

**Chatbot Interaktif Berbasis Transformer untuk Penyediaan Informasi
Dinamis di Industri Esports**

Calvin¹, Viny Christanti Mawardi², Irvan Lewenusa³

¹²³Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas
Tarumanagara
boeing999calvin@gmail.com¹

ABSTRACT

The esports industry is experiencing rapid growth and has given rise to the need for an information system that is able to present news quickly and relevantly. This study develops an interactive chatbot based on the BART (Bidirectional and Auto-Regressive Transformers) model to answer user questions directly based on information collected through web scraping. The system is designed using Python and Flask, and uses MongoDB as data storage. Evaluation is carried out through F1-Score and Conversation-Based Evaluation to assess system performance in terms of relevance, coherence, and fluency of answers. The results show that the chatbot is able to provide accurate and coherent answers with fast response times, although there are still challenges in terms of recall and understanding ambiguous contexts.

Keywords: chatbot, transformers, BART, NLP, esports, web scraping

ABSTRAK

Industri esports mengalami pertumbuhan yang pesat dan memunculkan kebutuhan akan sistem informasi yang mampu menyajikan berita secara cepat dan relevan. Penelitian ini mengembangkan chatbot interaktif berbasis model BART (Bidirectional and Auto-Regressive Transformers) untuk menjawab pertanyaan pengguna secara langsung berdasarkan informasi yang dikumpulkan melalui web scraping. Sistem dirancang menggunakan Python dan Flask, serta menggunakan MongoDB sebagai penyimpanan data. Evaluasi dilakukan melalui F1-Score dan Conversation-Based Evaluation untuk menilai performa sistem dalam hal relevansi, koherensi, dan kelancaran jawaban. Hasil menunjukkan bahwa chatbot mampu memberikan jawaban yang tepat dan koheren dengan waktu respons yang cepat, meskipun masih terdapat tantangan dalam hal recall dan pemahaman konteks yang ambigu.

Kata kunci: chatbot, transformers, BART, NLP, esports, web scraping

PENDAHULUAN

Perkembangan industri esports mendorong meningkatnya kebutuhan akan informasi yang cepat, relevan, dan akurat. Pengguna dan penggemar esports membutuhkan cara efisien untuk memperoleh berita terbaru mengenai pertandingan, tim, pemain, hingga perkembangan teknis lainnya. Tantangan utama dalam penyediaan informasi tersebut adalah tingginya volume berita yang tersebar di berbagai platform dan kurangnya sistem otomatis yang dapat merangkum serta menyajikan informasi secara personal [1].

Esports adalah istilah yang digunakan untuk menggambarkan kompetisi video game yang dimainkan secara profesional oleh individu atau tim. Jutaan orang telah menjadi penonton dan partisipan esports dalam beberapa tahun terakhir.

Esports mencakup berbagai jenis permainan, seperti permainan strategi, tembak-menembak, olahraga simulasi, dan masih banyak lagi, yang dimainkan dalam turnamen besar serta memberikan hadiah kepada juara [2]. Esports juga didukung oleh kemajuan internet dan teknologi, yang membuat partisipasi dan informasi semakin mudah diakses, termasuk di Indonesia, di mana komunitas esports berkembang pesat [3].

Chatbot adalah solusi potensial dalam permasalahan ini. Dengan kemampuan memahami bahasa alami dan memberikan respons secara langsung, chatbot yang dilengkapi teknologi pemrosesan bahasa alami (Natural Language Processing/NLP) dapat menyajikan informasi dengan cara yang lebih interaktif [4]. Penelitian ini merancang sebuah chatbot interaktif berbasis model BART (Bidirectional and Auto-Regressive Transformers) yang diintegrasikan dengan sistem pengumpulan data melalui web scraping dari situs esports [5].

Transformers merupakan arsitektur model pembelajaran mendalam yang digunakan secara luas dalam NLP karena efisiensinya dalam memahami konteks sekuens teks [8]. BART adalah salah satu model berbasis transformers yang dirancang oleh Facebook AI Research untuk tugas-tugas seperti summarization dan text generation [5]. Penelitian sebelumnya telah menggunakan BART untuk menghasilkan ringkasan dokumen dan menjawab pertanyaan berbasis konteks [8].

Dalam konteks chatbot, pendekatan berbasis model pembelajaran mendalam menunjukkan hasil yang menjanjikan dalam pemahaman dan respons terhadap pertanyaan pengguna [1]. Integrasi sistem scraping untuk pengumpulan berita esports juga terbukti efektif digunakan dalam sistem informasi real-time [6]. Evaluasi performa chatbot umumnya dilakukan dengan metrik seperti Precision, Recall, dan F1-Score, serta Conversation-Based Evaluation (CBE) yang menilai fluency, relevansi, dan koherensi [7].

METODE PENELITIAN

Penelitian ini bertujuan untuk membangun sebuah chatbot interaktif yang dapat menjawab pertanyaan seputar berita esports dengan ringkas dan relevan. Metodologi yang digunakan terdiri dari beberapa tahap utama, yaitu:

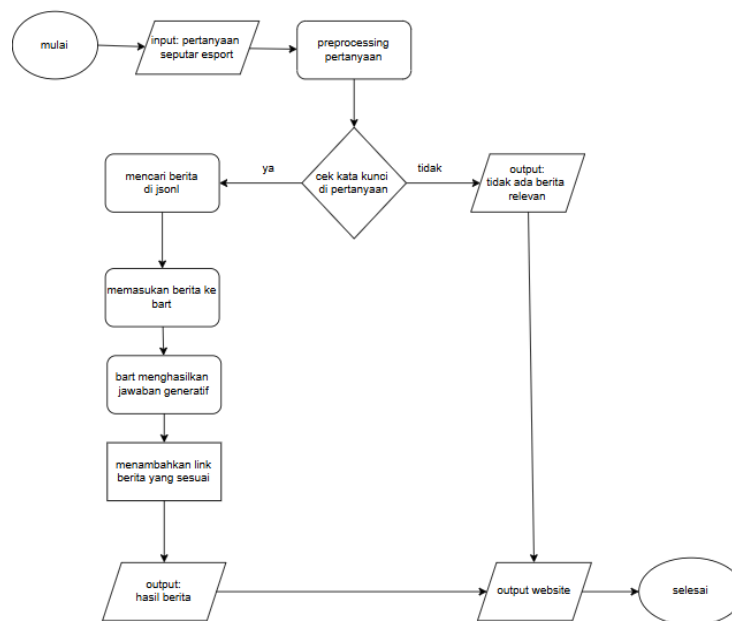
Arsitektur Sistem

Sistem yang dikembangkan terdiri dari beberapa komponen utama, yaitu:

- Web scraping module: bertugas mengambil data berita terbaru dari situs-situs esports terpercaya seperti esports.id, liquipedia.net, dan ligagame.tv menggunakan Python dan library BeautifulSoup.
- Preprocessing module: melakukan pembersihan dan normalisasi teks, termasuk case folding, tokenisasi, filtering, dan stemming agar data siap diproses oleh model NLP.
- Model NLP (BART): digunakan untuk menjawab pertanyaan pengguna. Model facebook/bart-base di-fine-tune dengan data berita esports yang telah dianotasi dalam format tanya-jawab.
- Backend: dibangun menggunakan Flask untuk mengatur komunikasi antara

frontend dan model.

- Frontend: antarmuka pengguna berbasis web dibangun dengan NextJS, dirancang menyerupai sistem chatbot interaktif.
- Database: menggunakan JsonL yydsebagai penyimpanan data berita yang telah dikumpulkan.



Gambar 1 Rancangan Proses

Pengumpulan dan Pengolahan Data

Data dikumpulkan dalam format .jsonl berisi elemen judul, isi, tanggal, dan sumber. Data ini dibersihkan dari elemen HTML, karakter tidak relevan, serta di-split ke dalam pasangan pertanyaan dan konteks berita. Setiap pertanyaan kemudian dihubungkan dengan jawaban ringkasan dari isi berita.

Fine-Tuning Model

Model BART di-fine-tune menggunakan HuggingFace Seq2SeqTrainer. Parameter pelatihan meliputi:

- learning_rate: 2e-5
- batch_size: 16
- epochs: 3

Evaluasi selama training dilakukan menggunakan metrik ROUGE dan Exact Match (EM).

Proses Tanya Jawab

Setiap pertanyaan pengguna diproses sebagai berikut:

- Sistem mencocokkan pertanyaan dengan konteks berita menggunakan cosine similarity (TF-IDF atau sentence embeddings).

Comit: Communication, Information and Technology Journal

Volume 3 Nomor 2 (2025) 345–353 E-ISSN 2986-5395

DOI: 10.47467/comit.v3i2.8906

- Konteks yang paling relevan dipilih sebagai input bersama pertanyaan ke model BART.
- Model menghasilkan ringkasan jawaban dan dikembalikan ke pengguna.

Evaluasi

Evaluasi dilakukan menggunakan dua pendekatan:

- Evaluasi otomatis: menggunakan metrik Precision, Recall, dan F1-Score berdasarkan kesamaan jawaban chatbot dan jawaban referensi.
- Evaluasi manual: dilakukan dengan pendekatan Conversation-Based Evaluation (CBE) terhadap tiga aspek: relevansi, koherensi, dan kelancaran bahasa (fluency).

HASIL DAN PEMBAHASAN

Setelah proses pembangunan dan fine-tuning model BART selesai, chatbot diuji menggunakan serangkaian pertanyaan yang telah disusun untuk menguji relevansi, keakuratan, dan keterbacaan respons sistem. Hasil pengujian dikategorikan ke dalam dua pendekatan utama:

Evaluasi otomatis

Menggunakan metrik Precision, Recall, dan F1-Score untuk menilai kualitas jawaban. Rata-rata hasil dari 30 pertanyaan uji menunjukkan:

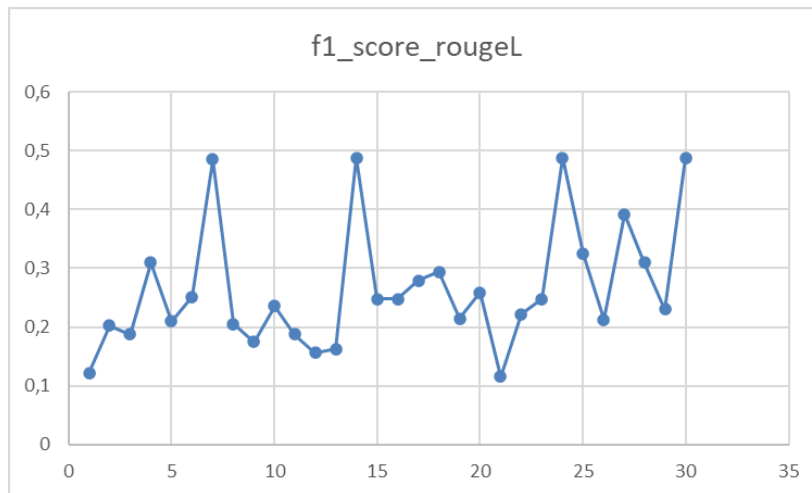
- Precision: 0,978
- Recall: 0,250
- F1-Score: 0,385

Hasil precision yang tinggi mengindikasikan bahwa sebagian besar kata yang muncul dalam jawaban chatbot relevan, namun nilai recall yang rendah menunjukkan masih banyak informasi penting dari referensi yang tidak muncul dalam jawaban chatbot. Hal ini menandakan bahwa sistem masih perlu ditingkatkan dalam menangkap informasi secara menyeluruh dari konteks berita.

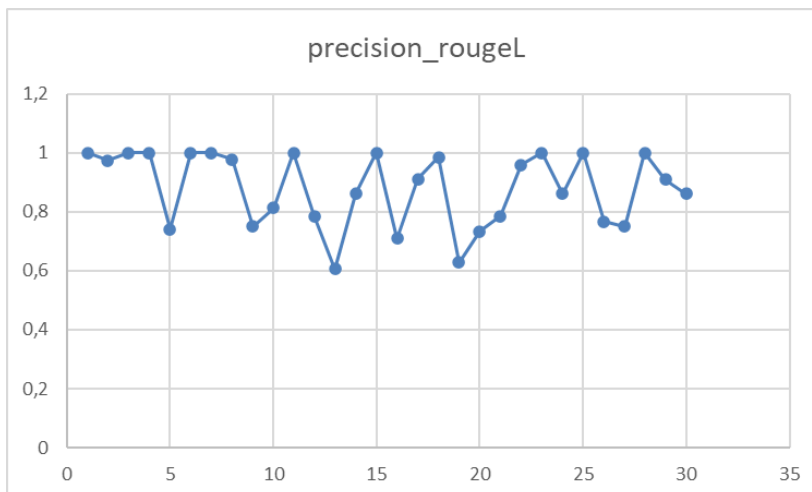
Evaluasi ROUGE-L

Selain metrik dasar, evaluasi juga menggunakan metrik ROUGE-L yang memperhatikan kesamaan urutan kata antara jawaban sistem dan referensi. Hasil evaluasi menunjukkan:

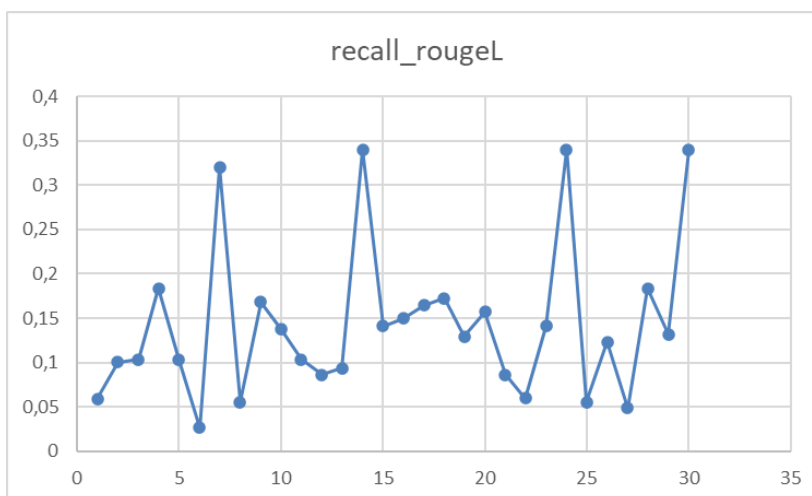
- ROUGE-L Precision: 0,6 – 1,0
- ROUGE-L Recall: 0,05 – 0,35
- ROUGE-L F1: antara 0,1157 hingga 0,4878



Gambar 2 grafik F1-Score



Gambar 3 Grafik Precision



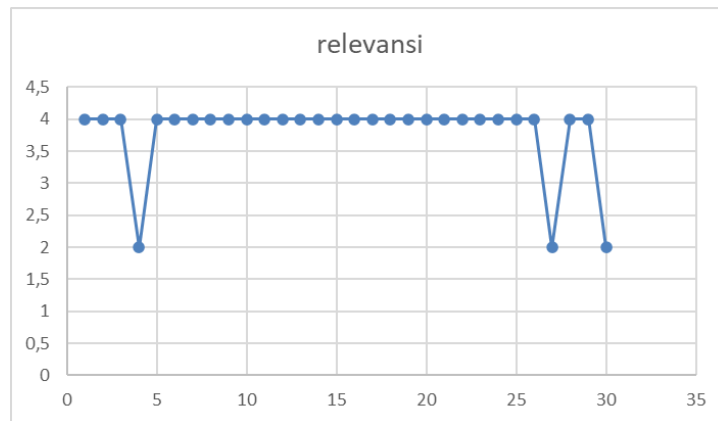
Gambar 4 Grafik Recall

Rentang skor ini mencerminkan bahwa meskipun sistem dapat menghasilkan ringkasan yang tepat, masih terdapat inkonsistensi dalam kelengkapan isi jawaban.

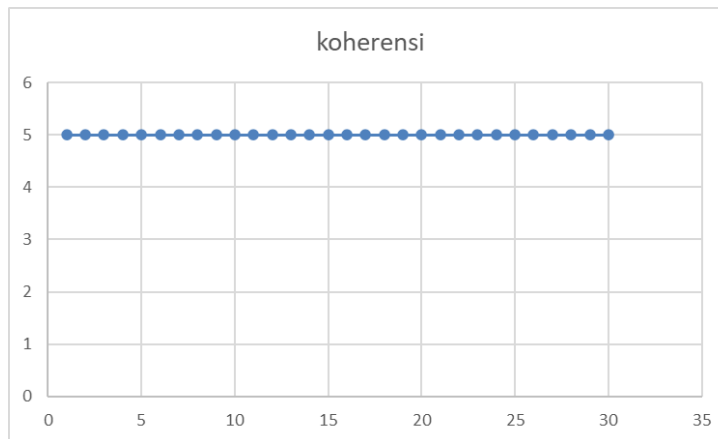
Evaluasi Conversation-Based Evaluation (CBE)

Evaluasi manual dilakukan terhadap 30 pasang pertanyaan-jawaban untuk mengukur:

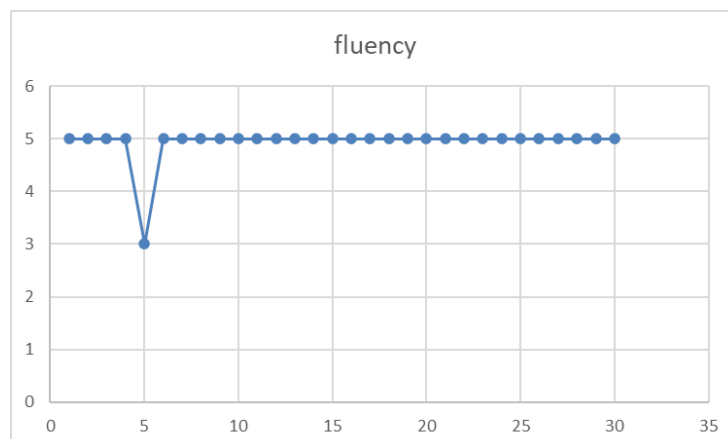
- Relevansi: rata-rata 3,80 dari skala 5
- Koherensi: rata-rata 5,00
- Fluency: rata-rata 4,93



Gambar 5 Grafik Relevansi



Gambar 6 Grafik Kohersi



Gambar 7 Grafik Fluency

Hasil ini menunjukkan bahwa jawaban chatbot umumnya memiliki struktur kalimat yang baik dan mudah dipahami. Namun, relevansi perlu ditingkatkan terutama untuk pertanyaan yang bersifat ambigu atau sangat spesifik.

Evaluasi True Positive, False Positive, False Negative

Pengujian lanjut dilakukan terhadap komponen n-gram jawaban untuk menilai keakuratan berdasarkan jumlah kata:

- TP (True Positive): banyak kata relevan yang dikenali dengan baik
- FP (False Positive): jumlah kecil, menunjukkan sedikit kata tidak relevan
- FN (False Negative): masih tinggi, menunjukkan masih banyak informasi penting belum muncul

	Prediksi: Jawaban Diberikan	Prediksi: Tidak Menjawab	
Relevan	TP = 16	FN = 3	19
Tidak Relevan	FP = 9	TN = 2	11
	25	5	

Tabel 1 Confusion matrix

Nilai TN (True Negative) tidak dihitung karena seluruh konteks merupakan bagian dari isi berita yang dianggap penting.

KESIMPULAN DAN SARAN

Berdasarkan hasil pengembangan dan evaluasi sistem Chatbot yang telah dilakukan, maka diperoleh beberapa kesimpulan sebagai berikut:

1. Chatbot berbasis Flask dengan model BART mampu merangkum berita esports secara relevan sesuai pertanyaan pengguna.

**Comit: Communication, Information and Technology
Journal**

Volume 3 Nomor 2 (2025) 345–353 E-ISSN 2986-5395

DOI: 10.47467/comit.v3i2.8906

2. Hasil metrik F1, precision, dan recall menunjukkan performa cukup baik, namun nilai recall yang rendah menandakan belum semua informasi penting tertangkap, terutama pada berita panjang.
3. Berdasarkan penilaian aspek relevansi, koherensi, dan kelancaran, respons sistem umumnya informatif dan mudah dipahami.
4. Model masih kesulitan menjawab pertanyaan terlalu umum atau spesifik akibat keterbatasan ukuran model dan panjang input. Proses pencarian berita juga masih perlu ditingkatkan.

Berdasarkan temuan dan kendala yang dihadapi dalam penelitian ini, beberapa saran yang dapat diberikan adalah:

1. Melakukan fine-tuning lanjutan pada model BART dengan dataset yang lebih beragam untuk meningkatkan kualitas ringkasan.
2. Menggunakan teknik retrieval yang lebih baik dan memperbarui koleksi berita secara rutin agar model mendapat konteks yang lebih relevan dan mutakhir.

**Comit: Communication, Information and Technology
Journal**

Volume 3 Nomor 2 (2025) 345–353 E-ISSN 2986-5395

DOI: 10.47467/comit.v3i2.8906

DAFTAR PUSTAKA

- Heinzerling, B., et al. (2021). Evaluating the Efficacy of Summarization Evaluation across Languages, arXiv preprint arXiv:2106.01478. [7]
- Firdaus, P. Y. S. S. V. A. H. (2020). Intelligence *Chatbot* for Indonesian Law on Electronic. [2]
- Lewis, M., et al. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension, arXiv preprint arXiv:1910.13461. [5]
- Qiang, Z., K. Taylor., and W. Wang. (2024). How Does A Text Preprocessing Pipeline Affect Ontology Syntactic Matching?, arXiv preprint arXiv:2411.03962. [6]
- Qotrunnida, E. S. d. R. N. A. Nabila. (2023). Penggunaan *Chatbot* Mela terhadap Peningkatan Kemampuan Kosakata Bahasa Indonesia Anak. pp. 448-459. [1]
- Santoso, I. R. N. K. M. R. A. A. S. Y. Pandu Dewonoto Laut. (2021). Penerapan Artificial Intelligence dalam Aplikasi *Chatbot* sebagai Media Informasi dan Pembelajaran mengenai Kebudayaan Bangsa. [4]
- Santoso, J. S. A. S. Gerald. (2023). Development of OpenAI API-Based *Chatbot* to Improve User Interaction on the JBMS *Website*. [3]
- Xu, Q., M. Zhang., X. Wu., et al. (2022). A survey of transformers. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9), 6449–6469. doi: 10.1109/TNNLS.2022.3143861. [8]