Jurnal Ilmu Komputer dan Sistem Informasi (JIRSI) **e-ISSN 2830-3954**
Volume: 5, Nomor: 3, September 2026: 1471-1482 **p-ISSN 2830-6031**
https://jurnal.unity-academy.sch.id/index.php/jirsi/index

Perancangan Chatbot Layanan Kependudukan Berbasis Hybrid Rule-Based dan NLP dengan Decision Layer

Design of a Hybrid Chatbot Based on Rules and NLP for Civil Registration Services with a Decision-Making Layer

Heru manudin wibowo[1], Saidi Ramadan Siregar [2], Sri Wanti Nasution, S.E, M.M[3]
[1,3] Program Studi Teknik Informatika, Universitas Budi Darma
Program Studi Teknik Informatika, Universitas Budi Darma
Email: [1]herumahasiswa95@gmail.com, [2]saidiramadan89@email.com, [3,]
sri.wanty.nasution@gmail.com

Abstrak

Layanan administrasi kependudukan memiliki volume permintaan informasi yang tinggi, namun akses masyarakat terhadap informasi prosedural pada Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil) Kota Medan masih terbatas pada jam operasional dan kehadiran petugas. Penelitian ini merancang chatbot layanan kependudukan berbasis pendekatan *hybrid* yang menggabungkan metode *Rule-Based* dan *Natural Language Processing* (NLP). Kebaruan penelitian terletak pada penerapan Decision Layer, yaitu lapisan pengambil keputusan yang merutekan pertanyaan secara berjenjang: pencocokan pola *Rule-Based*, resolusi konteks, klasifikasi NLP berbasis *TF-IDF* dan *Cosine Similarity*, serta *fallback response* apabila tingkat keyakinan berada di bawah ambang (threshold τ). Klasifikasi intent menggunakan pendekatan 1-*Nearest Neighbor* pada korpus 380 utterance yang terbagi ke dalam 19 intent. Dalam kondisi klasifikasi paksa (*forced-choice*), pengujian dengan confusion matrix menghasilkan *accuracy* 60,53%, yang mencerminkan keterbatasan pendekatan leksikal pada *utterance* pendek berbahasa Indonesia. Kalibrasi ambang berdasarkan indeks Youden menetapkan τ = 0,60. Dibandingkan konfigurasi tanpa Decision Layer yang memaksakan jawaban atas seluruh pertanyaan — termasuk menjawab keliru seluruh pertanyaan di luar cakupan — konfigurasi dengan *Decision Layer* mampu menolak 80% pertanyaan di luar cakupan sekaligus meloloskan 69,74% pertanyaan valid. Analisis lanjutan menunjukkan bahwa penyaringan berbasis ambang tidak meningkatkan akurasi jawaban di dalam cakupan, sehingga kontribusi Decision Layer terletak pada pengendalian terhadap pertanyaan di luar cakupan, bukan pada peningkatan akurasi klasifikasi. Pengujian penerimaan terhadap 31 responden menghasilkan indeks 83,92% (kategori Baik). Hasil menunjukkan bahwa Decision Layer berperan penting dalam mengendalikan penyampaian informasi sehingga sistem tidak memaksakan jawaban yang berpotensi keliru terhadap pertanyaan di luar cakupan. Penelitian ini berkontribusi pada pengembangan sistem layanan informasi publik yang adaptif dan terkendali.

***Kata kunci**:* chatbot, rule-based, natural language processing, decision layer, layanan kependudukan

Abstract

Population administration services face a high volume of information requests, yet public access to procedural information at the Population and Civil Registration Office (Disdukcapil) of Medan City remains limited to operating hours and staff availability. This study designs a

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

population service chatbot using a hybrid approach that combines Rule-Based methods and Natural Language Processing (NLP). Its novelty lies in a Decision Layer that routes queries hierarchically: Rule-Based pattern matching, context resolution, NLP classification based on TF-IDF and Cosine Similarity, and a fallback response when confidence falls below a threshold (τ). Intent classification uses a 1-Nearest Neighbor approach on a corpus of 380 utterances across 19 intents. Under forced-choice classification, testing with a confusion matrix yields an accuracy of 60.53%, reflecting the limitations of a lexical approach on short Indonesian-language utterances. Threshold calibration based on the Youden index sets τ = 0.60. Compared to a configuration without the Decision Layer—which forces an answer to every query and thus answers all out-of-scope queries incorrectly—the Decision Layer rejects 80% of out-of-scope queries while admitting 69.74% of valid ones. Further analysis shows that threshold-based filtering does not improve in-scope accuracy; the Decision Layer's contribution therefore lies in controlling out-of-scope queries rather than improving classification accuracy. User acceptance testing with 31 respondents produced an index of 83.92% (Good category). These results show that the Decision Layer is important for controlling information delivery, preventing the system from forcing potentially erroneous answers. This research contributes to adaptive, controlled public information service systems.

***Keywords**: chatbot, population services, Rule-Based, Natural Language Processing, Decision Layer, TF-IDF, Cosine Similarity*

1. PENDAHULUAN

Transformasi digital pada sektor pelayanan publik telah menjadi agenda strategis pemerintah Indonesia dalam kerangka Sistem Pemerintahan Berbasis Elektronik (SPBE) [1]. Salah satu layanan yang bersinggungan langsung dengan seluruh warga negara adalah administrasi kependudukan, yang mencakup pengurusan dokumen seperti Kartu Tanda Penduduk Elektronik (KTP-el), Kartu Keluarga (KK), serta akta pencatatan sipil [2]. Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil) Kota Medan menghadapi tingginya kebutuhan informasi masyarakat, namun akses terhadap informasi prosedural masih terbatas pada jam operasional dan kehadiran petugas. Kondisi ini menyebabkan masyarakat kesulitan memperoleh informasi di luar jam kerja, sekaligus membebani petugas dengan pertanyaan rutin yang berulang[3], [4].

Pemanfaatan *chatbot* menjadi solusi yang banyak dikembangkan untuk meningkatkan aksesibilitas layanan publik. Dalam pengembangannya, terdapat dua pendekatan utama [5]. Pendekatan *Rule-Based* unggul dalam merespons pertanyaan terstruktur secara presisi, tetapi kaku terhadap variasi Bahasa [4] . Sebaliknya, pendekatan *Natural Language Processing* (NLP) mampu memahami keberagaman bahasa, tetapi berisiko memberikan respons yang kurang konsisten pada informasi administratif resmi [6]. Karakteristik yang saling melengkapi inilah yang mendasari kebutuhan akan pendekatan *hybrid* [7][8].

Beberapa penelitian terkait telah dilakukan. Penelitian[6]. mengkaji persepsi masyarakat terhadap *chatbot* layanan publik kota. Penelitian [9] mengembangkan *chatbot context-aware* untuk percakapan *multi-turn*. Penelitian [10] menerapkan NLP pada layanan pemerintahan digital. Namun, telaah terhadap penelitian terdahulu menunjukkan beberapa kesenjangan (research gap). Pertama, sebagian besar penelitian masih menerapkan Rule-Based atau NLP secara terpisah, sehingga tidak dapat sekaligus menjaga presisi pada pertanyaan baku dan keluwesan

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

terhadap variasi bahasa. Kedua, penelitian yang mengadopsi pendekatan hybrid pada umumnya menggabungkan kedua modul tanpa mekanisme pengambilan keputusan yang eksplisit; perutean antar-modul kerap bersifat statis dan tidak memiliki kendali terhadap tingkat keyakinan, sehingga sistem tetap memaksakan jawaban meskipun pertanyaan berada di luar cakupan. Ketiga, evaluasi empiris penerapan chatbot pada konteks pemerintahan daerah, khususnya yang mengukur kemampuan sistem menolak pertanyaan di luar cakupan, masih terbatas.

Berdasarkan uraian tersebut, penelitian ini merancang dan mengimplementasikan chatbot layanan kependudukan berbasis Rule-Based dan NLP dengan kebaruan pada penerapan Decision Layer. Berbeda dari perutean hybrid statis pada penelitian terdahulu, Decision Layer merupakan lapisan yang mengorkestrasi kedua pendekatan secara berjenjang dengan pengambilan keputusan berbasis ambang keyakinan (threshold τ): sistem meneruskan jawaban hanya ketika keyakinan melampaui ambang, dan mengalihkannya ke fallback ketika berada di bawah ambang. Mekanisme inilah yang menjadi justifikasi kebaruan penelitian, karena menjawab langsung celah kedua, yakni ketiadaan kendali eksplisit atas perutean antar-modul

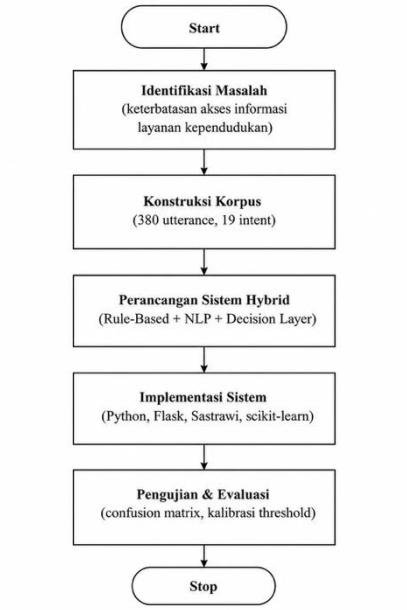
Tujuan penelitian ini adalah menghasilkan sistem yang mampu menjaga presisi pada pertanyaan baku, tetap luwes terhadap variasi bahasa, sekaligus terkendali ketika berhadapan dengan pertanyaan di luar cakupan. Kontribusi utama penelitian dinilai bukan pada peningkatan akurasi klasifikasi, melainkan pada kemampuan pengendalian yang diukur secara empiris melalui kalibrasi ambang (indeks Youden) serta tingkat penolakan pertanyaan di luar cakupan, sehingga sistem tidak memaksakan jawaban yang berpotensi keliru.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan *Design Science Research* yang berorientasi pada perancangan dan evaluasi artefak. Tahapan penelitian ditunjukkan pada Gambar 1, yang meliputi identifikasi masalah, konstruksi korpus, perancangan sistem, implementasi, dan pengujian.

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)



Gambar 1 Tahapan Penelitian

Tahap pertama adalah identifikasi masalah, yaitu keterbatasan akses masyarakat terhadap informasi layanan kependudukan [11]. Tahap kedua adalah konstruksi korpus, yaitu penyusunan data ungkapan (*utterance*) berdasarkan pertanyaan layanan kependudukan. Tahap ketiga adalah perancangan sistem *hybrid* yang terdiri atas modul *Rule-Based*, modul NLP, dan *Decision Layer [12]*. Tahap keempat adalah implementasi sistem menggunakan bahasa pemrograman Python. Tahap kelima adalah pengujian akurasi klasifikasi *intent* menggunakan *confusion matrix* serta kalibrasi ambang keyakinan [13].

2.2 Konstruksi Korpus dan Sampel Data

Sampel data dikumpulkan menggunakan teknik *purposive sampling* berdasarkan kriteria pertanyaan yang berkaitan langsung dengan layanan informasi kependudukan [14] . Korpus terdiri atas 380 *utterance* yang dikelompokkan ke dalam 19 *intent*, dengan setiap *intent* diwakili oleh 20 variasi ungkapan. Untuk keperluan evaluasi, korpus dibagi menjadi data latih dan data uji dengan rasio 80:20 [15] secara *stratified*, sehingga diperoleh 304 *utterance* data latih dan 76 *utterance* data uji. Selain itu, disiapkan 30 *utterance* di luar cakupan (*out-of-scope*) untuk kalibrasi ambang keyakinan. Distribusi *intent* disajikan pada Tabel 1 [16].

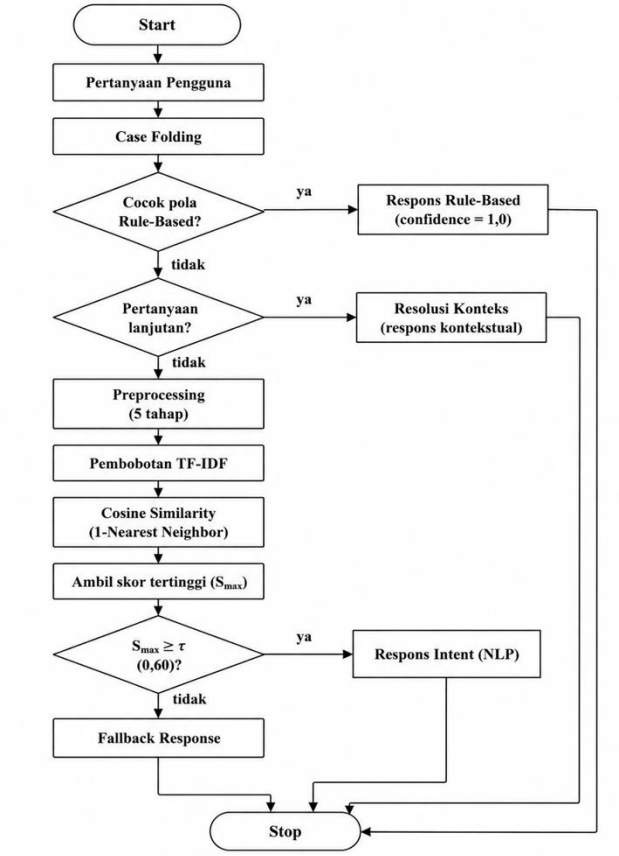
Tabel 1 Distribusi Sampel Data Berdasarkan Intent

Kelompok	Jumlah Intent	Jumlah Utterance
Percakapan Umum	3	60
Informasi Instansi	1	20
Kartu Keluarga	3	60
KTP-el	3	60
KIA	2	40
Akta Pencatatan Sipil	7	140
Total	**19**	**380**

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

2.3. Arsitektur Sistem Hybrid

Sistem dirancang dengan arsitektur *hybrid* yang mengorkestrasi modul *Rule-Based* dan NLP melalui *Decision Layer [17]*. Alur pemrosesan pertanyaan ditunjukkan pada Gambar 2. Pertanyaan pengguna terlebih dahulu dicocokkan dengan pola deterministik pada modul *Rule-Based [18]*. Apabila tidak ditemukan kecocokan, sistem memeriksa apakah pertanyaan merupakan pertanyaan lanjutan melalui resolusi konteks. Apabila bukan, pertanyaan diproses oleh modul NLP [19].



Gambar 2 Arsitektur Sistem Hybrid dengan Decision Layer

Pada modul NLP, teks masukan melalui tahap pra-pemrosesan yang meliputi *case folding*, tokenisasi, normalisasi, *stopword removal*, dan *stemming* menggunakan pustaka Sastrawi [10]. Teks hasil pra-pemrosesan direpresentasikan sebagai vektor menggunakan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF), sebagaimana Persamaan (1) dan (2) [20]

$$W(t,d)=tf(t,d)\times idf(t) \quad (1)$$

$$Id f(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

Dengan *tf(t,d)* adalah frekuensi term *t* pada dokumen *d*, *N* adalah jumlah dokumen, dan *df(t)* adalah jumlah dokumen yang memuat term *t*.

Klasifikasi *intent* dilakukan dengan menghitung kemiripan antara vektor pertanyaan dan seluruh vektor korpus menggunakan *Cosine Similarity*, sebagaimana Persamaan (3) [21].

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

$$\text{sim}(Q, D_i) = \frac{\sum_{j=1}^{n} WQ_j \times W_{d_i j}}{\sqrt{\sum_{j=1}^{n} W_{Q,j}^2} \times \sqrt{\sum_{j=1}^{n} W_{D_i,J}^w}} \quad (3)$$

Intent dari *utterance* dengan nilai kemiripan tertinggi dipilih sebagai hasil prediksi menggunakan pendekatan *1-Nearest Neighbor*. Nilai kemiripan tertinggi tersebut menjadi tingkat keyakinan (*confidence*) yang dievaluasi oleh *Decision Layer* menggunakan fungsi keputusan pada Persamaan (4).

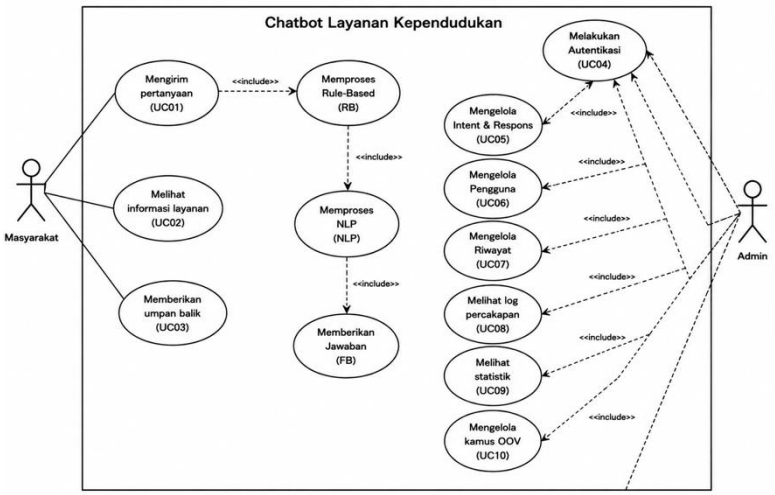
$$Respons \begin{cases} RES(intent(D_{max})), JikaS_{max} \geq \tau \\ RES - FBK \quad , JikaS_{max} \geq \tau \end{cases} \quad (4)$$

Apabila nilai keyakinan tertinggi (*S_max*) memenuhi atau melampaui ambang τ, sistem mengembalikan respons *intent*. Sebaliknya, sistem mengaktifkan *fallback response* sebagai pengendali agar tidak memberikan informasi yang berpotensi keliru.

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Sistem

Sistem diimplementasikan sebagai aplikasi web menggunakan bahasa pemrograman *Python* dengan kerangka kerja *Flask*, sehingga dapat diakses masyarakat secara daring. Pra-pemrosesan teks Bahasa Indonesia memanfaatkan pustaka Sastrawi, sedangkan pembobotan TF-IDF dan perhitungan *Cosine Similarity* menggunakan pustaka *scikit-learn*. Antarmuka sistem menampilkan kolom masukan pertanyaan, respons chatbot beserta nilai keyakinan, serta pintasan kategori layanan, sebagaimana ditunjukkan pada Gambar 3.



Gambar 3 Antarmuka Chatbot Layanan Kependudukan

Sistem juga dilengkapi fitur *Mode Audit* yang menampilkan sumber jawaban (*rule-based*, *context*, *nlp*, atau *fallback*) beserta nilai keyakinannya, sebagai wujud transparansi dalam penelusuran keputusan sistem.

3.2 Pengujian Akurasi Klasifikasi

Pengujian akurasi dilakukan pada 76 *utterance* data uji menggunakan pendekatan *1-Nearest Neighbor* berbasis *Cosine Similarity*. Pembentukan vektor *TF-*

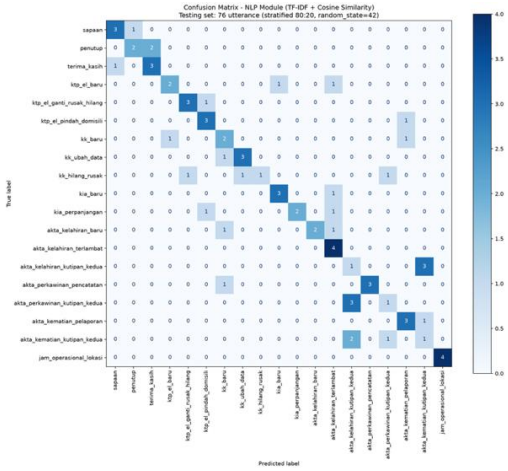
© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

IDF dilakukan hanya pada data latih untuk menghindari kebocoran data (*data leakage*). Hasil pengujian disajikan pada Tabel 2.

Tabel 2 Hasil Pengujian Akurasi Model NLP

Metrik	Nilai
Accuracy	0,6053
Precision (macro average)	0,6080
Recall (macro average)	0,6080
F1-Score (macro average)	0,6080

Berdasarkan Tabel 2, model mencapai akurasi sebesar 60,53%. Distribusi hasil klasifikasi terhadap 19 *intent* direpresentasikan dalam bentuk *confusion matrix* pada Gambar 4.



Gambar 4 Confusion Matrix Hasil Pengujian Model NLP

Analisis terhadap *confusion matrix* menunjukkan bahwa kesalahan klasifikasi terkonsentrasi pada kelompok *intent* pengurusan kutipan kedua akta, yaitu akta kelahiran, akta perkawinan, dan akta kematian. Ketiga *intent* tersebut saling tertukar karena memiliki struktur kalimat yang hampir identik dan hanya dibedakan oleh satu kata penanda jenis akta

Nilai *accuracy* sebesar 60,53% diperoleh dalam kondisi klasifikasi paksa (*forced-choice*), yakni ketika sistem diwajibkan memilih salah satu dari 19 intent untuk setiap utterance uji tanpa opsi menahan jawaban. Angka ini mengukur mutu klasifikator mentah (TF-IDF + *Cosine Similarity* + 1-*Nearest Neighbor*) dan mencerminkan keterbatasan pendekatan leksikal terhadap utterance pendek berbahasa Indonesia yang memiliki variasi kata tinggi. Dengan demikian, 60,53% merepresentasikan kemampuan klasifikasi, bukan keandalan jawaban akhir yang disampaikan sistem kepada pengguna; kedua besaran ini dianalisis secara terpisah pada subbab berikut.

3.3 Kalibrasi Threshold τ

Nilai ambang keyakinan pada *Decision Layer* ditetapkan secara empiris dengan membandingkan distribusi skor *Cosine Similarity* tertinggi antara data uji *in-scope* (76 *utterance*) dan data *out-of-scope* (30 *utterance*). Efektivitas setiap nilai kandidat τ diukur menggunakan indeks Youden. Hasil kalibrasi disajikan pada Tabel 3.

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

Tabel 3 Kalibrasi Nilai Threshold τ

τ	In-scope Lolos (%)	Out-of-scope Ditolak (%)	Indeks Youden
0,45	93,4	36,7	0,301
0,50	81,6	40,0	0,216
0,55	75,0	60,0	0,350
0,60	69,7	80,0	0,497
0,65	51,3	86,7	0,380

Berdasarkan Tabel 3, nilai τ = 0,60 menghasilkan indeks Youden tertinggi (0,497), sehingga ditetapkan sebagai ambang keyakinan sistem. Pada nilai ini, sistem mampu menolak 80% pertanyaan di luar cakupan sekaligus meloloskan 69,7% pertanyaan valid.

3.4 Perbandingan Sistem Tanpa dan Dengan Decision Layer

Untuk mengevaluasi kontribusi Decision Layer terhadap keandalan sistem, dilakukan perbandingan antara dua konfigurasi, yaitu sistem tanpa Decision Layer dan sistem dengan Decision Layer pada ambang τ = 0,60. Pada konfigurasi tanpa Decision Layer, setiap pertanyaan dipaksa dipetakan ke intent terdekat tanpa mempertimbangkan tingkat keyakinan, sedangkan pada konfigurasi dengan Decision Layer, jawaban dengan keyakinan di bawah ambang dialihkan ke *fallback response*. Perbandingan disajikan pada Tabel 4.

Tabel 4 Perbandingan Kemampuan Pengendalian Pertanyaan di Luar Cakupan

Aspek	Tanpa Decision Layer	Dengan Decision Layer (τ = 0,60)
Penolakan pertanyaan di luar cakupan	0%	80,0%
Peloloskan pertanyaan valid	100%	69,7%
Perlakuan terhadap keyakinan rendah	Dijawab dengan yakin	Dialihkan ke *fallback*

Tabel 4 menunjukkan bahwa tanpa Decision Layer, seluruh pertanyaan di luar cakupan dipaksa dipetakan ke intent terdekat sehingga menghasilkan jawaban yang keliru. Sebaliknya, dengan Decision Layer, sistem mampu menolak 80% pertanyaan di luar cakupan dan mengalihkannya ke *fallback response*, meskipun dengan konsekuensi 30,3% pertanyaan valid ikut tertahan. Pertukaran (*trade-off*) ini menunjukkan bahwa Decision Layer memprioritaskan pencegahan penyampaian informasi yang keliru terhadap pertanyaan di luar cakupan, yang sesuai dengan tuntutan konteks layanan publik akan keakuratan informasi.

3.5 Pengujian penerimaan pengguna

Pengujian penerimaan pengguna (*User Acceptance Testing*) dilakukan untuk mengukur tingkat penerimaan masyarakat terhadap sistem. Kuesioner disebarkan kepada 31 responden yang telah mencoba chatbot secara daring, dengan latar belakang pekerjaan beragam meliputi mahasiswa, karyawan swasta, wiraswasta, dan pegawai negeri. Kuesioner terdiri atas 12 pernyataan yang mencakup lima aspek, yaitu kemudahan penggunaan, kualitas jawaban, kecepatan dan keandalan, kemanfaatan, serta kepuasan keseluruhan, diukur menggunakan skala Likert lima tingkat. Tingkat penerimaan dihitung menggunakan indeks persentase, sebagaimana Persamaan (5).

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

$$Indeks(\%) = \frac{Total\ Skor}{Skor\ Maksimum} \times 100\% \quad (5)$$

Hasil rekapitulasi pengujian disajikan pada Tabel 5.

Tabel 5 Hasil Pengujian *User Acceptance Testing*

Aspek	Indeks (%)	Kategori
Kemudahan Penggunaan	84,1	Sangat Baik
Kualitas Jawaban	83,9	Baik
Kecepatan dan Keandalan	83,2	Baik
Kemanfaatan	83,9	Baik
Kepuasan Keseluruhan	84,6	Sangat Baik
Indeks Keseluruhan	**83,92**	**Baik**

Berdasarkan Tabel 5, sistem memperoleh indeks penerimaan keseluruhan sebesar 83,92%, yang termasuk kategori Baik. Aspek kemanfaatan dan kepuasan memperoleh penilaian tinggi, yang menunjukkan bahwa meskipun akurasi klasifikasi model masih terbatas, keberadaan Decision Layer yang menjaga sistem tetap terkendali membuat pengalaman pengguna secara keseluruhan tetap positif. Seluruh aspek memperoleh indeks yang relatif seragam pada rentang 83,2%–84,6%. Aspek kecepatan dan keandalan memperoleh indeks terendah (83,2%), sedangkan aspek kualitas jawaban memperoleh 83,9%. Nilai pada aspek kualitas jawaban tersebut konsisten dengan keterbatasan metode berbasis kemiripan leksikal yang teridentifikasi pada pengujian akurasi.

3.6 Pembahasan

Hasil pengujian menunjukkan bahwa akurasi klasifikasi sebesar 60,53% mencerminkan keterbatasan pendekatan berbasis kemiripan leksikal (TF-IDF dan *Cosine Similarity*) dalam membedakan *intent* dengan struktur kalimat yang sangat mirip. Metode ini mengukur kemiripan berdasarkan kesamaan kata, bukan makna, sehingga kata penanda pembeda memperoleh bobot yang tidak cukup besar untuk menentukan klasifikasi secara tepat.

Analisis lanjutan dilakukan untuk menguji apakah penyaringan berbasis ambang turut meningkatkan akurasi jawaban di dalam cakupan. Akurasi dihitung ulang hanya pada subset data uji yang lolos ambang (keyakinan ≥ 0,60), dan diperoleh nilai sebesar 56,60%, yaitu lebih rendah dibandingkan akurasi pada seluruh data uji (60,53%). Hasil ini mengindikasikan bahwa skor *Cosine Similarity* belum berkorelasi kuat dengan kebenaran prediksi pada utterance pendek berbahasa Indonesia, sehingga penyaringan berbasis ambang tidak serta-merta meningkatkan akurasi jawaban di dalam cakupan. Temuan ini mempertegas bahwa kontribusi Decision Layer terletak pada kemampuan pengendalian terhadap pertanyaan di luar cakupan, bukan pada peningkatan akurasi klasifikasi di dalam cakupan.

Meskipun demikian, keterbatasan pada modul NLP tersebut tidak sepenuhnya menurunkan kualitas layanan sistem. Hal ini disebabkan oleh peran *Decision Layer* sebagai mekanisme pengendali. Ketika nilai keyakinan berada di bawah ambang τ, sistem tidak memaksakan jawaban yang berpotensi keliru, melainkan mengaktifkan *fallback response*. Dengan demikian, risiko penyampaian informasi administratif yang salah dapat ditekan. Temuan ini menegaskan bahwa kontribusi utama penelitian bukan semata pada akurasi klasifikasi, melainkan pada

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

mekanisme *Decision Layer* yang memungkinkan sistem menyampaikan informasi secara terkendali, sesuai konteks layanan publik yang menuntut keakuratan informasi.

KESIMPULAN

Penelitian ini berhasil merancang dan mengimplementasikan chatbot layanan kependudukan Kota Medan berbasis pendekatan *hybrid* yang menggabungkan metode *Rule-Based* dan *Natural Language Processing* melalui sebuah *Decision Layer*. Sistem mengorkestrasi kedua pendekatan secara berjenjang, yaitu pencocokan pola deterministik pada modul *Rule-Based*, resolusi konteks, klasifikasi *intent* berbasis TF-IDF dan *Cosine Similarity*, serta *fallback response* apabila tingkat keyakinan berada di bawah ambang. Pengujian akurasi menghasilkan nilai *accuracy* sebesar 60,53%, dengan kesalahan klasifikasi yang terkonsentrasi pada *intent* yang memiliki struktur kalimat sangat mirip. Kalibrasi empiris berdasarkan indeks Youden menetapkan nilai ambang $\tau = 0{,}60$, yang mampu menolak 80% pertanyaan di luar cakupan sekaligus meloloskan 69,7% pertanyaan valid. Hasil penelitian menunjukkan bahwa *Decision Layer* berperan penting dalam mengendalikan penyampaian informasi sehingga sistem tidak memaksakan jawaban yang berpotensi keliru. Penelitian selanjutnya disarankan menggunakan representasi teks berbasis makna (*semantic representation*), seperti *word embedding*, untuk mengatasi keterbatasan dalam membedakan *intent* dengan struktur kalimat yang mirip.

DAFTAR PUSTAKA

[1] R. S. Anjani and A. Abdal, "Implementasi Kebijakan Sistem Pemerintahan Berbasis Elektronik (SPBE) dalam Mendukung Transformasi Digital Pelayanan Publik di Indonesia," *Indonesian Journal of Multidisciplinary on Social and Technology*, vol. 4, no. 2, pp. 2571–2579, Jun. 2026, doi: 10.69693/ijmst.v4i2.10558.

[2] T. B. M. F. Suryana, A. A. Rahmawati, N. S. Ramdanti, and A. N. Safitri, "Transformasi Digital dalam Pelayanan Publik: Tinjauan Yuridis terhadap SPBE di Indonesia," *CONSTITUO : Journal of State and Political Law Research*, vol. 4, no. 1, pp. 37–53, Jun. 2025, doi: 10.47498/constituo.v4i1.4944.

[3] E. L. Wargadinata and N. R. P. Tendean, "Amidst the Euphoria of Digitalization Public Services in Municipal Government: Raising Public Acceptance," *Jurnal Bina Praja*, vol. 16, no. 2, pp. 335–346, Aug. 2024, doi: 10.21787/jbp.16.2024.335-346.

[4] J. Khatib Sulaiman, M. Abid Nadzif, and U. Stikubank Semarang, "Penggunaan Teknologi Natural Language Processing dalam Sistem Chatbot Untuk Peningkatan Layanan Informasi Administrasi Publik," *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 1, pp. 2024–1227.

[5] R. Aulia and A. Purnama, "Development of an Intent-Classification Chatbot to Support Operational Services at Kadin Indonesia," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 1171–1180, Dec. 2025, doi: 10.47709/brilliance.v5i2.7438.

[6] R. D. Safitri[1] and D. Casmiwati[2], "Policy And Maritime Review 50."

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

[7] R. Aulia and A. Purnama, "Development of an Intent-Classification Chatbot to Support Operational Services at Kadin Indonesia," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 1171–1180, Dec. 2025, doi: 10.47709/brilliance.v5i2.7438.

[8] Z. A. Annisa, R. S. Perdana, and P. P. Adikara, "Kombinasi Intent Classification dan Named Entity Recognition pada Data Berbahasa Indonesia dengan Metode Dual Intent and Entity Transformer," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 5, pp. 1017–1024, Oct. 2024, doi: 10.25126/jtiik.2024117985.

[9] M. Zulkarnain, M. Zulkarnain Lubis, and M. Khaibar Putra Adithia, "Desain dan Implementasi Inovasi Pelayanan Publik Berbasis Website dengan Bantuan AI di Kota Medan," vol. 4, no. 5, pp. 1232–1242, 2025, [Online]. Available: https://ojs.trigunadharma.ac.id/index.php/jsi

[10] M. Huda, A. M. Kusumaningrum, and G. Aninditiyah, "Pemanfaatan Chatbot Untuk Meningkatkan Layanan Informasi Publik Di Instansi Pemerintah Tingkat Kecamatan," vol. 15, no. 2, pp. 413–424, 2024, [Online]. Available: http://ejurnal.provisi.ac.id/index.php/JTIKP

[11] M. Akbar Nasution *et al.*, "Implementasi NLP Dalam Pembuatan Chatbot Customer Service Publisher Jurnal Studi Kasus LARISMA," 2024.

[12] S. Ernawati, I. Maulana, and N. Ruhyana, "Simpatik: Jurnal Sistem Informasi dan Informatika Penerapan Natural Language Processing Pada Model Chatbot Hybrid Untuk Optimasi Pelayanan Open Journal Systems." [Online]. Available: http://jurnal.bsi.ac.id/index.php/simpatik

[13] W. Bunandi and V. C. Mawardi, "Chatbot Virtual Assistant Help Desk Dengan Metode Aiml," *Jurnal Serina Sains, Teknik dan Kedokteran*, vol. 2, no. 1, pp. 51–60, Mar. 2024, doi: 10.24912/jsstk.v2i1.32051.

[14] R. Nur Aziza *et al.*, "Pembangunan Aplikasi dan Klasifikasi Pertanyaan Chatbot Informasi Akademik Menggunakan Metode Cosine Similarity dan Naïve Bayes," vol. 12, no. 2, doi: 10.33322/kilat.v12i2.1921.

[15] M. T. Hidayat, M. Arifin, and S. Muzid, "Prediction Sentiment Analysis Grab Reviews Using SVM Linear Based Streamlit," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems*, vol. 19, no. 2, pp. 189–200, 2025, doi: 10.22146/ijccs.104924.

[16] C. B. Chandrakala, R. Bhardwaj, and C. Pujari, "An intent recognition pipeline for conversational AI," *International Journal of Information Technology (Singapore)*, vol. 16, no. 2, pp. 731–743, Feb. 2024, doi: 10.1007/s41870-023-01642-8.

[17] A. Fauzan Zaky and C. Indah Ratnasari, "Developing Rule-Based and AI Hybrid Chatbot for Academic Information Services," *International Journal of Informatics and Computation (IJICOM)*, vol. 7, no. 2, p. 2025, doi: 10.35842/ijicom.

[18] N. Khumairoh and W. Syaihul Huda, "Implementation Of Hybrid Chatbot Based On Dialogflow And Keyword Matching To Improve Website Service Efficiency," 2025.

[19] L. Grassi, C. T. Recchiuto, and A. Sgorbissa, "Knowledge-Grounded Dialogue Flow Management for Social Robots and Conversational Agents," *Int. J. Soc.*

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)

Robot., vol. 14, no. 5, pp. 1273–1293, Jul. 2022, doi: 10.1007/s12369-022-00868-z.

[20] D. Kristanto, R. A. Ramadhani, and A. B. Setiawan, "Pengembangan Chatbot Layanan Informasi Kampus Menggunakan TF-IDF," *Jurnal Algoritma*, vol. 22, no. 2, pp. 103–115, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2350.

[21] A. Rinaldi, L. Hiryanto,) Viny, and C. Mawardi, "Jurnal Ilmu Komputer dan Sistem Informasi Penerapan Tf-Idf Dan Cosine Similarity Pada Question And Answering System Terkait Informasi Konsol Dan Game Playstation."

© 2026 The Author(s). Published by UNITY ACADEMY . This is an open access article under the CC BY-SA license (http://creativecommons.org/licenses/by-sa/4.0/)