

Pemanfaatan AI sebagai alat Terjemahan Bahasa Arab : Kajian kualitas berdasarkan metrik otomatis dan penilaian manusia

Oleh:

Rahma Auliya Izzati

Farikh Marzuki Ammar

Pendidikan Bahasa Arab

Universitas Muhammadiyah Sidoarjo

Juni 2026



Pendahuluan

- Latar belakang penelitian ini berangkat dari kompleksitas bahasa Arab yang cukup tinggi baik dari aspek morfologi, sintaksis maupun makna. Hal ini sering menjadi kendala dalam memahami dan menerjemahkan teks Arab terutama yang bersifat idiomatik dan kontekstual.
- Di sisi lain perkembangan teknologi AI seperti ChatGPT mulai banyak dimanfaatkan sebagai alat bantu penerjemahan. Namun meskipun terlihat praktis, kualitas terjemahan dari ChatGPT tetap perlu diuji baik secara otomatis maupun melalui penilaian manusia.

Pertanyaan Penelitian (Rumusan Masalah)

1. Mengetahui kualitas hasil terjemahan ChatGPT dari teks Bahasa Arab ke Bahasa Indonesia jika dievaluasi menggunakan metode gabungan antara evaluasi otomatis dan evaluasi manusia
2. Bagaimana Tingkat kesesuaian kedua metode tersebut dalam menilai kualitas terjemahan ChatGPT sebagai alat terjemahan bahasa Arab.

Metode

- Penelitian ini menggunakan metode kuantitatif deskriptif. Data berupa teks bahasa Arab yang diterjemahkan menggunakan ChatGPT kemudian hasilnya dinilai oleh 30 mahasiswa sebagai evaluator.
- Instrumen yang digunakan adalah skala Likert dengan tiga aspek penilaian yaitu keakuratan, keberterimaan dan keterbacaan. Selanjutnya analisis dilakukan menggunakan metrik BLEU dan statistik deskriptif untuk melihat kecenderungan data.

Hasil

Hasil BLEU

Hasil analisis menunjukkan bahwa rata-rata skor BLEU yang diperoleh sebesar 0,4123 yang mengindikasikan tingkat kesesuaian yang moderat. Namun skor ini bersifat fluktuatif dengan nilai terendah 0,0764 dan tertinggi 0,6004. Hal ini menunjukkan bahwa kualitas terjemahan ChatGPT belum konsisten dan sangat dipengaruhi oleh kompleksitas struktur kalimat dalam bahasa Arab.

Hasil Evaluasi Manusia

Sementara itu berdasarkan evaluasi manusia, aspek keakuratan dan keterbacaan berada pada kategori tinggi sedangkan keberterimaan berada pada kategori sedang. Artinya, secara umum terjemahan ChatGPT sudah mampu menyampaikan makna dan cukup mudah dipahami tetapi masih terdapat kekurangan pada aspek kealamiah bahasa sehingga terkesan kaku.

Pembahasan

- Dari hasil tersebut dapat disimpulkan bahwa skor BLEU tidak selalu sejalan dengan penilaian manusia. Hal ini karena BLEU hanya mengukur kesamaan teks secara numerik sedangkan manusia menilai makna, konteks dan kealamiahannya bahasa.
- Evaluasi kualitas terjemahan sebaiknya dilakukan secara komplementer dengan menggabungkan kedua pendekatan tersebut

Temuan Penting Penelitian

Penelitian menunjukkan bahwa terjemahan ChatGPT dari Bahasa Arab ke Bahasa Indonesia memiliki kualitas yang cukup baik tetapi masih fluktuatif. Skor rata-rata BLEU sebesar 0,4123 menunjukkan adanya variasi kesesuaian antara terjemahan ChatGPT dan terjemahan rujukan. Berdasarkan penilaian manusia, aspek keakuratan dan keterbacaan berada pada kategori tinggi, sedangkan keberterimaan berada pada kategori sedang. Selain itu, skor BLEU dan penilaian manusia tidak selalu menunjukkan hasil yang selaras, karena BLEU berfokus pada kesamaan n-gram, sedangkan manusia juga mempertimbangkan makna, konteks dan kealamiahannya bahasa. Dengan demikian ChatGPT berpotensi menjadi alat bantu penerjemahan, tetapi tetap memerlukan evaluasi dan penyuntingan manusia untuk memperoleh hasil yang optimal.

Manfaat Penelitian

- Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan evaluasi terjemahan berbasis AI khususnya dalam pembelajaran bahasa Arab.
- Selain itu penelitian ini juga mendorong penggunaan ChatGPT secara lebih kritis tidak hanya sebagai alat penerjemah instan tetapi juga sebagai sarana untuk melatih analisis dan evaluasi bahasa.

Referensi

1. K. Peneliti, *Bahasa arab di Indonesia*. 1436. [Online]. Available: <https://ketabpedia.com/في-العربية-اللغة/تحميل/إندونيسيا/>
2. L. Judijanto et al., *Optimalisasi ChatGPT: Panduan dan Penerapan untuk Belajar, Mengajar, dan Membuat Konten Tanpa Batas*. PT. Green Pustaka Indonesia, 2025. [Online]. Available: <https://books.google.co.id/books?id=mhBLEQAAQBAJ>
3. A. Ruhmadi and M. Z. Al Farisi, "Analisis Kesalahan Morfologi Penerjemahan Arab-Indonesia pada ChatGPT," *Aphorisme J. Arab. Lang. Lit. Educ.*, vol. 4, no. 1, pp. 55–75, 2023, doi: 10.37680/aphorisme.v4i1.3148.
4. S. N. Latifah and W. I. F. Djamilah, "Penggunaan chatgpt dalam penerjemahan bahasa arab," *Al Dirosah J. Pendidik. Agama Islam*, vol. 01, no. 2, pp. 1–13, 2024, [Online]. Available: <http://ojs.iai-darussalam.ac.id/index.php/pai>.
5. M. Waruwu, S. N. Pu`at, P. R. Utami, E. Yanti, and M. Rusydiana, "Metode Penelitian Kuantitatif: Konsep, Jenis, Tahapan dan Kelebihan," *J. Ilm. Profesi Pendidik.*, vol. 10, no. 1, pp. 917–932, 2025, doi: 10.29303/jipp.v10i1.3057.
6. A. Wijanarko and A. N. Al Haura, "Model Recurrent Neural Network-Gated Recurrent Unit untuk Membangun Mesin Penerjemah Bahasa Indonesia-Banyumasan," *J. Eksplora Inform.*, vol. 13, no. 2, pp. 218–226, 2024, doi: 10.30864/eksplora.v13i2.977.

