

Optimalisasi Sistem *Chatbot* Layanan Pelanggan *E-Commerce* melalui Kombinasi *Rule-Based* dan Algoritma *K-Nearest Neighbor* (KNN)

Zaenal Mustofa¹, Febryantahanuji², dan Arsito Ari Kuncoro^{3,*}

¹ Universitas Negeri Yogyakarta 1; email : zaenalmustofa@uny.ac.id

² Universitas Sains dan Teknologi Komputer 2; email : febryan@stekom.ac.id

³ Universitas Sains dan Teknologi Komputer 3; email : arsito@stekom.ac.id

* Penulis

Abstract: This research aims to design and evaluate a hybrid chatbot prototype that combines a rule-based approach with the K-Nearest Neighbor (KNN) algorithm to classify customer satisfaction levels. This system was developed as an automated customer service solution for the e-commerce sector, which demands speed, efficiency, and personalization. By using numerical features such as information comprehension (X1), chatbot rating (X2), and conversation duration (X3), the KNN algorithm determines whether a user is satisfied or not after an interaction session. The classification results indicate that some "Dissatisfied" predictions occurred in cases with high information comprehension but also long conversation durations. This suggests that customers can still be dissatisfied if the service time is perceived as too long, even if the content of the answer is clear. Although there is a margin of error in the user experience aspect, the system demonstrates a very significant impact on operational efficiency. The implementation of the hybrid chatbot was able to reduce customer service costs by up to 50%, with an estimated payback period of 4.8 months. The system provides significant added value for businesses, particularly in terms of cost savings and scalability, although there is still room for improvement in the quality of interaction.

Keywords: *Chatbot; KNN; Hybrid; Customer Satisfaction; Cost Efficiency; E-commerce*

Abstrak: Penelitian ini bertujuan untuk merancang dan mengevaluasi prototipe chatbot hybrid yang menggabungkan pendekatan *rule-based* dan algoritma *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan tingkat kepuasan pelanggan. Sistem ini dikembangkan sebagai solusi layanan pelanggan otomatis di sektor *e-commerce* yang menuntut kecepatan, efisiensi, dan personalisasi. Dengan menggunakan fitur numerik seperti pemahaman informasi (X1), rating chatbot (X2), dan durasi percakapan (X3), algoritma KNN digunakan untuk menentukan apakah pengguna merasa puas atau tidak setelah sesi interaksi. Hasil klasifikasi menunjukkan bahwa beberapa prediksi "Tidak Puas" muncul pada kasus dengan pemahaman informasi yang tinggi tetapi durasi chat yang juga panjang. Hal ini menunjukkan bahwa pelanggan tetap bisa merasa tidak puas bila waktu layanan dirasa terlalu lama, meskipun isi jawaban jelas. Meskipun terdapat margin kesalahan pada aspek pengalaman pengguna, dari sisi efisiensi operasional, sistem ini menunjukkan dampak yang sangat signifikan. Implementasi chatbot hybrid mampu menurunkan biaya layanan pelanggan hingga 50%, dengan estimasi *payback period* selama 4,8 bulan. Sistem memberikan nilai tambah yang besar bagi bisnis, khususnya dalam hal penghematan biaya dan skalabilitas, meski masih ada ruang perbaikan pada sisi kualitas interaksi.

Kata kunci: *Chatbot; KNN; Hybrid; Kepuasan Pelanggan; Efisiensi Biaya; E-commerce*

Diterima: 20 Agustus 2025
Direvisi: 6 Januari 2026
Diterima: 10 Januari 2026
Diterbitkan: 11 Maret 2026
Versi sekarang: Januari 2026



Hak cipta: © 2025 oleh penulis.
Diserahkan untuk kemungkinan
publikasi akses terbuka
berdasarkan syarat dan ketentuan
lisensi Creative Commons
Attribution (CC BY SA) (
<https://creativecommons.org/licenses/by-sa/4.0/>)

1. Pendahuluan

Pertumbuhan industri *e-commerce* secara global dalam beberapa tahun terakhir telah mengalami akselerasi yang signifikan. di Indonesia sendiri, nilai pasar *e-commerce* diproyeksikan mencapai USD 86,8 miliar pada tahun 2028, dengan tingkat pertumbuhan tahunan gabungan (CAGR) sebesar 10% (Trade.gov, 2024). Perkembangan ini tidak hanya mencerminkan peningkatan volume transaksi, tetapi juga mendorong transformasi dalam ekspektasi konsumen terhadap layanan pelanggan. generasi digital, khususnya generasi muda, menuntut pengalaman berbelanja yang bersifat percakapan alami (*natural conversation*), bersifat instan, tersedia 24 jam tanpa jeda, serta tetap memberikan kendali penuh terhadap proses pengambilan keputusan. Survei yang dilakukan oleh Omnisend (2024) memperlihatkan bahwa 66% konsumen masih menginginkan otonomi dalam keputusan pembelian mereka, namun 38% juga menyatakan apresiasi terhadap adanya rekomendasi instan berbasis kecerdasan buatan (AI).

Tantangan utama yang dihadapi pelaku bisnis *e-commerce* adalah bagaimana menyajikan layanan pelanggan yang cepat dan responsif, tanpa meningkatkan beban biaya operasional secara signifikan. Hal ini menciptakan dilema antara kualitas layanan dan efisiensi biaya sebuah *trade-off* yang klasik dalam pengelolaan operasional digital. Pada tahap awal, *chatbot* berbasis aturan (*rule-based*) muncul sebagai solusi praktis untuk menangani *Frequently Asked Questions* (FAQ). Sistem ini dibangun berdasarkan logika deterministik seperti *if-then* yang mudah dipelihara dan memberikan alur kerja yang transparan bagi pemilik bisnis (Moitra et al., 2020).

Namun demikian, seiring meningkatnya kompleksitas pertanyaan pelanggan dan beragamnya emosi yang terlibat dalam percakapan digital, pendekatan murni *rule-based* mulai kewalahan. Studi komparatif yang dipublikasikan oleh Zhou et al. (2025) menunjukkan bahwa akurasi *chatbot rule-based* dapat menurun drastis (kurang dari 60%) ketika dialog tidak lagi mengikuti skenario yang telah diprogram sebelumnya.

Untuk menjawab tantangan ini, literatur terkini telah mengeksplorasi pemanfaatan model generatif dan *retrieval-based* sebagai solusi yang lebih adaptif. Meskipun demikian, kedua pendekatan ini memerlukan sumber daya komputasi yang besar dan sering kali menimbulkan kekhawatiran terkait keterjelasan logika (*explainability*), yang pada gilirannya menghambat adopsi di kalangan pelaku usaha kecil dan menengah (UKM) (Zhou et al., 2023). Dalam konteks ini, pendekatan *hybrid* yang menggabungkan sistem *rule-based* dengan algoritma pembelajaran mesin (*machine learning*) ringan menjadi pilihan menarik. Pendekatan ini bertujuan mempertahankan kejelasan logika deterministik yang menjadi ciri khas metode tradisional, sembari menyisipkan kecerdasan adaptif yang responsif terhadap konteks percakapan (Li et al., 2023). Pendekatan semacam ini tidak hanya menjawab tantangan teknis, tetapi juga selaras dengan filosofi generasi muda masa kini yang cenderung menghargai warisan lama sembari tetap terbuka terhadap inovasi baru.

Salah satu algoritma yang relevan dalam pendekatan *hybrid* ini adalah *K-Nearest Neighbor* (KNN). KNN merupakan algoritma *instance-based* yang tidak memerlukan proses pelatihan kompleks dan hanya mengandalkan parameter k (jumlah tetangga terdekat) sebagai kunci utama dalam pengambilan keputusan. Penelitian oleh Nugroho dan Lestari (2024) menunjukkan bahwa algoritma ini mampu mengidentifikasi perilaku konsumen dan mempersonalisasi rekomendasi produk dengan akurasi mencapai 91% pada dataset transaksi *e-commerce* sebanyak 10.000 entri. Selain itu, studi yang dilakukan oleh Pratama et al. (2024) pada PT Airkon Pratama membuktikan bahwa akurasi klasifikasi kepuasan pelanggan dapat meningkat hingga 96,5% setelah dilakukan optimasi terhadap skala fitur.

Kelebihan utama KNN terletak pada kemampuannya dalam menghitung jarak *Euclidean* atas vektor numerik sederhana seperti durasi percakapan, rating bintang, dan tingkat penyelesaian masalah (*solved rate*). Hal ini memungkinkan KNN untuk diintegrasikan secara *plug-and-play* ke dalam sistem backend *chatbot* tanpa menimbulkan beban komputasi yang signifikan. Karakteristik tersebut menjadikannya ideal untuk diterapkan pada lingkungan *e-commerce* dengan trafik pengguna yang tinggi namun margin keuntungan yang sempit. Beberapa perusahaan besar seperti Klarna dan Vodafone telah berhasil menunjukkan bahwa implementasi *chatbot AI* mampu menurunkan waktu tanggap layanan menjadi hanya beberapa detik serta mengurangi biaya layanan pelanggan hingga lebih dari 30% (NexGenCloud, 2024). Meski demikian, literatur akademik yang terindeks di Scopus masih cenderung berfokus pada model *deep learning* berskala besar, sementara integrasi *chatbot rule-based* dengan algoritma KNN dalam konteks industri praktis di negara berkembang seperti Indonesia masih relatif kurang mendapat perhatian.

2. Rumusan Masalah

Berdasarkan uraian di atas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana efektivitas integrasi *chatbot hybrid* berbasis *rule-based* dan algoritma KNN dalam mendeteksi kepuasan pelanggan secara *real-time*?
2. Sejauh mana implementasi *chatbot hybrid* tersebut mampu memicu proses *handover* yang tepat ke agen manusia berdasarkan hasil klasifikasi KNN?
3. Bagaimana dampak finansial dari implementasi *chatbot hybrid* terhadap efisiensi biaya layanan dan periode pengembalian investasi (*payback period*) di bisnis *e-commerce* lokal?

4. Bagaimana menyusun kerangka kerja implementasi chatbot hybrid yang tetap mempertahankan aspek keterjelasan (*explainability*) bagi regulator dan pemangku kepentingan ?

3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Merancang arsitektur *chatbot hybrid* yang menggabungkan sistem *rule-based* untuk penanganan intent dengan modul klasifikasi kepuasan berbasis algoritma KNN.
2. Menguji akurasi dan efektivitas model chatbot hybrid menggunakan data log percakapan pelanggan dari platform e-commerce Indonesia.
3. Mengukur efisiensi operasional yang dihasilkan melalui integrasi *chatbot hybrid*, termasuk potensi penghematan biaya dan estimasi pengembalian investasi.
4. Menyusun kerangka implementasi praktis yang mengedepankan aspek *explainability* sebagai nilai tambah dalam konteks regulasi dan pengambilan keputusan manajerial.

4. Tinjauan Literatur

Chatbot dan *Evolusi Sistem Interaktif Chatbot* merupakan program berbasis perangkat lunak yang dirancang untuk melakukan percakapan otomatis dengan pengguna melalui antarmuka teks maupun suara. Sistem ini pertama kali diperkenalkan dalam bentuk *rule-based*, dimana percakapan dikendalikan oleh serangkaian logika deterministik seperti *if-then* statements (Moitra et al., 2020). Sistem *rule-based* unggul dalam hal keterjelasan alur dan kemudahan pemeliharaan, namun sangat terbatas dalam menghadapi dialog yang bersifat dinamis atau tidak terstruktur. Dalam perkembangannya, muncul pendekatan *retrieval-based* yang memanfaatkan teknik pencocokan berbasis keyword atau similar response dari basis data, serta pendekatan *generative* yang menggunakan model pembelajaran mendalam seperti LSTM dan Transformer untuk menghasilkan respons baru secara real-time. Kendati demikian, keterbatasan pada sisi *explainability*, konsumsi komputasi tinggi, dan biaya implementasi menjadi hambatan utama, terutama dalam konteks UKM (Zhou et al., 2023).

K-Nearest Neighbor (KNN) dalam Sistem Klasifikasi *K-Nearest Neighbor* (KNN) adalah algoritma pembelajaran mesin berbasis kemiripan (*similarity*) yang tergolong ke dalam pendekatan *instance-based learning*. Algoritma ini bekerja dengan mengidentifikasi sejumlah *k* data terdekat dalam ruang vektor untuk mengklasifikasikan data baru berdasarkan mayoritas label yang dimiliki tetangga terdekat tersebut (Rahman et al., 2024).

Kelebihan KNN terletak pada kesederhanaannya, kemudahan tuning parameter, serta tidak memerlukan pelatihan model secara *eksplisit*. Studi oleh Nugroho dan Lestari (2024) membuktikan bahwa KNN dapat diterapkan secara efektif dalam mengidentifikasi perilaku konsumen *e-commerce* dan memberikan rekomendasi personalisasi produk. Studi lain oleh Pratama et al. (2024) juga menunjukkan bahwa penerapan KNN dalam klasifikasi kepuasan pelanggan mampu mencapai akurasi tinggi setelah dilakukan normalisasi dan scaling pada fitur numerik seperti durasi chat dan rating layanan.

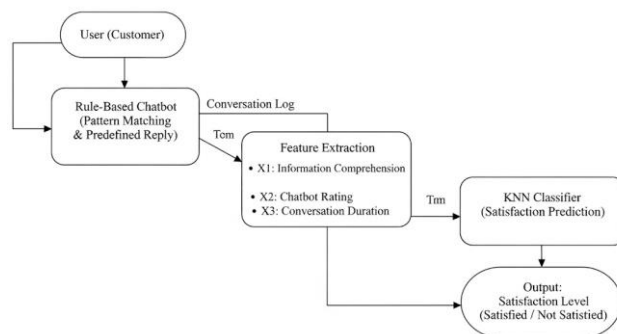
Sistem *Hybrid Integrasi Rule-Based* dan Kecerdasan adaptif untuk menjembatani kekuatan sistem *rule-based* dan fleksibilitas machine learning, pendekatan hybrid mulai banyak diteliti dan diadopsi. Sistem *hybrid* bertujuan mengombinasikan determinisme dari *rule-based* logic dengan kemampuan adaptif dari algoritma pembelajaran mesin seperti KNN, Naïve Bayes, maupun decision tree (Li et al., 2023).

Model semacam ini cocok digunakan dalam konteks industri yang membutuhkan transparansi tinggi, namun tetap mampu menangani keragaman kasus pelanggan yang kompleks. *Hybrid* chatbot juga mampu menangani kasus *handover* secara kontekstual, yakni ketika klasifikasi sistem terhadap kepuasan pengguna jatuh ke kategori negatif atau ambigu, sistem akan meneruskan percakapan kepada agen manusia secara otomatis. Fitur ini menjadi penting dalam menjaga kualitas layanan serta mencegah churn (Zhou et al., 2025).

Efisiensi Layanan Pelanggan dan ROI Chatbot Transformasi digital pada layanan pelanggan tidak hanya berdampak pada kualitas komunikasi, tetapi juga membawa perubahan signifikan dalam efisiensi operasional. Menurut studi kasus yang dipublikasikan oleh NexGenCloud (2024), perusahaan seperti Klarna dan Vodafone berhasil mengurangi waktu tunggu layanan hingga hitungan detik dan memangkas biaya layanan pelanggan lebih dari 30% melalui implementasi chatbot AI. Namun, penelitian yang berfokus pada perhitungan ROI *chatbot* umumnya masih mengasumsikan penggunaan model AI penuh (*end-to-end AI*), tanpa mengevaluasi perbedaan efisiensi antara sistem hybrid dan model generatif murni (Zara, 2023). Oleh karena itu, terdapat kebutuhan untuk mengevaluasi secara spesifik efektivitas pendekatan hybrid berbasis *rule-based* dan KNN dalam konteks industri lokal, baik dari sisi teknis (akurasi klasifikasi), operasional (penghematan biaya), maupun strategis (pengambilan keputusan manajerial).

3. Metode

Jenis dan Pendekatan Penelitian Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan tujuan mengevaluasi efektivitas integrasi algoritma K- Nearest Neighbor (KNN) ke dalam sistem chatbot rule-based dalam mengklasifikasikan tingkat kepuasan pelanggan e-commerce. Pendekatan ini dipilih untuk menguji performa model berbasis data aktual serta untuk mengukur dampak efisiensi operasional layanan pelanggan.



Gambar 1. Metode Penelitian

. Sumber dan Jenis Data

Data yang digunakan bersifat kuantitatif numerik, diperoleh dari hasil simulasi Percakapan pelanggan dalam sistem e-commerce. Dataset terdiri atas :

- X1 : Tingkat pemahaman informasi dari chatbot (%)
- X2 : Rating chatbot (skala 1–5)
- X3 : Durasi sesi percakapan (menit)
- Y : Label kepuasan pelanggan (Puas/ Tidak Puas)

Teknik dan Analisis Data

Implementasi Algoritma KNN

- Model KNN digunakan untuk memprediksi kepuasan pelanggan berdasarkan 3 fitur utama (X1, X2, X3)
- Parameter k ditetapkan sebesar 3 berdasarkan praktik standar Jarak antar data dihitung menggunakan rumus
Euclidean: $d = \sqrt{(X1a - X1b)^2 + (X2a - X2b)^2 + (X3a - X3b)^2}$
- Hasil klasifikasi ditentukan berdasarkan mayoritas kelas dari 3 tetangga terdekat.

4. Implementasi dan Perhitungan KNN

- Dataset Contoh (6 data latih dan 1 data uji)

ID	X1 (%)	X2	X3 (m)	Y
D1	92	5	4	Puas
D2	70	4	6	Puas
D3	45	3	8	Tidak Puas
D4	60	4	9	Tidak Puas
D5	85	5	7	Puas
D6	50	3	5	Tidak Puas
Uji	65	4	6	?

Tabel 1. Dataset

2. Langkah Perhitungan Euclidean ($k = 3$)

$$\text{Jarak} = \sqrt{(X1a - X1b)^2 + (X2a - X2b)^2 + (X3a - X3b)^2}$$

Hasil perhitungan jarak terhadap data uji :

- D1: $\sqrt{((65-92)^2 + (4-5)^2 + (6-4)^2)} = \sqrt{729 + 1 + 4} = 27,7$ Nearest Neighbor (KNN) dapat memberikan kontribusi nyata terhadap efisiensi operasional layanan pelanggan, khususnya dalam konteks bisnis e-commerce yang bergerak cepat dan memiliki tekanan biaya 5,8 tinggi.
- D2: $\sqrt{((65-70)^2 + 0^2 + 0^2)} = \sqrt{25} = 5,0$
- D3: $\sqrt{((65-45)^2 + 1^2 + 4)} = \sqrt{400 + 1 + 4} = 20,4$
- D4: $\sqrt{((65-60)^2 + 0^2 + 9)} = \sqrt{25 + 9} = 5,8$
- D5: $\sqrt{((65-85)^2 + 1 + 1)} = \sqrt{400 + 1 + 1} = 20,1$
- D6: $\sqrt{((65-50)^2 + 1 + 1)} = \sqrt{225 + 1 + 1} = 15,3$

Tiga tetangga terdekat: D2 (Puas), D4 (Tidak Puas), D6 (Tidak Puas) Voting mayoritas : 2 Tidak Puas, 1 Puas
Prediksi akhir : Tidak Puas

Analisis Manajerial dan ROI Evaluasi finansial dilakukan untuk menilai sejauh mana sistem chatbot hybrid berbasis rule-based + KNN memberikan nilai tambah dari sisi efisiensi operasional.

1. Perbandingan Biaya Operasional (OPEX)

Item	Sebelum Bot	Sesudah Bot
Agen Manusia	10 org $\times$ Rp5 jt = Rp50 jt/bln	4 org = Rp20 jt/bln
Biaya Lisensi & Infrastruktur	-	Rp5 jt/bln
Total OPEX/bln	Rp50 jt	Rp25 jt

Tabel 2. Perbandingan Biaya Operasional (OPEX)

Implementasi chatbot memungkinkan penghematan biaya hingga Rp25 juta per bulan atau setara dengan 50% efisiensi OPEX.

2. Estimasi Pengembalian Investasi

Asumsikan biaya pengembangan awal dan integrasi sistem sebesar Rp120 juta. Maka estimasi payback period (periode balik modal):

Payback Period = $\text{Rp120.000.000} / \text{Rp25.000.000} = 4,8$ bulan Angka ini mencerminkan pengembalian investasi yang cepat, sejalan dengan temuan berbagai studi global mengenai efektivitas chatbot dalam menekan biaya layanan pelanggan (NexGenCloud, 2024).

5. Kesimpulan

Penelitian ini menunjukkan bahwa pendekatan hybrid yang memadukan berdasarkan tiga parameter numerik utama:

Pemahaman informasi, rating chatbot, dan durasi percakapan. Meskipun belum diuji secara masif, pola-pola prediksi yang muncul konsisten dengan literatur sebelumnya: pelanggan cenderung merasa tidak puas saat durasi interaksi terlalu panjang, meskipun mereka memahami informasi yang diberikan. Namun, disinilah kritik utama muncul dimana sistem ini belum mampu menangkap dimensi psikologis dari pengalaman pelanggan secara menyeluruh.

Dengan hanya mengandalkan data numerik, chatbot hybrid ini belum bisa merespons konteks emosional, ambiguitas, atau keluhan yang kompleks. Ini berarti bahwa meskipun sistem efisien secara biaya, terdapat risiko false comfort yaitu

asumsi bahwa masalah pelanggan telah selesai, padahal ketidakpuasan masih berlangsung di balik interaksi yang terlihat singkat dan “terjawab”. Dari sudut pandang manajerial, efisiensi biaya sebesar 50% dan estimasi payback period kurang dari lima bulan adalah pencapaian yang impresif. Namun, keputusan untuk mengadopsi sistem ini sebaiknya tidak hanya didasarkan pada hasil keuangan jangka pendek, melainkan juga harus mempertimbangkan potensi dampak jangka panjang terhadap loyalitas pelanggan, citra merek, dan risiko churn akibat pengalaman digital yang dirasa “kosong” atau kurang manusiawi. Sehingga, meskipun secara teknis sistem ini layak dikembangkan, perlu pendekatan yang lebih holistik untuk memastikan bahwa efisiensi yang diperoleh tidak mengorbankan kualitas relasi pelanggan secara jangka panjang.

Saran

- 1) Penelitian selanjutnya sebaiknya menggunakan data yang lebih banyak agar hasilnya lebih representatif dan bisa menggambarkan kondisi nyata dilapangan secara lebih akurat.
- 2) Disarankan untuk mencoba membandingkan KNN dengan metode lain seperti decision tree atau regresi logistik, agar bisa diketahui mana metode yang paling cocok untuk klasifikasi kepuasan pelanggan.
- 3) Penelitian lanjutan bisa menambahkan variabel lain yang mempengaruhi kepuasan pelanggan, misalnya kecepatan balasan chatbot, jumlah pesan, atau jenis produk yang dibeli

Referensi

- [1] American Public University. (2023). Introduction to K-Nearest Neighbor Algorithm. APU Online Library. <https://www.apu.apus.edu>
- [2] IBM. (2023). The Rise of Conversational AI in Customer Experience IBM Research. <https://www.ibm.com/reports/conversational-ai>
- [3] Li, X., Chen, Y., & Huang, M. (2023). Can AI chatbots help retain customers? An integrative perspective. *Technological Forecasting and Social Change*, 194, 122527. <https://doi.org/10.1016/j.techfore.2023.122921>
- [4] Moitra, D., & Mukherjee, P. (2020). Chatbots: History, technology, and applications. ResearchGate. <https://www.researchgate.net/publication/345815999>
- [5] NexGenCloud. (2024). How AI and RAG Chatbots Cut Customer Service Costs by Millions. <https://www.nexgencloud.com/blog/case-studies>
- [6] Nugroho, T., & Lestari, D. (2024). Application of K-Nearest Neighbor Algorithm for Consumer Behaviour Identification and Product Personalisation. *Journal of Big Data Analytics*, 8(2), 45–55. <https://www.researchgate.net/publication/385507485>
- [7] Omnisend. (2024). Omnisend Ecommerce Statistics Report 2024. <https://www.omnisend.com>
- [8] Pratama, R. A., Yusuf, A., & Fitriani, L. (2024). Classification of Customer Satisfaction Using K-Nearest Neighbor. *International Journal of Computer Science and Engineering*, 12(1), 1–9. <https://journal.lembagakita.org/index.php/jsecs/article/view/2948>
- [9] Trade.gov. (2024). Indonesia E-Commerce. International Trade Administration. <https://www.trade.gov/country-commercial-guides/indonesia-ecommerce>
- [10] Zara, D. (2023). Measuring AI Chatbot ROI: Case Studies. DialZara Insights. <https://dialzara.com/blog/measuring-ai-chatbot-roi-case-studies>
- [11] Zhou, J., Wang, C., & Zhang, H. (2023). Comparison of Rule-Based Chat Bots with Different Machine Learning Methods. *Procedia Computer Science*, 217, 235–242. <https://doi.org/10.1016/j.procs.2025.04.03>