

STUDI PERBANDINGAN AKURASI SERTA RELEVANSI JAWABAN PADA APLIKASI *CHATGPT* DAN *PERPLEXITY* DALAM PENCARIAN INFORMASI MENGGUNAKAN METODE AHP

ABSTRAK

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) telah menghasilkan berbagai aplikasi *chatbot* yang berperan penting dalam pencarian informasi. ChatGPT dan Perplexity merupakan dua aplikasi yang sering digunakan untuk tujuan tersebut. Penelitian ini bertujuan untuk membandingkan tingkat akurasi dan relevansi jawaban yang dihasilkan oleh kedua aplikasi menggunakan dua metode evaluasi, yaitu *BLEU Score* dan *Analytical Hierarchy Process (AHP)*. *BLEU Score* digunakan untuk mengukur kesamaan jawaban terhadap referensi manusia secara linguistik dan kontekstual, sedangkan AHP digunakan untuk mengevaluasi performa berdasarkan kriteria akurasi, relevansi, waktu respons, dan kesamaan kata. Hasil pengujian menunjukkan bahwa ChatGPT memiliki nilai *BLEU Score* rata-rata sebesar 0,77 dengan waktu respons 2,85 detik, sedangkan Perplexity memperoleh skor 0,75 dan waktu respons 3,13 detik. Evaluasi AHP juga menunjukkan bahwa ChatGPT lebih unggul dalam hal akurasi (nilai 4,533), sementara Perplexity unggul tipis pada relevansi (nilai 4,533). Namun, skor akhir AHP menunjukkan Perplexity sedikit lebih tinggi dengan skor global 4,50 dibandingkan ChatGPT yang memperoleh 4,42. Meskipun demikian, berdasarkan konsistensi jawaban dan waktu respons, ChatGPT dinilai lebih efisien dan stabil dalam mendukung kebutuhan pencarian informasi. Dengan demikian, penelitian ini menyimpulkan bahwa ChatGPT secara umum lebih unggul dalam aspek akurasi dan efisiensi, sedangkan Perplexity memiliki keunggulan dalam relevansi isi. Hasil ini dapat menjadi acuan dalam memilih aplikasi AI yang sesuai dengan kebutuhan pengguna, baik dalam konteks akademik maupun profesional.

Kata Kunci: Akurasi, Relevansi, ChatGpt, Perplexity, AHP

A STUDY COMPARING THE ACCURACY AND RELEVANCE OF ANSWERS ON CHATGPT AND PERPLEXITY APPLICATIONS IN SEARCHING FOR INFORMATION USING THE AHP METHOD

ABSTRACT

The advancement of artificial intelligence (AI) technology has led to the emergence of various chatbot applications that play an important role in information retrieval. ChatGPT and Perplexity are two widely used platforms in this domain. This study aims to compare the accuracy and relevance of the answers generated by both applications using two evaluation methods: BLEU Score and Analytical Hierarchy Process (AHP). BLEU Score is used to measure linguistic and contextual similarity with human reference answers, while AHP is employed to assess performance based on criteria such as accuracy, relevance, response time, and word similarity. The experimental results show that ChatGPT achieved an average BLEU Score of 0.77 with an average response time of 2.85 seconds, while Perplexity scored 0.75 with a response time of 3.13 seconds. AHP evaluation further indicates that ChatGPT scored higher in terms of accuracy (4.533), while Perplexity scored slightly higher in relevance (4.533). However, the final AHP global score favored Perplexity (4.50) over ChatGPT (4.42). Despite this, ChatGPT demonstrated greater consistency and faster response time, making it more efficient in supporting information-seeking tasks. In conclusion, the study finds that ChatGPT is generally superior in terms of accuracy and efficiency, while Perplexity offers an edge in relevance. These findings provide practical insights for users in selecting the most appropriate AI-based application, particularly for academic and professional information needs.

Keywords: Accuracy, Relevance, ChatGpt, Perplexity, AHP