

ANALISIS RELIABILITAS CHATBOT PADA PENDAFTARAN PELANGGAN BARU DI CYB MEDIA MENGGUNAKAN METODE TF-IDF DAN LOGISTIC REGRESSION

Ridwan Tri Cahya Kusuma, Afu Ichsan Pradana, Agustina Purwatiningsih

Teknik Informatika, Universitas Duta Bangsa Surakarta

Jl. Bhayangkara No.55, Tipes, Kec. Serengan, Kota Surakarta, Jawa Tengah 57154

ridwancahyakusuma@gmail.com

ABSTRAK

Pelayanan pelanggan yang cepat dan akurat menjadi faktor penting bagi penyedia layanan internet, termasuk CYB Media, dalam meningkatkan kepuasan pelanggan dan efisiensi operasional. Proses pendaftaran pelanggan baru yang masih dilakukan secara manual menyebabkan penumpukan antrean pesan, meningkatnya beban kerja staf, serta lambatnya respons layanan. Penelitian ini bertujuan menganalisis keandalan dan efektivitas chatbot berbasis Telegram menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan Logistic Regression untuk mengklasifikasikan intent pengguna pada proses pendaftaran pelanggan. Model dibangun menggunakan representasi fitur TF-IDF dan algoritma Logistic Regression, kemudian dievaluasi menggunakan 40 data uji dengan metrik akurasi, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model mencapai akurasi sebesar 95,00%, dengan intent register_wifi memperoleh nilai F1-score tertinggi sebesar 0,9714. Hasil tersebut menunjukkan bahwa chatbot yang dikembangkan mampu mengklasifikasikan intent pengguna dengan baik dan layak diterapkan sebagai sistem otomatisasi pendaftaran pelanggan berbasis Telegram.

Kata kunci : *automasi, chatbot, layanan pelanggan, TF-IDF, logistic regression.*

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mendorong perusahaan penyedia layanan internet untuk meningkatkan kualitas pelayanan pelanggan, khususnya dalam proses pendaftaran pelanggan baru. Di era digital, ekspektasi pelanggan terhadap kecepatan dan akurasi layanan meningkat secara signifikan [1].

Banyak penyedia layanan masih mengandalkan proses manual dalam menangani pesan pendaftaran, sehingga menyebabkan penumpukan antrean pesan, beban kerja customer service yang tinggi, serta penurunan respons layanan secara keseluruhan. Ketergantungan pada proses manual juga berkontribusi terhadap rendahnya efisiensi operasional dan potensi kesalahan manusia dalam pemeriksaan data pelanggan [2].

Untuk menjawab tantangan tersebut, sejumlah studi menunjukkan bahwa teknologi kecerdasan buatan (Artificial Intelligence) khususnya chatbot dapat menjadi solusi dalam otomatisasi layanan pelanggan [1].

Chatbot merupakan program atau agen percakapan yang dapat berinteraksi dengan pengguna secara otomatis melalui bahasa alami (natural language), sehingga mampu memberikan respons real time, konsisten, dan tanpa batasan jam operasional [3].

Di Indonesia sendiri, adopsi chatbot oleh perusahaan dan organisasi sebagai alat bantu pelayanan pelanggan terus berkembang, baik pada sektor layanan publik, bisnis, maupun institusi pendidikan [4], [5].

Salah satu pendekatan yang banyak diterapkan dalam pengembangan chatbot generasi dasar adalah penggunaan teknik Term Frequency–Inverse

Document Frequency (TF IDF) dalam pemrosesan teks untuk menangkap kata kunci pertanyaan pelanggan dan menghubungkannya dengan respons yang sesuai dalam basis pengetahuan. Metode TF IDF telah diterapkan pada beberapa aplikasi chatbot di Indonesia, seperti pada pengembangan chatbot layanan akademik di Universitas Nusantara PGRI Kediri dengan tingkat akurasi yang menunjukkan efektivitas awal dalam klasifikasi pertanyaan [6].

Selain itu, implementasi chatbot sebagai virtual assistant dengan metode TF IDF di Universitas Panca Marga Probolinggo menunjukkan bahwa pembobotan teks mampu menghasilkan sistem respons otomatis yang layak sebagai alat bantu layanan informasi dasar kepada pengguna [7].

Meskipun metode TF IDF terbukti efektif dalam mengidentifikasi kata kunci dan memetakan respons dasar, metode ini memiliki keterbatasan dalam menangkap konteks keseluruhan pesan atau variasi bahasa alami yang kompleks [6].

Permasalahan ini sering muncul ketika chatbot dihadapkan pada pertanyaan yang ambigu atau frasa yang tidak secara langsung terkait dengan vektor kata kunci yang telah dipetakan sebelumnya, sehingga akurasi respons dapat menurun secara signifikan.

Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan chatbot berbasis Telegram yang memanfaatkan TF-IDF sebagai metode ekstraksi fitur dan Logistic Regression sebagai algoritma klasifikasi intent. Kombinasi kedua metode tersebut digunakan untuk mengidentifikasi intent pengguna berdasarkan pesan yang dikirimkan, sehingga chatbot dapat memberikan respons yang sesuai dan mengotomatisasi proses pendaftaran pelanggan baru di CYB Media.

Penelitian ini bertujuan untuk menganalisis kemampuan chatbot dalam mengklasifikasikan intent pengguna pada proses pendaftaran pelanggan baru, mengukur tingkat akurasi model yang dihasilkan, serta mengevaluasi fungsionalitas sistem dalam mendukung layanan pendaftaran pelanggan secara otomatis melalui Telegram. Hasil penelitian diharapkan dapat memberikan gambaran mengenai efektivitas penerapan TF-IDF dan Logistic Regression dalam meningkatkan kualitas layanan pelanggan pada penyedia layanan internet.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian terdahulu digunakan sebagai landasan dan referensi dalam pengembangan penelitian ini. Penelitian terdahulu dikaji untuk memberikan gambaran mengenai perkembangan penelitian yang berkaitan dengan chatbot Telegram, klasifikasi intent, dan penerapan metode pemrosesan teks. Kajian ini juga digunakan untuk mengidentifikasi pendekatan yang telah digunakan serta menemukan celah penelitian yang dapat dikembangkan lebih lanjut.

Penelitian yang dilakukan oleh Maulana dkk. (2024) berjudul Implementasi Pembobotan TF-IDF pada Chatbot Telegram untuk Sistem Layanan Informasi menerapkan metode TF-IDF pada chatbot Telegram menggunakan 94 dokumen dan 300 data uji. Hasil penelitian menunjukkan bahwa chatbot menghasilkan 268 respons yang relevan dan akurat, 12 respons tidak relevan, serta 32 respons yang seharusnya relevan namun tidak ditemukan. Penelitian tersebut membuktikan bahwa TF-IDF mampu memberikan respons yang cukup baik pada sistem layanan informasi [8].

Penelitian oleh Pratama dkk. (2025) berjudul Penerapan Metode TF-IDF dan Cosine Similarity pada Chatbot Informasi Pilkada Bekasi 2024 Berbasis Telegram menggunakan metode TF-IDF dan Cosine Similarity untuk memberikan informasi Pilkada melalui Telegram. Hasil penelitian menunjukkan tingkat akurasi jawaban sebesar 93,33%. Fokus penelitian tersebut adalah pencarian jawaban berdasarkan kemiripan teks terhadap basis pengetahuan yang tersedia [9].

Penelitian oleh Aulia dan Purnama (2025) berjudul *Development of an Intent-Classification Chatbot to Support Operational Services at Kadin Indonesia* mengembangkan chatbot klasifikasi intent menggunakan metode Long Short-Term Memory (LSTM). Hasil penelitian menunjukkan bahwa chatbot mampu mengotomatisasi layanan operasional organisasi dan memberikan respons secara otomatis kepada pengguna dengan performa yang baik dalam mengenali variasi kalimat [10].

Berdasarkan penelitian terdahulu, penerapan chatbot Telegram terbukti mampu meningkatkan efisiensi layanan dan otomatisasi respons pengguna. Namun, sebagian besar penelitian masih berfokus pada layanan informasi atau menggunakan metode

klasifikasi yang berbeda. Oleh karena itu, penelitian ini menerapkan kombinasi TF-IDF dan Logistic Regression untuk mengklasifikasikan intent pengguna pada chatbot Telegram yang digunakan dalam proses pendaftaran layanan internet.

2.2. ChatBot

Chatbot adalah program komputer yang dirancang untuk meniru percakapan manusia dengan memanfaatkan kecerdasan buatan (*Artificial Intelligence*) dan pemrosesan bahasa alami (*Natural Language Processing*) [6].

Chatbot mampu memahami masukan pengguna dan memberikan respons yang sesuai dengan konteks. Teknologi ini banyak digunakan untuk mengotomatisasi layanan pelanggan, sehingga dapat meningkatkan efisiensi operasional dan mengurangi kebutuhan tenaga kerja manusia [11].

2.3. Artificial Intelligence (AI)

Kecerdasan Buatan (*Artificial Intelligence*) merupakan bidang dalam ilmu komputer yang berfokus pada pengembangan sistem yang mampu menjalankan tugas-tugas yang biasanya memerlukan kecerdasan manusia. Kemampuan tersebut diperoleh melalui pengetahuan dan pengalaman yang ditanamkan ke dalam sistem. Oleh karena itu, AI dirancang dengan kemampuan penalaran untuk mengolah informasi dan menghasilkan solusi atau keputusan secara cerdas. Selain itu, AI juga digunakan untuk memodelkan dan menguji teori-teori kecerdasan yang dapat diimplementasikan melalui bahasa pemrograman dan dijalankan pada komputer [12].

2.4. Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. Teknologi ini bertujuan agar komputer dapat memahami, menganalisis, dan mengolah bahasa manusia secara otomatis [13].

NLP menggabungkan konsep linguistik dan ilmu komputer untuk memungkinkan sistem mengenali makna dari teks maupun ucapan. Dengan NLP, komputer dapat melakukan berbagai tugas seperti penerjemahan bahasa, analisis sentimen, klasifikasi teks, pencarian informasi, hingga pengembangan chatbot yang mampu berkomunikasi secara lebih alami dengan pengguna.

2.5. Machine Learning

Machine Learning (ML) merupakan cabang kecerdasan buatan yang memungkinkan komputer belajar dari data dan pengalaman tanpa perlu diprogram secara eksplisit untuk setiap tugas. Teknologi ini memanfaatkan konsep statistika, matematika, dan data mining untuk mengenali pola serta menghasilkan prediksi atau keputusan secara otomatis. Machine Learning banyak digunakan dalam

berbagai aplikasi, seperti klasifikasi data, pengenalan pola, sistem rekomendasi, dan analisis prediktif [11].

2.6. TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan metode pembobotan kata (*term weighting*) yang digunakan dalam text mining dan Natural Language Processing (NLP) untuk menentukan tingkat kepentingan suatu kata dalam sebuah dokumen relatif terhadap kumpulan dokumen [14].

Metode ini bekerja dengan memberikan bobot berdasarkan frekuensi kemunculan kata pada suatu dokumen (*term frequency*) serta mempertimbangkan seberapa jarang kata tersebut muncul pada dokumen lain (*inverse document frequency*) [15].

Dalam proses klasifikasi teks, TF-IDF sering digunakan sebagai metode ekstraksi fitur untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning.

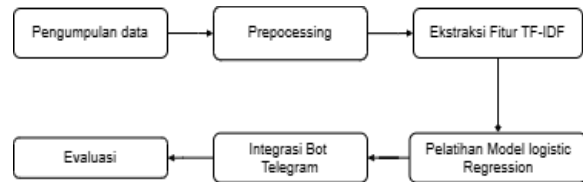
2.7. Logistic Regression

Logistic Regression merupakan salah satu algoritma supervised learning yang digunakan untuk menyelesaikan permasalahan klasifikasi [16].

Algoritma ini bekerja dengan memprediksi probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan hubungan antara variabel input dan output yang diamati selama proses pelatihan. Logistic Regression termasuk metode yang sederhana namun efektif sehingga banyak diterapkan dalam berbagai bidang, khususnya pada *predictive analytics* dan klasifikasi teks [17].

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan fokus pada pengembangan dan evaluasi sebuah sistem chatbot otomatis berbasis Telegram untuk layanan pendaftaran pelanggan. Alur metode penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Metode Penelitian

3.1. Pengumpulan Data Dan Data Labeling

Tahap awal penelitian bertujuan untuk mengumpulkan semua data teks percakapan yang relevan sebagai bahan pengembangan chatbot. Dataset yang digunakan merupakan data berbahasa Indonesia yang telah diberi label intent [18].

Labeling ini merupakan bagian dari proses *ground truth*, yang menetapkan intent sebenarnya dari tiap pesan untuk mendukung metode pembelajaran *supervised learning*. Sumber data mencakup log percakapan pelanggan sebelumnya dan log percakapan Telegram, yang disimpan dalam bentuk file berformat CSV. Dataset yang digunakan dalam penelitian ini berjumlah 130 yang dibagi menjadi dua subset yaitu data training untuk melatih model TF-IDF dan data testing untuk mengevaluasi performa model [19]. Dalam penelitian ini, intent yang digunakan dijelaskan pada tabel berikut.

Tabel 1. Label Intent

Intent	Deskripsi	Contoh
register_wifi	Pesan yang bertujuan mendaftar paket wifi	"mau daftar wifi", "daftar paket wifi"
inquiry	Pesan yang berisi pertanyaan mengenai paket, harga, atau layanan WiFi	"berapa harga paket wifi?", "info paket wifi terbaru?"
Greeting	Pesan sapaan atau salam dari pengguna	"halo", "hai", "selamat siang"
unknown	Pesan yang tidak termasuk ke dalam intent yang sudah ditentukan atau ambigu	"blah blah", "ok", "siap"

Pada tabel label intent menunjukkan tujuan atau maksud dari tiap pesan pengguna sehingga model dapat belajar memetakan pesan ke kategori yang sesuai.

3.2. Tahapan Preprocessing

Teknik awal dalam data mining yang berfungsi mengubah data mentah (*raw data*) yang diperoleh dari berbagai sumber menjadi informasi yang lebih bersih dan siap untuk diproses lebih lanjut [7].

Dalam penelitian ini, proses preprocessing dilakukan melalui beberapa tahapan [20], meliputi:

- Pembersihan Teks:** Tahap cleaning dilakukan untuk menghapus elemen teks yang tidak relevan, seperti URL, mention, hashtag, emotikon, simbol, angka, dan tanda baca yang berpotensi mengganggu proses analisis.

- lowercase:** pada tahap case folding mengubah semua kata yang diinputkan oleh pengguna menjadi menjadi huruf kecil. Case folding hanya memproses huruf dari 'a' hingga 'z', sedangkan karakter selain huruf akan dihapus atau diabaikan.
- Tokenizing:** Proses memecah kalimat menjadi satuan kata atau token, agar setiap kata dapat di analisis secara individual.
- Stopword Removal:** menghilangkan kata-kata umum yang tidak memberikan pengaruh maupun nilai analitis penting dalam proses pengolahan data, seperti "dan", "yang", "atau".

3.3. Ekstraksi Fitur TF-IDF

Tahap ini merupakan langkah penting untuk menganalisis teks. Proses ini mengubah kata-kata

menjadi representasi numerik, baik berupa bilangan *integer* maupun *float*, sesuai dengan algoritma yang digunakan dalam pembelajaran mesin [21].

Proses ekstraksi fitur teks dilakukan menggunakan *TfidfVectorizer* dari *scikit-learn* untuk mengonversi data teks hasil pelabelan intent menjadi representasi numerik berdasarkan bobot kata yang dihitung menggunakan formula TF-IDF [22].

Pada tahap ini, setiap kalimat pengguna diubah menjadi vektor fitur yang merepresentasikan tingkat kepentingan kata atau frasa dalam korpus. Hasil representasi numerik tersebut kemudian digunakan sebagai input bagi algoritma *Logistic Regression* untuk melakukan klasifikasi intent pengguna. Secara umum, TF-IDF dihitung dengan rumus:

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

Dengan:

$$TF(t, d) = \frac{\text{Jumlah kata } t \text{ dalam dokumen } d}{\text{Total kata dalam dokumen } d} \quad (2)$$

$$IDF(t) = \log \frac{N}{1+n_t} \quad (3)$$

Rumus TF-IDF yang digunakan oleh *TfidfVectorizer* *scikit-learn* secara default menggunakan Smoothed IDF sebagai berikut:

$$IDF(t) = \log \left(\frac{1+N}{1+df(t)} \right) + 1 \quad (4)$$

TfidfVectorizer dari *scikit-learn* secara default juga melakukan normalisasi L2, sehingga nilai akhir TF-IDF menjadi:

$$TF - IDF_{norm}(t, d) = \frac{TF-IDF(t, d)}{\sqrt{\sum_{i=1}^n TF-IDF_i^2}} \quad (5)$$

3.4. Pemodelan Logistic Regression

Logistic Regression merupakan salah satu algoritma supervised learning yang digunakan untuk menyelesaikan permasalahan klasifikasi [16].

Pemodelan pada penelitian ini dilakukan menggunakan algoritma *Logistic Regression* sebagai model klasifikasi intent. Hasil ekstraksi fitur teks menggunakan TF-IDF digunakan sebagai input bagi model untuk mempelajari pola hubungan antara kata atau frasa dalam pesan pengguna dengan label intent yang telah ditentukan. *Logistic Regression* dipilih karena efektif digunakan pada data teks berdimensi tinggi, relatif ringan secara komputasi, dan banyak diterapkan pada tugas klasifikasi teks berbasis TF-IDF [23].

Karena penelitian ini memiliki lebih dari dua kelas intent, *Logistic Regression* digunakan dalam bentuk multiclass classification. Pada skema ini, model menghitung skor untuk setiap kelas intent, kemudian mengubah skor tersebut menjadi probabilitas menggunakan fungsi softmax. Kelas dengan probabilitas tertinggi akan dipilih sebagai hasil prediksi [24].

Secara matematis, probabilitas suatu pesan X termasuk ke dalam kelas j dapat dituliskan sebagai berikut.

Nilai skor awal atau logit untuk setiap kelas dihitung dengan rumus:

$$z_j = (X_1 W_{1j} + X_2 W_{2j} + \dots + (X_{n1} W_{nj}) + b_j \quad (6)$$

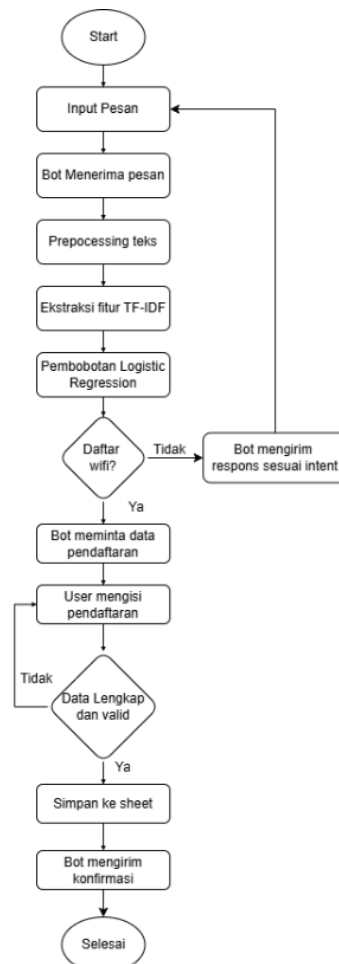
Setelah logit diperoleh, probabilitas setiap kelas dihitung menggunakan fungsi softmax:

$$P(y = j|X) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (7)$$

3.5. Integrasi Bot Telegram

Setelah model *Logistic Regression* berhasil dilatih menggunakan fitur TF-IDF, model tersebut diintegrasikan ke dalam platform Bot Telegram sebagai antarmuka komunikasi berbasis percakapan (*conversational interface*). Telegram dipilih sebagai platform karena menyediakan *Bot API* yang bersifat terbuka, ringan, dan mendukung komunikasi dua arah secara *real-time* melalui protokol HTTP [11].

Berikut merupakan alur integrasi aplikasi Telegram dengan sistem chatbot menggunakan Telegram Bot API.



Gambar 2. Alur integrasi telegram

Gambar di atas menggambarkan alur proses chatbot Telegram yang dimulai dari penerimaan pesan, preprocessing teks, ekstraksi fitur TF-IDF, hingga klasifikasi intent menggunakan Logistic Regression. Jika intent yang terdeteksi adalah pendaftaran WiFi, sistem akan mengumpulkan data pengguna, melakukan validasi, menyimpan data ke Google Sheet, dan mengirimkan konfirmasi secara otomatis.

3.6. Evaluasi Model

Tahap Evaluasi bertujuan mengukur kinerja model yang dikembangkan, untuk memastikan kesesuaiannya dengan tujuan penelitian serta relevansinya terhadap permasalahan bisnis yang telah ditetapkan [25].

Evaluasi sistem chatbot dilakukan dengan memanfaatkan Confusion Matrix, yang menyediakan metrik seperti akurasi, presisi, recall, dan F1 Score. Confusion Matrix sendiri adalah alat evaluasi berbentuk matriks yang menampilkan jumlah prediksi yang benar maupun salah untuk setiap kelas dalam sistem klasifikasi [26].

Ground truth ditetapkan secara manual oleh peneliti untuk memberi relevansi terhadap hasil pencarian. Tujuan evaluasi ini adalah untuk mengukur kemampuan sistem dalam mengklasifikasikan serta merespons query pengguna secara tepat, sekaligus menemukan aspek yang masih perlu diperbaiki agar akurasi dan relevansi sistem dapat meningkat pada pengembangan berikutnya [6].

Perhitungan dari Confusion Matrix dijelaskan pada bagian berikut.

a. Accuracy

Akurasi (Accuracy) adalah metrik yang menunjukkan proporsi jumlah prediksi yang benar terhadap keseluruhan prediksi yang dilakukan oleh model [27].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

b. Precision

Presisi (Precision) mencerminkan sejauh mana prediksi positif yang dihasilkan model benar-benar sesuai dengan kasus positif sebenarnya.

$$precision = \frac{TP}{TP+FP} \quad (9)$$

c. Recall

Recall (Sensitivitas) mengukur kemampuan model dalam menemukan semua kasus positif yang sebenarnya ada dalam dataset.

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

d. F1 Score

F1-Score adalah rata-rata harmonik antara presisi dan recall, yang menyatukan kedua metrik tersebut ke dalam satu nilai tunggal [28].

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (11)$$

4. HASIL DAN PEMBAHASAN

4.1. Integrasi Bot Telegram

Telegram dipilih sebagai platform karena menyediakan Bot API yang bersifat terbuka, ringan, dan mendukung komunikasi dua arah secara real-time melalui protokol HTTP. Sistem chatbot ini mengotomatisasi pendaftaran pelanggan baru melalui Telegram. Proses dimulai dengan bot menerima pesan dari pengguna, yang kemudian diproses melalui tahap preprocessing teks untuk menghapus simbol dan stopwords. Selanjutnya, fitur teks yang telah diproses dikonversi ke representasi numerik menggunakan TF-IDF, yang mengukur relevansi kata dalam konteks

data. Model *Logistic Regression* kemudian digunakan untuk mengklasifikasikan intent dari pesan, seperti pendaftaran WiFi. Jika intent terdeteksi sebagai “Daftar WiFi”, bot meminta data pendaftaran dari pengguna. Setelah data valid, informasi tersebut disimpan dan bot mengirimkan konfirmasi pendaftaran, menyelesaikan proses secara otomatis. Berikut adalah tampilan sistem chatbot yang telah terintegrasi dengan aplikasi Telegram menggunakan Telegram Bot API.



Gambar 3. Integrasi telegram

Gambar tersebut menunjukkan implementasi chatbot Telegram yang digunakan untuk mengotomatisasi proses pendaftaran pelanggan baru. Sistem mampu melakukan interaksi, validasi data, dan penyampaian konfirmasi secara otomatis sehingga mendukung efisiensi layanan pelanggan.

4.2. Contoh Perhitungan TF-IDF Dan Pembobotan Logistic Regression

a. Data awal dan Hasil Preprocessing

Tahap awal dilakukan dengan membersihkan teks melalui proses *lowercase*, penghapusan tanda baca, penghapusan *stopword*, normalisasi spasi, dan tokenisasi. Hasil preprocessing ditunjukkan pada tabel berikut.

Tabel 2. Contoh Dataset Untuk Perhitungan

Dokumen	Teks Asli	Hasil Preprocessing	Ground Truth
D1	Mau daftar wifi dong	“daftar”, “wifi”	register_wifi
D2	Mau nanya paket wifi yang tersedia?	“nanya”, “paket”, “wifi”, “tersedia”	inquiry

D3	Selamat siang min	“selamat”, “siang”	greeting
----	-------------------------	-----------------------	----------

Pada tahap ini, kata seperti “mau”, “dong”, dan “yang” dihapus karena dianggap sebagai *stopword*. Dengan demikian, model hanya memproses kata yang lebih relevan terhadap intent, seperti daftar, wifi, nanya, paket, tersedia, selamat, dan siang.

b. Perhitungan TF-IDF

Tabel 3. Perhitungan Nilai IDF Pada Term

Term	df(t)	Rumus IDF	Nilai IDF
daftar	1	$\log \frac{1+3}{1+1} + 1$	1.693
wifi	2	$\log \frac{1+3}{1+2} + 1$	1.288
nanya	1	$\log \frac{1+3}{1+1} + 1$	1.693
paket	1	$\log \frac{1+3}{1+1} + 1$	1.693
tersedia	1	$\log \frac{1+3}{1+1} + 1$	1.693
selamat	1	$\log \frac{1+3}{1+1} + 1$	1.693
siang	1	$\log \frac{1+3}{1+1} + 1$	1.693

Pada Tabel 3 menunjukkan bahwa kata wifi memiliki nilai IDF lebih rendah karena muncul pada dua dokumen, yaitu D1 dan D2. Sementara itu, kata seperti daftar, nanya, paket, tersedia, selamat, dan siang hanya muncul pada satu dokumen sehingga nilai IDF-nya lebih tinggi. Setelah perhitungan IDF, Langkah berikutnya adalah menghitung nilai TF-IDF untuk mendapatkan bobot masing-masing term yang terdapat pada ketiga dokumen dalam dataset.

Tabel 4. Perhitungan nilai TF-IDF Pada Dokumen 1

Term	TF	IDF	TF-IDF (TFxIDF)	TF-IDF Normalized
daftar	1	1.693	1.693	0.796
wifi	1	1.288	1.288	0.605

Tabel 5. Perhitungan nilai TF-IDF Pada Dokumen 2

Term	TF	IDF	TF-IDF (TFxIDF)	TF-IDF Normalized
nanya	1	1.693	1.693	0.529
paket	1	1.693	1.693	0.529
tersedia	1	1.693	1.693	0.529
wifi	1	1.288	1.288	0.402

Tabel 6. Perhitungan Nilai TF-IDF Pada Dokumen 3

Term	TF	IDF	TF-IDF (TFxIDF)	TF-IDF Normalized
selamat	1	1.693	1.693	0.707
siang	1	1.693	1.693	0.707

Kolom TF-IDF normalisasi menunjukkan nilai akhir dari pembobotan TF-IDF setelah dilakukan normalisasi L2. Normalisasi ini bertujuan untuk

menyamakan skala representasi setiap dokumen, sehingga dokumen yang lebih panjang tidak secara otomatis memiliki bobot lebih besar dibandingkan dokumen yang lebih pendek. Dalam implementasi *TfidfVectorizer*, nilai TF-IDF awal dihitung dari hasil perkalian TF dan IDF, kemudian setiap nilai pada vektor dokumen dibagi dengan panjang vektor tersebut. Nilai hasil normalisasi inilah yang digunakan sebagai fitur masukan pada model *Logistic Regression* untuk melakukan klasifikasi intent.

c. Pembobotan Logistic Regression

Pada tahap ini nilai hasil vektorisasi TF-IDF digunakan sebagai masukan pada model *Logistic Regression* untuk melakukan klasifikasi intent. Model menghitung nilai logit untuk setiap kelas intent. Nilai logit tersebut selanjutnya dikonversi menjadi probabilitas menggunakan fungsi softmax. Kelas yang memiliki probabilitas paling tinggi ditetapkan sebagai hasil prediksi akhir. Bobot model yang dihasilkan adalah sebagai berikut.

Tabel 7. Perhitungan Nilai Probabilitas Dokumen 1

Kelas	Logit $z_j = (X_1W_{1j} + X_2W_{2j} + \dots + (X_{n1}W_{nj}) + b_j$	Probabilitas $P(y = j X) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$
greeting	-0.273	0.2412
inquiry	-0.159	0.2703
Register_wifi	0.432	0.4885

Tabel 8. Perhitungan Nilai Probabilitas Dokumen 2

Kelas	Logit $z_j = (X_1W_{1j} + X_2W_{2j} + \dots + (X_{n1}W_{nj}) + b_j$	Probabilitas $P(y = j X) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$
greeting	-0.273	0.2412
inquiry	0.432	0.4885
Register_wifi	-0.159	0.2703

Tabel 9. Perhitungan Nilai Probabilitas Dokumen 3

Kelas	Logit $z_j = (X_1W_{1j} + X_2W_{2j} + \dots + (X_{n1}W_{nj}) + b_j$	Probabilitas $P(y = j X) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$
greeting	0.509	0.5177
inquiry	-0.255	0.2412
Register_wifi	-0.255	0.2412

Hasil pengujian menunjukkan bahwa Dokumen 1 diklasifikasikan sebagai *register_wifi* dengan probabilitas 0.4885, Dokumen 2 diklasifikasikan sebagai *inquiry* dengan probabilitas 0.4885, dan Dokumen 3 diklasifikasikan sebagai *greeting* dengan probabilitas 0.5177. Meskipun semua dokumen diklasifikasikan dengan tingkat kepercayaan yang

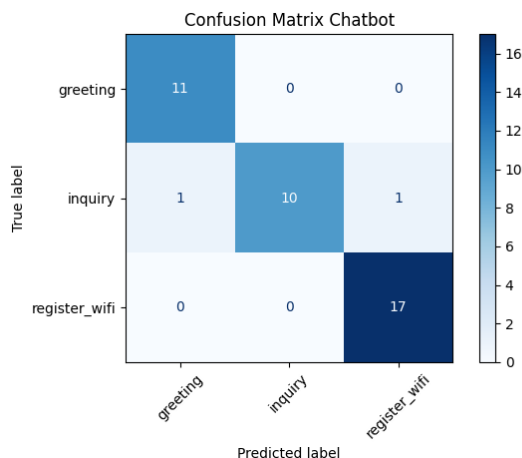
baik, hasil ini mengindikasikan adanya variasi dalam tingkat probabilitas untuk setiap kelas intent yang diuji. Hal ini memberikan gambaran bahwa model memiliki kemampuan untuk menangani variasi dalam cara pengguna menyampaikan pesan, meskipun masih ada ruang untuk perbaikan, terutama dalam menangani kalimat yang ambigu atau dengan konteks yang mirip.

4.3. Pengujian Model

Pengujian model dilakukan menggunakan data log percakapan Telegram yang telah diberi label ground truth. Data uji terdiri dari 40 pesan pengguna yang terbagi ke dalam tiga kelas intent, yaitu greeting, inquiry, dan register_wifi. Model yang diuji merupakan kombinasi ekstraksi fitur TF-IDF dan algoritma *Logistic Regression*. Setiap pesan pengguna diproses melalui tahap text cleaning, stopword removal, pembobotan TF-IDF, kemudian diklasifikasikan menggunakan *Logistic Regression* untuk memperoleh prediksi intent.

4.4. Evaluasi Confusion Matrix

Confusion matrix digunakan untuk melihat pola kesalahan klasifikasi antar kelas intent. Hasil confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan 11 data greeting, 10 data inquiry, dan 17 data register_wifi secara benar. Hasil evaluasi menggunakan diagram *Confusion Matrix* dapat dilihat pada gambar 3 berikut.



Gambar 3. Diagram Confusion Matrix

Berdasarkan diagram confusion matrix tersebut, diperoleh nilai evaluasi model yang mencakup metrik

accuracy, precision, recall, dan F1-score sebagai berikut.

Tabel 11. Perhitungan Confusion Matrix

Intent	Precision	Recall	F1-Score	Support
greeting	0.9167	1.0000	0.9565	11
inquiry	1.0000	0.8333	0.9091	12
register_wifi	0.9444	1.0000	0.9714	17
Accuracy			0.9500	40

Berdasarkan hasil confusion matrix, model berhasil mengklasifikasikan 38 dari 40 data uji dengan benar, sehingga diperoleh nilai akurasi sebesar 95%. Pada kelas greeting, terdapat 11 data yang seluruhnya berhasil dikenali dengan benar. Namun, terdapat satu data dari kelas inquiry yang salah diprediksi sebagai greeting, sehingga precision kelas greeting sebesar 0,9167, sedangkan recall mencapai 1,0000.

Pada kelas inquiry, model berhasil mengklasifikasikan 10 dari 12 data dengan benar. Dua data lainnya mengalami kesalahan klasifikasi, yaitu satu data diprediksi sebagai greeting dan satu data diprediksi sebagai register_wifi. Hal ini menyebabkan recall kelas inquiry menjadi 0,8333, meskipun precision-nya mencapai 1,0000 karena seluruh data yang diprediksi sebagai inquiry benar-benar berasal dari kelas tersebut.

Sementara itu, kelas register_wifi menunjukkan performa yang sangat baik dengan 17 data berhasil diklasifikasikan secara benar dan tidak terdapat data register_wifi yang salah diklasifikasikan ke kelas lain. Nilai precision kelas ini sebesar 0,9444 karena terdapat satu data dari kelas inquiry yang salah diprediksi sebagai register_wifi, sedangkan nilai recall mencapai 1,0000. Secara keseluruhan, hasil ini menunjukkan bahwa model TF-IDF dan *Logistic Regression* mampu melakukan klasifikasi intent dengan baik dan layak digunakan sebagai dasar sistem chatbot pendaftaran pelanggan baru melalui Telegram, meskipun masih terdapat kesalahan pada intent inquiry yang memiliki variasi kalimat lebih beragam.

4.5. Pengujian Sistem

Pengujian sistem menggunakan metode Black Box Testing dilakukan untuk memverifikasi fungsionalitas chatbot Telegram berdasarkan kesesuaian antara input yang diberikan dan output yang dihasilkan. Pengujian ini bertujuan memastikan seluruh fitur sistem dapat beroperasi sesuai dengan kebutuhan yang telah ditetapkan.

Tabel 12. Pengujian Sistem

Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Status
Memulai chatbot	Bot menampilkan pesan pembuka dan menu layanan	Bot menampilkan pesan pembuka dan menu layanan	Berhasil
Identifikasi intent pengguna	Bot memberikan respons sesuai dengan intent pengguna	Bot memberikan respons yang sesuai	Berhasil
Pengiriman pesan di luar intent yang tersedia	Bot memberikan respons default	Respons diberikan dengan benar	Berhasil

Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Status
Memilih layanan pendaftaran WiFi	Bot memulai proses pendaftaran pelanggan baru	Bot memulai proses pendaftaran pelanggan baru	Berhasil
Mengisi data pelanggan lengkap	Bot menerima dan menyimpan data pendaftaran	Data berhasil diterima dan diproses	Berhasil
Validasi data pendaftaran	Bot menampilkan ringkasan data untuk konfirmasi	Ringkasan data ditampilkan dengan benar	Berhasil
Konfirmasi pendaftaran	Bot mengirimkan pesan bahwa pendaftaran berhasil	Bot menampilkan pesan konfirmasi bahwa pendaftaran berhasil dikirim	Berhasil
Pembatalan pendaftaran	Bot membatalkan proses pendaftaran	Proses pendaftaran dibatalkan	Berhasil
Pengulangan proses pendaftaran	Bot memulai proses pendaftaran baru	Proses pendaftaran baru berhasil dijalankan	Berhasil
Penyimpanan data pendaftaran ke Google Sheets	Data pelanggan tersimpan pada Google Sheets	Data berhasil tersimpan pada Google Sheets	Berhasil

Berdasarkan hasil Black Box Testing, seluruh fungsi chatbot berjalan sesuai dengan hasil yang diharapkan. Sistem mampu mengidentifikasi intent pengguna, memproses pendaftaran pelanggan baru, serta menyimpan data pendaftaran ke Google Sheets tanpa ditemukan kesalahan fungsional. Hasil tersebut menunjukkan bahwa chatbot telah memenuhi kebutuhan fungsional sistem dan layak digunakan sebagai media otomasi pendaftaran pelanggan melalui Telegram.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa kombinasi TF-IDF dan *Logistic Regression* telah berhasil menghasilkan model dengan akurasi tinggi sebesar 95,00%. Model ini mampu mengklasifikasikan intent pengguna dengan baik, khususnya pada kelas *register_wifi* yang menunjukkan performa terbaik dengan F1-score sebesar 0,9714. Kelas *greeting* juga memberikan hasil yang sangat baik dengan F1-score sebesar 0,9565, sedangkan kelas *inquiry* memiliki sedikit penurunan performa dengan F1-score sebesar 0,9091. Meskipun model menunjukkan performa yang memuaskan secara keseluruhan, kesalahan klasifikasi terutama terjadi pada kelas *inquiry*, yang menunjukkan bahwa variasi data latih pada kelas ini perlu ditingkatkan. Selain itu, penggunaan bahasa informal dan variasi kalimat dalam kelas *inquiry* mengindikasikan perlunya normalisasi bahasa untuk meningkatkan kemampuan model dalam mengklasifikasikan pertanyaan pengguna dengan lebih akurat. Secara keseluruhan, model yang dihasilkan telah cukup andal dalam mengidentifikasi intent pengguna dan dapat digunakan sebagai dasar untuk sistem chatbot pendaftaran pelanggan baru melalui Telegram. Pengembangan lebih lanjut pada data latih dan preprocessing di kelas *inquiry* dapat lebih menyempurnakan performa sistem chatbot ini, khususnya dalam menghadapi variasi kalimat dan bahasa yang lebih kompleks. Selain itu, metode-metode lain yang dapat digunakan untuk ekstraksi fitur berdimensi tinggi dan mengatasi imbalance class, seperti penerapan BERT (Bidirectional Encoder Representations from Transformers) dan ADASYN

(Adaptive Synthetic Sampling), dapat membantu meningkatkan kualitas model dalam klasifikasi pesan dengan lebih baik.

DAFTAR PUSTAKA

- [1] L. Nicolescu and M. T. Tudorache, "Human-Computer Interaction in Customer Service: The Experience with AI Chatbots—A Systematic Literature Review," May 01, 2022, *MDPI*. doi: 10.3390/electronics11101579.
- [2] J. Yun and J. Park, "The Effects of Chatbot Service Recovery With Emotion Words on Customer Satisfaction, Repurchase Intention, and Positive Word-Of-Mouth," *Front. Psychol.*, vol. 13, May 2022, doi: 10.3389/fpsyg.2022.922503.
- [3] T. Zhao, J. Cui, J. Hu, Y. Dai, and Y. Zhou, "Is Artificial Intelligence Customer Service Satisfactory? Insights Based on Microblog Data and User Interviews," *Cyberpsychology, Behav. Soc. Netw.*, vol. 25, no. 2, pp. 110–117, Feb. 2022, doi: 10.1089/cyber.2021.0155.
- [4] R. F. Hudiyo, "Exploring Indonesian Companies' Chatbots to Gather Customer Experience," *MDPI AG*, Dec. 2022, p. 1. doi: 10.3390/proceedings2022083001.
- [5] R. H. Ardiansyah and A. G. Sulaksono, "Layanan pelanggan berbasis Natural Language Processing melalui chatbot pada aplikasi pesan," *J. Inf. Syst. Appl. Dev.*, vol. 1, no. 1, pp. 29–37, Mar. 2023, doi: 10.26905/jisad.v1i1.9858.
- [6] D. Kristanto, R. A. Ramadhani, and A. B. Setiawan, "Pengembangan Chatbot Layanan Informasi Kampus Menggunakan TF-IDF," *J. Algoritma*, vol. 22, no. 2, pp. 103–115, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2350.
- [7] Nuzul Hikmah, Dyah Ariyanti, and Ferry Agus Pratama, "Implementasi Chatbot Sebagai Virtual Assistant di Universitas Panca Marga Probolinggo menggunakan Metode TF-IDF," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 2, pp. 133–148, Aug. 2022, doi: 10.35746/jtim.v4i2.225.
- [8] M. S. Maulana, Y. Anshori, R. Azhar, R. Laila, and N. T. Lapatta, "Implementasi Pembobotan

- Tf-Idf Pada Chatbot Telegram Untuk Sistem Layanan Informasi,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 3, pp. 1869–1877, 2025, doi: 10.29100/jupi.v10i3.6314.
- [9] N. Pratama¹, Z. L. Ramadan, Y. Putra, B. Satria, and D. A. Asy’ari, “Penerapan Metode TF-IDF dan Cosine Similarity pada ChatbotInformasi Pilkada Bekasi 2024 Berbasis Telegram,” vol. 4, no. 1, pp. 93–100, 2025.
- [10] R. Aulia and A. Purnama, “Development of an Intent-Classification Chatbot to Support Operational Services at Kadin Indonesia,” *Brill. Res. Artif. Intell.*, vol. 5, no. 2, pp. 1171–1180, 2025, doi: 10.47709/brilliance.v5i2.7438.
- [11] R. Y. Pratama and Supatman, “Pengembangan Sistem Chatbot Cerdas Berbasis Natural Language Processing (Nlp) Untuk Peningkatan Layanan Informasi Hotel,” *J. Inform. dan Tek. Elektro Terap.*, vol. 14, no. 1, 2026, doi: 10.23960/jitet.v14i1.8528.
- [12] A. Syahra, P. Mecca, W. A. Hidayat, H. Tuasikal, and U. M. Sorong, “Pemanfaatan Teknologi Kecerdasan Buatan (Artificial Intelligence) Dalam Sistem Peradilan Pidana di Indonesia,” *J. Sos. Teknol.*, vol. 5, no. 6, pp. 1–17, 2025.
- [13] D. O. Sihombing, “Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis,” *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 396, 2022, doi: 10.30865/json.v4i2.5374.
- [14] A. H. Dani, E. Y. Puspaningrum, and R. Mumpuni, “Studi Performa TF-IDF dan Word2Vec Pada Analisis Sentimen Cyberbullying,” *J. Tek. Inform. dan Terap.*, vol. 2, no. 2, pp. 94–106, 2024, [Online]. Available: <https://doi.org/10.62951/router.v2i2.76>
- [15] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *Sintesia*, vol. 1, pp. 81–88, 2022.
- [16] A. Firizkiansah, A. Muhammad, and I. R. Maulana, “Optimasi Klasifikasi Data Teks Menggunakan Algoritma Logistic Regression dengan TF-IDF dan SMOTE,” *JIKOMTI J. Ilm. Ilmu Komput. dan Teknol. Inf.*, vol. 2, no. 1, pp. 29–38, 2025.
- [17] M. Brusco, “Logistic Regression via Excel Spreadsheets: Mechanics, Model Selection, and Relative Predictor Importance,” *INFORMS Trans. Educ.*, vol. 23, no. 1, pp. 1–11, 2022, doi: 10.1287/ited.2021.0263.
- [18] M. D. Ubaidillah, A. Yusril, I. Sya, and L. Cahyani, “Klasifikasi Berita Hoaks Berbahasa Indonesia Menggunakan TF- IDF dengan Jumlah Fitur Tinggi dan Algoritma K-Nearest Neighbors sebagai Model Baseline,” vol. 6, no. 1, pp. 63–71, 2026.
- [19] A. Haris, T. Tukino, A. Hananto, and F. Nurapriani, “Penerapan Algoritma Naive Bayes Untuk Klasifikasi Sentimen Ulasan Pengguna Aplikasi Pidi Umkm Dengan Pembobotan Fitur Tf-Idf,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 5, pp. 8457–8462, 2025, doi: 10.36040/jati.v9i5.15019.
- [20] Q. Budiman, S. Mouton, L. Veenhoff, and A. Boersma, “CHATBOT PADA SMART COOPERATIVE ORIENTED PROBLEM MENGGUNAKAN NATURAL LANGUAGE PROCESSING DAN NAIVE BAYES CLASSIFIER CHATBOT,” *J. Inov. Penelit.*, vol. 1, no. 0.1101/2021.02.25.432866, pp. 1–15, 2021.
- [21] M. A. R. Reynaldhi and Y. Sibaroni, “Analisis Sentimen Review Film pada Twitter menggunakan Metode Klasifikasi Hybrid SVM, Naïve Bayes, dan Decision Tree,” *e-Proceeding Eng.*, vol. 8, no. 5, p. 10127, 2021.
- [22] J. E. Br Sinulingga and H. C. K. Sitorus, “Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF,” *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.
- [23] Andi Riswawan, “Implementing TF-IDF and Logistic Regression for Sentiment Analysis of YouTube Comments on the iPhone 16,” *J. Teknol. Dan Open Source*, vol. 8, no. 2, pp. 604–611, 2025, doi: 10.36378/jtos.v7i2.4753.
- [24] I. Mustafa and A. Allam, “Multinomial Logistic Regression in C: A Comprehensive Implementation for Multi-Class Classification,” 2026, doi: <https://doi.org/10.13140/RG.2.2.35362.59841>.
- [25] I. Alpiana, W. Yustanti, and Y. Yamasari, “Otomatisasi Klasifikasi Tingkat Urgensi Keluhan E-Layanan Unesa Berbasis TF-IDF dan Logistic Regression,” *R2J*, vol. 8, no. 1, 2025, doi: 10.38035/rj.v8i1.
- [26] I. L. Sebayang, Alvin Christoper Kharisma, A. Sujjada, and K. Kamdan, “Implementasi Chatbot dengan pendekatan Natural Language Processing dan Naïve Bayes dalam meningkatkan layanan perusahaan,” *LPPM STIKOM Tunas Bangsa*, vol. Vol 5, No, 2024.
- [27] B. Stojanović *et al.*, “Follow the trail: Machine learning for fraud detection in fintech applications,” *Sensors*, vol. 21, no. 5, pp. 1–43, 2021, doi: 10.3390/s21051594.
- [28] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, “Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems,” *PLOS Digit. Heal.*, vol. 2, no. 11 November, pp. 1–19, 2023, doi: 10.1371/journal.pdig.0000290.