaban referensi. Hasil penelitian menunjukkan bahwa pendekatan ini efektif dalam mengidentifikasi jawaban yang relevan dari konteks panjang dengan efisiensi dan tingkat ketepatan yang tinggi. Sistem ini berpotensi diterapkan dalam berbagai kebutuhan pencarian informasi otomatis, seperti layanan bantuan digital, chatbot, dan aplikasi pendidikan.

Kata kunci: Question Answering, BERT, Start–End Logit, F1 Score, Multi-head Attention.

MPLEMENTATION OF TRANSFER LEARNING WITH INDOBERT IN INDONESIAN LANGUAGE QUESTION ANSWERING

ABSTRACT

This research discusses the development of a text-based Question Answering (QA) system capable of automatically answering questions based on user-provided context. The system is designed using a pretrained BERT (Bidirectional Encoder Representations from Transformers) model, which understands the relationship between questions and context to determine the correct answer. The system's workflow includes tokenization, the implementation of multi-head self-attention through the Transformer architecture, and the calculation of start and end logit values to determine the most likely text range as an answer.

The system not only displays a single best answer but also presents a number of other possible answers, ranked by accuracy score. Evaluation was conducted using the Exact Match (EM) and F1 Score metrics to measure the agreement between the system's answers and reference answers. The results show that this approach is effective in identifying relevant answers from long contexts with high efficiency and accuracy. This system has the potential to be applied to various automated information retrieval needs, such as digital assistance services, chatbots, and educational applications.

Keywords: Question Answering, BERT, Start–End Logit, F1 Score, Multi-head Attention.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

Salah satu fokus penting dalam pemrosesan bahasa alami (NLP) adalah pengembangan sistem *Question Answering* (QA) dalam Bahasa Indonesia. Sistem *Question Answering* merupakan suatu sistem yang didalamnya menggunakan *Natural Language Processing* (NLP) sebagai pemrosesan bahasa alami manusia untuk menggali informasi kemudian secara otomatis menjawab pertanyaan yang telah diajukan (Chandra & Suyanto, 2019). Kemajuan dalam teknologi *transfer learning*, khususnya dengan model besar seperti *Bidirectional Encoder Representations for Transformers* (BERT), telah membawa perubahan signifikan dalam pendekatan pemrosesan bahasa alami. *IndoBERT*, sebagai model BERT yang dirancang khusus untuk Bahasa Indonesia, dilatih menggunakan korpus teks besar dari berbagai domain, memungkinkan model ini memahami konteks bahasa dengan lebih baik dibandingkan model generik seperti BERT multibahasa (Koto et al., 2020).

Dengan kemajuan yang cepat dari teknologi dan informasi pada saat ini, hingga kita semakin sulit untuk terlepas dari perkembangannya. Hal ini menyebabkan arus informasi berkembang dengan cepat (Topsakal & Akinci, 2023). Pengiriman informasi saat ini dapat dilakukan dengan cepat dan mudah, salah satunya melalui aplikasi tanya jawab atau Question Answering System (QAS) terkait materi yang ingin diketahui oleh pengguna (Apriliani et al., 2023). QAS adalah suatu sistem yang memungkinkan pengguna untuk mengajukan pertanyaan dalam bahasa sehari-hari dan mendapatkan jawaban secara cepat dan

ringkas, kadang-kadang disertai dengan informasi tambahan untuk mendukung kebenaran jawaban tersebut (Agrawal et al., 2021).

Dalam konteks analisis sentimen teks Bahasa Indonesia, penelitian terkait penerapan metode transfer learning juga pernah dilakukan sebelumnya. Dalam penelitian tersebut ditemukan bahwa model yang telah dilakukan proses pre-train seperti Vader, Textblob, dan Flair memiliki performa yang lebih unggul dibandingkan dengan model klasifikasi tradisional dalam kasus analisis sentimen (Arifiyanti et al., 2022). Selain itu, sebelumnya juga pernah dilakukan beberapa penelitian terkait penerapan metode transfer learning pada model IndoBERT. Sebuah penelitian membahas tentang implementasi model IndoBERT dalam analisis sentimen untuk dataset review aplikasi mobile di Indonesia. Temuan penelitian menunjukkan bahwa model IndoBERT mencapai rata-rata akurasi yang lebih tinggi dalam memproses data berbasis leksikon dibandingkan dengan model BERT (Nugroho et al., 2021). Pada penelitian lain, penerapan metode transfer learning pada model IndoBERT untuk analisis sentimen pada data review aplikasi kesehatan juga pernah dilakukan. Hasil penelitian menunjukkan bahwa penerapan metode transfer learning dengan model IndoBERT pada data kesehatan menghasilkan tingkat akurasi yang tinggi, mencapai 96% akurasi, 95% presisi, dan 95% F1-score (Imaduddin et al., 2023).

Metode transfer learning merupakan teknik pembelajaran mesin yang memanfaatkan pengetahuan yang telah dipelajari dari model sebelumnya untuk menyelesaikan tugas baru (Dhyani, 2021). Disini ide untuk menerapkan metode transfer learning sebagai evaluasi kemampuan model IndoBERT dalam implementasi Sistem Question Answering dalam bahasa indonesia. Pengetahuan

yang akan ditransfer adalah kemampuan IndoBERT dalam memahami konteks dan representasi kata. Model ini telah terbukti efektif dalam tugas analisis sentimen teks Bahasa Indonesia (Nugroho et al., 2021). Maka dari itu, dengan menerapkan metode transfer learning model ini memberikan peluang baru untuk mengatasi tantangan QA dalam Bahasa Indonesia. Dalam penelitian ini, dataset IndoLEM (*Indonesian Language Evaluation and Benchmark*) digunakan sebagai sumber data utama. IndoLEM menyediakan data terstruktur untuk tugas-tugas NLP, termasuk QA, yang dirancang untuk mengevaluasi kemampuan model memahami teks dan menjawab pertanyaan dalam Bahasa Indonesia.

Dataset ini sangat relevan untuk membangun dan mengevaluasi sistem QA berbasis IndoBERT. Penelitian ini membantu dalam membangun sebuah aplikasi QA berbahasa Indonesia dengan memanfaatkan model transfer learning IndoBERT dan dataset IndoLEM. Aplikasi ini dirancang untuk memiliki antarmuka yang *user-friendly* dan mampu secara otomatis menjawab pertanyaan berbasis teks dengan akurat. Untuk memastikan kualitas jawaban yang dihasilkan, kinerja model dievaluasi menggunakan metrik seperti skor *Exact Match* (EM) dan F1. Dengan menggabungkan metode IndoBERT dan dataset IndoLEM, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan sistem pemrosesan bahasa alami yang praktis, akurat, dan dapat diterapkan.

1.2 Perumusan Masalah

Berdasarkan sumber literature tantangan utama dalam pembuatan sistem jawaban pertanyaan (QA) berbasis Bahasa Indonesia adalah bagaimana merancang dan menerapkan sistem QA yang dapat memberikan jawaban yang akurat dan relevan terhadap pertanyaan yang diajukan oleh pengguna. Penelitian

ini dibuat untuk mengoptimalkan pemrosesan bahasa alami agar sistem dapat memahami konteks pertanyaan dan memberikan jawaban yang lebih akurat. Ini dicapai melalui penerapan pendekatan transfer learning pada model IndoBERT.

1.3 Batasan Masalah

Untuk membuat penelitian ini fokus dalam permasalahan yang diteliti, penulis menetapkan beberapa batasan:

1. Model yang Digunakan: Untuk membangun sistem Answering Question (QA), kami menggunakan IndoBERT sebagai model transfer learning.
2. Dataset: Data utama untuk pelatihan, pengujian, dan evaluasi model QA adalah dataset IndoLEM.
3. Metode Evaluasi: Kinerja sistem QA diukur dengan metrik evaluasi Exact Match (EM) dan skor F1.
4. Implementasi Sistem: Aplikasi pemeriksaan sistem hanya memiliki antarmuka pengujian sistem sederhana.
5. Librari Pemrograman: Menggunakan bahasa pemrograman Python untuk implementasi model NLP dengan framework NLP seperti TensorFlow.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengembangkan sistem Question Answering (QA) berbasis Bahasa Indonesia menggunakan pendekatan transfer learning dengan model IndoBERT, yang dirancang untuk menjawab pertanyaan secara akurat dan relevan.

1.5 Manfaat penelitian

Manfaat yang diperoleh dari penelitian ini adalah :

1. Pengguna: Pengguna dapat dengan mudah mengajukan pertanyaan dan mendapatkan jawaban yang akurat dan relevan dengan bantuan sistem Question Answering yang menggunakan model IndoBERT. Pengguna mendapatkan jawaban yang disajikan dalam format yang jelas, ringkas, dan mudah dipahami, sehingga mempermudah mereka dalam mendapatkan informasi yang mereka butuhkan.
2. Tim Pengembang: Dengan mengurangi waktu yang dihabiskan untuk mencari informasi, karyawan dapat lebih fokus pada tugas yang lebih penting serta dengan memberikan jawaban yang cepat dan akurat, pelanggan lebih puas, yang dapat meningkatkan loyalitas.

BAB II

LANDASAN TEORI

Bab ini membahas dasar teori yang diperlukan untuk penelitian yang membahas tulisan yang telah dibuat dan yang akan datang, serta dasar teori yang diperlukan untuk menyusun tulisan yang diusulkan.

2.1 Pengenalan NLP dan Transfer Learning

Pemrosesan Bahasa Alami (NLP) adalah cabang dari ilmu komputer dan kecerdasan buatan yang memungkinkan komputer untuk memahami dan berkomunikasi dengan bahasa manusia. Dengan NLP, komputer dapat mengenali, memahami, dan menghasilkan teks serta ucapan. Ini dilakukan dengan menggabungkan teknik linguistik, aturan bahasa, dan metode statistik serta machine learning. Penelitian di bidang NLP telah mendorong kemajuan dalam AI generatif, seperti model bahasa besar yang dapat berkomunikasi dan model yang dapat membuat gambar berdasarkan permintaan. Saat ini, NLP sudah menjadi bagian dari kehidupan sehari-hari kita, digunakan dalam mesin pencari, chatbot untuk layanan pelanggan, sistem GPS yang bisa dioperasikan dengan suara, dan asisten digital di smartphone seperti Amazon Alexa, Apple Siri, dan Cortana dari Microsoft. Selain itu, NLP juga semakin penting dalam dunia bisnis, membantu perusahaan untuk meningkatkan efisiensi, mengotomatiskan proses, dan meningkatkan produktivitas karyawan (Muliyono, 2021).

Transfer Learning adalah teknik pembelajaran mesin yang efisien, dimana model dilatih terlebih dahulu pada kumpulan data besar untuk tugas umum dan kemudian disempurnakan pada data yang lebih kecil dan spesifik. Dalam konteks

Natural Language Processing (NLP), teknik ini memanfaatkan model bahasa besar yang telah dilatih sebelumnya, seperti BERT dan GPT, untuk meningkatkan kinerja pada tugas-tugas tertentu, seperti analisis sentimen dan penerjemahan mesin, sambil mengurangi waktu dan sumber daya yang diperlukan untuk pelatihan.

2.2 Pengolahan awal teks (*Preprocessing Text)*

Langkah pengolahan awal teks diperlukan sebagai proses pertama pengolahan data, hal itu harus dilakukan untuk membersihkan data yang akan diolah agar data tidak mengandung *noise*, sehingga memiliki dimensi yang lebih kecil dan lebih terstruktur. Data akan diolah dengan lebih baik dan menginterpretasi hasil pengamatan. Berikut pengolahan awal teks yang akan dilakukan dalam penelitian ini:

1. *Remove Duplicate*

   Remove duplicate atau menghilangkan duplikasi merupakan proses menghapus baris dengan *instance* yang sama. Proses ini perlu dilakukan karena data pada media sosial sangatlah *noisy*, sehingga jika terjadi duplikasi pada data dan duplikasi tersebut tidak dihilangkan, interpretasi serta pengambilan kesimpulan pada data dapat menjadi bias (Agrawal et al., 2021).

2. *Case Folding*

   Case Folding merupakan proses menyeragamkan karakter pada suatu data, jika data merupakan dokumen maka akan diseragamkan ke dalam bentuk huruf kecil (Rosid et al., 2020). Proses ini perlu dilakukan karena data pada teks seringkali memiliki perbedaan karakter, dapat disebabkan oleh

kesalahan penulisan atau sengaja oleh pengguna.

3. Menghapus Tanda baca, angka, dan Simbol

   Proses menghapus tanda baca, angka serta simbol dilakukan pada *tweet* yang mengandungnya, hal ini dilakukan agar dapat memudahkan dalam pengolahan data.

4. *Tokenizing*

   Tokenizing adalah proses pemisahan kata-kata dari sebuah kalimat yang mana hasil dari pemisahan tersebut disebut sebagai token (Rosid et al., 2020). Proses ini perlu dilakukan karena data berupa dokumen teks seringkali terdiri dari kumpulan kalimat.

5. *Remove Stopwords*

   Remove Stopwords atau penghapusan *stopwords* adalah proses menghilangkan kata-kata yang sering muncul namun tidak terdapat makna pada kata-kata tersebut (Rosid et al., 2020). Proses ini perlu dilakukan karena data pada teks pada umumnya terdapat *stopwords* yang digunakan seperti kata hubung, kata ganti, dan kata depan.

6. *Stemming*

   Stemming adalah proses memetakan dan memecahkan suatu bentuk kata menjadi kata dasar (Rosid et al., 2020). Proses *stemming* dilakukan dengan cara menghilangkan imbuhan baik imbuhan awal maupun imbuhan akhir.

2.3 Visualisasi

Visualisasi merupakan penggunaan teknologi komputer untuk mengkonfirmasipengamatan dengan menampilkan data visual interaktif (Arabnia, 1999). Definisi lain terkait visualisasi adalah metode penggunaan komputer dalam

melakukan transformasi simbol ke dalam bentuk geometri, serta dapat memberikan kemungkinan dalam proses penemuan ilmiah dengan mengamati simulasi dan komputasi sehingga dapat digunakan dalam pengembangan pemahaman yang lebih dalam dan tak terduga (Marefat et al., 1997).

Pada penggunaan visualisasi, ada beberapa karakteristik yang menjadi pembeda dalam visualisasi satu dengan lainnya. Menurut Mc Cormick, visualisasi informasi terbagi menjadi empat karakteristik yaitu menggunakan pola, menggunakan perbandingan gambar, menggunakan perbandingan animasi, dan menggunakan perbandingan warna (Marefat et al., 1997).

Karakteristik dalam menggunakan pola dapat digunakan dalam melakukan *scanning*, *recognizing*, *remembering* terhadap suatu informasi yang terlihat dan dapat bernalar dalam penentuan kesimpulan be kan perbedaan pola satu dengan pola lainnya. Dalam menggunakan perbandingan gambar, dapat dibandingkan dalam beberapa macam seperti panjang, bentuk, orientasi, gradasi warna, serta tekstur yang kemudian akan diambil rdasar kesimpulan dari informasi yang didapatkan pada gambar. Karakteristik dengan melakukan perbandingan animasi dilakukan dengan membedakan berdasarkan perjalanan waktu yang terjadi dimana tidak dapat digambarkan secara jelas dengan menggunakan gambar yang diam. Pada perbandingan terakhir yaitu perbandingan warna dapat membantu dalam mendapatkan perbedaan informasi berdasarkan deskripsi warna yang digunakan.

2.4 Web Scraping

Web Scraping merupakan teknik untuk memperoleh informasi dari website secara otomatis tanpa harus menyalin informasi secara manual. Tujuan dilakukannya *web scraping* adalah untuk mendapatkan informasi tertentu yang

kemudian dikumpulkan dalam web baru. Manfaat dalam menggunakan *web scraping* ialah agar informasi yang diperoleh menjadi lebih terfokus sehingga memudahkan dalam melakukan pencarian tertentu, karena *web scraping* hanya fokus pada cara memperoleh data melalui pengambilan serta ekstraksi data dengan ukuran data yang bervariasi. Beberapa langkah dalam melakukan *web scraping* yaitu (A. Yani et al., 2019) :

1. *Create Scraping Template*

   Pada langkah ini, pembuat program mempelajari dokumen HTML dari website yang akan diambil informasinya untuk tag HTML yang mengapit informasi yang akan diperoleh.

2. *Explore Site Navigation*

   Pembuat program perlu mempelajari teknik navigasi pada website yang akan diambil informasinya untuk ditiru pada aplikasi *web scraper* yang akan dibuat.

3. *Automate Navigation and Extraction*

   Setelah memperoleh informasi terkait dokumen HTML serta teknik navigasi pada website, aplikasi *web scraper* akan dibuat untuk mengotomatisasi pengambilan informasi dari website yang sudah ditentukan.

4. *Extracted Data and Package History*

   Langkah terakhir dalam *web scraping* adalah menyimpan seluruh informasi yang diperoleh dari aplikasi *web scraper* dalam tabel database.

2.5 Question Answering

Sistem question answering adalah pencapaian kecerdasan sintetis yang bertahan lama dan kerumitan proses pengolahan bahasa natural (Agrawal et al.,

2021) Struktur question answering memungkinkan seseorang mengajukan pertanyaan khusus dalam bahasa alami dan mendapatkan jawaban langsung dan tanggapan singkat. Sekarang, sistem QA ditemukan di mesin pencari seperti Google dan antarmuka ercakapan telepon, yang menunjukkan bahwa mereka cukup mahir dalam menanggapi potongan statistik sederhana.Namun, pertanyaan yang lebih menantang biasanya paling mudah untuk dilewati karena pelanggan harus mengembalikan daftar cuplikan yang mereka berikan dan kemudian menelusuri untuk menemukan jawaban pertanyaan mereka.(Agrawal et al., 2021).

QA System didasarkan pada komponen inti yaitu: analisa pertanyaan (question analysis) meliputi komponen analisa pertanyaan berasal dari informasi sintaksis atau semantik dari pertanyaan, pembentukan query (query generation) yaitu informasi diubah menjadi sebuah set querypencarian yang akan diteruskan ke komponen pencarian untuk menemu kembalikan konten yang relevan dari kumpulan sumber pengetahuan, hasil pencarian akan diproses oleh candidate generationyang mengekstrak kandidat dari jawaban, pada tahapan pembentukan kandidat jawaban dan penilaian jawaban (candidate answer generation and answer scoring) akan memperkirakan nilai kepercayaan dari kandidat jawaban, kemudian merangkingnya, dan menggabungkan kandidat jawaban yang sama (Verberne et al., 2011).

Terdapat beberapa tipe QA berdasarkan jawaban yang dihasilkannya. Tipe pertanyaan yang ditangani sebuah QA system terbagi atas 5 jenis pertanyaan yaitu factoid, non-factoid, yes no, list, dan opini. Untuk domain Bahasa Indonesia, sistem QA yang sudah ada hanya menangani pertanyaan factoid yaitu

pertanyaan yang jawabannya berupa frase singkat dari orang, lokasi,

organisasi, tanggal, angka, dan jenis jawaban singkat lainnya (Juliane et al., 2018). Dalam membangun QA system dengan pendekatan pola bahasa Indonesia, terdapat membangun pola pertanyaan yang terdiri dari tipe pertanyaan yaitu orang, waktu, tempat, organisasi, ukuran, angka, dll (Juliane et al., 2018).

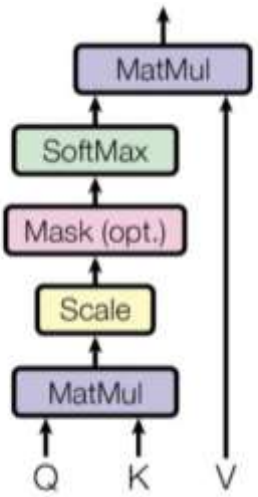
2.6 Transformer

Sebuah algoritma jaringan saraf tiruan yang revolusioner dirancang untuk mengatasi kendala dalam memahami konteks jarak jauh dan memodelkan hubungan antar kata dalam teks yang panjang, arsitektur ini telah berkontribusi besar dalam model-model NLP yang berkembang belakangan dan nama arsitektur ini adalah *Transformer* (Vaswani et al., 2017). Arsitektur ini digunakan untuk memproses data dalam bentuk urutan, seperti teks atau suara, dan telah terbukti sangat efektif dalam tugas- tugas seperti penerjemahan mesin dan pengenalan ucapan. Arsitektur *Transformer* memperkenalkan mekanisme *Attention* yang kuat yang memungkinkan modeluntuk memberikan bobot pada kata-kata yang relevan dalam teks input saat menghasilkan representasi yang diperkaya. Ini memungkinkan pemodelan paralel yang efisien dan mengatasi masalah dependensi jarak jauh yang dihadapi oleh model berbasis rekurensi sebelumnya. Selain itu, *Transformer* juga memperkenalkan blok penyandian (encoder) dan blok dekoder yang memungkinkan model untuk menerima input dan menghasilkan output dengan panjang yang berbeda. *Transformer* menggunakan *self-attention* sebagai pengganti proses sequential yang ada pada RNN untuk mengatasi permasalahan tersebut.

Self-attention sendiri merupakan mekanisme *attention* yang menggunakan posisi dari masing-masing kata untuk mengekstraksi hubungan antar kata dalam

suatu kalimat. Dalam neural network, lapisan encoder pada model *Transformer* akan memproses nilai input untuk menghasilkan vektor yang berisi informasi dan fitur yang merepresentasikan nilai input. Pada model *Transformer*, setiap lapisan encoder berisi *self-attention* dan lapisan *feed-forward*. Berbeda dengan *Convolutional Neural Network* (CNN) yangbiasanya memperkecil fitur dalam tiap lapisannya, *Transformer* akan selalu memberikan panjang output sesuai dengan panjang input yang diberikan (Vaswani et al., 2017).

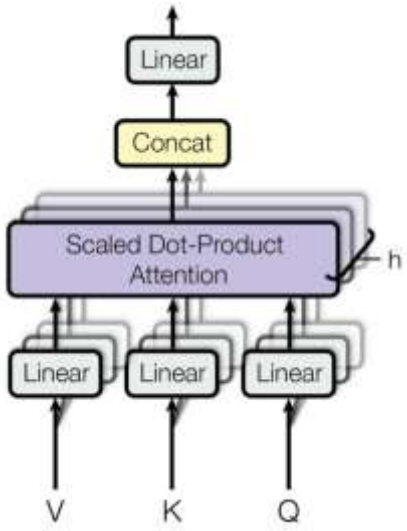
Komponen *penyusun* utama arsitektur *Transformer* adalah *multi-head self-attention mechanism*. Pada arsitektur ini input dipandang sebagai sebuah pasangan kunci - nilai (key dan value) dengan notasi (*K,V*) dan masukan terdiri dari kueri dan kunci dimensi d_k (panjang urutan input) proses ini terjadi didalam encoder. Pada decoder kemudian fungsi *attention* dihitung dengan sekumpulan queri kedalam matriks *Q* dan keluaran selanjutnya dihasilkan dengan memetakan Q ke K dan V. *Transformer* sendiri mengadopsi *scaled dot-product* yang mana keluarannya adalah jumlah nilai yang terboboti dengan bobot yang ditetapkan untuk setiap nilai diperoleh dari *dot-product* (Vaswani et al., 2017).



Gambar 2. 1 Scaled dot product attention

$$\text{Attention (Q,K,V)} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Multi-head mechanism dijalankan beberapa kali secara paralel sepanjang *scaled dot-product attention*. Keluaran *attention* yang independen secara linear digabungkan dan ditransformasikan kedalam dimensi yang diharapkan. Model dimungkinkan oleh mekanisme *multi-head attention* untuk memperhatikan informasi dari representasi subruang yang berbeda pada posisi yang berbeda secara bersama sama.



Gambar 2. 2 Multi-head self attention

MultiHead (Q, K, V) = $[head_1, \dots, head_h]W_0$

Dimana

$$head_i = Attention\ (QW_i^Q, KW_i^K, VW_i^V)$$

Karena QW^Q, KW^K, VW^V adalah matriks parameter yang akan dipelajari.

Contoh perhitungan:

Kita akan menggunakan kalimat: **"Akses akun Anda sekarang!"** yang mungkin muncul dalam email *phishing*. Kita fokus pada bagaimana kata **"akun"** mendapatkan representasi kontekstualnya dari kata-kata lain, menggunakan dua *attention heads* (Head 1 dan Head 2).

Asumsi Penyederhanaan:

- **Dimensi *Embedding* (d_model):** 4
- **Jumlah *Attention Heads* (h):** 2

- **Dimensi Per Kepala (d_k = d_v):** dmodel/h=4/2=2
- Kita akan mengabaikan *Positional Encoding* dan *Layer Normalization* untuk fokus pada inti perhitungan *attention*.

Setiap token (kata) diubah menjadi vektor *embedding*:

- **Akses:** XA=[1.0,0.5,0.2,0.8]
- **akun:** XK=[0.8,1.2,0.3,0.7]
- **Anda:** XU=[0.1,0.4,0.9,0.6]
- **sekarang!:** XS=[0.7,0.3,0.6,0.9]

Kita susun ini dalam satu matriks input X:

$$X = \begin{pmatrix} X_A \\ X_K \\ X_U \\ X_S \end{pmatrix} = \begin{pmatrix} 1.0 & 0.5 & 0.2 & 0.8 \\ 0.8 & 1.2 & 0.3 & 0.7 \\ 0.1 & 0.4 & 0.9 & 0.6 \\ 0.7 & 0.3 & 0.6 & 0.9 \end{pmatrix}$$

Setiap *attention head* memiliki set matriks bobot W^Q,W^K , W^V sendiri. Karena d_k=2, matriks ini akan berukuran $d_{model} \times d_k$.

a. Head 1:

$$W_1^Q = \begin{pmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \\ 0.7 & 0.8 \end{pmatrix}, W_1^K = \begin{pmatrix} 0.2 & 0.1 \\ 0.4 & 0.3 \\ 0.6 & 0.5 \\ 0.8 & 0.7 \end{pmatrix}, W_1^V = \begin{pmatrix} 0.3 & 0.4 \\ 0.5 & 0.6 \\ 0.7 & 0.8 \\ 0.1 & 0.2 \end{pmatrix}$$

b. Head 2:

$$W_2^Q = \begin{pmatrix} 0.8 & 0.7 \\ 0.6 & 0.5 \\ 0.4 & 0.3 \\ 0.2 & 0.1 \end{pmatrix}, W_2^K = \begin{pmatrix} 0.7 & 0.8 \\ 0.5 & 0.6 \\ 0.3 & 0.4 \\ 0.1 & 0.2 \end{pmatrix}, W_2^V = \begin{pmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \\ 0.7 & 0.8 \end{pmatrix}$$

Menghitung Q, K, V setiap Head:

$$Q_1 = XW_1^Q = \begin{pmatrix} 1.0 & 0.5 & 0.2 & 0.8 \\ 0.8 & 1.2 & 0.3 & 0.7 \\ 0.1 & 0.4 & 0.9 & 0.6 \\ 0.7 & 0.3 & 0.6 & 0.9 \end{pmatrix} \begin{pmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \\ 0.7 & 0.8 \end{pmatrix} \approx \begin{pmatrix} 0.91 & 1.34 \\ 1.21 & 1.62 \\ 0.98 & 1.22 \\ 1.06 & 1.36 \end{pmatrix}$$

$$K_1 = XW_1^K = \begin{pmatrix} 1.0 & 0.5 & 0.2 & 0.8 \\ 0.8 & 1.2 & 0.3 & 0.7 \\ 0.1 & 0.4 & 0.9 & 0.6 \\ 0.7 & 0.3 & 0.6 & 0.9 \end{pmatrix} \begin{pmatrix} 0.2 & 0.1 \\ 0.4 & 0.3 \\ 0.6 & 0.5 \\ 0.8 & 0.7 \end{pmatrix} \approx \begin{pmatrix} 1.02 & 0.92 \\ 1.22 & 1.08 \\ 1.00 & 0.88 \\ 1.10 & 0.98 \end{pmatrix}$$

$$V_1 = XW_1^V = \begin{pmatrix} 1.0 & 0.5 & 0.2 & 0.8 \\ 0.8 & 1.2 & 0.3 & 0.7 \\ 0.1 & 0.4 & 0.9 & 0.6 \\ 0.7 & 0.3 & 0.6 & 0.9 \end{pmatrix} \begin{pmatrix} 0.3 & 0.4 \\ 0.5 & 0.6 \\ 0.7 & 0.8 \\ 0.1 & 0.2 \end{pmatrix} \approx \begin{pmatrix} 0.91 & 1.26 \\ 1.22 & 1.54 \\ 1.10 & 1.34 \\ 1.08 & 1.38 \end{pmatrix}$$

2.6.1 Hitung Skor Perhatian (Scaled Dot-Product Attention) untuk Setiap Head

Untuk Head 1, kita hitung $Q_1 K_1^T / \sqrt{d_k}$. Fokus pada **baris kedua** dari Q1 (yaitu Q1[akun]=[1.21,1.62]) dan kalikan dengan transpose dari setiap baris K1 . Akar kuadrat dari dimensi Key (dk) adalah $\sqrt{2} \approx 1.414$.

- **Skor "akun" ke "Akses":** $([1.21,1.62] \cdot [1.02,0.92]^T)$ / 1.414 = (1.21·1.02+1.62·0.92) / 1.414 = (1.234+1.490) / 1.414 ≈ 1.92
- **Skor "akun" ke "akun":** $([1.21,1.62] \cdot [1.22,1.08]^T)$ / 1.414 = (1.476+1.750) / 1.414 ≈ 2.28
- **Skor "akun" ke "Anda":** $([1.21,1.62] \cdot [1.00,0.88]^T)$ / 1.414 = (1.210+1.426) / 1.414 ≈ 1.86
- **Skor "akun" ke "sekarang!":** $([1.21,1.62] \cdot [1.10,0.98]^T)$ / 1.414 = (1.331+1.587) / 1.414 ≈ 2.06

Jadi, baris skor mentah untuk "akun" dari Head adalah: [1.92,2.28,1.86,2.06]

2.7 IndoBERT (Indo Bidirectional Encoder Representations From Transformers)

IndoBERT adalah metode modifikasi dari BERT-base yang sudah ada dengan mengikuti konfigurasi dari BERT-base (*uncased*) yang memiliki 12 *hidden layers* dengan masing-masing memiliki 768 dimensi, 12 *attention heads*, dan 3.072 dimensi *feed-forward hidden layers*. Secara total, IndoBERT sudah dilatih lebih dari 220 juta kata Bahasa Indonesia dengan 3 sumber utama yaitu Indonesia Wikipedia (74 juta kata), artikel seperti Kompas, Tempo, dan Liputan6 (55 juta kata), dan Indonesia Web Corpus (90 juta kata) (Koto et al., 2020). IndoBERT dilatih dengan dataset Indo4B dan TPUv3-8 dalam dua fase (Wilie et al., 2020). Fase pertama pelatihan dilakukan dengan input data sebesar 128 dan waktu pelatihannya selama 35, 38, 89 dan 134 jam untuk model $\text{IndoBERT}_{\text{BASE}}$, $\text{IndoBERT-lite}_{\text{BASE}}$, $\text{IndoBERT}_{\text{LARGE}}$ dan $\text{IndoBERT-lite}_{\text{LARGE}}$. Pada fase kedua jumlah sekuens ditingkatkan menjadi 512 dan memerlukan waktu pelatihan selama 9, 23, 32 dan 45 jam untuk model diatas. Pada kedua fase $\text{IndoBERT}_{\text{BASE}}$ memiliki *batch size* sebesar 256 dan *learning rate* sebesar 2e-5 sedangkan $\text{IndoBERT}_{\text{LARGE}}$ pada fase pertama memiliki *batch size* sebesar 256 dan *learning rate* sebesar 1e-4 dan pada fase kedua diturunkan menjadi *batch size* sebesar 128 dan *learning rate* sebesar 8e-5.

Model	Maximum Sequence Length = 128				Maximum Sequence Length = 512			
	Batch Size	Learning Rate	Steps	Duration (Hr.)	Batch Size	Learning Rate	Steps	Duration (Hr.)
IndoBERT-lite$_{BASE}$	4096	0.00176	112.5 K	38	1024	0.00088	50 K	23
IndoBERT$_{BASE}$	256	0.00002	1 M	35	256	0.00002	68 K	9
IndoBERT-lite$_{LARGE}$	1024	0.00044	500 K	134	256	0.00044	129 K	45
IndoBERT$_{LARGE}$	256	0.0001	1 M	89	128	0.00008	120 K	32

Tabel 2. 1 Summary Pelatihan IndoBert

IndoBERT menerapkan *max multi-head attention* pada mekanismenya, seperti yang telah dijelaskan *multi-head attention* memungkinkan model untuk memperhatikan informasi dari beberapa representasi subspace yang berbeda pada posisi yang berbeda dalam teks. *Max multi-head attention* sendiri adalah jenis *multi-head attention* yang mengambil nilai maksimum dari output dari setiap *attention head*, artinya output dari setiap *attention head* dihitung, dan nilai maksimum dari output-output ini diambil sebagai output akhir (Voita et al., 2019). Pada algoritma IndoBERT penggunaan mekanisme *max multi-head attention* ini adalah untuk meningkatkan kinerja model dan memperkuat representasi dari teks dalam pemrosesan NLP. Selain itu IndoBERT adalah modifikasi dari BERT-base tepatnya menggunakan konfigurasi BERT-base-uncased. Model ini merupakan salah satu model yang baru muncul dan berkembang diantara model *neural network* lainnya seperti *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN), dan *Long Short-Term Memory* (LSTM). Tabel berikut ini menyajikan perbandinganantara model BERT-base dengan beberapa model neural network lainnya.

Tabel 2. 2 Perbedaan Model-Model Neural Network

Model	Arsitektur	Penggunaan Utama	Pre-Training	Kelebihan	Keterbatasan

BERT-base	*Transformer*	Pemrosesan bahasa alami, tugas NLP	MLM dan NSP	Memahami konteks kata dengan representasi bidirectional	Memerlukan daya komputasi dan memori yang besar
CNN	Tumpukan layer konvolusi	Pengenalan pola dalam data grid, pengolahan citra	Tidak	Efisien untuk pemrosesan datagrid dan citra	Tidak dapat menangkap ketergantungan urutan antar kata
RNN	Sekuensial	Pemrosesan urutan, pengenalan ucapan, tugas NLP	Tidak	Mampu memodelkan ketergantungan urutan antar elemen data	Rentan pada masalah menghilangnya atau meledaknya gradien
LSTM	Sekuensial (dengan unit LSTM)	Pemrosesan urutan dengan ketergantungan jangka panjang	Tidak	Mengatasi masalah vanishing / exploding gradient pada RNN	Kompleksitas perhitungan dan memori yang tinggi
Bi-LSTM	Sekuensial (Bi-LSTM)	Memodelkan ketergantungan urutan dari kedua arah	Tidak	Pemrosesan urutan dengan ketergantungan jangka panjang	Komputasi yang lebih mahal dibandingkan dengan LSTM

Tabel 2.2 merupakan rangkuman dari beberapa model *neural network*.

Gambaran singkat tentang arsitektur masing-masing model dan fokus utama dari kelebihan dan keterbatasannya. BERT-base, yang didasarkan pada *Transformer*, dikenal karena kemampuannya dalam memahami konteks kata dengan representasi bidirectional. Sementara itu, CNN efisien dalam memproses data grid dan citra, tetapi tidak mampu menangkap ketergantungan urutan antar kata. RNN, yang merupakan model sekuen, mampu memodelkan ketergantungan urutan antar elemen data, tetapi rentan terhadap masalah menghilangnya atau meledaknya gradien. LSTM, yang merupakan variasi dari RNN, berhasil mengatasi masalah gradien tersebut dan lebih cocok untuk memproses urutan data dengan ketergantungan jangka panjang. Bidirectional Long Short-Term Memory (Bi-LSTM) adalah arsitektur neural network yang digunakan untuk memodelkan ketergantungan urutan dalam data. Keunggulannya adalah kemampuannya untuk memahami urutan data dari kedua arah, dari kiri ke kanan dan dari kanan ke kiri. Salah satu kekurangan Bi-LSTM adalah komputasi yang lebih mahal dibandingkan dengan model-model yang lebih sederhana seperti CNN.

Tabel 2. 3 Penelitian Terdahulu

No.	Nama	Judul	Kesimpulan	Tahun
1.	Setyo Nugroho kuncahyo, Yullian Sukmadewa Anantha, Wuswilahaken DW Haftittah, Abdurrachman Bachtiar Fitra, Yudistira Novanto	BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews	Meneliti penerapan IndoBERT dalam analisis sentimen ulasan aplikasi mobile di Indonesia dan menemukan bahwa model ini memiliki akurasi lebih tinggi dibandingkan metode lain.	2021

2.	Helmi Imaduddin, Fiddin Yusfida A'la, Yususf Sulistyo Nugroho	Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach	Dalam bidang kesehatan mengaplikasikan IndoBERT untuk analisis sentimen dalam aplikasi kesehatan,mencapai akurasi hingga 96%.	2023
3.	Bryan Willie, Karissa Vincentio, Genta Indra Winata, Samuel Cahyawijaya, Xiaohong Li, Zhi Yuan Lim, Sidik Soleman, Rahmad Mahendra, Pascale Fung, Syafri Bahar, Ayu Purwarianti	IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding.	memperkenalkan IndoNLU sebagai benchmark dataset yang mengevaluasi model NLP untuk berbagai tugas Bahasa Indonesia, termasuk QA.	2023
4.	Fajri Koto, Afshin Rahimi, Jey Han Lau, Timothy Baldwin	IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP.	dataset IndoLEM sebagai sumber data utama untuk mengevaluasi model IndoBERT dalam tugas NLP, menunjukkan keunggulan dibandingkan model lain.	2020
5.	Gieldy Akhyat Affandi Utama	Analisis Pemahaman Quality Assurance pada Mahasiswa Bisnis Digital	dalam dunia bisnis digital. Hasil penelitian menunjukkan bahwa mahasiswa memiliki pemahaman yang baik tentang konsep dasar QA	2024

		Universitas Pendidikan Indonesia.	tetapi kurang dalam penerapannya di industri. Studi ini merekomendasikan integrasi elemen praktis seperti simulasi dan studi kasus dalam kurikulum QA.	

BAB III

METODOLOGI PENELITIAN

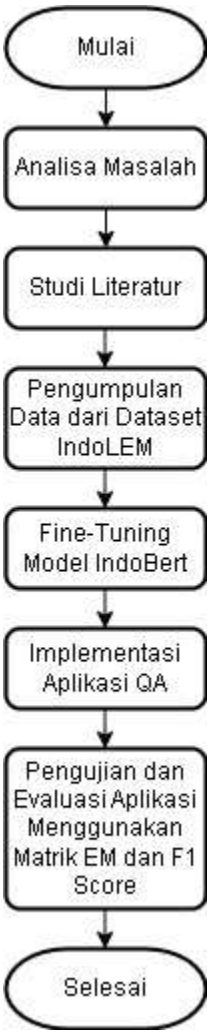
Bab ini akan membahas mengenai metodelogi yang digunakan dalam penyusunan riset ini serta apa saja tahapan yang menggambarkan proses penelitian secara sistematis dan terstruktur. Metodelogi merupakan suatu sistem praktik, prinsip, dan tata cara yang diterapkan pada cabang pengetahuan untuk dapat digunakan pada suatu proses tertentu.

3.1 Metode Usulan

Metode yang diusulkan dalam penelitian ini bertujuan untuk membangun sistem Question Answering (QA) dalam Bahasa Indonesia dengan memanfaatkan transfer learning menggunakan IndoBERT dan dataset IndoLEM Proses dimulai dengan analisis masalah untuk memahami kebutuhan sistem QA dan tantangan dalam menangani struktur Bahasa Indonesia. Selanjutnya, dilakukan studi literatur yang mencakup kajian terhadap model deep learning berbasis transformer, seperti BERT, serta implementasinya dalam domain QA. Setelah itu, data dikumpulkan dari dataset IndoLEM, yang dirancang khusus untuk evaluasi performa model NLP dalam Bahasa Indonesia.

Proses inti melibatkan fine-tuning model IndoBERT menggunakan dataset IndoLEM untuk menyesuaikan model dengan tugas QA. Hasil fine-tuning ini diimplementasikan dalam sebuah aplikasi berbasis web, yang dirancang untuk memberikan jawaban atas pertanyaan berbasis teks secara akurat. Aplikasi ini

kemudian diuji menggunakan metrik evaluasi seperti Exact Match (EM) dan F1 Score untuk mengukur akurasi dan relevansi jawaban yang dihasilkan. Akhirnya, penelitian ini memberikan kesimpulan terkait efektivitas model yang diusulkan dan kontribusinya dalam pengembangan sistem QA berbasis Bahasa Indonesia. Arsitektur skema umum penelitian ini dapat dilihat pada Gambar 3.1.



Gambar 3. 1 Flowchart Alur Penelitian

3.1.1 Pengembangan Sistem

Perancangan aplikasi ini menggunakan pipeline sebagai framework utamanya dan akan menampilkan sistem berbasis web. Proses "preprocessing

teks" yaitu membersihkan dan mengubah teks, lalu analisis oleh "fine-tuned IndoBERT", hingga "evaluasi jawaban" dan "menampilkan jawaban" secara bersama-sama membentuk sebuah *pipeline.* Tahapan-tahapan tersebut adalah data yang berupa pertanyaan dan kalimat sumber yang mengalir secara berurutan. Proses ini melibatkan fine-tuning model IndoBERT menggunakan dataset IndoLEM untuk menyesuaikan model dengan tugas QA. Hasil fine-tuning ini diimplementasikan dalam sebuah aplikasi berbasis web, yang dirancang untuk memberikan jawaban atas pertanyaan berbasis teks secara akurat.

Penerepan sistem QA yang dapat memberikan jawaban yang akurat dan relevan dapat dilakukan dengan mengunduh dataset IndoLEM dari repository GitHub dalam melatih model. Lalu data pertanyaan dan konteks akan melalui tahapan pra-pemrosesan seperti case folding, tokenization, dan stopwords removal. Selanjutnya pelatihan model dimana model IndoBert digunakan sebagai speasialis dalam menjawab pertanyaan. Setelah itu dapat di implementasikan dalam sebuah aplikasi berbentuk web agar dapat berfungsi dan bisa digunakan.

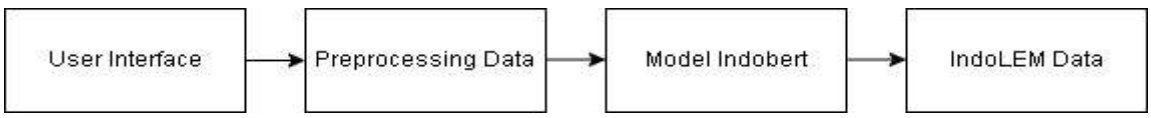
Evaluasi sistem digunakan untuk memberikan dan mengukur seberapa akurat jawaban yang dihasilkan. Evaluasi ini menggunakan metric standar seperti Exact Match (EM) dan F1-Score. Exact Match bertugas memahami konteks dan hubungan antar kata dalam teks dan F1-Score mampu mengukur kesimbanngan antara pesisi dan recall dalam menjawab pertanyaan sehingga dapat memberikan jawaban terbaik terhadap tingkat nilai yang diberikan.

3.2 Sumber Data

Pada penelitian ini data yang diambil bersumber dari IndoLEM, IndoLEM sendiri merupakan kumpulan data yang dibuat khusus untuk menguji dan menilai model-model NLP dalam Bahasa Indonesia. IndoLEM merupakan data terstruktur dan mencakup berbagai tugas NLP salah satunya Question Answering sehingga mudah dalam melatih model. Dalam penggunaan data IndoLEM bertujuan untuk memastikan model IndoBert dalam penelitian ini benar-benar mampu memahami teks Bahasa Indonesia dan memberikan jawaban yang akurat. Pengumpulan data dengan IndoLEM dapat diunduh dari repository yang menyediakan akses ke data dan melibatkan kunjungan ke web repository yang menyimpan dataset serta mengunduh file dataset dengan format yang sesuai kebutuhan misalnya JSON dan CSV.

3.3 Arsitektur sistem

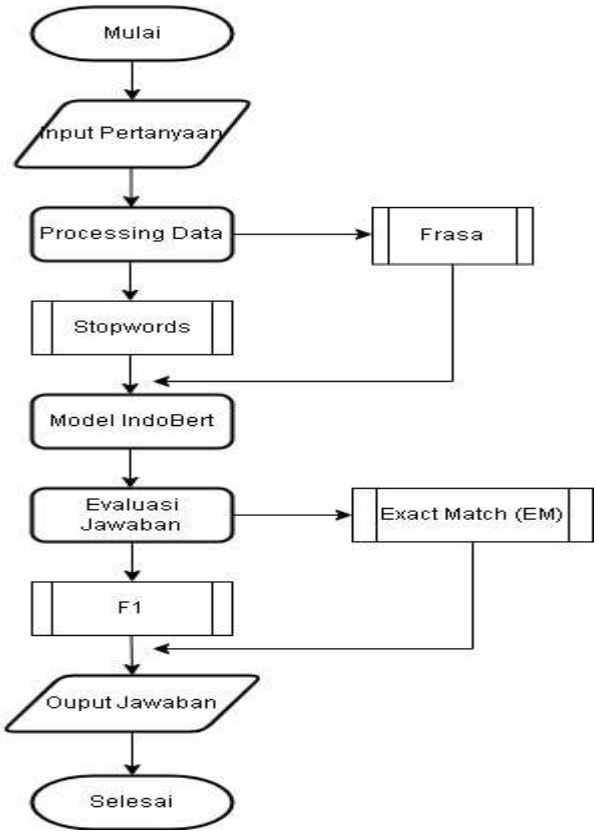
Sistem *Question* Answering berbasis IndoBERT ini terdiri dari beberapa komponen utama yang bekerja secara berurutan. Proses dimulai dengan pengguna mengajukan pertanyaan melalui antarmuka aplikasi. Pertanyaan ini kemudian diproses untuk membersihkan teks sebelum dikirim ke model IndoBERT, yang telah dilatih menggunakan dataset IndoLEM. Model ini akan menganalisis pertanyaan, memahami konteksnya, dan menghasilkan jawaban yang paling sesuai. Terakhir, aplikasi menampilkan jawaban kepada pengguna.



Gambar 3. 2 Diagram Arsitektur Skema Proses

3.4 Flowchart Algoritma dan Sistem

Di balik layar, sistem ini mengandalkan algoritma yang dirancang khusus untuk memahami dan menjawab pertanyaan pengguna secara efektif. Setelah menerima pertanyaan, sistem menjalankan serangkaian langkah penting. Pertama, teks diproses menggunakan teknik Natural Language Processing (NLP) seperti tokenizing dan penghapusan stopwords agar lebih terstruktur. Selanjutnya, model IndoBERT yang telah di-fine-tune digunakan untuk mencari jawaban yang paling relevan. Setelah jawaban ditemukan, sistem mengevaluasi akurasinya menggunakan metrik seperti Exact Match (EM) dan F1-score. Jika jawaban telah memenuhi standar kualitas, sistem akan menampilkannya kepada pengguna dalam format yang jelas dan informatif.



Gambar 3. 3 Flowchart Sistem

Penjelasan :

Diagram ini menggambarkan alur kerja sistem *Question Answering* berbasis IndoBERT. Berikut adalah penjelasannya:

a. Sistem memulai proses pemrosesan data untuk menangani pertanyaan yang diberikan pengguna.

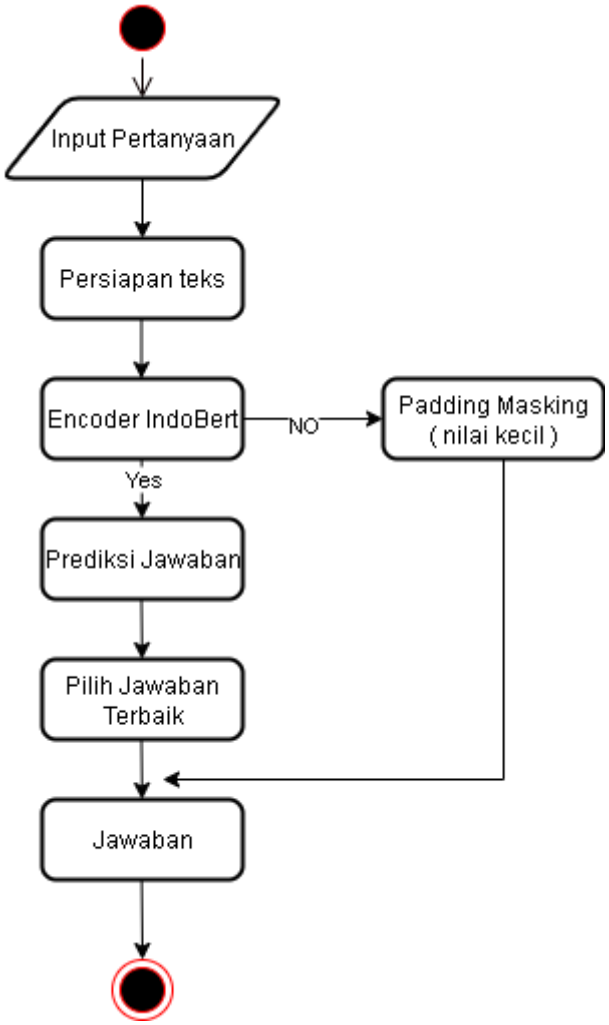
b. Input Pertanyaan: Pengguna memasukkan pertanyaan dalam bentuk teks melalui antarmuka aplikasi.

c. Preprocessing (Tokenizing, Stopwords Removal) Sistem membersihkan teks dari elemen yang tidak relevan (noise), seperti kata-kata umum (stopwords), serta mempersiapkannya agar lebih mudah dipahami oleh model.

d. Fine-Tuned IndoBERT Mengekstrak Jawaban: Model IndoBERT yang telah dilatih secara khusus akan menganalisis pertanyaan dan mencari jawaban yang paling sesuai berdasarkan data yang tersedia.

e. Evaluasi dengan EM & F1-score: Jawaban yang dihasilkan dievaluasi menggunakan metrik Exact Match (EM) dan F1-score untuk memastikan tingkat akurasinya.

f. Setelah melalui tahap evaluasi, jawaban yang telah diverifikasi akan ditampilkan kembali kepada pengguna dalam format yang jelas dan mudah dipahami.



Gambar 3. 4 Flowchart Algoritma

Penjelasan:

Diagram ini menggambarkan alur kerja algoritma *Question Answering* berbasis IndoBERT. Berikut adalah penjelasannya:

a. Mulai: Proses dimulai.

b. Input pertanyaan: pengguna akan memasukkan data utama ke dalam sistem terdiri dari pertanyaan dan konteks.

c. Preprocessing Data (persiapan teks): Ini adalah tahap di mana teks konteks dan pertanyaan "dibersihkan" dan dipersiapkan agar dapat dipahami oleh model.

d. Prediksi jawaban: Ini adalah inti dari sistem QA Anda. Pada tahap ini, model IndoBERT yang sudah di-*fine-tune* akan mengambil alih

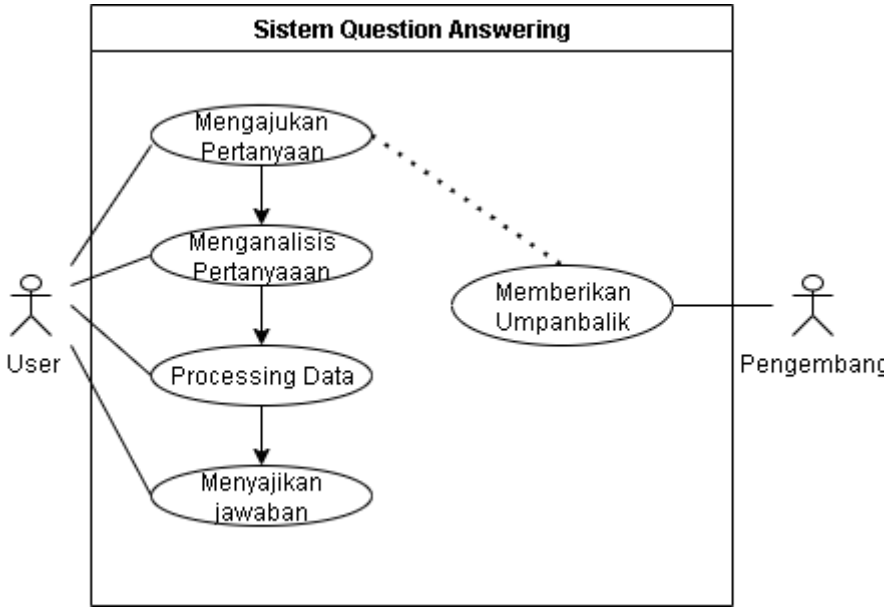
e. Pilih jawaban terbaik: Berdasarkan hasil dari "Proses Indobert" (yaitu, prediksi posisi awal dan akhir), sistem akan mengekstraksi segmen teks yang spesifik dari konteks asli. Segmen inilah yang dianggap sebagai jawaban yang paling relevan.

f. Jawaban: Ini adalah keluaran dari sistem. Jawaban yang berhasil diekstrak oleh model akan ditampilkan kepada pengguna, biasanya melalui antarmuka aplikasi.

i. Selesai (end): Proses berakhir.

3.5 Use case Diagram

Sistem ini dirancang untuk menjadi mudah digunakan oleh siapa pun. Pada dasarnya, pengguna adalah aktor utama sistem. Mereka hanya perlu memasukkan pertanyaan melalui antarmuka aplikasi, dan sistem akan melakukan berbagai proses teknis, seperti membersihkan data, mengolah data menggunakan model IndoBERT, dan menampilkan jawaban yang tepat. Pengguna tidak perlu memahami kompleksitas ini; mereka hanya perlu fokus pada pertanyaan yang ingin diajukan dan menikmati kemudahan mendapatkan jawaban yang cepat dan akurat.



Gambar 3. 5 Diagram Usecase

Penjelasan:

Diagram ini menggambarkan bagaimana pengguna berinteraksi dengan sistem Question Answering. Berikut adalah penjelasannya:

a. Pengguna berinteraksi dengan sistem untuk mengajukan pertanyaan.

b. Pengguna memasukkan pertanyaan melalui antarmuka aplikasi yang disediakan.

c. Preprocessing Data: Sistem membersihkan dan mempersiapkan teks pertanyaan agar lebih mudah diproses oleh model, termasuk tahap seperti tokenizing dan penghapusan stopwords.

d. Model IndoBERT menganalisis pertanyaan dan mencari jawaban yang paling sesuai berdasarkan data yang tersedia.

e. Setelah diproses, sistem menyajikan jawaban kepada pengguna dalam format yang jelas, ringkas, dan mudah dipahami.

Tabel 3. 1 Rencana Kegiatan

No	Kegiatan	Desember				Januari				Februari				Maret				April			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Menyusun Proposal																				
2	Seminar Proposal																				
3	Perbaikan Proposal																				
4	Pelaksanaan Penelitian																				
5	Pengolahan dan Penyusunan Laporan Hasil Penelitian																				
6	Seminar Hasil Penelitian Sidang																				
7	Perbaikan Hasil Penelitian																				

BAB IV
HASIL DAN PEMBAHASAN

Bab ini akan dibahas hasil implementasi dan evaluasi dari sistem Question Answering (QA) berbasis IndoBERT-QA yang telah dikembangkan pada penelitian ini. Pembahasan meliputi gambaran umum implementasi aplikasi, proses pengujian sistem, serta analisis performa model berdasarkan data uji yang telah disiapkan.

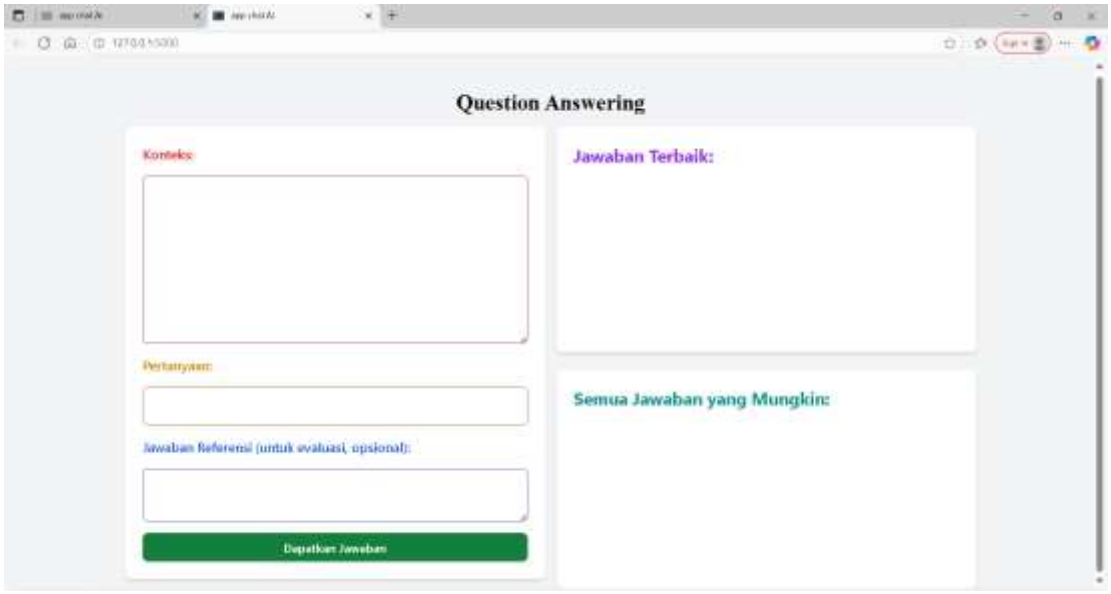
4.1 Hasil Implementasi Sistem

Pada tahap ini, aplikasi Question Answering (QA) berbasis IndoBERT-QA telah berhasil direalisasikan sesuai dengan desain yang telah dijelaskan pada BAB III. Pengembangan aplikasi dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan framework Flask, serta mengintegrasikan model IndoBERT-QA pre-trained secara langsung tanpa melalui proses fine-tuning tambahan. Sistem ini dapat dijalankan secara lokal (offline), sehingga memberikan fleksibilitas tinggi untuk digunakan dalam kegiatan penelitian maupun keperluan edukasi.

4.1.1 Tampilan Antar Muka Aplikasi

Aplikasi Question Answering (QA) berbasis IndoBERT-QA yang dikembangkan dalam penelitian ini dilengkapi dengan antarmuka web yang sederhana namun tetap intuitif. Pada halaman utama, pengguna disajikan dua bidang input utama, yaitu kolom context untuk memasukkan potongan teks atau artikel yang relevan dengan pertanyaan, serta kolom pertanyaan untuk mengetikkan pertanyaan yang ingin diajukan. Setelah kedua kolom tersebut

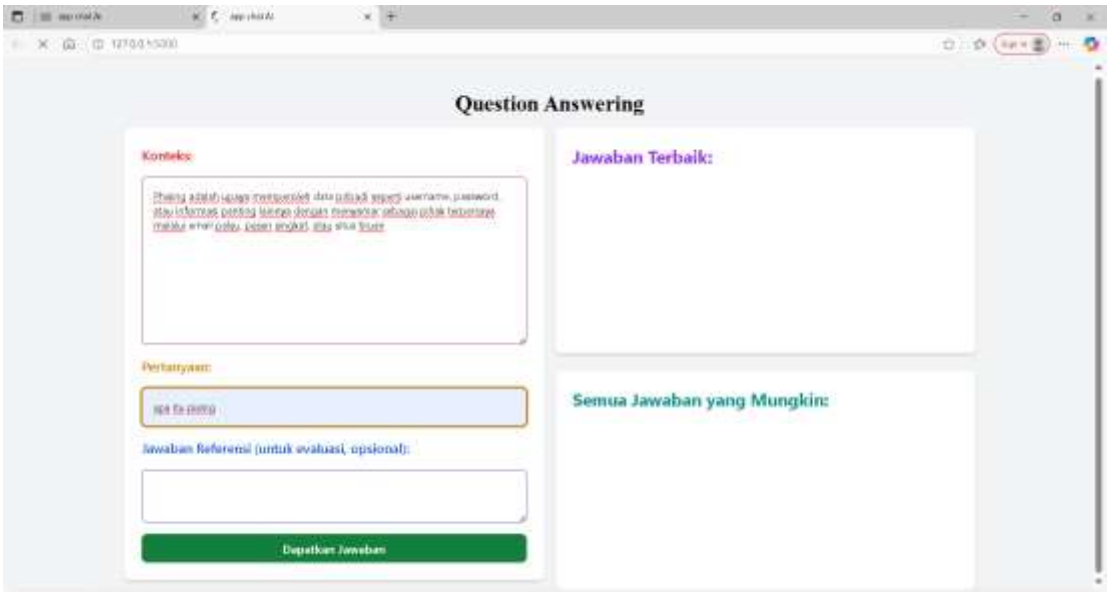
diisi, pengguna cukup menekan tombol “Cari Jawaban” untuk memulai proses pencarian dan mendapatkan respons dari sistem. Gambar 4.1 berikut memperlihatkan tampilan utama aplikasi QA yang menampilkan kedua kolom input beserta tombol eksekusi.



Gambar 4. 1 Tampilan halaman web

4.1.2 Pengisian Konteks dan Pertanyaan

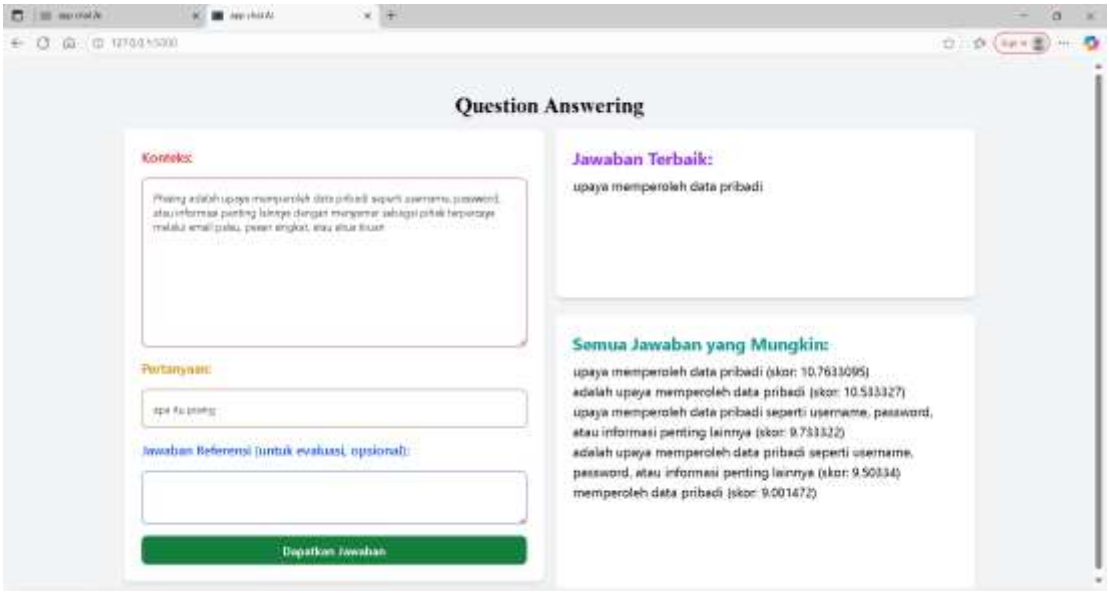
Setelah pengguna mengisi kolom context dan pertanyaan, aplikasi akan memproses data tersebut dengan memanfaatkan model IndoBERT-QA yang telah dilatih sebelumnya. Jawaban paling relevan yang diidentifikasi oleh sistem akan secara otomatis ditampilkan pada area output.



Gambar 4. 2 Tampilan membuat pertanyaan

4.1.3 Mendapatkan Jawaban

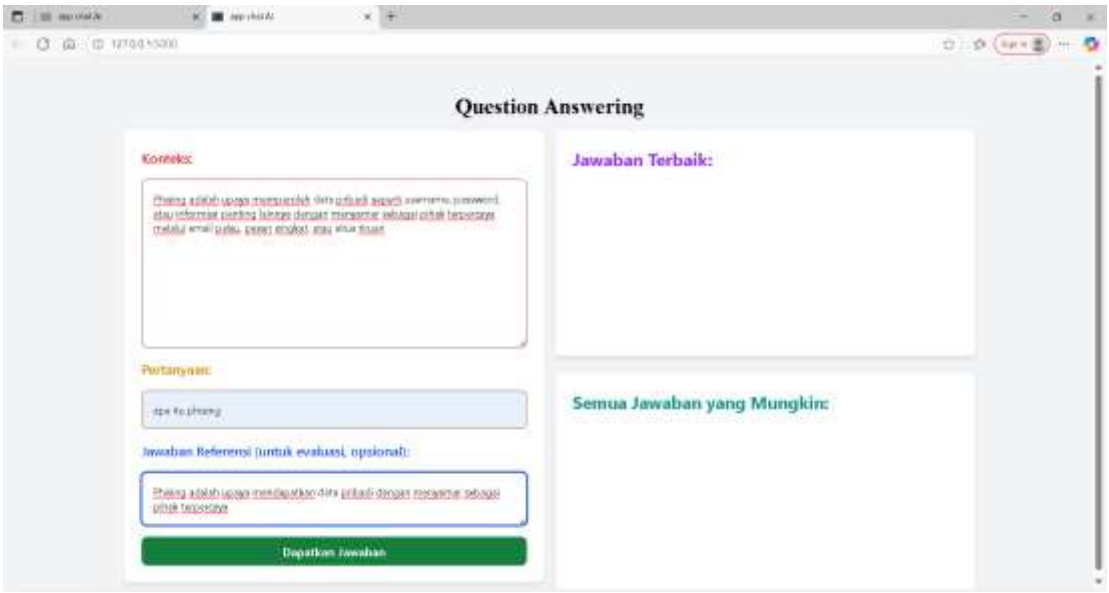
Selain menampilkan jawaban utama, sistem juga dapat memberikan alternatif jawaban beserta skor keyakinan dari model, apabila fitur ini diaktifkan. Ilustrasi hasil keluaran aplikasi setelah proses inferensi ditampilkan pada Gambar 4.3 berikut.



Gambar 4. 3 Tampilan mendapatkan jawaban

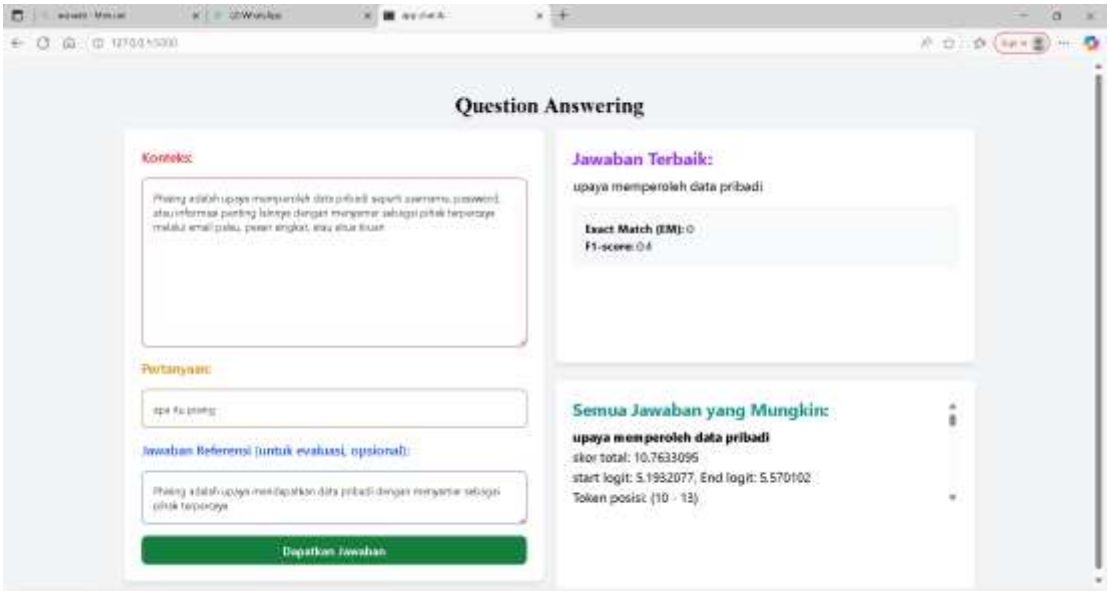
4.1.4 Mengisi Kolom Referensi

Pada saat pengguna mengisi kolom referensi maka sistem akan memberikan jawaban dan menampilkan nilai dari evaluasi EM dan F1-score. Fungsi ini bertujuan menampilkan seberapa identik dan sama kalimat yang diberikan pengguna dengan jawaban yang diberikan oleh sistem .



Gambar 4. 4 Tampilan mengisi jawaban referensi

Seperti pada gambar 4.6 dibawah ini EM memberikan hasil 0 yang berarti jawaban referensi dan jawaban model 100% tidak sama persis, namun F1 memberikan hasil 0.4 yang mempunyai arti meskipun tidak sama persis 100% namun kedua masih memiliki kesamaan.



Gambar 4. 5 Tampilan mendapatkan jawaban beserta hasil evaluasi

4.2 Pengujian Sistem

Pengujian sistem dilakukan dengan menerapkan sejumlah skenario kasus nyata yang diambil dari artikel dan berita seputar cybercrime. Langkah ini bertujuan untuk mengevaluasi kemampuan sistem dalam memberikan jawaban yang akurat dan relevan terhadap pertanyaan yang diajukan, asalkan context yang diberikan memang memuat informasi yang dibutuhkan.

4.2.1 Skema Pengujian

Tahapan pengujian yang dilakukan meliputi:

- Menyiapkan context berupa artikel singkat atau kutipan berita yang berkaitan dengan kasus cybercrime.

- Menyusun daftar pertanyaan yang sesuai dan relevan dengan isi context yang telah dipilih.
- Memasukkan context dan pertanyaan ke dalam aplikasi Question Answering.
- Mendokumentasikan jawaban yang dihasilkan oleh model.
- Membandingkan hasil jawaban model dengan jawaban referensi atau jawaban manual.
- Mengevaluasi hasil berdasarkan tingkat kecocokan (Exact Match/EM), relevansi, dan tingkat keterpahaman jawaban yang diberikan.

4.2.2 Pengujian Perhitungan

1.) Perhitungan evaluasi Exact Math dan F1-score

Berikut ini merupakan penjelasan contoh dari perhitungan evaluasi EM dan F1-Score yang digunakan model untuk memberikan keakuratan pada jawaban.

Contoh:

- Jawaban referensi (Ground Truth): “Phising adalah upaya mendapatkan data pribadi dengan menyamar sebagai pihak terpercaya”.
- Jawaban model: “Upaya memperoleh data pribadi”.

1. Exact Match

   - EM = 1 jika jawaban model 100% sama persis dengan jawaban referensi.
   - EM = 0 jika 100% tidak sama persis.

   Maka hasil EM dari jawaban refensi dan jawaban model diatas adalah 0, karena jawaban referensi dan jawaban model 100% tidak sama persis.

2. F1-score

$$F1 = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

- Precision (jawaban model) = $\frac{jumlah\ overlap}{jumlah\ kata} = \frac{3}{4} = 0.75$
- Recall (jawaban referensi) = $\frac{jumlah\ overlap}{jumlah\ kata} = \frac{3}{11} = 0.273$

Maka:

$$F1 = \frac{2 * (0.75 * 0.273)}{0.75 + 0.273} = \frac{0.4095}{1.023} = 0.4$$

2.) Perhitungan Score

Pada perhitungan score ini mennggunakan penjumlahan dari nilai start logit dan end logit yang merupakan representasi dari nilai Multihead untuk menemukan rentang jawaban (span) dalam konteks yang menjawab pertanyaan.

$$Score = start\_logit + end\_logit$$

Contoh: **Upaya memperoleh data pribadi**

$$Score = 5.193 + 5.570 = 10.763$$

4.2.3 Contoh Kasus Uji Coba

Berikut adalah beberapa contoh kasus uji beserta hasil keluaran model dan jawaban referensi sebagai pembanding:

Tabel 4. 1 Tabel Jawaban menggunakan Jawaban Referensi

NO	Context	Pertanyaan	Jawaban Model	Referensi

1	Phising adalah upaya memperoleh data pribadi seperti username, password, atau informasi penting lainnya dengan menyamar sebagai pihak terpercaya melalui email palsu, pesan singkat, atau situs tiruan	Apa itu phising ?	upaya memperoleh data pribadi	Phising adalah upaya mendapatkan data pribadi dengan menyamar sebagai pihak terpercaya.
2	Serangan DDoS (Distributed Denial of Service) bertujuan membuat layanan tidak dapat diakses pengguna sah dengan membanjiri server menggunakan lalu lintas palsu secara masif dan serentak	Jelaskan tujuan serangan DDoS!	membuat layanan tidak dapat diakses pengguna sah	Membuat layanan tidak dapat diakses pengguna dengan membanjiri server secara masif.
3	Penipuan online melalui media sosial umumnya dilakukan dengan memanfaatkan akun palsu, penawaran barang murah, atau meminta transfer dana dengan alasan	Bagaimana modus penipuan online di media sosial?	memanfaatkan akun palsu, penawaran barang murah, atau meminta transfer dana dengan alasan tertentu	Dengan menggunakan akun palsu atau penawaran palsu lalu meminta transfer dana pada korban.

	tertentu kepada korban.			
4	Malware adalah perangkat lunak berbahaya yang dirancang untuk merusak, mengambil data, atau mengganggu sistem komputer korban tanpa izin.	Apa yang di maksud dengan malware	perangkat lunak berbahaya	Perangkat lunak berbahaya yang dapat merusak atau mencuri data komputer korban.
5	phising merupakan teknik untuk ‘memancing’ informasi dan data rahasia dari para korban melalui umpan atau data palsu yang dibuat semenarik mungkin dan semirip mungkin dengan aslinya.	apa harfiah yang dimiliki phising?	memancing	mencuri informasi dan data pribadi seseorang
6	Phising adalah penipuan online yang dilakukan melalui email, link, website, atau telepon palsu yang dibuat semirip mungkin dengan aslinya.	bagaimana phising dilakukan?	melalui email, link, website, atau telepon palsu yang dibuat semirip mungkin dengan aslinya	dengan mengambil data pribadi korban
7	Pelaku kejahatan cyber melakukan	apa saja cara-cara yang	melalui link penipuan	menjalankan kejahatannya

	phising dengan berbagai cara. Biasanya, pelaku phising menjalankan kejahatannya melalui link penipuan dalam email atau SMS, atau dengan suara via telepon.Mereka menyamarkan identitasnya seolah-olah berasal dari perusahaan	dilakukan pelaku phising?	dalam email atau SMS, atau dengan suara via telepon	melalui link penipuan dalam email atau SMS, atau dengan suara via telepon mereka menyamarkan identitasnya seolah-olah berasal dari perusahaan
8	Tujuannya yaitu untuk mendapatkan data dan informasi sensitif, seperti rekening bank atau username dan password.	apa tujuan phising?	untuk mendapatkan data dan informasi sensitif	untuk mendapatkan data dan informasi sensitif
9	Mereka menyamarkan identitasnya seolah-olah berasal dari perusahaan yang valid untuk menarik dan membujuk calon korban agar memberikan informasi sensitif seperti nomor kartu kredit, informasi	informasi seperti apa yang akan diberikan korban kepada pelaku?	nomor kartu kredit, informasi login, dan nomor KTP	informasi sensitif

	login, dan nomor KTP.			
10	Pelaku kejahatan cyber melakukan phising dengan berbagai cara.	jenis kejahatan?	kejahatan cyber melakukan phising dengan berbagai cara.	phising

4.3 Evaluasi Hasil

Evaluasi hasil dilakukan dengan membandingkan jawaban model dan jawaban referensi menggunakan metrik Exact Match (EM) dan F1 Score. Exact Match (EM) mengukur persentase jawaban model yang identik dengan jawaban referensi. F1 Score mengukur tingkat kesamaan (overlap) antara jawaban model dan referensi, dengan memperhitungkan presisi dan recall. Berikut tabel hasil evaluasi:

Tabel 4. 2 Tabel Hasil Evaluasi EM dan F1

NO	EM	F1 Score	Keterangan
1	0	0.4	Jawaban model tidak identik, namun inti jawaban sudah sesuai.
2	0	0.87	Jawaban hampir sesuai dengan referensi.
3	0	0.5	Jawaban model relevan, namun tidak identik.
4	0	0.43	Jawaban model sesuai inti, namun

			belum lengkap.
5	0	0.0	Jawaban model sangat tidak identik dan tidak relevan.
6	0	0.11	Jawaban tidak relevan, namun sesuai inti.
7	0	0.71	Jawaban hampir identik dengan referensi
8	1	1.0	Jawaban identik dan sesuai referensi.
9	0	0.22	Jawaban tidak identik dan relevan, namun sesuai inti
10	0	0.9	Jawaban tidak relevan, namun identik.

4.4 Analisis Dan Pembahasan

Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa sistem Question Answering berbasis IndoBERT-QA mampu menghasilkan jawaban yang cukup relevan pada berbagai kasus nyata, terutama ketika context yang diberikan secara eksplisit memuat informasi yang dibutuhkan untuk menjawab pertanyaan. Secara umum, jawaban yang dihasilkan oleh model bersifat singkat, padat, dan langsung menyoroti inti informasi, meskipun

terkadang tidak sepenuhnya identik dengan jawaban referensi manual. Beberapa temuan penting selama proses pengujian antara lain:

- Model IndoBERT-QA cenderung mengekstraksi frase utama dari context berdasarkan kata kunci yang terdapat pada pertanyaan.
- Apabila context yang tersedia kurang informatif atau mengandung ambiguitas, model sesekali memberikan jawaban yang kurang spesifik atau kurang tepat.
- Nilai Exact Match (EM) relatif lebih rendah akibat variasi dalam redaksi kalimat, namun F1 Score tetap dapat menunjukkan hasil yang baik selama substansi jawaban tetap relevan.

Secara keseluruhan, aplikasi Question Answering berbasis IndoBERT-QA telah berhasil memenuhi tujuan penelitian, yakni memberikan jawaban yang akurat dan relevan untuk pertanyaan berbasis context dalam Bahasa Indonesia. Hasil pengujian memperlihatkan bahwa sistem ini cukup layak digunakan sebagai alat bantu dalam menjawab pertanyaan berbasis teks pada domain yang telah diuji.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Penelitian ini bertujuan mengembangkan sistem *Question Answering* dengan menerapkan metode *transfer learning* menggunakan model IndoBert yang berfungsi memberikan jawaban yang akurat dan relevan terhadap pertanyaan pengguna. Data yang digunakan adalah dataset IndoLem untuk melatih dan menguji sistem dan dibuat khusus untuk mengevaluasi model NLP dalam Bahasa Indonesia.

2. Penggunaan sistem ini hanya dengan memasukkan konteks (teks sumber) pada interface yang tersedia, pada tampilan webnya sistem menggunakan framework yang disebut dengan Flask. Kemudian memasukkan pertayaan yang berkaitan dari konteks tersebut dan sistem akan bersihin dulu datanya. Lalu kita akan mendapatkan jawaban yang paling sesuai dari beberapa jawaban yang telah diukur seberapa kemiripan jawaban dari sistem dan jawaban yang benar dengan evaluasi kedua metrik yaitu Exact Match dan F1 Score. Jawaban diambil oleh model IndoBert dengan cara dihubungkan pakai *Framework Pipeline Hugging Face*. Pipeline ini adalah alat bantu atau fitur dari Hugging Face yang membantu mengambil jawaban dari model tanpa harus mengoding dengan susah payah.

3. Penelitian ini menunjukkan bahwa model IndoBert ini dapat digunakan dalam pembuatan dan pengembangan sistem tanya jawab otomatis dalam Bahasa

Indonesia, yang bermanfaat untuk pendidikan atau mencari informasi secara cepat.

5.2 Saran

Dalam pengembangan sistem lebih lanjut ada beberapa poin yang dapat dilakukan selanjutnya agar sistem menjadi lebih maksimal dan mencakup beberapa manfaat yang luas:

1. Gunakan model lain atau mengkombinasikan model-model canggih untuk meningkatkan akurasi jawaban, karna menggunakan satu model tidak cukup untuk memberikan jawaban yang terbaik.
2. Gunakan dataset yang lebih beragam dan besar agar model yang digunakan dapat dibaca dengan banyak pola bahasa yang dapat di pelajarinya.
3. Pada konteks dapat ditambah option upload dalam bentuk file.
4. Selanjutnya bisa dapat membandingkan model yang lebih efektif dengan model NLP lainnya seperti MUltigual Bert atau GPT.

DAFTAR PUSTAKA

A. Yani, D. D., Pratiwi, H. S., & Muhardi, H. (2019). Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, *7*(4), 257. https://doi.org/10.26418/justin.v7i4.30930

Agrawal, A., Atri, A., Chowdhury, A., Koneru, R., Batchu, K., & Mallavaram, S. (2021). Question Answering System Using Natural Language Processing. *International Journal of Research in Engineering, Science and Management*, *4*(12).

Apriliani, D., Handayani, S. F., Anugrahaeni, T. N., Miftahudin, A., Nurarifiah, L., & Saputra, I. T. (2023). APLIKASI QUESTION ANSWER SEBAGAI MEDIA PEMBELAJARAN INTERAKTIF UNTUK MATA PELAJARAN AKUNTANSI. *JMM (Jurnal Masyarakat Mandiri)*, *7*(2), 2003. https://doi.org/10.31764/jmm.v7i2.13867

Arabnia, H. (1999). Reading in information visualization: using vision to Think [Media Review]. *IEEE Multimedia*, *6*(4), 93–93. https://doi.org/10.1109/MMUL.1999.809241

Arifiyanti, A. A., Kartika, D. S. Y., & Prawiro, C. J. (2022). Using Pre-Trained Models for Sentiment Analysis in Indonesian Tweets. *2022 6th International Conference on Informatics and Computational Sciences (ICICoS)*, 78–83. https://doi.org/10.1109/ICICoS56336.2022.9930599

Chandra, Y. W., & Suyanto, S. (2019). Indonesian Chatbot of University Admission Using a Question Answering System Based on Sequence-to-Sequence Model. *Procedia Computer Science*, *157*, 367–374. https://doi.org/10.1016/j.procs.2019.08.179

Dhyani, B. (2021). Transfer Learning in Natural Language Processing: A Survey. *Mathematical Statistician and Engineering Applications*, *70*(1), 303–311. https://doi.org/10.17762/msea.v70i1.2312 Gieldy Akhyat Affandi Utama, Adam Hermawan, & Adi Prehanto. (2024). Analisis Pemahaman Quality

Assurance pada Mahasiswa Bisnis Digital Universitas Pendidikan Indonesia. *Jurnal Kewirausahaan Cerdas Dan Digital*, *1*(4), 51–61. https://doi.org/10.61132/jukerdi.v1i4.312

Imaduddin, H., A'la, F. Y., & Nugroho, Y. S. (2023). Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach. *International Journal of Advanced Computer Science and Applications*, *14*(8). https://doi.org/10.14569/IJACSA.2023.0140813

Juliane, C., Armant, A. A., Sastramihardja, H. S., & Supriana, I. (2018). Question-answer pair templates based on bloom's revised taxonomy. *IOP Conference Series: Materials Science and Engineering*, *434*, 012281. https://doi.org/10.1088/1757-899X/434/1/012281

Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770. https://doi.org/10.18653/v1/2020.coling-main.66

Marefat, M. M., Varecka, A. F., & Yost, J. (1997). An Intelligent Visualization Agent for simulation-based decision support. *IEEE Computational Science and Engineering*, *4*(3), 72–82. https://doi.org/10.1109/99.615433

Muliyono, M. (2021). *Identifikasi Chatbot dalam Meningkatkan Pelayanan Online Menggunakan Metode Natural Language Processing* (Doctoral dissertation, Universitas Putra Indonesia YPTK). http://repository.upiyptk.ac.id/4399/

Nugroho, K. S., Sukmadewa, A. Y., Wuswilahaken DW, H., Bachtiar, F. A., & Yudistira, N. (2021). BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews. *6th International Conference on Sustainable Information Engineering and Technology 2021*, 258–264. https://doi.org/10.1145/3479645.3479679

Rosid, M. A., Fitrani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science*

and Engineering, *874*(1), 012017. https://doi.org/10.1088/1757-899X/874/1/012017

Siregar, F. A., Siregar, A. F., & Setiadi, E. W. N. (2024). Sistem Pendukung Keputusan Menentukan E-Commerce Terbaik Menggunakan Metode Topsis. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, *5*(3), 935-947. https://pkm.tunasbangsa.ac.id/index.php/kesatria/article/view/419

Topsakal, O., & Akinci, T. C. (2023). Creating Large Language Model Applications Utilizing LangChain: A Primer on Developing LLM Apps Fast. *International Conference on Applied Engineering and Natural Sciences*, *1*(1), 1050–1056. https://doi.org/10.59287/icaens.1127

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *2017-Decem*(Nips), 5999–6009.

Verberne, S., van Halteren, H., Theijssen, D., Raaijmakers, S., & Boves, L. (2011). Learning to rank for why-question answering. *Information Retrieval*, *14*(2), 107–132. https://doi.org/10.1007/s10791-010-9136-6

Voita, E., Talbot, D., Moiseev, F., Sennrich, R., & Titov, I. (2019). Analyzing Multi-Head Self-Attention: Specialized Heads Do the Heavy Lifting, the Rest Can Be Pruned. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 5797–5808. https://doi.org/10.18653/v1/P19-1580

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). *IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*. *October*. https://doi.org/10.48550/arXiv.2009.05387