



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kecerdasan buatan merupakan bidang penelitian yang populer di ilmu komputer karena telah meningkatkan cara hidup manusia di banyak bidang (Pannu, 2015). Salah satu penerapan teknologi kecerdasan buatan adalah *voice chatbot*. Alumni Universitas Multimedia Nusantara melakukan penelitian mengenai aplikasi *voice chatbot* bernama Jacob (Wijaya dan Wicaksana, 2019). Aplikasi ini dapat memberikan informasi mengenai program *dual degree* di UMN dengan menggunakan Bahasa Inggris. Jawaban yang diberikan oleh Jacob berasal dari database yang menyimpan jawaban-jawaban untuk pertanyaan tertentu. Jacob masih memiliki banyak ruang untuk berkembang. Salah satunya adalah kemampuan untuk mengetahui jawaban dari pertanyaan yang bukan mengenai program *dual degree* UMN.

Internet menyediakan hampir semua informasi yang kita butuhkan. Dengan memanfaatkan internet, informasi dari suatu *keyword* dapat ditemukan. Untuk mendapatkan jawaban dari informasi yang ditemukan perlu peranan dari *text mining*. Hasil pencarian dari internet berupa data mentah *website* yang ditelusuri dan harus diolah terlebih dahulu sebelum diterima sebagai jawaban. Proses pengolahan data tersebut dapat dilakukan dengan teknik *text mining*. ALBERT merupakan salah satu arsitektur *neural network* yang dapat digunakan pada *text mining*.

Lan dkk (2019) melakukan penelitian mengenai arsitektur *neural network* bernama ALBERT. Penelitian didasari oleh arsitektur *neural network* BERT

(*Bidirectional Encoder Representation from Transformers*) dan membuat BERT menjadi lebih ringan. Hasil dari penelitian menunjukkan ALBERT memiliki nilai *F-score* tertinggi sebesar 92,2% dari proses *testing* dengan *dataset* SQuAD v2 dan terdapat peningkatan sebesar 17,4% dibandingkan dengan penelitian mengenai arsitektur *neural network* BERT yang dilakukan oleh Devlin dkk (2018).

Arsitektur ALBERT menunjukkan performa yang baik pada kasus *question answering*. Zhang dkk (2020) mendapatkan nilai *exact match* sebesar 87% dan *F-score* sebesar 90,2% dalam penelitian mengenai *machine reading comprehension*. Dengan didasari latar belakang masalah di atas penelitian ini dilakukan untuk mengimplementasikan ALBERT sebagai *web service* pada aplikasi *voice chatbot* Jacob untuk *text mining*. *Web service* akan dikembangkan menggunakan bahasa pemrograman Python 3.5 dan *framework web service* Python Flask 1.0.2. *Whitebox testing* akan digunakan untuk pengujian *web service*. Evaluasi performa *web service* akan dilihat dari nilai akurasi, *F-score* hasil implementasi ALBERT, dan *user time* dari *web service* yang dihitung menggunakan *time.perf\_counter()* pada bahasa pemrograman Python (Wicaksana dan Tang, 2017; Poole dan Mackworth, 2020).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka dapat diambil suatu rumusan masalah sebagai berikut.

1. Bagaimana mengimplementasikan ALBERT sebagai *web service* untuk *text mining* pada aplikasi *voice chatbot* Jacob?

2. Berapakah nilai akurasi dan *F-Score* untuk kecocokan jawaban *web service* dengan pertanyaan *user*?
3. Berapakah nilai *user time* dari *web service*?

### 1.3 Batasan Masalah

Dalam penelitian ini, beberapa batasan masalah sudah ditentukan sebagai berikut.

1. *Text Mining* hanya dalam Bahasa Inggris.
2. *Dataset* berasal dari Stanford Question Answering Dataset (SQuAD) v2 (El-geish, 2020).
3. *Pre-trained model* ALBERT yang digunakan adalah *albert-base-v2* dan *albert-xxlarge-v1* (El-geish, 2020).
4. *Region* diisi dengan “*All Regions*” pada pengaturan Google Custom Search API.

### 1.4 Tujuan Penelitian

Tujuan yang diharapkan dapat tercapai dalam penelitian ini adalah sebagai berikut.

1. Mengimplementasikan ALBERT sebagai *web service* untuk *text mining* pada aplikasi *voice chatbot* Jacob.
2. Mengetahui nilai akurasi dan *F-Score* untuk kecocokan jawaban *web service* dengan pertanyaan *user*.
3. Mengetahui nilai *user time* dari *web service*.

## 1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Memberikan kemampuan kepada aplikasi *voice chatbot* Jacob untuk mencari tahu sendiri jawaban dari sebuah pertanyaan yang tidak ada di basis pengetahuan Jacob.
2. Membantu *voice chatbot* Jacob menjadi kecerdasan buatan dengan pendekatan *bottom-up*.

## 1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan ini adalah sebagai berikut.

### BAB I PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

### BAB II LANDASAN TEORI

Bab ini menjelaskan teori-teori dan konsep dasar yang mendukung penelitian terkait, yaitu teori dasar mengenai *text mining*, ALBERT, *voice chatbot* Jacob, SQuAD, akurasi, *F-score*, dan *user time*.

### BAB III METODOLOGI PENELITIAN DAN PERANCANGAN SISTEM

Bab ini menjelaskan metodologi, analisis, dan rancangan dari *web service* yang dibuat. Gambaran umum, *flowchart*, dan diagram perancangan *web service* dijabarkan pada bab ini.

#### BAB IV IMPLEMENTASI DAN ANALISIS

Bab ini berisikan implementasi *web service*, pengujian, dan data hasil penelitian maupun hasil analisis untuk menjawab rumusan masalah pada penelitian ini.

#### BAB V SIMPULAN DAN SARAN

Bab ini berisikan simpulan dari hasil penelitian terhadap tujuan yang ingin dicapai dalam penelitian dan saran untuk pengembangan sistem lebih lanjut.