

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Setiap manusia memerlukan suri teladan dalam kehidupannya. Orang yang dapat dijadikan suri tauladan tentunya adalah orang yang terbaik. Sebagai seorang muslim, kita meyakini bahwa manusia terbaik setelah Rasulullah adalah para sahabat.

Definisi “sahabat” yang dimaksud adalah sebagaimana yang dikatakan oleh Al-Hafidz Ibn Hajar Al-Asqalani rahimahullah, “Sahabat adalah setiap yang bertemu dengan Nabi *Shallallaahu'alaihi wasallam*, beriman kepada beliau, meninggal dalam Islam, sekalipun pernah murtad”. Kata “setiap” dalam hadits tersebut mencakup pria dan wanita, manusia dan jin. [1]

Rasulullah *shallallaahu'alaihi wasallam* bersabda,

“Sebaik-baik manusia adalah generasiku, kemudian yang dating sesudah mereka, kemudian yang dating sesudah mereka, kemudian dating kaum yang persaksian seorang dari mereka mendahului sumpahnya dan sumpahnya mendahului persaksiannya” (HR. Bukhari Muslim)

Berdasarkan hadits tersebut maka sangat penting bagi seorang muslim untuk mengetahui bagaimana kehidupan mereka agar setiap dari kita dapat mengikuti mereka dalam menjalani kehidupan ini. Salah satu cara untuk mengetahui bagaimana kehidupan mereka adalah dengan mengetahui sejarah. Oleh karena itu dalam penelitian ini akan digunakan sejarah shahabiyah atau sejarah sahabat wanita sebagai datanya. Hal tersebut karena wanita memiliki peranan yang penting dalam kehidupan manusia, baik sebagai seorang anak, ibu, saudara, istri ataupun bagi laki-laki untuk mengajari keluarga perempuan mereka agar bisa mengikuti kehidupan para shahabiyah.

Disisi lain, kehidupan manusia tidak terlepas dari pertanyaan serta keinginan untuk memperoleh jawaban yang benar dan tepat. Jenis jawaban dapat berupa audio, gambar maupun teks. Saat ini yang paling umum adalah jawaban berupa teks. Dahulu, manusia mencari jawaban dari buku, tetapi mencari jawaban melalui buku memiliki keterbatasan waktu yang tidak efisien dan sumber yang

terbatas. Dengan adanya sistem *search engine* sumber yang terbatas tersebut dapat teratasi karena *systems search engine* mengembalikan dokumen-dokumen, tetapi system *search engine* tidak mengatasi keefisienan waktu karena manusia tetap harus mencari jawaban yang paling sesuai dari dokumen-dokumen yang diperoleh. Oleh karena itu dibuatlah *question answering system*, dimana sistem ini dapat mengembalikan jawaban atas pertanyaan yang diinputkan oleh *user*.

Sistem *Question Answering* dapat mengembalikan jawaban yang spesifik dan tepat dari dokumen, sehingga manusia dapat memperoleh jawaban yang diinginkan dengan lebih cepat dan efisien. Sistem ini bekerja dengan mengembalikan kalimat atau potongan teks tertentu yang paling relevan dengan pertanyaan yang diajukan. Dengan adanya sistem *Question Answering*, manusia dapat memperoleh jawaban yang sesuai dengan pertanyaan yang diajukan dengan lebih mudah dan efektif berdasarkan sumber yang sudah disediakan.

Dalam membuat sebuah sistem *Question Answering* dibutuhkan suatu model yang dapat memahami pertanyaan serta konteks agar dapat memberikan jawaban yang benar dan tepat kepada pengguna. Salah satu model NLP yang dapat digunakan untuk tugas *question answering systems* adalah model BERT. Model ini merupakan model yang dikembangkan oleh para peneliti Google pada tahun 2018.

Sebelumnya, Jacob Devlin, Ming Weng Chang, Kenton Lee, dan Kristina Toutanova telah melakukan penelitian BERT menggunakan dataset SQuAD dengan hasil yang diperoleh mencapai nilai F1 mencapai 83.1 pada pengujian *question answering systems*. [2] Berdasarkan penelitian tersebut, pada Skripsi ini akan dilakukan percobaan *question answering system* untuk mengetahui bagaimana kinerja model BERT dalam menjawab pertanyaan berdasarkan kompleksitas konteks pada data sirah shahabiyah yang bersumber dari buku (Kisah Keteledanan Wanita-Wanita Surga di Zaman Rasulullah Shallallahu'alaihi wasallam) [3].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka rumusan masalah pada penelitian ini, diantaranya:

1. Dibutuhkan suatu sistem yang dapat memberikan jawaban yang tepat mengenai pertanyaan yang diajukan
2. Teks yang di gunakan sebagai sumber informasi biasanya berupa kumpulan beberapa kalimat dari yang sederhana hingga kompleks sehingga diperlukan penelitian untuk mengetahui bagaimana kinerja model dalam memahami setiap jenis kalimat tersebut

1.3 Batasan Masalah

1. Konteks sebagai sumber informasi untuk menjawab pertanyaan dibuat berdasarkan informasi yang diperoleh dari buku (Kisah Keteledanan Wanita-Wanita Surga di Zaman Rasulullah Shallallaahu'alaihi wasallam)
2. Pertanyaan dibuat berdasarkan isi konteks yang telah dibuat
3. Pertanyaan yang digunakan dalam penelitian ini hanya kata tanya “apa” dan kata tanya “siapa”
4. Model yang digunakan adalah BERT yang sudah tersedia untuk Bahasa Indonesia, yaitu Indobert-QA.

1.4 Tujuan Penelitian

Tujuan dari penelitian skripsi ini adalah sebagai berikut :

1. Mengetahui kinerja model BERT dalam *question answering system* dalam menjawab pertanyaan berdasarkan kompleksitas konteks yang diberikan
2. Menganalisis kinerja model BERT dalam *question answering system* dalam menjawab pertanyaan berdasarkan kata tanya dari kompleksitas konteks yang diberikan

1.5 Metode Penelitian

1. Studi Literatur

Tahap Studi Literatur merupakan tahap untuk mengumpulkan materi, dan informasi mengenai question answering dan model BERT dari berbagai sumber, diantaranya buku, jurnal, artikel, dan lain sebagainya.

2. Analisis dan Persiapan Penelitian

Pada tahap ini, penulis mengkaji dan menganalisis hasil dari tahap Studi Literatur sesuai dengan masalah yang dipilih dalam penelitian ini. Kemudian di tahap ini juga dilakukan pembuatan dataset beserta pasangan pertanyaan dan jawabannya.

3. Simulasi

Pada tahap ini penulis melakukan simulasi model BERT menggunakan data Sirah Shahabiyah serta menggunakan bahasa pemrograman *python*. Kemudian dari hasil simulasi tersebut akan dianalisis bagaimana kinerja model dalam menjawab pertanyaan.

1.6 Sistematika Penelitian

Sistematika penulisan pada skripsi ini terdiri dari lima bab.

BAB 1 : PENDAHULUAN

Pada bagian ini berisi tentang latar belakang masalah dari penelitian ini, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian, dan sistematika penelitian.

BAB 2 : LANDASAN TEORI

Pada bagian ini berisi penjelasan mengenai teori-teori yang berhubungan dengan penelitian dan penulisan laporan Skripsi, yaitu *Natural Language Processing, Question Answering Systems, Transformer, Bidirectional Encoder Representations from Transformers (BERT)* dan Kalimat.

BAB 3 : QUESTION ANSWERING SYSTEM MENGGUNAKAN MODEL BERT

Pada bagian ini berisi pembahasan mengenai penelitian yang dilakukan, yaitu pembuatan data, *data preprocessing, fine tuning*, dan *testing*.

BAB 4 : QUESTION ANSWERING SYSTEMS PADA MODEL BERT BERDASARKAN KOMPLEKSITAS KONTEKS

Pada bagian ini berisi penjelasan mengenai hasil penelitian pada dataset Sirah Shahabiyah menggunakan program yang sudah dibuat yaitu *Question Answering System (QAS)* menggunakan model BERT. Penelitian dilakukan dengan menggunakan dua skenario, yaitu kinerja BERT

berdasarkan kompleksitas kalimat dalam konteks dan kinerja BERT berdasarkan jenis kata tanya “apa” dan “siapa”.

BAB 5 : PENUTUP

Pada bab ini dipaparkan kesimpulan sebagai jawaban dari rumusan masalah serta saran untuk penelitian selanjutnya pada topik yang serupa.

