

Pengembangan Chatbot Berbasis Large Language Model (LLM) pada Platform E-Commerce

1st Yoga Raditya Nugraha Sukma
Pradana
Prodi Sistem Informasi
Fakultas Rekayasa Industri
Bandung, Indonesia
yogaraditya@telkomuniversity.ac.id

2nd Nur Ichsan Utama
Prodi Sistem Informasi
Fakultas Rekayasa Industri
Bandung, Indonesia
nichsan@telkomuniversity.ac.id

3rd Sinung Suakanto
Prodi Sistem Informasi
Fakultas Rekayasa Industri
Bandung, Indonesia
sinung@telkomuniversity.ac.id

Abstrak—Perkembangan teknologi digital di Indonesia yang pesat telah mendorong adopsi sistem perdagangan elektronik pada berbagai institusi, termasuk platform e-commerce internal seperti TokoPoin di Koperasi Telkom University. Namun, implementasi sistem komunikasi pelanggan pada platform tersebut masih menghadapi tantangan signifikan, yaitu inefisiensi pengelolaan komunikasi customer service akibat keterbatasan akses penjual terhadap platform komunikasi dan pola komunikasi pelanggan yang tidak efisien dengan pertanyaan-pertanyaan repetitif. Penelitian ini bertujuan mengembangkan sistem chatbot berbasis Large Language Model (LLM) yang terintegrasi dengan WhatsApp untuk mengatasi masalah tersebut. Metode penelitian menggunakan pendekatan kuantitatif dengan implementasi model llama3.2:3b yang diintegrasikan melalui multi-channel antara platform WhatsApp dan antarmuka web. Hasil penelitian menunjukkan bahwa sistem berhasil mencapai akurasi pengenalan intent sebesar 93,75% dengan rata-rata waktu respons 8,4 detik. Integrasi multi-channel berfungsi optimal dengan pengiriman pesan real-time rata-rata 2,412 detik, berada di bawah standar 3 detik. Kualitas jawaban chatbot mencapai skor rata-rata 4,10 dari skala 5. Pengujian usability menggunakan System Usability Scale (SUS) terhadap 36 responden menghasilkan skor rata-rata 77,85, melampaui ambang batas standar 68 dan masuk kategori usability baik. Penelitian ini membuktikan bahwa implementasi chatbot berbasis LLM dapat secara efektif mengurangi beban komunikasi repetitif penjual, meningkatkan efisiensi operasional, dan memberikan pengalaman komunikasi yang responsif dalam konteks e-commerce internal.

Kata kunci: chatbot, customer service, e-commerce, Large Language Model, multi-channel integration, TokoPoin, WhatsApp

I. PENDAHULUAN

Perkembangan teknologi digital di Indonesia mengalami kemajuan pesat, termasuk dalam adopsi sistem perdagangan elektronik di institusi pendidikan. TokoPoin sebagai platform e-commerce internal Koperasi Telkom University merupakan inovasi penerapan teknologi digital untuk memfasilitasi transaksi antar anggota komunitas kampus.

Implementasi sistem komunikasi pelanggan pada platform e-commerce seperti TokoPoin menghadapi tantangan operasional signifikan. Masalah utama adalah inefisiensi pengelolaan komunikasi customer service, dimana penjual kesulitan mengelola dan merespons pertanyaan pelanggan secara efektif dan tepat waktu. Keterbatasan akses penjual terhadap platform komunikasi berbasis web menyebabkan keterlambatan respons yang berpengaruh pada kepuasan pelanggan dan reputasi penjual.

Pola komunikasi pelanggan yang tidak efisien juga menjadi tantangan tersendiri. Pembeli sering mengajukan pertanyaan dasar seperti ketersediaan stok dan spesifikasi

produk yang sebenarnya telah tersedia dalam deskripsi produk. Hal ini menciptakan beban komunikasi tidak perlu dan mengganggu fokus penjual dari aktivitas strategis.

Penelitian Nugraha dkk. [1] menunjukkan bahwa responsivitas chatbot AI berperan signifikan dalam membentuk kepuasan pelanggan melalui pengaruhnya terhadap *extrinsic value*, *intrinsic value*, dan *online convenience*. Jupri & Fasa [2] membuktikan bahwa implementasi chatbot pada Shopee Indonesia meningkatkan kepuasan pelanggan dan efisiensi layanan. Namun, implementasi chatbot existing masih fokus pada layanan customer service umum untuk pembeli, sehingga masih ada celah pengembangan sistem yang membantu penjual mengelola komunikasi pelanggan.

Pengembangan chatbot berbasis LLM menjadi solusi relevan untuk mengatasi inefisiensi komunikasi customer service. Chatbot ini dirancang menjawab pertanyaan umum pelanggan berdasarkan informasi produk tersedia, memungkinkan penjual fokus pada aktivitas strategis. Teknologi LLM memungkinkan sistem memahami konteks percakapan dan memberikan respons natural dibandingkan chatbot konvensional berbasis rule-based system.

II. TINJAUAN PUSTAKA

A. E-Commerce dan Platform Digital

E-Commerce adalah kegiatan jual-beli elektronik melalui media digital yang melibatkan berbagai pihak seperti penjual, pembeli, dan instansi pemerintah [3]. Platform ini menggunakan teknologi digital untuk mempercepat transaksi, mempermudah komunikasi, dan memperluas jangkauan pasar tanpa batasan lokasi.

Jenis e-commerce dibedakan berdasarkan pola transaksi yaitu *Business to Business (B2B)*, *Business to Consumer (B2C)*, *Customer to Customer (C2C)*, *Business to Government (B2G)*, dan *Government to Government (G2G)*. Di Indonesia, e-commerce berkembang pesat karena meningkatnya akses internet, pengguna *smartphone*, dan kepercayaan masyarakat terhadap belanja online [1].

Faktor keberhasilan platform e-commerce meliputi aspek teknologi (keandalan sistem dan keamanan data), aspek bisnis (model pendapatan dan kualitas layanan), serta aspek sosial (kepercayaan pembeli dan kesiapan teknologi). Kesuksesan juga ditentukan oleh *user experience*, kemudahan pembayaran, dan strategi pemasaran yang tepat.

B. Teknologi Chatbot dan Kecerdasan Buatan

Chatbot adalah sistem yang dapat melakukan percakapan panjang untuk meniru interaksi informal manusia [4]. Perkembangan chatbot telah mengalami evolusi dari sistem

sederhana seperti ELIZA hingga *chatbot* neural modern yang menggunakan *Large Language Model* (LLM).

Jenis *chatbot* berdasarkan teknologi yang digunakan yaitu: *Menu/Button-Based Chatbots* yang menggunakan pohon keputusan, *Keyword Recognition-Based Chatbots* yang mengenali kata kunci spesifik, *Contextual Chatbots* yang memanfaatkan *Machine Learning* dan AI, serta neural *chatbots* modern dengan arsitektur *transformer* [5].

Sistem *chatbot* neural modern terdiri dari komponen arsitektur *causal language model* yang memprediksi kata berdasarkan konteks sebelumnya. LLM memiliki kemampuan pemrosesan bahasa alami canggih melalui *conditional generation* dan dapat menangani berbagai tugas NLP seperti *question answering*, *summarization*, dan *sentiment analysis* [4].

Natural Language Processing (NLP) dalam *chatbot* mencakup *text normalization*, *tokenization*, *parsing*, dan *entity recognition*. Kemampuan pemahaman konteks dan *intent recognition* memungkinkan *chatbot* memberikan respons yang relevan dan konsisten sepanjang percakapan.

C. Teknologi Integrasi Sistem

Application Programming Interface (API) adalah sekumpulan aturan dan protokol yang memungkinkan aplikasi berkomunikasi satu sama lain [6]. API berfungsi sebagai jembatan penghubung untuk pertukaran data menggunakan format standar seperti JSON atau XML. Jenis API yang umum digunakan adalah REST (paling populer karena sederhana dan *scalable*), GraphQL (untuk optimasi *query*), dan SOAP (untuk keamanan *robust*).

Webhook adalah mekanisme *callback* otomatis yang memungkinkan aplikasi mengirimkan informasi *real-time* ketika *event* tertentu terjadi. Berbeda dengan *API polling* yang melakukan *request* berkala, *webhook* menggunakan pendekatan *push-based* yang lebih efisien dalam penggunaan *bandwidth* dan *resource server*.

WhatsApp *Gateway* adalah layanan perantara yang memungkinkan integrasi sistem bisnis dengan platform WhatsApp untuk komunikasi otomatis [6]. Fonnte merupakan salah satu penyedia layanan WhatsApp *Gateway* di Indonesia yang menawarkan API mudah diintegrasikan dengan fitur seperti pengiriman pesan otomatis, *scheduling*, *auto-reply*, dan *webhook* untuk notifikasi *real-time*.

D. Komunikasi Real-Time dan WebSocket

Komunikasi *real-time* mengacu pada pertukaran informasi dengan penundaan minimal antara pengirim dan penerima. *WebSocket* adalah protokol yang menyediakan *full-duplex communication channel* melalui *single TCP connection*, memungkinkan server dan klien saling mengirim data kapan saja setelah koneksi terbentuk.

Keunggulan *WebSocket* dibandingkan *polling* tradisional adalah efisiensi *bandwidth* dan pengurangan server load. *WebSocket* API menyediakan *event-driven programming model* dengan *event* seperti *onopen*, *onmessage*, *onerror*, dan *onclose*.

Redis (*Remote Dictionary Server*) adalah *in-memory data structure store* yang berperan penting dalam aplikasi *real-time*. Redis menyediakan fitur *Pub/Sub messaging system* untuk *broadcasting message* ke *multiple clients*, serta

berbagai *data structure* untuk representasi data yang efisien. Redis *Streams* menyediakan *persistent log data structure* untuk *message history* dan *guaranteed delivery*.

E. Metodologi Pengembangan Perangkat Lunak

Software Development Life Cycle (SDLC) adalah kerangka kerja sistematis yang mendefinisikan tahapan pengembangan *software* dari konsepsi hingga pemeliharaan [7]. SDLC menyediakan struktur untuk meminimalkan risiko, mengontrol biaya, dan memastikan kebutuhan pengguna terpenuhi.

Model *Waterfall* merupakan pendekatan sekuensial linear yang terdiri dari tahapan berurutan tanpa tumpang tindih antar fase. Model ini dinamakan "waterfall" karena kemajuan mengalir seperti air terjun melalui fase-fase berurutan dengan dokumentasi menyeluruh di setiap tahap.

Keunggulan model *Waterfall* meliputi struktur yang jelas, dokumentasi lengkap, prediktabilitas waktu dan anggaran, serta cocok untuk proyek dengan kebutuhan stabil. Kelemahannya adalah kurang fleksibel terhadap perubahan, pengujian hanya di akhir, dan risiko produk tidak sesuai ekspektasi pengguna.

III. METODE

A. Metodologi Pengembangan

Metodologi *Waterfall* dipilih sebagai pendekatan. Penelitian ini menggunakan metodologi *Waterfall* sebagai pendekatan pengembangan sistem yang dipilih berdasarkan kesesuaian dengan kompleksitas sistem yang mengintegrasikan berbagai teknologi seperti *Large Language Model*, WhatsApp API, *WebSocket*, dan *Redis*. Metodologi *Waterfall* memungkinkan pengembangan yang terstruktur dan sistematis melalui tahapan-tahapan yang berurutan, dimana setiap fase harus diselesaikan sepenuhnya sebelum melanjutkan ke fase berikutnya. Pendekatan ini memungkinkan validasi menyeluruh pada setiap tahap pengembangan, sehingga dapat memastikan kualitas dan stabilitas sistem sebelum melanjutkan ke tahap berikutnya.

Selain itu, metodologi *Waterfall* sangat sesuai untuk penelitian akademik yang memerlukan dokumentasi yang lengkap dan *traceability* yang jelas dari setiap tahap pengembangan. Setiap fase dalam *Waterfall* menghasilkan *deliverable* yang dapat dievaluasi dan divalidasi secara komprehensif, sehingga mendukung proses penelitian yang sistematis dan terukur.

B. Metode Pengumpulan Data

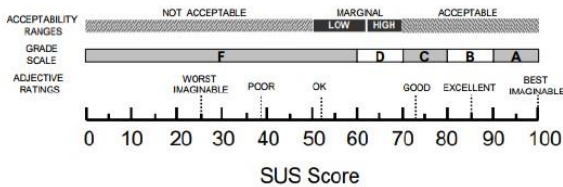
Pengumpulan data dalam penelitian ini menggunakan dua pendekatan utama yaitu data primer dan data sekunder untuk mendukung pengembangan sistem yang komprehensif. Data primer diperoleh melalui wawancara mendalam dengan pemilik dan pengelola platform TokoPoin untuk memahami kebutuhan sistem komunikasi pelanggan, tantangan operasional, ekspektasi sistem *chatbot*, serta *workflow* komunikasi yang diinginkan. Wawancara menggunakan *open-ended questions* untuk mendapatkan informasi yang mendalam dan relevan. Selain itu, observasi sistem dilakukan terhadap sistem *existing* TokoPoin untuk memahami arsitektur dan infrastruktur yang ada, *integration points* yang tersedia, serta *workflow* operasional pengelolaan platform.

Data sekunder diperoleh melalui dokumentasi teknis yang mencakup *API documentation* untuk *LLM providers*, Fonnte WhatsApp *API documentation*, *Laravel framework documentation*, serta teknologi pendukung seperti *Redis* dan *WebSocket*. *Literature review* juga dilakukan terhadap penelitian implementasi *Large Language Model* dalam *customer service*, arsitektur sistem komunikasi *real-time*, serta metode evaluasi sistem *chatbot* berbasis AI.

C. Metode System Usability Scale (SUS)

Metode *System Usability Scale* (SUS) digunakan untuk mengukur tingkat kegunaan (*usability*) dari sistem yang dikembangkan. SUS merupakan instrumen evaluasi standar yang terdiri dari 10 pernyataan, masing-masing menggunakan skala *Likert 5* poin dari "Sangat Tidak Setuju" (1) hingga "Sangat Setuju" (5). Pernyataan disusun secara bergantian antara positif dan negatif untuk menjaga konsistensi serta menghindari bias dalam penilaian responden. SUS dipilih karena metode ini mudah digunakan, cepat diterapkan, dan telah terbukti valid serta reliabel untuk berbagai jenis produk digital dan sistem interaktif.

Skor akhir SUS berada pada rentang 0 hingga 100, dengan interpretasi sebagai berikut: skor < 50 menunjukkan sistem perlu perbaikan serius, skor 50-70 menunjukkan *usability* dapat diterima namun masih bisa ditingkatkan, skor 70-85 menunjukkan sistem layak dan *usable*, dan skor > 85 menunjukkan sistem sangat mudah digunakan dan disukai pengguna. Secara statistik, skor rata-rata SUS dari berbagai studi adalah 68, sehingga jika hasil pengujian sistem menunjukkan skor di atas 68, berarti sistem tersebut memiliki *usability* yang di atas rata-rata dan dapat dianggap layak digunakan. Interpretasi skor SUS dan kategorinya dapat dilihat pada Gambar 1.

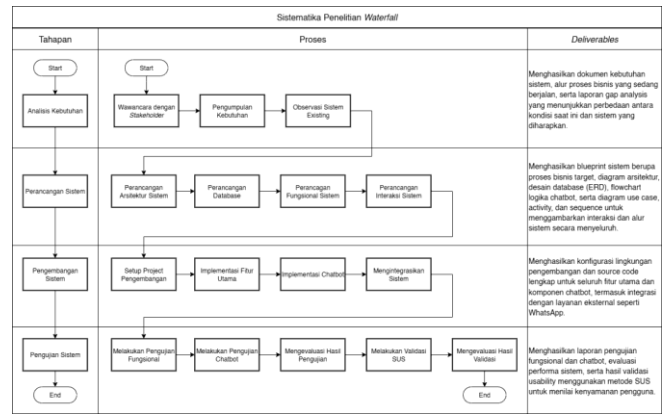


GAMBAR 1 (Interpretasi Skor SUS dan Kategorinya)

D. Sistematika Penyelesaian Masalah

Pengembangan sistem *chatbot* terintegrasi dalam penelitian ini menggunakan metodologi *Waterfall* karena menawarkan tahapan yang runtut, terdokumentasi, dan mudah dikontrol. Pendekatan ini dinilai sesuai untuk menangani proses pengembangan yang membutuhkan struktur jelas dan verifikasi di setiap fase.

Sistematika penyelesaian masalah disusun mengikuti fase-fase *Waterfall* yang telah disesuaikan dengan kebutuhan sistem *chatbot* berbasis AI. Setiap tahap memiliki tujuan spesifik, aktivitas terstruktur, dan *deliverable* yang dapat dievaluasi. Tahapan dimulai dari analisis kebutuhan, perancangan sistem, implementasi, hingga pengujian akhir. Alur keseluruhan ditampilkan pada Gambar 2 yang menggambarkan proses dari awal hingga akhir pengembangan.



GAMBAR 2 (Sistematika Penelitian)

Proses dimulai dengan analisis kebutuhan untuk memahami permasalahan komunikasi pelanggan melalui wawancara, observasi proses bisnis *existing*, dan studi literatur. Keluaran dari tahap ini berupa dokumen kebutuhan sistem, pemetaan proses bisnis saat ini, dan laporan *gap analysis*. Tahap berikutnya adalah perancangan sistem yang mengubah kebutuhan menjadi rancangan teknis, mencakup proses bisnis target, arsitektur sistem, ERD, *flowchart* logika *chatbot*, serta diagram *use case*, *activity*, dan *sequence* untuk menggambarkan interaksi dan alur sistem secara detail.

Pada tahap implementasi, sistem dikembangkan berdasarkan desain yang telah disusun melalui beberapa tahapan: setup dan konfigurasi lingkungan pengembangan, implementasi antarmuka *chat real-time* dengan *WebSocket* dan *Redis pub/sub*, implementasi deteksi kehadiran pengguna, implementasi sistem konfigurasi chatbot, integrasi WhatsApp melalui API Fonnte, dan integrasi *Large Language Model* menggunakan Ollama sebagai *engine* lokal. Terakhir, pengujian sistem dilakukan untuk memastikan seluruh fungsionalitas berjalan sesuai kebutuhan melalui *functional testing*, pengujian *chatbot* dengan *dataset* uji, dan validasi *usability* menggunakan metode *System Usability Scale* (SUS).

E. Metode Evaluasi

Pengujian sistem dilakukan secara menyeluruh untuk memastikan sistem berjalan sesuai dengan kebutuhan yang telah ditetapkan. Pengujian fungsional menggunakan pendekatan *black box testing* untuk memverifikasi bahwa setiap fungsi dalam sistem bekerja sesuai dengan spesifikasi *requirements*. Kategori pengujian fungsional mencakup pengujian komunikasi *real-time* melalui *WebSocket* dan *Redis pub/sub*, pengujian deteksi kehadiran untuk status *online/offline* pengguna, pengujian integrasi WhatsApp dengan Fonnte API, dan pengujian sistem notifikasi *real-time*. Sebelum melakukan pengujian terhadap *enhanced system*, dilakukan *assessment* terhadap kondisi *baseline system* untuk menetapkan titik acuan evaluasi sebagaimana ditunjukkan pada Tabel 1.

TABEL 1 (BASELINE SYSTEM)

Aspek	Kondisi saat ini
Komunikasi Pesan	Berbasis <i>polling/refresh</i> manual
Waktu <i>Delivery</i> Pesan	60-300 detik bergantung <i>refresh</i> manual
Notifikasi <i>Real-time</i>	Tidak tersedia

Aspek	Kondisi saat ini
Indikator Status	Tidak ada indikator <i>online/offline</i>
Tidak ada indikator <i>online/offline</i>	Hanya tersedia ketika penjual <i>online</i>
Platform Komunikasi	Hanya melalui web <i>interface</i>
Penanganan Pertanyaan Umum	Semua pertanyaan harus dijawab manual

Pengujian *chatbot* dilakukan untuk mengevaluasi kinerja sistem dalam tiga aspek utama: akurasi pendeteksian intent pengguna, waktu respons sistem, dan kesesuaian respons yang dihasilkan. Dua model *Large Language Model (LLM)* yang dijalankan menggunakan Ollama dibandingkan untuk mengetahui model mana yang menghasilkan performa terbaik. Pengujian akurasi menggunakan metrik *accuracy* yang dihitung dari rasio prediksi benar terhadap seluruh data uji, sedangkan waktu respons diukur dari durasi *input* pertanyaan hingga *chatbot* menghasilkan respons lengkap. Untuk pengujian skor kesesuaian respons, setiap respons diberikan skor dalam skala 1 hingga 5 berdasarkan kriteria yang dijelaskan pada Tabel 2.

TABEL 2
(KRITERIA KESesuaIAN RESPONS)

Skor	Kriteria Kesesuaian Respons
1	Tidak menjawab pertanyaan sama sekali, jawaban tidak relevan
2	Menjawab sebagian kecil, tidak jelas, atau keluar konteks
3	Menjawab sebagian besar poin penting, tetapi masih kurang lengkap
4	Menjawab dengan cukup baik, tetapi ada informasi yang terlewat atau ambigu
5	Menjawab pertanyaan secara lengkap, tepat, dan langsung ke intinya

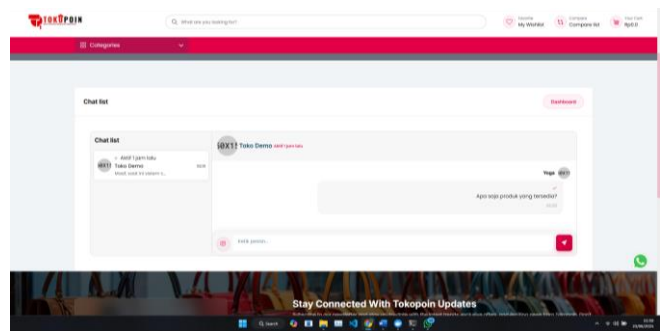
IV. HASIL DAN PEMBAHASAN

A. Lingkungan Pengembangan

Sistem chatbot berbasis *Large Language Model (LLM)* untuk platform TokoPoin dikembangkan menggunakan arsitektur full-stack dengan Node.js v22.15.1 sebagai runtime environment, MySQL 8.0 untuk database utama, dan Redis sebagai cache database. Implementasi menggunakan teknologi *WebSocket* untuk komunikasi real-time, integrasi *Fonnte API* untuk koneksi WhatsApp, dan platform Ollama untuk menjalankan model LLM. Lingkungan pengembangan menggunakan processor AMD Ryzen 5 5600U dengan 16 GB RAM untuk memastikan performa optimal dalam pemrosesan model bahasa.

B. Implementasi Fitur Utama

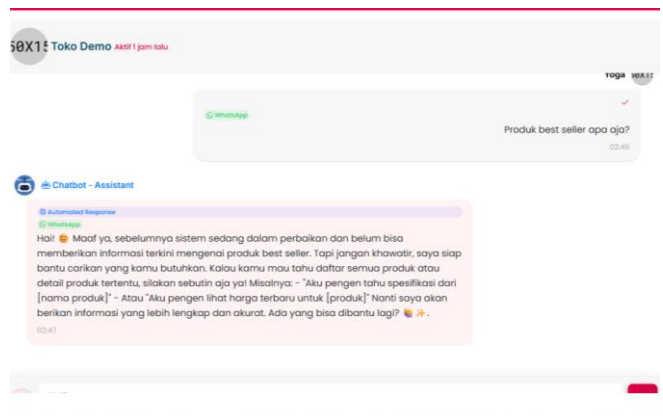
Sistem chatbot berbasis *Large Language Model (LLM)* Implementasi dimulai dengan pengembangan fitur komunikasi *real-time* menggunakan teknologi *WebSocket* yang memungkinkan pertukaran pesan instan antara pelanggan dan penjual. Sistem mengintegrasikan Redis sebagai *cache* untuk melacak status *online/offline* setiap pengguna secara akurat. Untuk meningkatkan efisiensi komunikasi, dikembangkan fitur *template chat* yang memungkinkan pengguna menggunakan pesan-pesan standar yang telah disiapkan sebelumnya, sehingga mengurangi waktu pengetikan untuk respons yang sering digunakan dalam konteks layanan pelanggan.



GAMBAR 2

(Antarmuka sistem chat real-time dengan indikator status *online/offline*)

Integrasi dengan platform WhatsApp dilakukan melalui layanan *Fonnte API* yang menyediakan komunikasi WhatsApp Business. Sistem menciptakan sinkronisasi dua arah yang memungkinkan pesan yang dikirim melalui WhatsApp muncul secara *real-time* di antarmuka web aplikasi. Ketika pelanggan mengirim pesan melalui WhatsApp, sistem menangkap pesan melalui *webhook* dan meneruskannya ke *endpoint* aplikasi dengan aktivasi otomatis sistem chatbot. Sistem notifikasi dirancang untuk memastikan penjual tidak melewatkan pesan penting dengan menampilkan notifikasi *toaster* pada antarmuka web dan mengirim pemberitahuan melalui WhatsApp.



GAMBAR 3

(Integrasi WhatsApp dengan antarmuka web aplikasi)

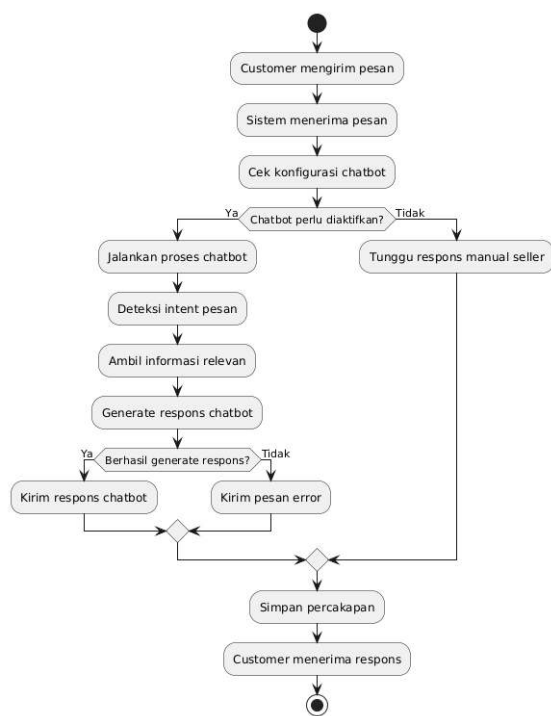
C. Implementasi Sistem Chatbot

Sistem chatbot berbasis *Large Language Model (LLM)* Implementasi dimulai dengan pengembangan fitur komunikasi *real-time* menggunakan teknologi *WebSocket* yang memungkinkan pertukaran pesan instan antara pelanggan dan penjual. Sistem mengintegrasikan Redis sebagai *cache* untuk melacak status *online/offline* setiap pengguna secara akurat. Untuk meningkatkan efisiensi komunikasi, dikembangkan fitur *template chat* yang memungkinkan pengguna menggunakan pesan-pesan standar yang telah disiapkan sebelumnya, sehingga mengurangi waktu pengetikan untuk respons yang sering digunakan dalam konteks layanan pelanggan.

Arsitektur sistem chatbot menggunakan pendekatan modular dengan dua model LLM utama: DeepSeek-R1:8B dan Llama 3.2:3B yang dijalankan melalui platform Ollama. Sistem dimulai dengan *IntentDetectionService* yang menggunakan pendekatan *fuzzy matching* dengan algoritma

perhitungan skor kemiripan untuk mengidentifikasi 20 kategori intent utama, dimana setiap kategori memiliki *keywords, patterns, confidence threshold*, dan *priority level*.

RAGService (Retrieval-Augmented Generation) berfungsi mengambil informasi relevan dari basis data berdasarkan *intent* yang terdeteksi. Meskipun tidak mengimplementasikan RAG dalam arti teknis sebenarnya dengan *vector embedding* dan *semantic search*, sistem menggunakan pendekatan yang disederhanakan namun efektif untuk menyediakan konteks yang tepat kepada LLM. Informasi yang diperoleh kemudian diformat menjadi *prompt* terstruktur dengan teknik *prompt engineering* yang disesuaikan untuk setiap kategori intent guna memastikan konsistensi dan relevansi respons.



GAMBAR 4
(Gambar Alur Logika Chatbot)

D. Pengujian dan Evaluasi

Pengujian fungsional dilakukan untuk memverifikasi bahwa setiap fitur inti sistem bekerja sesuai dengan kebutuhan dan skenario penggunaan. Pengujian mencakup enam test case utama yang dirancang dengan langkah-langkah terstruktur untuk mengevaluasi koneksi WebSocket, pengiriman pesan real-time, status online/offline pengguna, fungsi template chat, integrasi WhatsApp, dan sistem notifikasi real-time. Setiap test case memiliki tujuan spesifik dengan kriteria keberhasilan yang telah ditetapkan sebelumnya untuk memastikan objektivitas dalam penilaian performa sistem. Hasil pengujian menunjukkan bahwa semua komponen sistem berfungsi dengan baik dan memenuhi kriteria yang ditetapkan, sebagaimana ditampilkan dalam Tabel I. Koneksi WebSocket berhasil terbentuk tanpa hambatan, sementara pengiriman pesan real-time mencapai rata-rata waktu 2,412 detik, masih berada di bawah target maksimal 3 detik yang ditetapkan.

TABEL 3
(HASIL PERBANDINGAN PERFORMA MODEL LLM)

Pengujian	Hasil Didapatkan	Hasil Diharapkan	Status
Koneksi WebSocket	Koneksi berhasil terbentuk	Koneksi WebSocket berhasil	Sukses
Pengiriman Pesan Real-time	Pengiriman Pesan Real-time	≤ 3 detik	Sukses
Status Online/Offline	Status Online/Offline	Status Online/Offline	Sukses
Template Chat	Template mengisi textbox, dapat diedit/dihapus	Template mengisi textbox, dapat diedit	Sukses
Kirim Pesan WhatsApp	Respons chatbot diterima WhatsApp	Respons chatbot terkirim ke WhatsApp	Sukses
Notifikasi Real-time	Notifikasi terkirim dengan baik	Notifikasi diterima penjual pada WhatsApp	Sukses

Status online/offline pengguna dapat berubah secara real-time dan akurat, template chat berfungsi optimal dengan kemampuan mengisi textbox yang dapat diedit dan dihapus, integrasi WhatsApp berhasil mengirim respons chatbot ke pengguna, dan sistem notifikasi real-time beroperasi dengan baik dalam mengirimkan pemberitahuan ke penjual melalui WhatsApp. Keberhasilan semua test case pada pengujian fungsional mengindikasikan bahwa fitur-fitur utama sistem telah memenuhi spesifikasi yang ditetapkan dan siap untuk digunakan dalam kondisi operasional.

Evaluasi performa chatbot dilakukan menggunakan dataset 16 pertanyaan yang merepresentasikan berbagai skenario interaksi pelanggan, meliputi informasi produk, harga, stok, spesifikasi, rekomendasi, proses pemesanan, metode pembayaran, pengiriman, dan layanan pelanggan. Dataset dirancang untuk mencerminkan kondisi penggunaan nyata dengan setiap pertanyaan diberi ID unik (QC-001 hingga QC-016) dan dikaitkan dengan kategori ekspektasi jawaban yang merepresentasikan intent pengguna. Pengujian dilakukan terhadap dua model LLM yang dijalankan menggunakan Ollama, yaitu llama3.2:3b dan deepseek-r1:8b, dengan setiap model diuji menggunakan dataset yang sama untuk memastikan keadilan dalam perbandingan performa. Hasil evaluasi yang ditampilkan dalam Tabel II menunjukkan bahwa model deepseek-r1:8b memberikan akurasi deteksi intent tertinggi yaitu 95,84% (46 dari 48 percobaan) dan skor kualitas jawaban superior 4,48 dari skala 5, namun memerlukan waktu respons yang lebih lama yaitu 2 menit 9 detik.

TABEL 4
(HASIL PERBANDINGAN PERFORMA MODEL LLM)

Model	Akurasi	Rata-rata Waktu Respons	Skor Kualitas Jawaban
Deepseek-r1:8b	95,84% (46/48)	2 menit 9 detik	4,48/5,0
Llama3.2:3b	93,75% (45/48)	8,4 detik	4,48/5,0

Sebaliknya, model llama3.2:3b mencapai akurasi 93,75% (45 dari 48 percobaan) dengan waktu respons yang signifikan lebih cepat yaitu 8,4 detik, namun dengan skor kualitas jawaban 4,10. Dari segi distribusi kualitas jawaban, deepseek-r1:8b menghasilkan 31 jawaban dengan skor tertinggi (5), sementara llama3.2:3b hanya 16 jawaban dengan skor tertinggi, menunjukkan konsistensi kualitas yang lebih baik pada model deepseek-r1:8b meskipun dengan trade-off waktu respons yang lebih lama. Evaluasi perbandingan model LLM

mengkonfirmasi adanya trade-off antara kualitas dan kecepatan respons, dimana model llama3.2:3b menjadi pilihan optimal untuk implementasi karena mampu memberikan respons yang cepat dengan akurasi yang tetap tinggi, lebih sesuai untuk konteks komunikasi real-time dibandingkan deepseek-r1:8b yang meskipun lebih akurat namun memiliki waktu respons yang tidak praktis untuk penggunaan real-time.

Validasi usability dilakukan menggunakan metode System Usability Scale (SUS) yang terdiri dari 10 pertanyaan dengan skala Likert 1-5, dimana pernyataan ganjil bersifat positif dan pernyataan genap bersifat negatif. Sebanyak 36 responden terlibat dalam validasi ini, terdiri dari 4 penjual dan 32 pembeli yang menggunakan sistem dalam kondisi operasional. Perhitungan skor SUS mengikuti formula standar dimana untuk pernyataan positif, skor diperoleh dengan mengurangi nilai jawaban dengan 1, sedangkan untuk pernyataan negatif, skor diperoleh dengan mengurangi 5 dengan nilai jawaban, kemudian total skor dikalikan dengan 2,5 untuk mendapatkan skor SUS yang berkisar antara 0-100. Hasil validasi menunjukkan bahwa sistem memperoleh skor rata-rata SUS sebesar 77,85 dengan standar deviasi 16,51, yang berada dalam kategori usability baik dan melampaui ambang batas standar 68.

Analisis berdasarkan peran pengguna menunjukkan perbedaan yang menarik dimana penjual memiliki rata-rata skor 71,25 dengan standar deviasi rendah (9,46) yang menunjukkan penilaian konsisten meskipun sedikit lebih rendah, sementara pembeli memiliki rata-rata skor 78,59 dengan variasi yang lebih besar (standar deviasi 17,12). Perbedaan ini dapat disebabkan oleh kompleksitas fitur yang digunakan penjual lebih tinggi seperti konfigurasi chatbot dan pengelolaan chat, serta ekspektasi yang berbeda dimana pembeli fokus pada kemudahan mencari dan membeli produk sementara penjual membutuhkan efisiensi dalam mengelola bisnis. Hasil validasi SUS yang mencapai skor 77,85 menunjukkan bahwa sistem telah memenuhi standar usability yang baik dan siap untuk implementasi operasional, dengan sistem berhasil mengurangi beban kerja penjual dalam menangani pertanyaan repetitif pelanggan, menyediakan layanan 24/7, dan meningkatkan efisiensi komunikasi melalui integrasi multi-channel yang memungkinkan akses fleksibel bagi pengguna.

V. KESIMPULAN

Implementasi sistem chatbot berbasis LLM untuk platform e-commerce TokoPoin berhasil dikembangkan dengan mengintegrasikan teknologi WebSocket untuk komunikasi real-time, Fonnte API untuk koneksi WhatsApp, dan platform Ollama untuk menjalankan model bahasa. Hasil pengujian fungsional menunjukkan bahwa semua komponen sistem beroperasi dengan baik, mencapai waktu respons komunikasi real-time rata-rata 2,412 detik yang berada di bawah target maksimal 3 detik, serta berhasil mengimplementasikan integrasi multi-channel yang seamless antara platform web dan WhatsApp. Validasi usability menggunakan metode System Usability Scale (SUS) dengan 36 responden menghasilkan skor rata-rata 77,85, melampaui ambang batas standar 68 dan menunjukkan bahwa sistem memiliki tingkat usability yang baik untuk kedua kelompok pengguna (penjual dan pembeli).

Evaluasi perbandingan model LLM mengkonfirmasi adanya trade-off antara kualitas dan kecepatan respons. Model

deepseek-r1:8b menunjukkan akurasi deteksi intent tertinggi sebesar 95,84% dengan kualitas jawaban superior (4,48/5,0) namun memerlukan waktu pemrosesan yang signifikan lebih lama yaitu 2 menit 9 detik. Sebaliknya, model llama3.2:3b memberikan waktu respons yang jauh lebih cepat (8,4 detik) dengan akurasi 93,75% dan kualitas jawaban (4,10/5,0) yang tetap memadai untuk kebutuhan komunikasi real-time. Berdasarkan karakteristik aplikasi yang membutuhkan responsivitas tinggi, llama3.2:3b terbukti menjadi pilihan optimal untuk implementasi operasional karena mampu memberikan pengalaman pengguna yang lebih baik dalam konteks layanan pelanggan real-time.

Sistem berhasil mencapai tujuan utama dalam mengurangi beban kerja penjual melalui otomatisasi respons terhadap pertanyaan repetitif pelanggan, menyediakan layanan 24/7 yang konsisten, dan meningkatkan efisiensi komunikasi pada platform e-commerce internal Koperasi Telkom University. Kontribusi utama penelitian ini adalah pengembangan solusi chatbot yang spesifik dirancang untuk karakteristik komunitas tertutup dengan implementasi integrasi multi-channel yang memungkinkan pengguna mengakses layanan melalui berbagai platform sesuai preferensi mereka. Keberhasilan implementasi ini memberikan fondasi yang kuat untuk pengembangan sistem serupa di lingkungan e-commerce skala kecil hingga menengah, dengan fokus pada optimisasi keseimbangan antara kualitas respons dan kecepatan pemrosesan untuk meningkatkan kepuasan pengguna secara keseluruhan.

REFERENSI

- [1] Y. Nugraha, Y. Masnita, and K. Kurniawati, "Peran Responsiveness Chatbot Artificial Intelligence Dalam Membentuk Customer Satisfaction," *JURNAL MANAJEMEN DAN BISNIS SRIWIJAYA*, vol. 20, no. 3, 2022, doi: 10.29259/jmbs.v20i3.18528.
- [2] M. F. Umam Al Jupri and M. I. Fasa, "PENGARUH PENGGUNAAN TEKNOLOGI CHATBOT DALAM MENINGKATKAN LAYANAN PELANGGAN PADA PLATFORM E-COMMERCE: STUDI KASUS PADA SHOPEE INDONESIA," *Jurnal Media Akademik (JMA)*, vol. 3, no. 3, Apr. 2025, doi: 10.62281/v3i3.1688.
- [3] Q. Zheng, *Introduction to E-commerce*. 2009. doi: 10.1007/978-3-540-49645-8.
- [4] D. Jurafsky and J. H. Martin, *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, Third Edition draft. 2024.
- [5] A. Gupta, "Introduction to AI Chatbots," *International Journal of Engineering Research and*, vol. V9, no. 07, 2020, doi: 10.17577/ijertv9is070143.
- [6] M. Reddy, *API Design for C++*. 2011. doi: 10.1016/C2010-0-65832-9.
- [7] D. Jacobson, D. Woods, and G. Brail, *APIs: A Strategy Guide*. 2012.