LEMBAR PENGESAHAN

Nomor: 31/LP-LPBK/II/2025

Judul : Tidak Ada Judul

Nama : Fadhila Nur Aini

Menerangkan bahwa Abstrak ini telah diterjemahkan ke dalam Bahasa Inggris oleh Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK), Universitas Muhammadiyah Pekajangan Pekalongan.

Pekalongan, 14 Februari 2025

Disahkan oleh,
Kepala Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK)

Aida Rusmariana, S.Kep., Ns., MAN

Program Studi Sarjana Informatika
Fakultas Teknik dan Ilmu Komputer
Universitas Muhammadiyah Pekajangan Pekalongan
Februari, 2025

ABSTRAK

Fadhila Nur Aini

Penelitian ini bertujuan untuk menganalisis dan membandingkan *Performance & Accuracy* ChatGPT dan Gemini dalam menyelesaikan tugas mengenai *Internet of Things* (IoT). Penilaian menggunakan *Bleu Score Matrix*, metrik berbasis n-gram untuk mengevaluasi kualitas teks berdasarkan kesamaan dengan teks referensi. Data terdiri dari 50 pertanyaan mengenai IoT dari TechTarget.com dan Guru99.com. Pendekatan kuantitatif deskriptif diterapkan melalui implementasi Python dan Google Colab. Hasil analisis menunjukkan bahwa rata-rata *Bleu Score* ChatGPT adalah 0.0173, sedangkan Gemini memiliki skor lebih tinggi, yaitu 0.0190. Meskipun perbedaannya kecil, angka ini menunjukkan bahwa secara rata-rata, Gemini memiliki kesesuaian teks yang sedikit lebih baik dalam menjawab 50 pertanyaan terkait IoT. Uji *paired sample t-test* dilakukan untuk menentukan signifikansi perbedaan antara kedua model. Hasil analisis menunjukkan *t-statistic* sebesar -1.0046 dengan *p-value* 0.3322, yang lebih besar dari tingkat signifikansi 0.05. Dengan demikian, hipotesis nol (H_0) diterima, sedangkan hipotesis alternatif (H_1) ditolak, yang mengindikasikan tidak terdapat perbedaan signifikan antara *Performance & Accuracy* ChatGPT dan Gemini berdasarkan rata-rata *Bleu Score*. Penelitian ini menunjukkan bahwa kedua model memiliki tingkat akurasi yang serupa dalam menghasilkan teks sesuai referensi.

Kata Kunci: *Bleu Score Matrix,* ChatGPT, Gemini, IoT, *Performance* & *Accuracy Benchmarking*

Undergraduate Program in Informatics
Faculty of Engineering and Computer Sciences
University of Muhammadiyah Pekajangan Pekalongan

ABSTRACT

Fadhila Nur Aini

This study aims to analyze and compare the *Performance and Accuracy of* ChatGPT and Gemini in completing tasks regarding the *Internet of Things* (IoT). The assessment uses the *Bleu Score Matrix*, an n-gram-based metric to evaluate text quality based on similarity to reference text. The data consists of 50 questions regarding IoT from TechTarget.com and Guru99.com. A descriptive quantitative approach was applied through the implementation of Python and Google Colab. The results of the analysis show that the average ChatGPT *Bleu Score* is 0.0173, while Gemini has a higher score of 0.0190. Although the difference is small, this figure indicates that on average, Gemini has a slightly better text match in answering 50 IoT-related questions. A *paired sample t-test* was conducted to determine the significance of the difference between the two models. The analysis showed a *t-statistic* of -1.0046 with a *p-value* of 0.3322, which is greater than the significance level of 0.05. Thus, the null hypothesis (H_0) is accepted, while the alternative hypothesis (H_1) is rejected, which indicates that there is no significant difference between ChatGPT and Gemini's *Performance* & *Accuracy* based on the average *Bleu Score*. This study shows that both models have a similar level of accuracy in producing reference-based text.

Keywords: *Bleu Score Matrix,* ChatGPT, Gemini, IoT, *Performance* & *Accuracy Benchmarking*