

ABSTRAK

Perkembangan teknologi kecerdasan buatan, khususnya dalam bidang pemrosesan bahasa alami (*natural language processing* atau NLP), telah memungkinkan penerapan model bahasa besar (*large language model* atau LLM) dalam berbagai bidang, termasuk kesehatan. Penelitian ini bertujuan untuk mengembangkan *chatbot* cerdas yang dapat memberikan informasi seputar kesehatan gigi dan mulut dengan memanfaatkan model bahasa besar. Data pelatihan yang digunakan berasal dari situs Alodokter selama periode 2014 hingga 2025, yang mencakup pertanyaan dan jawaban terkait konsultasi kesehatan gigi dan mulut.

Model bahasa yang digunakan dalam penelitian ini adalah GPT-2 dan Llama 3.1 8B. Proses pelatihan dilakukan menggunakan metode *supervised learning*, di mana pasangan pertanyaan dan jawaban digunakan untuk mengarahkan model dalam memahami konteks dan menghasilkan respons yang relevan dan informatif. Eksperimen dilakukan untuk mengevaluasi kinerja model secara kuantitatif.

Hasil menunjukkan bahwa meskipun kedua model mampu menghasilkan kinerja yang baik secara kuantitatif, keduanya masih mengalami kesulitan pada pertanyaan kompleks misalnya pengaruh penyakit gigi terhadap kesehatan jantung serta cenderung mengulang jawaban. Keunggulan Llama 3.1 8B dibanding GPT-2 terlihat pada kemampuan merangkum informasi secara sederhana maupun detail, berkat *context length* yang lebih panjang dan efektivitas *prompt engineering*. Namun, peningkatan performa tersebut diikuti kebutuhan komputasi yang lebih tinggi.

Kata kunci: Kecerdasan Buatan, Model Bahasa Besar, Model Bahasa Generatif, Model Generatif, *Chatbot*.

ABSTRACT

The advancement of artificial intelligence technology, particularly in the field of natural language processing (NLP), has enabled the application of large language models (LLMs) across various domains, including healthcare. This study aims to develop an intelligent chatbot capable of providing information on dental and oral health by leveraging a large language model. The training data were sourced from the Alodokter website over the period 2014 to 2025, encompassing questions and answers related to dental and oral health consultations.

The language models employed in this research are GPT-2 and Llama 3.1 8B. Training was conducted using supervised learning, whereby question-answer pairs guide the model in understanding context and generating relevant, informative responses. Experiments were performed to quantitatively evaluate the performance of the models.

The results show that while both models performed well quantitatively, they still struggled with complex questions, such as the impact of dental disease on heart health, and tended to repeat answers. Llama 3.1 8B advantage over GPT-2 was seen in its ability to summarize information in both simple and detailed ways, thanks to its longer context length and effective prompt engineering. However, this improved performance came at the cost of higher computational requirements.

Keywords: Artificial Intelligence, Large Language Model, Generative Language Model, Generative Models, Chatbot.