

Implementasi Metode *Pattern-Based Natural Language Processing* (NLP) Pada *Chatbot*

¹Tri Astuti Rahayu, ²Agung Kharisma Hidayah, ³Hary Witriyono, ⁴Khairullah, ⁵RG Guntur Alam

^{1,2,3,4} Teknik Informatika, Universitas Muhammadiyah Bengkulu, Indonesia

⁵ Sistem Informasi, Universitas Muhammadiyah Bengkulu

tri.astutirahayu02@gmail.com, kharisma@umb.ac.id, harrywitriyono@umb.ac.id, khairullah@umb.ac.id,
datuak73@yahoo.com;

Article Info

Article history:

Received, 2024-07-24

Revised, 2024-10-10

Accepted, 2024-11-06

Kata kunci:

NLP,
Chatbot,
BLEU.

Keywords:

NLP,
Chatbot,
BLEU.

ABSTRAK

Chatbot berbasis *Natural Language Processing* (NLP) semakin banyak digunakan dalam berbagai aplikasi, mulai dari layanan pelanggan hingga pendidikan. Salah satu pendekatan yang sederhana namun efektif untuk mengimplementasikan NLP pada *chatbot* adalah metode *pattern-based*, yang menggunakan pola bahasa untuk memahami dan merespons masukan pengguna. Penelitian ini bertujuan untuk mengimplementasikan metode *pattern-based* NLP pada *chatbot* dan mengevaluasi kinerjanya menggunakan metrik BLEU. Dalam prosesnya, data percakapan dikumpulkan dan dianalisis untuk mengidentifikasi pola umum yang kemudian dipetakan menjadi *respons* yang sesuai. Pengujian dilakukan dengan mengukur tingkat kecocokan bigram antara *respons chatbot* dan *respons referensi*, di mana *precision BLEU-2* mencapai nilai 0.75 atau 75%. Hasil ini menunjukkan bahwa metode *pattern-based* mampu menghasilkan *respons* yang relevan, namun masih memerlukan penyempurnaan untuk mencapai akurasi yang lebih tinggi.

ABSTRACT

Natural Language Processing (NLP) based chatbots are increasingly used in various applications, ranging from customer service to education. One simple yet effective approach to implementing NLP on a chatbot is the pattern-based method, which uses language patterns to understand and respond to user input. This research aims to implement the pattern-based NLP method on a chatbot and evaluate its performance using the BLEU metric. Conversation data is collected and analyzed to identify common patterns, which are then mapped into appropriate responses. Testing was done by measuring the level of bigram match between the chatbot response and the reference response, where the BLEU-2 precision reached a value of 0.75 or 75%. These results show that the pattern-based method is capable of generating relevant responses but still requires refinement to achieve higher accuracy.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Tri Astuti Rahayu,
Program Studi Teknik Informatika,
Universitas Muhammadiyah Bengkulu, Indonesia
Email: tri.astutirahayu02@gmail.com

1. PENDAHULUAN

Chatbot merupakan salah satu teknologi *artificial intelligence* berbasis teks yang mampu meniru cara berbicara layaknya manusia sebagai mitra percakapan [1] [2]. Sehingga diperlukan perantara yang menghubungkan antara pengguna dan penyedia layanan. Teknologi informasi menjadi salah satu teknologi yang sangat berkembang pada saat ini [3]. Pelayanan informasi pada Penerimaan Mahasiswa Baru di Universitas Muhammadiyah Bengkulu masih menggunakan cara yang manual dan belum maksimal dalam menggunakan teknologi untuk memberikan pelayanan informasi yang mudah dan cepat kepada calon Mahasiswa Baru.

Saat ini, sistem pelayanan masih terbatas oleh jam kerja dan keterbatasan sumber daya manusia, menyebabkan respons terhadap pertanyaan calon mahasiswa baru menjadi lambat. Untuk mengatasi masalah ini, diperlukan solusi berbasis teknologi seperti *chatbot* yang dapat memberikan respons cepat dan akurat.

Berikut beberapa penelitian terdahulu yang menjadi pertimbangan dalam pemilihan judul ini yaitu pada penelitian Shita dan Lauw Li Hin melakukan penelitian yang berjudul “Aplikasi *Chatterbot* Memanfaatkan *Algoritma Boyer Moore* pada pendaftaran siswa baru SMP” dari hasil penelitian dengan melakukan pembuatan aplikasi *chatbot* dengan *Algoritma Boyer Moore* [4].

Kemudian pada penelitian Nuzul Hikmah, Dyah Ariyanti dan Ferry Agus Pratama melakukan penelitian yang berjudul “Implementasi *Chatbot* sebagai *Virtual Assistant* di Universitas Panca Marga Probolinggo menggunakan Metode TF-IDF” hasil penelitiannya adalah *Chatbot* dimanfaatkan sebagai *Virtual Assistant* untuk menggantikan *Customer Service* sebagai sebuah layanan informasi dan komunikasi [5].

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan yang berfokus pada pemahaman dan pemrosesan bahasa alami oleh computer [6], [7]. Salah satu pendekatan dalam NLP adalah metode *pattern-based*, yang menggunakan pola-pola linguistik untuk memahami struktur kalimat dan konteks percakapan [8]–[10]. Metode ini mengandalkan aturan yang dirancang untuk mengenali kata-kata atau frasa tertentu yang biasanya muncul dalam konteks tertentu, serta memetakan respons yang sesuai berdasarkan pola tersebut [11]–[13].

Pada *chatbot*, metode *pattern-based* dapat digunakan untuk mengidentifikasi niat pengguna (*intent*) dan memproses teks masukan untuk menghasilkan respons yang relevan [14]. Implementasi metode ini umumnya memerlukan definisi pola secara manual yang melibatkan identifikasi frasa kunci dan pemetaan terhadap berbagai respons yang telah disiapkan.

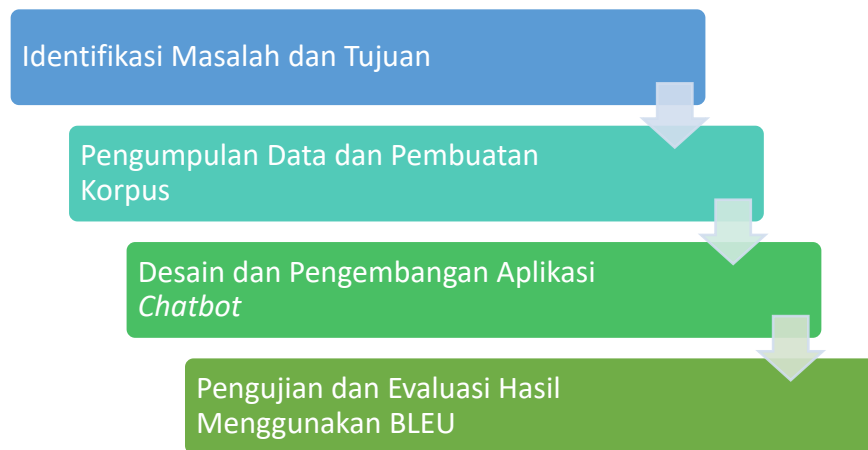
Keunggulan dari metode *pattern-based* adalah kesederhanaan dan efisiensinya, terutama untuk aplikasi *chatbot* yang ditujukan pada tugas-tugas sederhana seperti menjawab pertanyaan umum atau memberikan informasi yang terstruktur. Namun, metode ini memiliki keterbatasan dalam hal fleksibilitas dan kurang efektif jika diaplikasikan pada percakapan yang kompleks atau dinamis.

Tujuan dari implementasi metode *pattern-based NLP* pada *chatbot* adalah untuk meningkatkan akurasi dan relevansi respons yang dihasilkan berdasarkan pola kalimat dari pengguna. Dengan demikian, *chatbot* dapat memberikan pengalaman yang lebih interaktif dan responsif sesuai dengan konteks pertanyaan atau pernyataan yang diberikan pengguna.

Untuk memberikan tingkat keberhasilan yang akurat dalam penggunaan metode *pattern-based NLP*, akan dilakukan pengukuran menggunakan metode *Bilingual Evaluation Understudy* (BLEU) dimana perhitungan menggunakan BLEU Digunakan pada tugas terjemahan mesin atau tugas-tugas yang membutuhkan evaluasi kualitas teks yang dihasilkan, seperti *summarization* [8], [9]. BLEU mengukur seberapa mirip hasil terjemahan model dengan terjemahan referensi (*ideal*).

2. METODE PENELITIAN

Pada penelitian ini akan mengikuti alur penelitian sesuai dengan yang telah dilakukan oleh peneliti, dimana alur penelitian tersebut dapat dilihat pada gambar bagan dibawah ini:



Gambar 1. Alur Penelitian

Identifikasi Masalah dan Tujuan Penelitian

- Langkah pertama adalah mengidentifikasi masalah yang ingin diselesaikan oleh *chatbot* berbasis *pattern-based* NLP. Misalnya, *chatbot* ini dapat digunakan untuk menjawab pertanyaan umum atau memberikan informasi kepada pengguna dalam suatu bidang tertentu, seperti layanan pelanggan atau pendidikan.
- Tentukan tujuan spesifik dari *chatbot* ini, seperti meningkatkan akurasi dalam memahami pertanyaan pengguna atau mempercepat respons.

Pengumpulan Data dan Pembuatan Korpus

- Kumpulkan data percakapan yang relevan untuk mendukung pengembangan pola *pattern-based*. Data ini bisa berupa log percakapan dari sistem yang sudah ada atau dikumpulkan dari survei dan wawancara dengan pengguna.
- Setelah data terkumpul, buat korpus atau kumpulan teks yang berisi berbagai jenis pertanyaan, pernyataan, dan respons ideal dari *chatbot*. Korpus ini akan digunakan untuk mengidentifikasi pola bahasa yang sering digunakan.

Desain dan Pengembangan *Chatbot* Berbasis Pola

Bangun sistem *chatbot* menggunakan aturan *pattern-based* yang memetakan pola-pola bahasa yang diidentifikasi ke respons yang sesuai. Gunakan teknik seperti pencocokan pola (*pattern matching*) atau ekspresi reguler (*regular expression*) untuk mengenali pola-pola dalam teks masukan pengguna.

Jika diperlukan, implementasikan algoritma *rule-based* untuk memproses pola-pola yang lebih kompleks, seperti pola berurutan atau kombinasi dari beberapa kata kunci.

Pengujian dan Evaluasi *Chatbot*

Uji *chatbot* dengan data percakapan baru yang belum digunakan dalam tahap pengembangan. Data ini dapat berasal dari simulasi percakapan atau dari pengguna yang sebenarnya. Pada uji ini nantinya akan digunakan metode *Bilingual Evaluation Understudy* (BLEU) dimana perhitungan BLEU digunakan pada tugas terjemahan mesin atau tugas-tugas yang membutuhkan evaluasi kualitas teks yang dihasilkan, seperti summarization.

3. HASIL DAN ANALISIS

Chatbot dirancang untuk memberikan layanan cepat dan akurat kepada generasi milenial [15]. Pengembangan dilakukan dengan pendekatan berbasis pola (*Pattern-Based*), menggunakan Bahasa pemrograman *Python* dan *Framework Flask*. *Python* memudahkan pembacaan kode dengan sintaks sederhana [16], sementara *Flask*

adalah *micro-framework* yang mendukung ekstensi untuk fitur tambahan. *Chatbot* memahami dan merespons permintaan pengguna berdasarkan pola yang telah ditentukan, memberikan pengalaman interaksi yang terstruktur dan andal melalui antarmuka web responsif.

Uraian implementasi pada *pettern based* ini yaitu:

1. **Inisialisasi Aplikasi *Flask***
 Pada tahap ini, aplikasi *flask* diinisialisasi dengan menggunakan fungsi `Flask(__name__)` yang kemudian disimpan dalam variable `app`.
2. **Routing dan Pengolahan Masukan Pengguna**
 Aplikasi mendefinisikan dua rute, satu untuk halaman utama (`('/')`) dan satu lagi untuk menerima masukan pengguna (`('/get_response')`). Ketika pengguna mengirimkan masukan melalui formular, fungsi `get_response()` akan dipanggil untuk memprosesnya.
3. ***Case Folding***
Case Folding merupakan langkah merubah teks ke huruf kecil atau besar. Tujuannya untuk memudahkan pencarian karena tidak semua dokumen menggunakan huruf kapital secara konsisten. Dengan demikian, *case folding* bertugas untuk mengubah seluruh teks menjadi huruf kecil atau *lowercase* [17]. Tahapan dilakukan pada saat pengguna mengirimkan teks. Masukan pengguna diubah menjadi huruf kecil menggunakan metode `.lower()`. Hal ini dilakukan untuk memastikan bahwa pola percakapan yang telah ditentukan dapat dipidankan dengan masukan pengguna tanpa memperhatikan apakah hurufnya besar atau kecil.
4. **Pemberian Respons**
 Setelah masukan pengguna diproses dan diubah menjadi huruf kecil, *Chatbot* memberikan respons dengan menggunakan fungsi `.respond()` dari objek *Chatbot* yang telah diinisialisasi. Respons ini kemudian dikirimkan kembali ke pengguna melalui aplikasi web.
5. **Menjalankan Aplikasi**
 Aplikasi *Flask* dijalankan dengan menggunakan metode `app.run(debug=True)`, yang membuatnya berjalan dalam metode debug sehingga mudah untuk melihat pesan kesalahan dan debug saat pengembangan.

Uraian implementasi pada NLP yaitu:

1. ***Case Folding***
case Folding merupakan proses merubah teks yang diberikan oleh pengguna menjadi huruf kecil. Ini dilakukan untuk memastikan bahwa analisis teks tidak terpengaruh oleh perbedaan huruf besar dan kecil, sehingga "Mahasiswa" dan "mahasiswa" dianggap sama.
2. ***Tokenizing***
 Teks harus dipecah menjadi token agar dapat dianalisis lebih lanjut. Token adalah unit dasar dari teks, yang dipisahkan berdasarkan spasi atau tanda baca. Misalnya, kalimat "Penerimaan mahasiswa baru" akan dipecah menjadi token "Penerimaan", "mahasiswa", dan "baru".
3. ***Stemming***
Stemming yaitu menghilangkan kalimat tambahan agar menjadi kata dasar. *Stemming* bertujuan untuk mengurangi kata-kata ke bentuk dasarnya agar varian dari kata yang sama dapat diidentifikasi sebagai satu kata dasar. Contohnya, kata "mendaftar", "mendaftarkan", dan "daftarkan" semuanya akan diubah menjadi kata dasar "daftar".

Aplikasi ini berfungsi seperti percakapan biasa antara pengguna, tetapi yang membalas pesan adalah sistem yang telah diprogram untuk merespons dengan tepat sesuai dengan topik yang di bicarakan oleh pengguna.

a. Contoh Perumusan Kata Kunci dan Respon

Beberapa pertanyaan umum yang sering diajukan telah dikumpulkan dan dirumuskan menjadi kata kunci beserta responnya.

Tabel 1. Contoh perumusan kata kunci dan respon

Kata Kunci	Respon
(.*) proses pendaftaran	Online dan bisa langsung datang ke PMB
(.*) cara mendaftar	
(.*) gimana daftar	
(.*)((proses) s*(.*?))s*(pendaftaran)	
(.*)((cara) s*(.*?))s*(mendaftar)	

(.*) (gimana) \s* (.*?) \s* (daftar)	
(.*) biaya pendaftaran	Biaya
(.*) (biaya) \s* (.*?) \s* (pendaftaran)	
(.*) syarat pendaftaran	(1) Fotocopy ijazah yang sudah di legalisir
(.*) di bawa mendaftar	(2) Fotocopy kartu keluarga
(.*) (syarat) \s* (.*?) \s* (pendaftaran)	(3) Pas foto 3x4 (4 lembar)
(.*) (di bawa) \s* (.*?) \s* (mendaftar)	

b. Respon Dengan Pertanyaan Yang Tidak Dimengerti

Respon dengan pertanyaan yang tidak dimengerti adalah *chat* yang tidak memberikan respon, misalnya chat yang penulisannya kurang tepat.

Tabel 2. Pertanyaan yang tidak dimengerti

Pertanyaan/Pattern-Based	Respon
.	Maaf, saya tidak mengerti pertanyaan anda. Silahkan pilih tema apa yang ingin anda tanyakan (1) Cara pendaftaran (2) Biaya pendaftaran (3) Persyaratan pendaftaran (4) Biaya UKT (5) Tempat pendaftaran (6) Program studi yang tersedia (7) Waktu pendaftaran

Hasil dari *Chatbot* respon dengan pertanyaan yang tidak dimengerti atau kesalahan pada penulisan seperti gambar dibawah ini.



Gambar 3 Respon dengan pertanyaan yang tidak dimengerti

c. Pertanyaan/Pernyataan dan Respon

Pertanyaan atau pernyataan yang diinputkan oleh pengguna bertujuan untuk mendapatkan informasi, dan setelah itu, *Chatbot* memberikan respon yang sesuai. Dapat dilihat seperti gambar berikut.



Gambar 4 Pertanyaan dan respon

Pengujian BLEU

Perhitungan BLEU dilakukan dengan menghitung tingkat kecocokan n-gram antara respons *chatbot* dan *respons* referensi. Mari kita hitung BLEU-2 (bigram) sebagai contoh:

Bigram Respons Chatbot:

" proses pendaftaran ", " cara mendaftar ", " gimana daftar ", "pendaftarannya seperti apad", "pendaftaran untuk", "untuk pendaftaran", "meniru pendaftaran", "pendaftaran".

Bigram Respons Referensi:

"online", "bisa datang", "datang langsung", "secara online", "datang kekampus", "silakan online", "silakan datang", "langsung online", "datang saja".

Tingkat Kecocokan Bigram:

Terdapat 6 bigram yang cocok dari total 8 bigram pada respons chatbot. Dengan demikian, precision BLEU-2 untuk bigram adalah

$$\text{Precision BLEU} - 2 = \frac{\text{Bigram yang Cocok}}{\text{Total bigram dalam Respons Chat}} = \frac{6}{8} = 0,75$$

Dengan demikian, precision BLEU-2 adalah 0.75 atau 75%. Ini berarti bahwa 75% dari bigram dalam respons chatbot cocok dengan bigram dalam respons referensi. Hasil ini menunjukkan bahwa respons chatbot memiliki kesamaan yang cukup besar dengan respons referensi pada tingkat bigram, meskipun ada ruang untuk peningkatan.

4. KESIMPULAN

Metode *pattern-based Natural Language Processing* (NLP) pada chatbot dapat memberikan respons yang relevan dan akurat dengan mengenali pola tertentu dalam teks masukan pengguna. Berdasarkan perhitungan BLEU-2 pada hasil dan pembahasan dengan nilai 75% dari bigram respons *chatbot* cocok dengan bigram dalam respons referensi. Hasil ini menunjukkan tingkat kesesuaian yang baik antara respons chatbot dan respons ideal, meskipun masih ada ruang untuk peningkatan agar *chatbot* lebih responsif terhadap pola yang lebih kompleks.

UCAPAN TERIMA KASIH

Terimakasih kepada orang tua dan keluarga yang selalu memberikan doa, dukungan dan semangat kepada penulis serta seluruh pihak yang telah berkontribusi dalam proses pembuatan artikel ini.

REFERENSI

- [1] G. Guntoro, Loneli Costaner, and L. Lisnawita, "Aplikasi Chatbot untuk Layanan Informasi dan Akademik Kampus Berbasis Artificial Intelligence Markup Language (AIML)," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 11, no. 2, pp. 291-300, 2020, doi: 10.31849/digitalzone.v11i2.5049.
- [2] M. A. Nugroho, A. Damayanti, M. F. Rifai, and S. Windarti, "PENGEMBANGAN APLIKASI QnA UNTUK PENDAFTARAN MAHASISWA BARU STMIK AKAKOM," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 18-23, 2021, doi: 10.24076/joism.2021v3i1.408.
- [3] M. R. Suherlan¹, Asriyanik², and Agung Pambudi³, "UMMIBOT SEBAGAI MEDIA LAYANAN INFORMASI PENERIMAAN MAHASISWA BARU UNIVERSITAS MUHAMMADIYAH SUKABUMI," *J. Inform. Terpadu*, vol. 6, no. 1, pp. 20-28, 2020.
- [4] R. T. Shita and L. L. Hin, "Aplikasi Chatterbot Memanfaatkan Algoritma Boyer Moore Pada Pendaftaran Siswa Baru," *J. BIT, Issn 1693-9166*, vol. 14, no. 2, pp. 30-37, 2017.
- [5] Nuzul Hikmah, Dyah Ariyanti, and Ferry Agus Pratama, "Implementasi Chatbot Sebagai Virtual Assistant di Universitas Panca Marga Probolinggo menggunakan Metode TF-IDF," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 2, pp. 133-148, 2022, doi: 10.35746/jtim.v4i2.225.

- [6] O. Sen, M. Fuad, M. N. Islam, J. Rabbi, M. K. Hasan, and ..., "Bangla natural language processing: A comprehensive review of classical machine learning and deep learning based methods," *CoRR*. academia.edu, 2021. [Online]. Available: <https://www.academia.edu/download/78447185/2105.14875v2.pdf>
- [7] X. Zhu, M. Zhang, Y. Hong, and R. He, *Natural Language Processing and Chinese Computing: 9th CCF International Conference, NLPCC 2020, Zhengzhou, China, October 14–18, 2020* books.google.com, 2020. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=XUQBEEAAQBAJ&oi=fnd&pg=PR5&dq=%22pattern+based%22+chatbot+bleu&ots=qYEO5js7_Z&sig=35ytoadc079Q1Hm9qZe4S23B0U
- [8] M. Maryamah, M. M. Irfani, E. B. T. Raharjo, and ..., "Chatbots in academia: a retrieval-augmented generation approach for improved efficient information access," ... *Knowl. ...*, 2024, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10499652/>
- [9] F. Khennouche, Y. Elmir, Y. Himeur, N. Djebari, and ..., "Revolutionizing generative pre-trained: Insights and challenges in deploying ChatGPT and generative chatbots for FAQs," *Expert Syst. with ...*, 2024, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417424000897>
- [10] P. Safari and M. Shamsfard, "Data augmentation and preparation process of PerInfEx: a Persian chatbot with the ability of information extraction," *IEEE Access*, 2024, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10418595/>
- [11] J. Tang, Q. Zhang, Y. Li, and J. Li, "Grapharena: Benchmarking large language models on graph computational problems," *arXiv Prepr. arXiv2407.00379*, 2024, [Online]. Available: <https://arxiv.org/abs/2407.00379>
- [12] S. Tahir, S. A. Shah, and J. Abu-Khalaf, "Artificial Empathy Classification: A Survey of Deep Learning Techniques, Datasets, and Evaluation Scales," *arXiv Prepr. arXiv2310.00010*, 2023, [Online]. Available: <https://arxiv.org/abs/2310.00010>
- [13] S. V Bhalshankar and R. R. Deshmukh, "CAAS (Chatbot as a Service): Sector-Wise Survey," *Int. Conf. ...*, 2023, [Online]. Available: <https://www.atlantis-press.com/proceedings/icamida-22/125986282>
- [14] M. Dymetman, W. F. Aziz, N. Cancedda, and ..., "Dynamic bi-phrases for statistical machine translation," *US Pat. ...*, 2017, [Online]. Available: <https://patents.google.com/patent/US9552355B2/en>
- [15] M. Murhadi, "Rancang Bangun Aplikasi Chatbot Sebagai Bentuk Pelayanan Prima Untuk Penerimaan Mahasiswa Baru," *INTEK J. Inform. dan Teknol. Inf.*, vol. 2, no. 1, pp. 10–16, 2019, doi: 10.37729/intek.v2i1.84.
- [16] H. A. Kharisma, A. R. Guntur, and P. Prihandoko, *STRUKTUR DATA DENGAN PYTHON*. Jambi: PT. Sonpedia Publishing Indonesia, 2023.
- [17] D. Alita and A. R. Isnain, "Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier," *J. Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.