

## Chatbot Berbasis NLP dan *Multinomial Naive Bayes* untuk Otomatisasi Layanan RJ Visual Stories

Muhamad Syarif Hidayatullah <sup>1\*</sup>, Muhammad Izman Herdiansyah <sup>2</sup>, Firamon Syakti <sup>3</sup>, Ade Putra <sup>4</sup>, Ahmad Syazili <sup>5</sup>

<sup>1,2,3,4,5</sup> Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Bina Darma, Kota Palembang, Provinsi Sumatera Selatan, Indonesia.

### ABSTRACT

RJ Visual Stories faces delayed customer service due to manual handling of repetitive inquiries regarding packages, pricing, schedules, and portfolios. This study develops a conversational chatbot using Natural Language Processing (NLP) and the Multinomial Naive Bayes algorithm to provide automated customer support operating continuously (24/7). Using a Research and Development (R&D) method with a prototyping approach, the classification model was trained on 575 conversation samples across 15 intent categories utilizing TF-IDF (unigram-bigram) feature extraction. The model achieved 72.17% accuracy and a 71% weighted F1-score on the test set, with a stable average accuracy of 67.48% under 5-fold stratified cross-validation. Functional evaluation through black-box testing demonstrated a 100% success rate across 20 test scenarios. The developed system successfully automates customer inquiry handling, centralizes business information, and simplifies booking within a single chat interface, effectively eliminating administrative response delays and enhancing real-time customer engagement.

### ABSTRAK

RJ Visual Stories mengalami kendala respons lambat akibat penanganan pesan pelanggan yang masih dilakukan secara manual, terutama untuk pertanyaan repetitif seputar paket, harga, jadwal, dan portofolio. Penelitian ini mengembangkan *chatbot* percakapan berbasis *Natural Language Processing* (NLP) dan algoritma *Multinomial Naive Bayes* untuk menyediakan layanan pelanggan otomatis selama 24 jam sehari, 7 hari seminggu. Menggunakan metode *Research and Development* (R&D) dengan pendekatan model *prototyping*, model klasifikasi dilatih menggunakan 575 sampel percakapan pada 15 kategori *intent* dengan ekstraksi fitur TF-IDF (*unigram-bigram*). Model mencapai akurasi 72,17% dan *F1-score* tertimbang 71% pada data uji, dengan rata-rata akurasi 67,48% pada pengujian *5-fold stratified cross-validation*. Pengujian kelayakan fungsional (*black-box testing*) mencatatkan tingkat keberhasilan 100% pada 20 skenario uji. *Chatbot* ini berhasil mengotomatisasi layanan pertanyaan repetitif, memusatkan informasi paket, harga, jadwal, dan portofolio, serta menyederhanakan proses pemesanan dalam satu jendela obrolan, sehingga mampu memangkas kelambatan penanganan manual menjadi respons komputasi seketika (*real-time*) dan meningkatkan keterlibatan pelanggan (*customer engagement*).

### ARTICLE HISTORY

Received: 21 August 2026.  
Revised: 22 August 2026.  
Accepted: 26 September 2026.  
Available Online: 10 October 2026.  
Published: 1 December 2026.

### KEYWORDS

Chatbot; Natural Language Processing; Naive Bayes; Smart Service Automation; Customer Engagement.

### KATA KUNCI

Chatbot; Natural Language Processing; Naive Bayes; Smart Service Automation; Customer Engagement.

## 1. Pendahuluan

Transformasi digital menuntut setiap pelaku bisnis menyediakan layanan yang cepat, akurat, dan tersedia setiap saat, mengingat kecepatan merespons pelanggan kini menjadi standar keunggulan operasional utama dan bukan sekadar nilai tambah. RJ Visual Stories, sebuah unit usaha penyedia jasa dokumentasi visual dan paket pernikahan (*wedding*), menghadapi kendala keterlambatan penanganan pertanyaan repetitif calon pelanggan mengenai rincian harga, jadwal, portofolio, dan spesifikasi paket akibat seluruh proses komunikasi masih dikelola secara manual oleh administrator. Ketergantungan penuh pada tenaga manusia ini mengakibatkan proses pengecekan data terdistribusi pada beberapa aplikasi yang terpisah (*siloed data*), sehingga waktu tanggap menjadi lambat, terutama untuk pesan yang masuk di luar jam kerja operasional. Di sisi lain, sistem penjawab otomatis berbasis aturan konvensional (*rule-based*) yang kerap diadopsi cenderung bersifat kaku serta gagal memahami maksud pengguna saat pertanyaan memuat variasi format, singkatan, maupun kesalahan ketik yang tidak terdaftar dalam basis aturan (Febriano, 2026). Secara konseptual, *chatbot* atau agen percakapan (*conversational agent*) merupakan program perangkat lunak yang dirancang untuk meniru komunikasi manusia melalui media teks, memungkinkan pengguna berinteraksi secara alami dengan sistem komputasi (Prasetyo *et al.*, 2021). Program ini berfungsi sebagai layanan pelanggan virtual yang mampu merespons pertanyaan secara instan, menyampaikan rincian barang dan jasa, hingga memandu calon pembeli menyelesaikan pemesanan secara mandiri (Mahendra *et al.*, 2023). Ditinjau dari mekanisme pemrosesannya, *chatbot* terbagi menjadi pendekatan berbasis aturan yang mencocokkan kata kunci secara kaku, dan pendekatan pemrosesan bahasa alami (*Natural Language Processing/NLP*) yang memprediksi maksud pengguna secara probabilistik sehingga lebih adaptif terhadap dinamika bahasa manusia (Febriansyah *et al.*, 2025). Melalui NLP, sistem komputasi tidak hanya membaca susunan karakter, melainkan juga mengurai konteks dan makna semantik dari masukan pengguna (Mulyatun *et al.*, 2021). Integrasi kemampuan NLP dengan algoritma klasifikasi probabilistik seperti Multinomial Naive Bayes menjadi solusi yang tepat, mengingat algoritma tersebut terbukti andal dalam klasifikasi teks dengan kebutuhan korpus latih yang relatif efisien serta waktu komputasi pelatihan yang cepat (Nimatul Mamuriyah *et al.*, 2024).

Urgensi otomatisasi layanan pelanggan diperkuat oleh temuan empiris pada studi tindakan terhadap 15 UMKM lintas sektor, yang menunjukkan bahwa pengoperasian agen percakapan berbasis kecerdasan buatan mampu memangkas waktu tanggap awal (*first response time*) dari 2,1 jam menjadi 3 detik, mencapai akurasi pengenalan intensi sebesar 89,2%, serta mereduksi beban kerja operasional hingga 40% (Rachman *et al.*, 2025). Peningkatan interaktivitas ini juga berdampak langsung pada kinerja bisnis, sejalan dengan temuan Muhidin *et al.* (2023) mengenai korelasi positif yang kuat antara intensitas interaksi pelanggan melalui saluran obrolan otomatis dan pertumbuhan volume transaksi, sehingga keterlibatan pelanggan seketika (*real-time customer engagement*) berfungsi sebagai sarana komunikasi sekaligus pendorong konversi penjualan. Sejumlah riset terdahulu telah membuktikan efektivitas algoritma Multinomial Naive Bayes pada beragam ranah pemrosesan teks, meliputi klasifikasi pertanyaan akademik berbasis kemiripan kosinus (Nur Aziza *et al.*, 2023), analisis sentimen opini publik terhadap kebijakan pemerintah (Yuyun *et al.*, 2021), hingga pengelompokan masukan evaluasi akademik menggunakan pembobotan TF-IDF (Hidayah & Dodiman, 2024). Meskipun demikian, penelitian-penelitian tersebut masih berfokus pada klasifikasi informasi satu arah dan belum mengintegrasikan mesin inferensi teks ke dalam alur otomatisasi layanan terpadu yang memadukan penyediaan informasi tarif, verifikasi ketersediaan jadwal, penampilan portofolio, serta pemesanan langsung dalam satu antarmuka percakapan tunggal bagi industri jasa dokumentasi kreatif. Penelitian ini bertujuan merancang dan mengimplementasikan *chatbot* percakapan berbasis NLP dan Multinomial Naive Bayes pada platform digital RJ Visual Stories guna mewujudkan otomatisasi layanan interaktif 24/7 yang mengintegrasikan pusat informasi dan reservasi terpadu secara efisien.

## 2. Metodologi

Penelitian ini menggunakan metode *Research and Development* (R&D) dengan pendekatan model *prototyping*, dipilih karena kebutuhan fungsional sistem memerlukan siklus perancangan, evaluasi, dan penyempurnaan secara berulang bersama pihak mitra (Maulana *et al.*, 2025). Kerangka kerja *prototyping*

dijalankan melalui lima tahapan terstruktur: tahap komunikasi (*communication*) dilakukan melalui observasi dan wawancara mendalam bersama pemilik RJ Visual Stories untuk memetakan kategori *intent* percakapan serta kebutuhan operasional; tahap perancangan cepat (*quick design*) memetakan arsitektur sistem, alur dialog (*conversation flow*), dan tata letak antarmuka obrolan; tahap pembangunan prototipe (*build prototype*) mewujudkan rancangan tersebut ke dalam sistem fungsional menggunakan *framework* Laravel yang terintegrasi dengan modul inferensi *intent*; tahap evaluasi (*evaluation*) melibatkan pihak manajemen mitra untuk menguji kinerja awal sistem; serta tahap penyempurnaan (*refinement*) yang menindaklanjuti masukan evaluasi melalui perbaikan berulang hingga perangkat lunak dinyatakan stabil dan siap diuji secara formal. Data penelitian mencakup data primer berupa hasil observasi dan wawancara operasional, serta data sekunder berupa korpus riwayat percakapan nyata pelanggan melalui WhatsApp dan media sosial yang dihimpun ke dalam berkas *dataset\_chatbot\_rj\_visual\_stories.csv* sebanyak 575 sampel yang terbagi ke dalam 15 kategori *intent*. Lingkungan pengembangan perangkat lunak memanfaatkan Visual Studio Code, *framework* Laravel untuk pengelolaan *backend* dan antarmuka obrolan, MySQL sebagai sistem manajemen basis data, XAMPP sebagai peladen lokal, serta Python beserta pustaka Scikit-learn dan Sastrawi untuk *text preprocessing*, ekstraksi fitur, dan pemodelan klasifikasi. Basis data MySQL dirancang dengan tabel-tabel relasional utama, mencakup *intents*, *training\_data*, *packages*, *schedules*, *portfolios*, *bookings*, dan *conversations*. Secara arsitektural, alur operasional dimulai saat pengguna mengirimkan pesan teks atau memilih tombol respons cepat (*quick reply*) pada antarmuka *frontend*. Masukan diteruskan melalui rute HTTP POST menuju pengontrol (*controller*) Laravel untuk divalidasi. Apabila pesan bukan instruksi pembatalan, Laravel memanggil mesin inferensi Python (*predict\_intent.py*) di latar belakang.

Setelah proses inferensi menghasilkan luaran prediksi *intent* dalam format JSON, generator respons pada Laravel menentukan pola balasan, mengeksekusi kueri ke basis data MySQL (seperti ketersediaan slot jadwal atau rincian paket), menyusun komponen antarmuka, mencatat riwayat percakapan, dan mengembalikan pesan jawaban ke antarmuka pengguna. Sistem memfasilitasi dua peran pengguna, yaitu Pelanggan yang dapat mengakses informasi layanan dan mengajukan reservasi secara mandiri, serta Administrator yang berwenang memperbarui korpus data latih, melatih ulang model secara terprogram, dan memantau rekaman percakapan yang tidak teridentifikasi. Korpus data latih diproses melalui lima tahapan prapemrosesan teks (*text preprocessing*), meliputi *case folding* untuk menyeragamkan huruf menjadi kecil, pembersihan (*cleaning*) dari tanda baca dan karakter khusus, pemisahan kata (*tokenizing*), penyeragaman kata tidak baku (*slang normalization*), serta penghapusan kata umum (*stopword removal*). Teks bersih selanjutnya ditransformasikan menjadi representasi vektor numerik melalui pembobotan TF-IDF dengan rentang *n*-gram unigram dan bigram (*ngram\_range*=(1,2)) agar frasa penting seperti "terima kasih" atau "harga paket" tetap dipertahankan konteks semantiknya. Pembobotan TF-IDF memperhitungkan frekuensi kemunculan term dalam satu pesan (*Term Frequency*) dan tingkat kelangkaan term tersebut pada keseluruhan dokumen korpus (*Inverse Document Frequency*), sehingga kata-kata yang terlalu umum diberikan bobot lebih rendah dibandingkan istilah yang bersifat distingtif (Hidayah & Dodiman, 2024). Representasi vektor TF-IDF digunakan untuk melatih model klasifikasi Multinomial Naive Bayes berdasarkan Teorema Bayes dengan pembagian data menggunakan rasio 80:20, menghasilkan 460 sampel data latih dan 115 sampel data uji. Algoritma Multinomial Naive Bayes dipilih karena memperhitungkan frekuensi kemunculan term dalam dokumen sehingga adaptif terhadap ragam kalimat percakapan. Guna mengatasi masalah peluang bernilai nol (*zero probability problem*) saat menjumpai kata asing yang tidak terdapat pada data latih suatu kelas, model menerapkan teknik *Laplace smoothing* dengan menambahkan konstanta penghalusan pada perhitungan probabilitas. Sistem juga menetapkan batas ambang keyakinan (*fallback threshold*) sebesar 0,25; jika probabilitas prediksi suatu *intent* berada di bawah nilai tersebut, sistem secara otomatis mengalihkan respons ke kategori cadangan (*fallback*) "tidak\_paham" guna mencegah pemberian balasan yang keliru. Evaluasi performa klasifikasi dilakukan melalui pengujian *5-fold stratified cross-validation*, matriks konfusi (*confusion matrix*), serta pengukuran nilai metrik akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score*. Sementara itu, kelayakan fungsional sistem diuji menggunakan metode *black-box testing* dengan teknik pembagian kelas ekuivalen (*equivalence partitioning*), analisis nilai batas (*boundary value analysis*), dan analisis di luar cakupan (*out-of-scope analysis*) pada 20 skenario uji. Pengujian fungsional tersebut dibagi ke dalam dua modul kerja: modul AI Chatbot yang berfokus pada ketepatan identifikasi *intent* dan keandalan mekanisme *fallback* (16 skenario), serta modul Sistem dan Dasbor Administrator yang menguji fungsi formulir reservasi, pelatihan ulang model otomatis, kendali hak akses keamanan (*access control*), dan responsivitas tata letak antarmuka pengguna (4 skenario).

### 3. Hasil

Hasil implementasi penelitian berupa sistem *chatbot* cerdas yang terintegrasi secara penuh pada situs web RJ Visual Stories, menggabungkan *framework* Laravel sebagai pengelola antarmuka dan API dengan Python sebagai mesin inferensi komputasi NLP. Pada sisi pengguna, *widget* percakapan dilengkapi dengan indikator mengetik (*typing indicator*), tombol jawaban cepat (*quick replies*), serta formulir reservasi daring terpadu. Pada sisi pengelola, dasbor administrator menyediakan sarana pengelolaan data korpus serta tombol eksekusi pembaruan model klasifikasi (*retrain model*) secara mandiri tanpa mematikan layanan peladen (*zero downtime*). Proses penanganan masukan teks diawali dengan tahapan prapemrosesan teks (*text preprocessing*) untuk menyaring derau (*noise*) bahasa alami. Sebagai contoh, masukan percakapan calon pelanggan berupa "Halo min, brp harga paket weddingnya ya? thx!!" diproses melalui lima tahapan berturut-turut hingga menghasilkan himpunan kata dasar dan baku yang siap diekstraksi bobotnya, sebagaimana disajikan pada Tabel 1.

Tabel 1. Tahapan Prapemrosesan Teks pada Masukan Pesan Pelanggan

| No | Tahapan                                            | Hasil Transformasi Teks                                                          |
|----|----------------------------------------------------|----------------------------------------------------------------------------------|
| -  | Teks mentah ( <i>raw text</i> )                    | "Halo min, brp harga paket weddingnya ya? thx!!"                                 |
| 1  | Penyeragaman huruf ( <i>case folding</i> )         | "halo min, brp harga paket weddingnya ya? thx!!"                                 |
| 2  | Pembersihan ( <i>cleaning</i> )                    | "halo min brp harga paket weddingnya ya thx"                                     |
| 3  | Pemisahan kata ( <i>tokenizing</i> )               | "halo", "min", "brp", "harga", "paket", "weddingnya", "ya", "thx"                |
| 4  | Normalisasi slang                                  | "halo", "admin", "berapa", "harga", "paket", "weddingnya", "iya", "terima kasih" |
| 5  | Penghapusan kata henti ( <i>stopword removal</i> ) | "admin", "berapa", "harga", "paket", "weddingnya", "terima", "kasih"             |

Catatan. Data penelitian, 2026.

Himpunan kata bersih pada Tabel 1 selanjutnya ditransformasikan ke dalam representasi vektor menggunakan *TF-IDF Vectorizer* dengan konfigurasi rentang *n*-gram unigram dan bigram (*ngram\_range=(1,2)*). Konfigurasi ini memungkinkan sistem mempertahankan keterikatan semantik antarkata, seperti pada frasa "terima kasih" atau "harga paket", yang berisiko kehilangan konteks maknanya apabila hanya dipecah menjadi lema tunggal. Model klasifikasi Multinomial Naive Bayes dilatih menggunakan keseluruhan korpus percakapan yang dihimpun. Pengujian stabilitas model melalui *5-fold stratified cross-validation* mencatatkan rata-rata akurasi sebesar 67,48% dengan simpangan baku (*standard deviation*) 0,0104, mengindikasikan bahwa model memiliki konsistensi tinggi dan terhindar dari indikasi *overfitting*. Pengujian lanjutan pada 115 data uji menghasilkan akurasi sebesar 72,17%, dengan nilai rata-rata tertimbang (*weighted average*) untuk presisi sebesar 81%, *recall* sebesar 72%, dan *F1-score* sebesar 71%. Pengujian kelayakan sistem melalui *black-box testing* pada 20 skenario uji juga mencatatkan tingkat keberhasilan fungsional 100%. Ringkasan metrik performa model dan hasil evaluasi sistem disajikan pada Tabel 2.

Tabel 2. Ringkasan Hasil Evaluasi Performa Model dan Pengujian Fungsional

| Metrik Evaluasi                 | Metode Pengujian                               | Nilai Capaian      |
|---------------------------------|------------------------------------------------|--------------------|
| Akurasi                         | 5-Fold stratified cross-validation (rata-rata) | 67,48% (SD 0,0104) |
| Akurasi                         | Data uji (test set, $n=115$ )                  | 72,17%             |
| Presisi (weighted average)      | Confusion matrix                               | 81%                |
| Recall (weighted average)       | Confusion matrix                               | 72%                |
| F1-score (weighted average)     | Confusion matrix                               | 71%                |
| Tingkat keberhasilan fungsional | Black-box testing (20 skenario uji)            | 100%               |

Catatan. Data penelitian, 2026.

Distribusi keseluruhan korpus percakapan yang dihimpun ( $n=575$ ) menunjukkan adanya ketimpangan frekuensi (*class imbalance*) antar-kategori *intent*. Kategori *tanya\_paket* merupakan kelas dengan proporsi data terbesar, yaitu 62 sampel (10,78%), disusul oleh *tanya\_harga* dengan 59 sampel (10,26%) dan *salam*

dengan 56 sampel (9,74%). Sebaliknya, kategori ucapan\_terima\_kasih memiliki frekuensi terendah, yaitu sebanyak 25 sampel (4,35%), disusul oleh tanya\_promo dan tanya\_proses\_editing yang masing-masing memiliki 26 sampel (4,52%). Rincian persebaran frekuensi dan proporsi dari keseluruhan 575 data percakapan pada ke-15 kategori *intent* disajikan pada Tabel 3.

Tabel 3. Distribusi Frekuensi Sampel Percakapan per Kategori Intent pada Korpus Penelitian

| Kategori <i>Intent</i> | Jumlah Sampel (nn) | Proporsi (%) |
|------------------------|--------------------|--------------|
| tanya_paket            | 62                 | 10,78%       |
| tanya_harga            | 59                 | 10,26%       |
| salam                  | 56                 | 9,74%        |
| tanya_lokasi           | 45                 | 7,83%        |
| tanya_kontak           | 40                 | 6,96%        |
| cara_pesan             | 40                 | 6,96%        |
| tanya_ketersediaan     | 38                 | 6,61%        |
| tanya_portofolio       | 36                 | 6,26%        |
| tanya_pembayaran       | 35                 | 6,09%        |
| komplain               | 30                 | 5,22%        |
| tanya_layanan          | 30                 | 5,22%        |
| selamat_tinggal        | 27                 | 4,70%        |
| tanya_promo            | 26                 | 4,52%        |
| tanya_proses_editing   | 26                 | 4,52%        |
| ucapan_terima_kasih    | 25                 | 4,35%        |
| Total Korpus           | 575                | 100,00%      |

Catatan. Data penelitian, 2026. Korpus dipartisi menggunakan rasio 80:20 menjadi 460 sampel data latih dan 115 sampel data uji untuk kebutuhan pemodelan klasifikasi.

Pengujian fungsional sistem melalui metode *black-box testing* dilakukan pada dua modul operasional, yaitu modul AI Chatbot yang berfokus pada ketepatan inferensi masukan pengguna dan modul Sistem serta Dasbor Administrator. Evaluasi terhadap modul AI Chatbot mencakup 16 skenario masukan pengguna, sebagaimana disajikan pada Tabel 4.

Tabel 4. Hasil Pengujian Black-Box pada Modul AI Chatbot

| No    | Skenario Pengujian      | Masukan Pengguna (Data Uji)                           | <i>Intent</i> Terprediksi | Probabilitas | Status Hasil |
|-------|-------------------------|-------------------------------------------------------|---------------------------|--------------|--------------|
| TC-01 | AI - Salam              | "Halo selamat pagi admin RJ Visual Stories"           | salam                     | 47,89%       | Valid        |
| TC-02 | AI - Tanya Paket        | "Ada paket wedding apa aja kak?"                      | tanya_paket               | 49,94%       | Valid        |
| TC-03 | AI - Tanya Harga        | "Berapa harga paket silver berapaan rupiah?"          | tanya_harga               | 64,12%       | Valid        |
| TC-04 | AI - Tanya Lokasi       | "Lokasi studio domisili di mana kak?"                 | tanya_lokasi              | 46,68%       | Valid        |
| TC-05 | AI - Tanya Kontak       | "Minta nomor whatsapp admin yang bisa dihubungi dong" | tanya_kontak              | 47,13%       | Valid        |
| TC-06 | AI - Cara Pesan         | "Mulai booking sekarang kak"                          | cara_pesan                | 38,52%       | Valid        |
| TC-07 | AI - Kata Kunci Singkat | "booking"                                             | cara_pesan                | 39,70%       | Valid        |
| TC-08 | AI - Form Booking       | "isi form booking"                                    | cara_pesan                | 43,47%       | Valid        |
| TC-09 | AI - Ketersediaan       | "Apakah slot tanggal 20 desember masih"               | tanya_ketersediaan        | 42,65%       | Valid        |

|       |                                              |                                                            |                                 |        |       |  |
|-------|----------------------------------------------|------------------------------------------------------------|---------------------------------|--------|-------|--|
|       |                                              | kosong?"                                                   |                                 |        |       |  |
| TC-10 | AI - Portofolio                              | "Bisa liat sampel contoh foto wedding sebelumnya?"         | tanya_portofolio                | 26,54% | Valid |  |
| TC-11 | AI - Proses Editing                          | "Berapa hari estimasi proses edit foto dan video selesai?" | tanya_proses_editing            | 28,45% | Valid |  |
| TC-12 | AI - Pembayaran                              | "Sistem pembayaran DP berapa persen dan lewat apa?"        | tanya_pembayaran                | 40,64% | Valid |  |
| TC-13 | AI - Komplain                                | "Hasil fotonya jelek dan pelayanan sangat mengecewakan"    | komplain                        | 35,74% | Valid |  |
| TC-14 | AI - Terima Kasih                            | "Terima kasih banyak infonya ya kak"                       | ucapan_terima_kasih             | 31,12% | Valid |  |
| TC-15 | AI - Selamat Tinggal                         | "Oke makasih dadah sampai jumpa"                           | selamat_tinggal                 | 31,63% | Valid |  |
| TC-16 | AI - Di Luar Cakupan ( <i>Out-of-Scope</i> ) | "Tolong jelaskan resep membuat nasi goreng pedas manis"    | tidak_paham ( <i>fallback</i> ) | 16,53% | Valid |  |

Catatan. Data penelitian, 2026. Nilai batas ambang keyakinan (*fallback threshold*) ditetapkan sebesar 25% (0,25).

Pada skenario TC-16, keandalan mekanisme batas ambang (*fallback threshold*) teruji secara efektif saat sistem menerima masukan yang berada di luar konteks layanan fotografi dan dokumentasi. Nilai probabilitas prediksi yang diperoleh hanya mencapai 16,53%, berada di bawah ambang batas keyakinan 0,25 (25%), sehingga sistem secara adaptif mengalihkan jawaban ke respons cadangan (*fallback*) berupa pernyataan belum memahami maksud pengguna, alih-alih memberikan tanggapan yang tidak relevan. Pengujian fungsional modul Sistem dan Dasbor Administrator dilakukan terhadap empat aspek logika operasional, keamanan rute, dan tata letak antarmuka, sebagaimana dirangkum pada Tabel 5.

Tabel 5. Hasil Pengujian Black-Box pada Modul Sistem dan Dasbor Administrator

| No    | Fitur yang Diuji     | Aksi Pengujian (Masukan)                                                              | Hasil yang Diharapkan                                                                                                           | Status Hasil |
|-------|----------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|--------------|
| TC-17 | Formulir reservasi   | Mengisi data pemesanan pada <i>widget</i> obrolan lalu menekan Kirim Reservasi        | Data tersimpan pada basis data dengan status menunggu ( <i>pending</i> ) dan mencatat jadwal secara otomatis                    | Valid        |
| TC-18 | Pelatihan ulang AI   | Menekan tombol pembaruan model pada dasbor administrator                              | Skrip inferensi Python dijalankan di latar belakang untuk memperbarui model tanpa menghentikan layanan ( <i>zero downtime</i> ) | Valid        |
| TC-19 | Hak akses admin      | Mengakses rute <i>/admin/chatbot</i> secara langsung tanpa melalui proses autentikasi | Sistem menolak akses dan secara otomatis mengalihkan pengguna ke halaman <i>login</i>                                           | Valid        |
| TC-20 | Tata letak antarmuka | Membuka halaman admin pada layar beresolusi lebar (1920 times 1080 piksel)            | Wadah antarmuka tertata presisi di tengah layar secara simetris dan responsif                                                   | Valid        |

Catatan. Data penelitian, 2026.

Berdasarkan hasil pengujian pada Tabel 4 dan Tabel 5, seluruh skenario uji (20 kasus uji) berhasil diselesaikan tanpa ditemukan kegagalan logika sistem maupun celah fungsional. Guna mengkaji dampak

operasional sebelum dan setelah otomatisasi diterapkan pada RJ Visual Stories, Tabel 6 menyajikan komparasi indikator layanan kunci.

Tabel 6. Perbandingan Indikator Layanan Sebelum dan Setelah Implementasi Chatbot

| Indikator Layanan                      | Sebelum Implementasi (Penanganan Manual)                                                                             | Setelah Implementasi ( <i>Chatbot</i> NLP)                                                            |
|----------------------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| Waktu tanggap ( <i>response time</i> ) | Bergantung pada ketersediaan dan antrean admin (rentan keterlambatan respons)                                        | Diproses secara komputasi seketika ( <i>real-time</i> ) saat pesan diterima peladen                   |
| Ketersediaan operasional               | Terbatas pada jam kerja staf admin (pukul 09.00–17.00 WIB)                                                           | Beroperasi secara kontinu selama 24 jam sehari, 7 hari seminggu (24/7)                                |
| Aksesibilitas informasi                | Admin mencari informasi harga, paket, dan portofolio secara terpisah pada berbagai aplikasi ( <i>siloeled data</i> ) | Seluruh data harga, jadwal, portofolio, dan formulir terintegrasi pada satu antarmuka obrolan tunggal |

Catatan. Data penelitian, 2026.

Perbandingan indikator pada Tabel 6 menegaskan adanya transformasi signifikan pada efisiensi operasional layanan RJ Visual Stories melalui penerapan sistem percakapan otomatis. Ketergantungan terhadap penanganan staf admin yang rentan terhadap penumpukan antrean pesan berhasil dialihkan menjadi pemrosesan komputasi seketika tanpa jeda waktu tunggu manual. Selain memperluas aksesibilitas interaksi pelanggan di luar jam operasional konvensional, integrasi basis data paket, jadwal, portofolio, dan reservasi pada satu jendela percakapan tunggal secara efektif meniadakan fragmentasi data (*siloeled data*) yang sebelumnya menghambat proses pelayanan.

#### 4. Pembahasan

Capaian nilai presisi rata-rata tertimbang (*weighted average precision*) sebesar 81% menunjukkan bahwa sistem memiliki keandalan tinggi dalam memastikan kebenaran klasifikasi saat menetapkan suatu *intent*. Tingkat ketepatan ini didukung secara efektif oleh penerapan batas ambang keyakinan (*fallback threshold*) sebesar 0,25, yang menahan sistem agar tidak memaksakan luaran spekulatif saat nilai probabilitas posterior berada pada tingkat rendah, sehingga meminimalkan risiko pemberian respons yang keliru. Namun demikian, selisih yang terbentuk antara presisi (81%) dan *recall* (72%) mengindikasikan adanya sejumlah pesan pengguna aktual yang luput teridentifikasi dengan tepat (*false negatives*). Kondisi ini paling rentan terjadi pada kategori-kategori yang memiliki irisan semantik leksikal tinggi, serta pada kelas minoritas dengan ketersediaan data latih yang terbatas, seperti ucapan\_terima\_kasih (25 sampel) jika dibandingkan dengan tanya\_paket (62 sampel). Berdasarkan telaah mendalam terhadap matriks konfusi (*confusion matrix*), kategori tanya\_paket cenderung bertindak sebagai kelas penampung (*majority sink*) dari kesalahan klasifikasi *intent* lain, menerima distribusi galat dari tanya\_kontak, tanya\_portofolio, tanya\_layanan, tanya\_pembayaran, hingga cara\_pesan. Pola kebingungan mesin (*machine confusion*) paling kentara muncul antara kelas tanya\_paket dan tanya\_harga, yang disebabkan oleh tingginya ko-okurensi leksikal dalam satu struktur kalimat pengguna, seperti kemunculan kata "harga" dan "paket" secara bersamaan. Di bawah asumsi independensi fitur algoritma Multinomial Naive Bayes, kemunculan lema-lema yang mendominasi kelas mayoritas akan melipatgandakan probabilitas apriori kelas tersebut. Sebaliknya, kategori dengan korpus kata kunci yang sangat spesifik dan minim tumpang tindih leksikal, seperti tanya\_ketersediaan yang didominasi oleh kata waktu dan tanggal, memperlihatkan akurasi klasifikasi yang jauh lebih stabil. Fenomena ini menegaskan bahwa sensitivitas inferensi Naive Bayes sangat bergantung pada kekhasan perbendaharaan kata (*lexical distinctiveness*) serta proporsionalitas sebaran data latih pada masing-masing kelas. Temuan ini selaras dengan penelitian Nimatul Mamuriyah *et al.* (2024) serta Syaefuloh & Sugianto (2025), yang mengonfirmasi bahwa kendati Multinomial Naive Bayes sangat efisien dalam tugas kategorisasi percakapan layanan informasi, algoritma ini memiliki kerentanan performa terhadap masalah ketidakseimbangan kelas (*class imbalance*). Dari perspektif rekayasa perangkat lunak, perolehan tingkat keberhasilan 100% pada pengujian *black-box testing* di seluruh 20 skenario uji

membuktikan bahwa arsitektur integrasi *end-to-end* telah beroperasi secara andal pada lingkungan produksi. Koordinasi antara antarmuka pengguna, pengontrol *backend* Laravel, pemanggilan skrip inferensi Python, pengambilan data relasional pada MySQL, hingga eksekusi pelatihan ulang model mandiri (*automated retraining*) pada modul dasbor administrator mampu berjalan tanpa memicu kegagalan sistem maupun waktu henti (*downtime*). Keberhasilan skenario TC-19 secara khusus memvalidasi ketahanan kendali hak akses (*access control*), memastikan bahwa antarmuka konfigurasi korpus dan pemodelan AI terlindungi dari manipulasi data latih oleh pengguna yang tidak terautentikasi.

Implementasi otomatisasi layanan cerdas (*smart service automation*) ini memberikan dampak nyata terhadap pola operasional RJ Visual Stories, sebagaimana ditunjukkan pada evaluasi komparatif Tabel 6. Ketergantungan terhadap penanganan staf secara manual yang rentan mengalami keterlambatan respons akibat antrean jam sibuk berhasil dieliminasi menjadi penanganan otomatis berkecepatan komputasi seketika (*real-time*), sekaligus meniadakan restriksi jam kerja konvensional menjadi operasional kontinu 24 jam sehari. Kendati penelitian ini belum mengukur latensi interaksi secara kuantitatif sebagaimana tolok ukur (*benchmark*) reduksi waktu tanggap dari 2,1 jam menjadi 3 detik yang dipaparkan dalam studi *action research* pada 15 UMKM oleh Rachman *et al.* (2025), pergeseran kualitatif yang dicapai RJ Visual Stories memperlihatkan arah reduksi beban kerja operator yang ekuivalen. Penyatuan fungsi interaksi percakapan dengan fitur jawaban cepat (*quick replies*) dan formulir pemesanan terpadu di dalam satu jendela tampilan (*single interface*) turut memperkuat keterlibatan pelanggan seketika (*real-time customer engagement*). Calon pelanggan tidak lagi diharuskan berpindah saluran komunikasi atau membuka formulir eksternal untuk melengkapi data reservasi. Dampak positif integrasi interaksi ini sejalan dengan temuan Muhidin *et al.* (2023), yang menegaskan korelasi positif antara kelancaran alur percakapan sistem otomatis dengan peningkatan konversi layanan bisnis. Dengan demikian, konvergensi antara ketepatan inferensi teks, kecepatan komputasi, dan kepraktisan antarmuka pengguna menjadi determinan utama dalam memperkuat persepsi keandalan layanan digital RJ Visual Stories di mata konsumen.

## 5. Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan sistem *AI-powered conversational chatbot* pada ekosistem digital RJ Visual Stories dengan memanfaatkan teknik pemrosesan bahasa alami (*Natural Language Processing*) dan algoritma Multinomial Naive Bayes. Melalui pemanfaatan korpus 575 sampel percakapan yang terdistribusi ke dalam 15 kategori *intent*, model klasifikasi menghasilkan performa yang konsisten dengan akurasi pengujian sebesar 72,17%, *F1-score* rata-rata tertimbang sebesar 71%, serta ketepatan presisi sebesar 81%. Mekanisme batas ambang keyakinan (*fallback threshold*) sebesar 0,25 terbukti efektif mencegah pemberian luaran spekulatif terhadap masukan di luar konteks layanan bisnis fotografi (*out-of-scope*). Pengujian fungsional sistem menggunakan metode *black-box testing* mencatatkan tingkat keberhasilan 100% pada 20 skenario uji, memvalidasi bahwa integrasi *end-to-end* antara antarmuka percakapan pengguna, pemanggilan mesin inferensi Python, manajemen data reservasi MySQL pada *framework* Laravel, hingga modul pembaruan model mandiri (*automated retraining*) pada dasbor administrator telah beroperasi sesuai spesifikasi fungsional. Implementasi otomatisasi layanan cerdas (*smart service automation*) ini secara kualitatif berhasil mentransformasikan layanan interaksi pelanggan dari ketergantungan terhadap antrean staf manual menjadi penanganan komputasi seketika (*real-time*) selama 24 jam penuh tanpa hambatan fragmentasi data. Meskipun kelayakan fungsional dan kestabilan inferensi telah tercapai, sistem masih memiliki keterbatasan pada ketimpangan distribusi data latih yang memengaruhi kepekaan (*recall*) pada kategori *intent* minoritas. Untuk pengembangan sistem selanjutnya, disarankan melakukan penyeimbangan korpus data latih melalui penambahan variasi leksikal pada kelas minoritas, mengimplementasikan teknik penanganan ketidakseimbangan kelas (*class re-weighting* atau SMOTE untuk teks), serta melaksanakan pengujian beban (*load testing*) dan audit keamanan data transaksi secara komprehensif guna menjamin keandalan sistem dalam menangani lonjakan lalu lintas pengguna berskala masif.

## Ucapan Terima Kasih

Penulis menyampaikan terima kasih kepada Ir. Muhammad Izman Herdiansyah, ST, MM, PhD, IPU selaku pembimbing atas arahan dan bimbingan selama penelitian, kepada Ratna Juwita selaku pemilik RJ Visual Stories atas kesempatan dan data yang diberikan untuk keperluan penelitian, serta kepada Program Studi Teknik Informatika, Fakultas Sains Teknologi, Universitas Bina Darma atas dukungan fasilitas akademik selama penyusunan artikel ini.

## Referensi

- Febriano, F. (2026). Penerapan machine learning pada chatbot sebagai customer support untuk memberikan pendekatan problem solving. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 10(3), 2415–2422.
- Febriansyah, M. R., Hertantyo, G. B., & Wilonotomo. (2025). Implementasi chatbot sebagai virtual assistant: Systematic literature review. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 9(4), 3840–3847.
- Hidayah, N., & Dodiman, D. (2024). Implementasi algoritma Multinomial Naïve Bayes, TF-IDF dan confusion matrix dalam pengklasifikasian saran monitoring dan evaluasi mahasiswa terhadap dosen Teknik Informatika Universitas Dayanu Ikhsanuddin. *Jurnal Akademik Pendidikan Matematika*, 10(1), 8–15. <https://doi.org/10.55340/japm.v10i1.1491>
- Mahendra, M. D., Rasywir, E., & Prasasti, L. (2023). Implementasi chatbot sebagai layanan customer service pada Kafe Duniawi Coffee. *SAINTEKS: Jurnal Teknologi Komputer dan Sains*, 1(1), 25–32.
- Maulana, M. S., Anshori, Y., Azhar, R., Laila, R., & Lapatta, N. T. (2025). Implementasi pembobotan TF-IDF pada chatbot Telegram untuk sistem layanan informasi. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 10(3), 1869–1877. <https://doi.org/10.29100/jupi.v10i3.6314>
- Muhidin, A., Danny, M., & Rilvani, E. (2023). Algoritme Multinomial Naïve Bayes pada aplikasi chatbot layanan informasi berbasis teks. *Jurnal Pelita Bangsa*, 5(2), 45–53.
- Mulyatun, S., Utama, H., & Mustopa, A. (2021). Pendekatan natural language processing pada aplikasi chatbot sebagai alat bantu customer service. *Jurnal Sistem Informasi dan Informatika*, 2(2), 88–95.
- Nimatul Mamuriyah, Haeruddin, & Hero, H. (2024). Pembangunan chatbot interaktif dengan menggunakan algoritma Naive Bayes. *Informatika: Jurnal Teknik Informatika dan Multimedia*, 4(2), 82–94. <https://doi.org/10.51903/informatika.v4i2.864>
- Nur Aziza, R., Yosrita, E., Rahma, F. N., Tiara, S. A., & Syeh, R. A. (2023). Pembangunan aplikasi dan klasifikasi pertanyaan chatbot informasi akademik menggunakan metode cosine similarity dan Naïve Bayes. *KILAT*, 12(2), 241–252. <https://doi.org/10.33322/kilat.v12i2.1921>
- Prasetyo, V. R., Benarkah, N., & Chrisintha, V. J. (2021). Implementasi natural language processing dalam pembuatan chatbot pada Program Information Technology Universitas Surabaya. *Teknika*, 10(2), 114–121. <https://doi.org/10.34148/teknika.v10i2.370>
- Rachman, M. R., Rosidin, M., & Sulistyo, W. Y. (2025). Implementasi metode Retrieval Augmented Generation pada chatbot untuk otomatisasi layanan pelanggan kontrakan. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 11(2), 1201–1208.

- Rahmat Hidayat, C., Sumaryana, Y., Anwar, D. S., Linda, A., Fadilah, N., & Randika Saputra, M. (2025). Pengembangan virtual assistant (chatbot) berbasis NLP untuk portal informasi terpadu pariwisata Tasikmalaya. *CSRID (Computer Science Research and Information Technology) Journal*, 17(1), 136–144. <https://doi.org/10.22303/csrid.v17i1.620>
- Rizqi, R. M., Harshavardana, D. A., & Osmond, S. V. (2025). Integrasi algoritma NLP untuk peningkatan kecerdasan chatbot: Kajian literatur dan analisis perkembangan terkini. *Jurnal Riset Teknik Komputer*, 2(4), 157–171. <https://doi.org/10.69714/c31tr541>
- Syaefuloh, H., & Sugianto, C. A. (2025). Aplikasi chatbot berbasis NLP untuk layanan akademik studi kasus Fakultas Hukum Unisba. *COMPETITIVE: Jurnal Inovasi Dan IPTEKS*, 10(1), 45–54.
- Yuyun, Hidayah, N., & Sahibu, S. (2021). Algoritma Multinomial Naïve Bayes untuk klasifikasi sentimen pemerintah terhadap penanganan Covid-19 menggunakan data Twitter. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(4), 820–826. <https://doi.org/10.29207/resti.v5i4.3146>