

BAB 5 PENUTUP

5.1. Kesimpulan

Penelitian ini bertujuan untuk merancang dan mengimplementasikan chatbot berbasis AI menggunakan framework Rasa pada website PT Orbit Eka Semesta untuk meningkatkan pengalaman pengguna dalam mengakses informasi layanan. Berdasarkan hasil pengembangan dan pengujian yang telah dilakukan, dapat disimpulkan bahwa:

1. Terkait proses pengembangan fitur *chatbot* dalam menangani berbagai pertanyaan, *chatbot* berbasis AI ini dikembangkan menggunakan framework *Rasa Open Source*. Fitur-fitur yang disediakan meliputi intent untuk pertanyaan umum, respons otomatis, fallback, serta interaksi berbasis pilihan intent. Seluruh fitur tersebut dirancang berdasarkan hasil wawancara serta analisis data riwayat percakapan antara customer service dan pelanggan. Dataset yang digunakan untuk *Natural Language Understanding* (NLU) terdiri dari 507 contoh *examples* yang terbagi ke dalam 21 kelas *intent*. Selain itu, terdapat 20 kelas *stories* yang masing-masing merepresentasikan satu skenario dialog, di mana setiap *story* terdiri dari 2 hingga 4 langkah *steps*. Untuk mendukung alur percakapan yang lebih fleksibel, sistem juga dilengkapi dengan 21 *rules* yang telah ditentukan.

Seluruh konfigurasi seperti *intent*, *stories*, *domain*, *rules*, dan *config* diproses di sisi *backend* dan di deploy menggunakan server lokal yang dibuka ke publik dengan menggunakan *Ngrok*, serta diintegrasikan dengan website via widget berbasis JavaScript. Dengan demikian, sistem ini tidak hanya berfokus pada tampilan, tetapi juga mencakup backend yang mengelola logika percakapan secara menyeluruh. Selain itu pelatihan model chatbot juga dilakukan dari sisi backend.

2. Berkaitan dengan metode evaluasi yang dilakukan terhadap chatbot, pengujian Confusion Matrix untuk performa *Natural Language Understanding* (NLU) menunjukkan bahwa DIETClassifier memiliki performa lebih unggul dibandingkan SklearnIntentClassifier. DIETClassifier mencatatkan precision sebesar 92,4% (macro) dan 92,8% (weighted), serta F1-score sebesar 93,5% (macro) dan 94,0% (weighted). Sementara itu, SklearnIntentClassifier memperoleh precision 92,0% (macro) dan 92,2% (weighted), serta F1-score 93,3% (macro) dan 93,8% (weighted). Keunggulan DIETClassifier didukung oleh penggunaan pendekatan neural network yang mampu mengombinasikan fitur sparse dan dense secara bersamaan. Meskipun dalam pengujian ini hanya fitur sparse yang digunakan, model tetap mampu membangun representasi data yang

efektif dan mengenali pola bahasa dengan baik. Sebaliknya, SklearnIntentClassifier yang berbasis Linear SVM dengan optimasi grid search hanya mengandalkan fitur dense, sehingga kemampuannya lebih terbatas. Secara keseluruhan, DIETClassifier menawarkan optimasi yang lebih baik, generalisasi yang lebih kuat, dan dukungan pengembangan fitur yang lebih aktif.

Pengujian Confusion Matrix untuk performa dialog menggunakan perintah `test_core` menunjukkan bahwa kedua model DIETClassifier dan SklearnIntentClassifier berhasil mencapai akurasi sempurna sebesar 100% pada semua metrik evaluasi (precision, recall, dan F1-score) baik pada macro, micro, maupun weighted average. Hasil ini dicapai karena intent dapat diklasifikasikan dengan tepat dan skenario dialog yang digunakan bersifat sederhana sehingga jalur percakapan mudah diprediksi. Dengan demikian, performa dialog yang identik ini menunjukkan bahwa kedua algoritma mampu berfungsi optimal dalam kondisi pengujian yang konsisten.

Pengujian *Black Box Testing* dilakukan terhadap 22 fitur utama chatbot untuk memastikan seluruh fungsi berjalan sesuai spesifikasi. Hasil pengujian menunjukkan bahwa semua fitur berhasil dijalankan dengan tingkat keberhasilan 100%, tanpa error maupun malfungsi. Hal ini membuktikan bahwa sistem chatbot telah berfungsi secara optimal dalam merespons pertanyaan dan menjalankan tugas sesuai skenario yang ditentukan.

Selanjutnya, *User Acceptance Testing* (UAT) dilakukan terhadap 29 responden menggunakan skala likert. Hasil evaluasi menunjukkan bahwa chatbot memperoleh nilai rata-rata keseluruhan 4,2 dari 5,0, yang termasuk dalam kategori baik. Penilaian per indikator menunjukkan hasil positif, yaitu: kemudahan penggunaan (4,34), ketepatan respons (4,21), kemudahan memahami informasi (4,31), kemampuan menyarankan kontak ke CS (4,03), waktu respons (4,21), akurasi informasi (4,14), kemudahan navigasi percakapan (4,28), dan keinginan menggunakan kembali (4,28). Hasil ini menunjukkan bahwa chatbot telah memenuhi ekspektasi pengguna dari segi kegunaan, kecepatan, ketepatan informasi, dan kepuasan pengalaman. Dengan demikian, sistem yang dikembangkan dinilai efektif dalam memberikan layanan informasi yang efisien, responsif, dan meningkatkan kualitas interaksi antara perusahaan dan pengguna.

5.2. Saran

Berdasarkan hasil penelitian dan implementasi yang telah dilakukan, berikut beberapa saran yang dapat dipertimbangkan untuk pengembangan selanjutnya:

1. Chatbot dapat terus dikembangkan dengan menambahkan lebih banyak intent dan variasi frasa pengguna agar dapat menangani pertanyaan yang lebih kompleks dan bervariasi.
2. Mengembangkan integrasi chatbot dengan sistem database internal perusahaan agar chatbot dapat memberikan informasi secara dinamis dan real-time.
3. Model chatbot perlu dilakukan pelatihan ulang secara berkala menggunakan data terbaru dari interaksi pengguna untuk meningkatkan akurasi dan relevansi respons.
4. Untuk memperluas cakupan pengguna, disarankan agar chatbot dapat memahami dan merespons dalam lebih dari satu bahasa, terutama bahasa Inggris sebagai bahasa internasional.