

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara dengan populasi Muslim terbesar di dunia, di mana sekitar 87% penduduknya menganut agama Islam [1]. Keberagaman ini tercermin pada banyaknya destinasi wisata religi Islam yang tersebar di berbagai wilayah, seperti masjid bersejarah, kawasan ziarah, serta pusat pembelajaran sejarah Islam di Nusantara [2]. Wisata religi Islam tidak hanya menjadi aktivitas kunjungan ke tempat ibadah, tetapi juga mencakup pengalaman spiritual, edukasi sejarah, serta pengenalan nilai-nilai budaya dan keagamaan yang berkembang di masyarakat.

Sejalan dengan potensi tersebut, wisata religi Islam di Indonesia juga menunjukkan tren pertumbuhan yang positif dalam beberapa tahun terakhir. Berdasarkan data Badan Pusat Statistik (BPS), jumlah perjalanan wisatawan nusantara sepanjang tahun 2024 mencapai 1,02 miliar perjalanan, meningkat sebesar 21,61% dibandingkan tahun sebelumnya [3]. Tingginya mobilitas wisatawan, termasuk pada destinasi wisata religi, menunjukkan bahwa kebutuhan terhadap penyediaan informasi yang cepat, akurat, terstruktur, dan mudah diakses semakin meningkat.

Pertumbuhan minat terhadap wisata religi Islam ini semakin didukung oleh meningkatnya penggunaan teknologi digital dalam pencarian informasi wisata. Wisatawan modern cenderung mencari informasi yang cepat dan akurat melalui mesin pencari, media sosial, serta aplikasi perjalanan. Namun demikian, penelitian terkait wisata religi di Indonesia masih menunjukkan adanya ruang pengembangan, terutama pada topik-topik seperti wisata sejarah, wisata ziarah, dan desa wisata keagamaan yang masih memiliki intensitas penelitian yang relatif rendah [4]. Kondisi ini menunjukkan perlunya pengembangan sistem informasi yang mampu memberikan informasi secara lebih adaptif dan sesuai kebutuhan pengguna.

Perkembangan kecerdasan buatan, khususnya dalam bidang Pemrosesan Bahasa Alami (Natural Language Processing/NLP), menawarkan pendekatan yang potensial untuk mendukung penyediaan informasi tersebut. Salah satu model yang

banyak digunakan adalah Bidirectional Encoder Representations from Transformers (BERT), yang memiliki kemampuan memahami konteks bahasa secara dua arah dan telah menunjukkan performa yang baik pada berbagai tugas *Question answering (QA)* [5]. Pengembangan lebih lanjut melalui proses *fine-tuning* juga memungkinkan model digunakan pada berbagai domain dan bahasa, termasuk bahasa dengan sumber daya terbatas [6], [7], [8]. Dalam penelitian ini digunakan model IndoBERT karena lebih sesuai untuk memahami struktur bahasa, kosakata, dan konteks pertanyaan dalam Bahasa Indonesia. Sistem juga dibangun menggunakan basis pengetahuan yang disusun dari sumber informasi wisata religi Islam sehingga mampu menyediakan jawaban yang lebih akurat dan kontekstual.

Meskipun demikian, masih terdapat kesenjangan penelitian dalam pengembangan *Question answering* pada domain wisata religi. Penelitian sebelumnya seperti QASiNa [9], berfokus pada teks keagamaan dan belum diarahkan pada informasi wisata religi Islam di Indonesia. Di sisi lain, penelitian *chatbot* pariwisata menunjukkan bahwa sistem informasi wisata masih memiliki keterbatasan dalam memahami konteks dan memberikan jawaban yang fleksibel karena sebagian besar masih menggunakan pendekatan berbasis intent recognition atau FAQ [10], [11]. Hal ini menunjukkan bahwa masih diperlukan pengembangan sistem yang mampu memahami pertanyaan pengguna dan menghasilkan jawaban yang relevan sesuai konteks.

Pemanfaatan *chatbot* sebagai antarmuka layanan informasi juga menunjukkan potensi yang baik dalam meningkatkan aksesibilitas informasi wisata [12], [13]. Platform Telegram dipilih karena ringan, mudah diakses, serta mendukung penyampaian informasi secara *real-time*. Selain itu, pengembangan sistem informasi digital sejalan dengan Peraturan Pemerintah Nomor 50 Tahun 2011 tentang RIPPARNAS 2010–2025 yang menekankan pentingnya penyediaan informasi pariwisata yang terstruktur dan berbasis teknologi [14]. Berdasarkan uraian tersebut, penelitian ini bertujuan merancang *Knowledge-Based Question answering* System berbasis IndoBERT yang diintegrasikan dengan *chatbot* Telegram untuk menyediakan informasi wisata religi Islam di Indonesia secara cepat, akurat, dan kontekstual.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang dan membangun *Knowledge-Based Question answering System* berbasis model BERT untuk memberikan informasi mengenai destinasi wisata religi Islam di Indonesia?
2. Bagaimana hasil evaluasi dan perbandingan kinerja model BERT hasil *fine-tuning* pada beberapa jumlah *epoch* menggunakan metrik Exact Match (EM), F1-Score, ROUGE-L, dan BLEU dalam menjawab pertanyaan terkait wisata religi Islam di Indonesia?

1.3 Tujuan

Adapun tujuan dari penelitian ini yaitu:

1. Merancang dan membangun *Knowledge-Based Question answering System* berbasis model BERT yang mampu memberikan informasi yang akurat mengenai destinasi wisata religi Islam di Indonesia.
2. Mengevaluasi dan membandingkan kinerja model BERT hasil *fine-tuning* pada beberapa jumlah *epoch* menggunakan metrik Exact Match (EM), F1-Score, ROUGE-L, dan BLEU untuk memperoleh gambaran performa model secara menyeluruh serta menentukan model dengan performa terbaik dalam menjawab pertanyaan terkait wisata religi Islam di Indonesia.

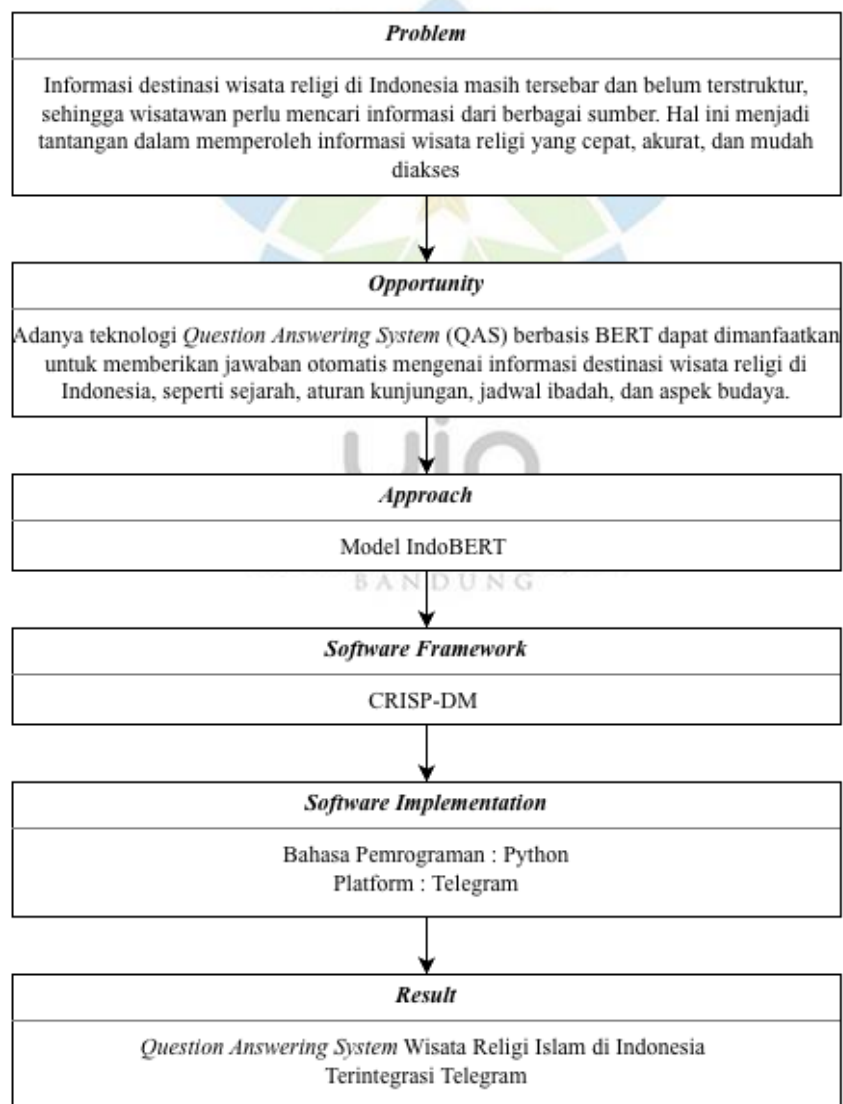
1.4 Batasan Masalah

Agar fokus dan ruang lingkup penelitian lebih terarah, maka batasan masalah dalam penelitian ini adalah:

1. Penelitian berfokus pada pengembangan *Knowledge-Based Question answering System* berbasis model IndoBERT untuk menjawab pertanyaan mengenai wisata religi Islam di Indonesia.
2. Objek penelitian dibatasi pada destinasi wisata religi Islam di Indonesia yang terdiri dari masjid dan makam, dengan informasi yang diperoleh dari sumber-sumber referensi yang kredibel.
3. Data yang digunakan meliputi informasi mengenai sejarah, lokasi, budaya atau tradisi, fasilitas, jadwal ibadah, serta informasi lain yang berkaitan dengan destinasi wisata religi.

4. Sistem menggunakan pendekatan *extractive question answering*, sehingga jawaban diambil langsung dari informasi yang terdapat pada basis pengetahuan dan tidak menghasilkan jawaban baru (generatif).
5. Sistem dibangun menggunakan pendekatan *knowledge-based* dan hanya melayani pertanyaan *single-turn* (bukan percakapan berkelanjutan).
6. Antarmuka sistem dibatasi pada platform Telegram Bot dan tidak mencakup pengembangan aplikasi berbasis web maupun mobile.
7. Evaluasi performa model dilakukan menggunakan metrik Exact Match (EM), F1-Score, ROUGE-L, dan BLEU.

1.5 Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran

Gambar 1.1 menunjukkan alur penelitian yang diawali dari permasalahan mengenai informasi destinasi wisata religi Islam di Indonesia yang masih tersebar di berbagai sumber dan belum tersusun secara terstruktur. Kondisi tersebut membuat wisatawan harus mencari informasi dari berbagai media sehingga proses memperoleh informasi menjadi kurang praktis. Akibatnya, penyediaan informasi wisata religi yang cepat, akurat, dan mudah diakses masih menjadi tantangan.

Sebagai peluang untuk mengatasi permasalahan tersebut, penelitian ini memanfaatkan teknologi *Question answering System (QAS)* berbasis BERT yang mampu memberikan jawaban secara otomatis berdasarkan pertanyaan pengguna. Informasi yang disediakan meliputi sejarah destinasi, aturan kunjungan, jadwal ibadah, serta aspek budaya yang berkaitan dengan wisata religi Islam di Indonesia.

Pendekatan yang digunakan dalam penelitian ini adalah model IndoBERT yang di-*fine-tune* untuk tugas *extractive question answering*. Proses pengembangan sistem mengacu pada kerangka kerja CRISP-DM, mulai dari pengumpulan data, penyusunan basis pengetahuan, proses pemodelan, hingga evaluasi hasil model.

Sistem kemudian diimplementasikan menggunakan bahasa pemrograman Python dan diintegrasikan dengan Telegram Bot sebagai media interaksi antara pengguna dan sistem. Hasil dari penelitian ini berupa *Knowledge-Based Question answering System* untuk informasi wisata religi Islam di Indonesia yang dapat diakses melalui Telegram sehingga pengguna dapat memperoleh informasi secara lebih cepat, akurat, dan terstruktur.

1.6 Sistematika Penulisan

Sistematika penulisan dalam laporan tugas akhir ini disusun untuk mempermudah pembaca dalam memahami isi penelitian secara terstruktur. Laporan ini terdiri atas lima bab dengan uraian sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang penelitian yang menjelaskan permasalahan terkait keterbatasan akses informasi wisata religi Islam di Indonesia, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, serta kerangka pemikiran yang digunakan sebagai dasar penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini membahas penelitian terdahulu (*state of the art*), landasan teori, serta konsep-konsep yang mendukung penelitian, seperti *Natural Language Processing (NLP)*, *Question answering System (QA)*, model BERT dan IndoBERT, serta konsep wisata religi Islam.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, yaitu pendekatan berbasis CRISP-DM yang meliputi tahapan *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Selain itu, dijelaskan pula kebutuhan data, alat, serta perancangan sistem yang dikembangkan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi sistem *Knowledge-Based Question answering* berbasis BERT, termasuk proses pelatihan model, integrasi dengan *chatbot* Telegram, serta hasil pengujian sistem berdasarkan akurasi dan relevansi jawaban. Pada bab ini juga dilakukan analisis terhadap hasil yang diperoleh sebagai jawaban dari rumusan masalah penelitian.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya, khususnya dalam pengembangan sistem QA berbasis kecerdasan buatan pada domain wisata religi Islam