

Chatbot WhatsApp Asisten *Workout* dengan *Integrasi Large Language Model Gemini dan spaCy*

Muhammad Aris Khuzaini¹, Ika Ratna Indra Astutik*², Arif Senja Fitriani³, Yulian Findawati⁴

^{1,2,3,4}Informatika, Universitas Muhammadiyah Sidoarjo

e-mail: ¹211080200086@umsida.ac.id, *²ikaratna@umsida.ac.id, ³asfjim@umsida.ac.id,

⁴yulianfindawati@umsida.ac.id

Abstrak

Olahraga memiliki peran penting dalam meningkatkan kesehatan fisik dan mental serta mencegah risiko penyakit kronis. Namun, berdasarkan laporan *World Health Organization* (WHO) tahun 2022 menunjukkan bahwa sekitar 1,4 miliar orang dewasa di dunia masih mengalami kurangnya aktivitas fisik, yang meningkatkan risiko penyakit kronis. Penelitian ini bertujuan mengembangkan chatbot asistensi olahraga untuk membantu individu menjaga rutinitas olahraga yang konsisten. Metode pengembangan sistem menggunakan *Software Development Life Cycle* (SDLC) model *iterative incremental*, sedangkan *Natural Language Processing* (NLP) berbasis spaCy digunakan sebagai metode pemrosesan teks untuk memahami input pengguna. Selain itu, chatbot diintegrasikan dengan API *Large Language Model* (LLM) Gemini untuk menghasilkan respons berdasarkan basis pengetahuan chatbot. Hasil evaluasi menunjukkan kinerja chatbot yang baik, dengan akurasi 79%, presisi 85%, recall 79%, dan F1-score 77%. Pengujian black-box mencatat tingkat keberhasilan 91,67% dari 12 skenario. Dari hasil tersebut, dapat disimpulkan bahwa chatbot asistensi olahraga ini efektif dalam mendukung pola hidup sehat dan dapat menjadi pendekatan teknologi untuk membantu meningkatkan aktivitas fisik secara mandiri.

Kata kunci—Olahraga, Chatbot, *Natural Language Processing*, *Large Language Model*, Gemini

Abstract

Exercise plays an important role in improving both physical and mental health while also preventing the risk of chronic diseases. However, according to the 2022 World Health Organization (WHO) report, approximately 1.4 billion adults worldwide still experience physical inactivity, which increases the risk of chronic diseases. This research aims to develop an exercise assistance chatbot to help individuals maintain a consistent workout routine. The system development process follows the Software Development Life Cycle (SDLC) iterative incremental model, while Natural Language Processing (NLP) using spaCy is employed for text processing to understand user input. Additionally, the chatbot is integrated with Gemini's Large Language Model (LLM) API to generate responses based on its knowledge base. Evaluation results indicate good chatbot performance, with an accuracy of 79%, precision of 85%, recall of 79%, and an F1-score of 77%. Black-box testing recorded a 91.67% success rate across 12 test scenarios. These results suggest that the exercise assistance chatbot is effective in promoting a healthy lifestyle and serves as a technological solution to encourage independent physical activity.

Keywords—*Workout, Chatbot, Natural Language Processing, Large Language Model, Gemini*

1. PENDAHULUAN

Olahraga atau *workout*, suatu aktivitas fisik dengan peran yang cukup krusial dalam menunjang kesehatan manusia[1]. Olahraga

mencakup berbagai aktivitas fisik seperti kardio dan latihan kekuatan juga dirasakan sebagai kegiatan positif yang berkontribusi pada peningkatan rasa koherensi, kesehatan mental, dan kesejahteraan[2]. Aktivitas ini

juga terbukti secara ilmiah dapat mencegah risiko penyakit jantung koroner, osteoporosis, diabetes, obesitas, depresi, serta meningkatkan daya tahan dan kekuatan otot[3]. Merujuk pada laporan *World Health Organization* (WHO) tahun 2022, sekitar 1,4 miliar orang dewasa di seluruh dunia mengalami kekurangan aktivitas fisik, yang berkontribusi pada peningkatan risiko berbagai penyakit kronis dan kondisi kesehatan lainnya[4]. Berdasarkan laporan sebelumnya, meskipun memiliki manfaat yang jelas, ternyata masih banyak individu yang masih kesulitan dalam mempertahankan rutinitas olahraga yang konsisten. Hal ini mungkin disebabkan karena aktivitas fisik masih dianggap sebagai beban tambahan yang harus dilakukan saat sedang menghadapi berbagai tantangan kesehatan, sosial, dan finansial[5]. Situasi ini mencerminkan bahwa persepsi dan hambatan terhadap aktivitas fisik menjadi tantangan dalam upaya meningkatkan kebiasaan hidup sehat.

Hambatan dari aktivitas fisik ini antara lain kebingungan dalam pemilihan latihan yang tepat karena kurangnya pengetahuan serta kondisi tiap orang yang berbeda-beda, sehingga banyak yang tidak dapat memulai rutinitas olahraga mereka. Tantangan ini juga berlaku untuk mereka yang pergi ke gym, di mana biaya sering menjadi kendala[6]. Meskipun mampu pergi ke gym, terdapat penelitian yang menyatakan bahwa 67% orang tidak rutin pergi ke gym karena merasa tidak nyaman dengan lingkungan yang asing, kesulitan menggunakan peralatan fitness, dan kurangnya interaksi dengan pihak gym[7]. Oleh karena itu, personalisasi dalam olahraga juga memiliki peran penting untuk mengoptimalkan rencana latihan sesuai preferensi dan kebutuhan individu[8]. Melalui kondisi tersebut, maka terlihat diperlukannya sistem yang dapat memberi asistensi olahraga, baik untuk penggunaan di rumah maupun di gym. Salah satu sistem yang dapat memenuhi kriteria tersebut adalah chatbot, implementasi chatbot terbukti mampu memberikan arahan dan informasi yang disesuaikan dengan kebutuhan setiap pengguna secara cepat dan akurat[9].

Terdapat beberapa penelitian terdahulu yang berkaitan dengan permasalahan ini. Pada penelitian yang dilakukan oleh Lola pada

tahun 2021 mengusulkan sebuah chatbot manajemen kebugaran dengan memanfaatkan teknologi Natural Language Understanding (NLU) dan Natural Language Processing (NLP) disertai dengan integrasi IBM Watson[7]. NLP sendiri merupakan teknologi yang dapat menawarkan kemampuan untuk memahami, merespons, dan mengidentifikasi perintah maupun pertanyaan dalam bahasa alami[10]. Penelitian tersebut menunjukkan bagaimana chatbot NLP dapat menawarkan rencana diet, latihan di rumah, sesi konseling, dan rekomendasi kebugaran. Pembeda penelitian ini terletak pada belum adanya integrasi chatbot ke dalam aplikasi pada penelitian tersebut, serta perbedaan dalam pengintegrasian model yakni penggabungan NLP dengan *large language model* (LLM).

Penelitian selanjutnya adalah tentang pemanfaatan API (*Application Programming Interface*) Gemini. Penelitian yang dilakukan oleh Hariyadi pada tahun 2024 mengimplementasikan Gemini AI sebagai pembuat teks deskripsi karya secara otomatis pada aplikasi pameran seni. Penelitian tersebut hanya bergantung pada teks *prompt* dalam pembuatan deskripsi karya berdasarkan judul karya dan gambar melalui Gemini AI[11].

Kemudian terdapat penelitian lain yang dilakukan oleh Aprilia pada tahun 2024 yang menerapkan API Gemini pada chatbot layanan peminjaman novel secara online berbasis website. Sistem chatbot pada penelitian tersebut cenderung bergantung pada data umum dari Gemini sehingga mengurangi tingkat akurasi dan relevansi informasi yang diberikan kepada pengguna jika terdapat novel yang tidak diketahui oleh Gemini[12].

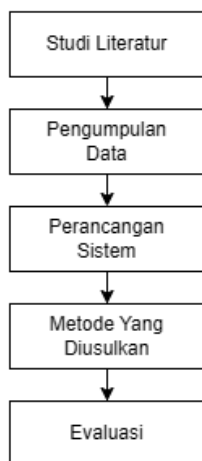
Penelitian yang mirip juga dilakukan oleh Alia pada tahun 2024, pada penelitian tersebut dilakukan implementasi NLP pada chatbot WhatsApp. Pembeda dari penelitian ini adalah metode yang diterapkan dan tahapan proses NLP pada penelitian tersebut yang hanya sampai pada tahap *stemming*[13].

Merujuk pada permasalahan dan penelitian terdahulu yang telah diuraikan sebelumnya, maka peneliti berinisiatif untuk mengembangkan chatbot ini agar dapat digunakan oleh masyarakat umum serta pihak gym. Penelitian ini menggabungkan NLP dengan API *Large Language Model* (LLM) Gemini dengan tujuan untuk menyediakan *query* jawaban yang lebih baik. Gemini

memungkinkan chatbot untuk memberikan respons yang lebih canggih, akurat, dan relevan[14]. Chatbot akan dilengkapi dengan rekomendasi latihan fisik, fitur pemahaman alat olahraga, kalkulator BMI, serta konsultasi protein, dengan harapan dapat membantu pengguna dalam asistensi olahraga.

2. METODE PENELITIAN

Alur penelitian dibuat sebagai tahapan pelaksanaan yang berurutan dalam menyelesaikan penelitian yang ditunjukkan pada Gambar 1. Tahap pertama adalah studi literatur, proses pengumpulan data dari berbagai referensi yang relevan dengan topik penelitian untuk memberikan landasan teori dalam pelaksanaan penelitian[15]. Studi literatur dilakukan dengan cara melakukan pencarian di buku, jurnal, dan artikel[16]. Selanjutnya, tahap pengumpulan data untuk kebutuhan pelatihan model dan analisis kebutuhan sistem. Peneliti menggunakan dataset "*Fitness Exercises Dataset*" dari Kaggle yang berisi sekitar 1300 data latihan kebugaran, termasuk nama latihan, target otot, dan alat yang digunakan. Ditambah dengan data pendukung lebih dari 200 kata kunci relevan. Selain itu, sekitar 100 data contoh pertanyaan atau perintah dari 9 kategori dikumpulkan untuk melatih model intent. Semua data dianalisis untuk memastikan validitas dan relevansi sebagai basis chatbot asisten *workout*. Tahap ketiga perancangan sistem, dibuatlah desain arsitektur sistem chatbot. Kemudian dilanjutkan dengan penerapan metode pengembangan yang diusulkan, hingga berakhir pada tahap evaluasi.



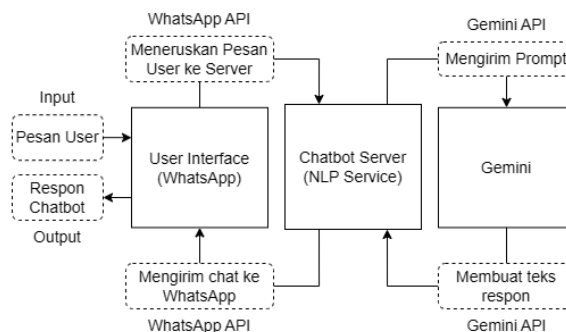
Gambar 1. Alur Penelitian

2.1 Arsitektur Sistem

Arsitektur sistem adalah struktur keseluruhan sistem mencakup hubungan antar komponen yang mengatur desain sistem untuk memastikan integrasi dan efisiensi dalam mencapai tujuan[17]. Desain arsitektur sistem dibagi menjadi dua bagian, yaitu desain arsitektur utama chatbot yang menggambarkan alur sederhana interaksi antara pengguna dan sistem, serta desain arsitektur server chatbot yang menjelaskan secara rinci proses internal, termasuk pemrosesan input, analisis NLP, dan integrasi dengan API eksternal.

a. Desain Arsitektur Sistem

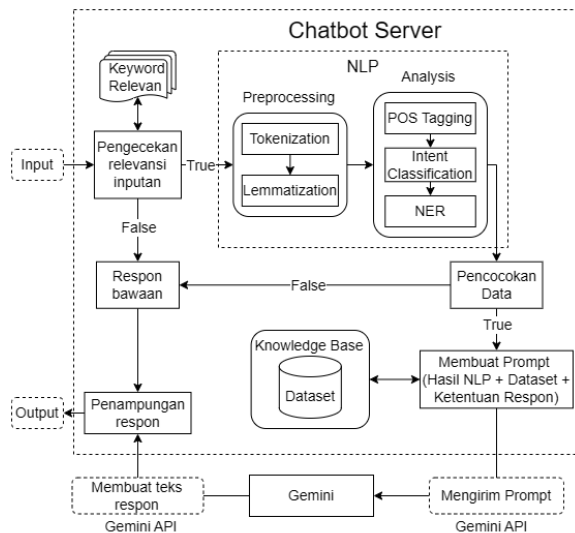
Desain arsitektur sistem chatbot ditunjukkan pada Gambar 2. Melalui API WhatsApp pesan dikirim ke server, kemudian dilakukan pengecekan relevansi input untuk diproses pada tahap layanan NLP dengan spaCy. Hasil analisis digunakan untuk membentuk *prompt* yang mencakup hasil analisis, informasi dataset, dan ketentuan respons. Respons dari API Gemini selanjutnya akan ditampung dan dikirimkan ke pengguna.



Gambar 2. Arsitektur Sistem

b. Desain Arsitektur Server Chatbot

Pada chatbot server terdapat sebuah proses NLP. Alur dari proses tersebut ditunjukkan pada Gambar 3 di bawah ini. Proses NLP meliputi 2 tahap utama, *preprocessing* yang terdiri dari proses tokenisasi, normalisasi, kemudian pada tahap *analysis* terdiri dari proses ekstraksi fitur kata, klasifikasi intent, dan pengenalan entitas bernama. Semua tahap tersebut bergantung pada basis pengetahuan (*knowledge base*) chatbot yang berasal dari dataset *workout* kemudian dilengkapi dengan data kata kunci dan data pertanyaan atau perintah yang relevan dengan *workout*.

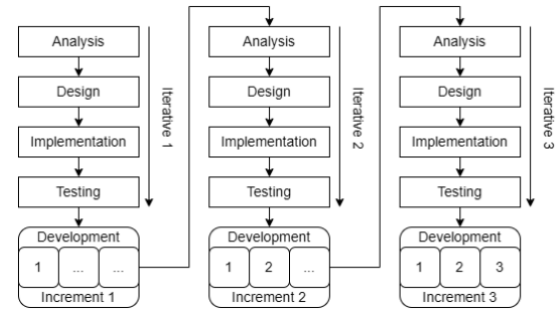


Gambar 3. Arsitektur Server Chatbot

2.2 Metode Pengembangan

Chatbot dapat dikembangkan secara mandiri atau melalui aplikasi percakapan dengan memanfaatkan metode pemrosesan bahasa alami (NLP) untuk memahami dan merespons percakapan[18]. Pengembangannya mengacu pada *System Development Life Cycle* (SDLC), sebuah proses bertahap untuk menganalisis, merancang, dan membangun sistem informasi secara efisien sebelum diserahkan kepada pengguna[19]. Model pengembangan yang digunakan pada penelitian ini adalah pengembangan *Iterative Incremental* yang merupakan gabungan dari model *Iterative* dan *Incremental*.

Model *iterative* sendiri merupakan strategi pengembangan perangkat lunak yang berbasis pada pengulangan siklus[20]. Sedangkan model *incremental* merupakan pendekatan pengembangan perangkat lunak secara bertahap melalui beberapa tahap rilis, dengan setiap tahap menambahkan fungsionalitas baru[21]. Model *Iterative Incremental* sendiri merupakan metode pengembangan sistem yang berasal dari permasalahan pada model *waterfall*[22]. Model ini dirancang untuk mengatasi kelemahan model *waterfall* yang tidak mendukung iterasi serta kekurangan model prototipe yang memiliki durasi pengembangan yang terlalu singkat[23]. Penggambaran dari model *Iterative Incremental* ini ditunjukkan pada Gambar 4 berikut.

Gambar 4. Model *Iterative Incremental*

2.3 Pengembangan Sistem

Tahapan pengembangan sistem melibatkan sejumlah proses utama yang dijelaskan sebagai berikut.

a. Teknologi Sistem

Chatbot ini menggunakan LLM Gemini dan spaCy sebagai teknologi utama. LLM sendiri merupakan teknologi kecerdasan buatan yang mampu memahami dan menghasilkan teks secara kontekstual berdasarkan model bahasa berskala besar, salah satu dari LLM yang digunakan pada penelitian ini adalah Gemini. spaCy adalah sebuah library open-source yang dirancang untuk pemrosesan bahasa alami berbasis Python[24]. Bahasa pemrograman utama yang dipakai pada layanan NLP adalah bahasa Python, kemudian menggunakan *framework* Flask yang berperan pada bagian *back-end*. Selanjutnya thefuzz merupakan *library* pencocokan string fuzzy. Kedua layanan dengan bahasa yang berbeda tersebut kemudian dihubungkan dengan REST API.

b. Implementasi NLP

Teknik NLP seperti tokenisasi, normalisasi, ekstraksi fitur, klasifikasi intent, dan identifikasi entitas diterapkan menggunakan spaCy. Pada proses *preprocessing* dilakukan pelatihan model intent menggunakan data sampel pertanyaan atau perintah berdasarkan kategori seperti *ask-exercise*, *muscle-target*, *ask-equipment*, dan sebagainya. Tahap *preprocessing* merupakan tahap penting dalam teknik ini sebelum melakukan operasi klasifikasi data[25]. Data kata kunci relevan dipecah menjadi beberapa kategori kemudian digunakan pada pengecekan relevansi input menggunakan pencocokan fuzzy serta mekanisme penambahan skor keyakinan.

Pada fungsi tokenisasi, penghapusan *stopwords* dilakukan sedikit penyesuaian untuk menghindari penghapusan nama latihan yang seharusnya tidak dihapus seperti push-up, sit-up, pull-up, dan lain sebagainya. Tahap NER (*Named Entity Recognition*) dilakukan pelatihan model menggunakan dataset untuk menambahkan entitas seperti nama latihan, otot, bagian tubuh, dan alat olahraga.

Tahap terakhir, pembuatan prompt dibuat berdasarkan hasil analisis NLP, *query* data dari dataset, dan ketentuan hasil respons. Ketentuan ini meliputi pencegahan olahraga yang berbahaya (membutuhkan pengawasan ahli), aturan bahasa yang konsisten, dan penyederhanaan respons untuk memastikan Gemini membuat respons yang sesuai.

c. Pengujian

Pengujian yang dilakukan pada sistem chatbot meliputi pengujian evaluasi model yang digunakan untuk mengevaluasi performa model melalui pengukuran metrik (Akurasi, Presisi, Recall, dan F1-score)[26]. Perhitungan dari evaluasi model akan dijelaskan dalam bentuk rumus sebagai berikut.

1. Akurasi (*Accuracy*) mengukur prediksi benar yang ditunjukkan oleh rumus (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2. Presisi (*Precision*) mengukur proporsi prediksi positif ada pada rumus (2).

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3. Recall mengukur proporsi data positif yang terdeteksi dengan benar oleh model ditunjukkan pada rumus (3).

$$Precision = \frac{TP}{TP + FN} \quad (3)$$

4. F1-score menghitung ukuran rata-rata harmonik berdasarkan Presisi dan Recall ditunjukkan oleh rumus (4).

$$F1 = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

Keterangan:

TP (True Positive): Sampel positif diprediksi sebagai positif.

FP (False Positive): Sampel negatif diprediksi sebagai negatif.

TN (True Negative): Sampel negatif diprediksi sebagai positif.

FN (False Negative): Sampel positif diprediksi sebagai negatif.

Pengujian terakhir dilakukan menggunakan metode *Black-box Testing*. Pengujian dilakukan dengan mengamati hasil operasi dan memverifikasi fungsionalitas perangkat lunak menggunakan data uji, dengan fokus pada kinerja sistem[27]. Dengan melakukan pengujian, kelemahan fitur dapat diidentifikasi sehingga aplikasi dapat memiliki kinerja yang optimal[28].

3. HASIL DAN PEMBAHASAN

3.1 Hasil Metode *Iterative Incremental*

Penerapan metode ini menghasilkan tiga *incremental* yang dijelaskan sebagai berikut.

a. *Incremental 1*

Tahap awal berfokus pada integrasi API WhatsApp berbasis Node.js dan pengembangan chatbot sederhana berbasis aturan (*rule-based*) menggunakan logika *if-else*. Hasilnya, chatbot dapat merespons pesan pengguna sesuai aturan yang telah ditentukan, tetapi masih terbatas dalam memahami variasi input.

b. *Incremental 2*

Pada tahap ini dilakukan penggabungan API WhatsApp yang telah diimplementasikan sebelumnya, dilanjut dengan penambahan LLM Gemini dan NLP spaCy berbasis Python. Hasil dari tahap ini, kedua layanan yang berbeda antara API WhatsApp dan server NLP dapat berkomunikasi melalui REST API. Chatbot dapat merespons pengguna dengan jawaban yang sesuai berdasarkan basis pengetahuan dalam model NLP kustom.

c. *Incremental 3*

Tahap terakhir mencakup penambahan fitur workout seperti kalkulator BMI, penghitung protein, rekomendasi latihan, dan pencarian alat workout dengan gambar. Setiap fitur dikembangkan secara terpisah untuk memudahkan pengujian tiap fitur, kemudian dihubungkan dalam sistem utama agar semua fitur tetap dapat saling berkomunikasi dan beroperasi sebagai satu kesatuan dalam chatbot.

3.2 Hasil Implementasi

Implementasi API Gemini dan NLP ditunjukkan pada Gambar 5 dan 6.

```

GNU nano 4.8
load_dotenv()
genai.configure(api_key=os.getenv("GEMINI_API_KEY"))
model = genai.GenerativeModel("gemini-1.5-flash")

def get_gemini_response_text(prompt):
    try:
        response = model.generate_content([prompt])

        if response and response.text:
            return {"response": response.text}
        else:
            return {"error": "No response from Gemini."}
    except Exception as e:
        return {"error": str(e)}

```

Gambar 5. Implementasi API Gemini

```

# Check if the input is related to workouts
if is_related_to_workout(translated_input, KEYWORDS):
    # Step 1: Classify input and get prompt info
    prompt_info = classify_input_with_intent_model(translated_input, df, KEYWORDS)

    # Step 2: Conditional processing based on refined label
    refined_label = prompt_info["refined_label"]

    if refined_label in ["calc_bmi", "protein_intake"]:
        # Directly call a specific function instead of constructing a prompt
        height, weight = extract_data(translated_input)
        response = bmi_response(refined_label, height=height, weight=weight)
    else:
        # Construct prompt and get Gemini response for other labels
        result = construct_prompt(
            user_input=translated_input,
            predicted_label=prompt_info["predicted_label"],
            confidence_score=prompt_info["confidence_score"],
            matched_entities=prompt_info["matched_entities"],
            matched_keywords=prompt_info["matched_keywords"],
            refined_label=prompt_info["refined_label"],
            prompt_info=prompt_info.get("prompt_info", {}) # Handle missing key gracefully
        )

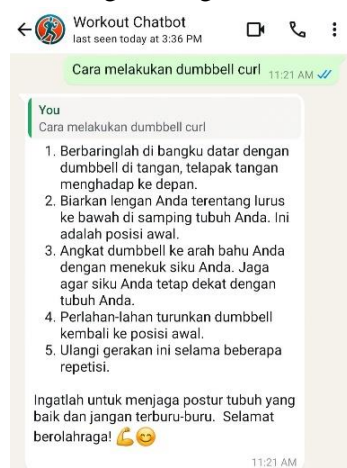
        # Use the constructed prompt to get the response
        prompt = result
        response = get_gemini_response_text(prompt)
    else:
        # Fallback response for unrelated inputs
        response = "Saya tidak yakin bagaimana menjawabnya. Coba tanyakan sesuatu terkait olahraga."

```

Gambar 6. Implementasi NLP

a. Tanya Seputar *Workout*

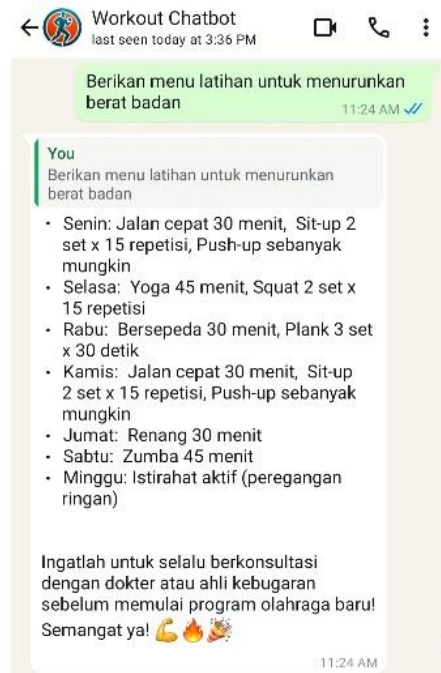
Pada Gambar 7 ditunjukkan bagaimana chatbot memberi tutorial dan menjawab pertanyaan tentang olahraga.



Gambar 7. Tanya Seputar *Workout*

b. Rekomendasi Menu Latihan

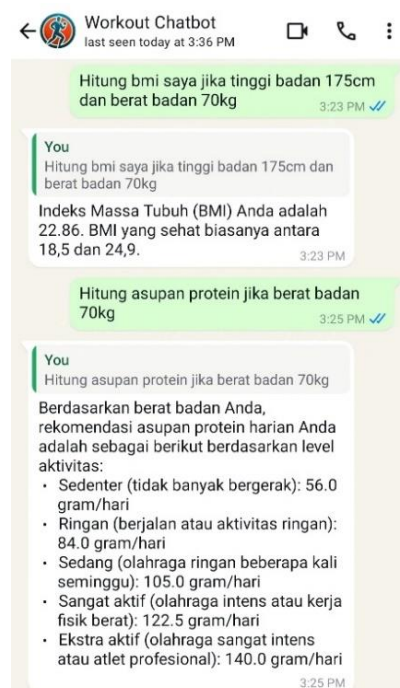
Gambar 8 menunjukkan chatbot memberikan menu latihan kepada pengguna.



Gambar 8. Rekomendasi Menu Latihan

c. Hitung BMI dan Asupan Protein

Gambar 9 menunjukkan bagaimana chatbot dapat berfungsi sebagai kalkulator BMI dan asupan protein.



Gambar 9. BMI dan Asupan Protein

d. Pertanyaan Tidak Relevan

Gambar 10 menunjukkan bagaimana chatbot memberikan respon bawaan ketika pengguna menanyakan hal yang tidak relevan atau tidak sesuai dengan ketentuan chatbot.



Gambar 10. Pertanyaan Tidak Relevan

e. Tanya Alat Workout dengan Gambar

Gambar 11 menunjukkan bahwa chatbot dapat menerima pesan media berupa gambar yang digunakan untuk menanyakan informasi tentang alat workout.



Gambar 11. Tanya dengan Gambar

3.3 Hasil Pengujian

Terdapat 2 pengujian pada penelitian ini, yakni pengujian evaluasi model dan pengujian Black-box yang dijabarkan sebagai berikut.

a. Evaluasi Model

Pada pengujian ini dilakukan evaluasi terhadap kinerja model NLP yang digunakan dalam chatbot, khususnya dalam mengenali intent dan entity dari input pengguna. Evaluasi ini menggunakan metrik seperti yang telah dijelaskan sebelumnya untuk mengukur seberapa baik model dapat memproses pesan.

Jumlah data sampel yang digunakan pada pengujian ini sebanyak lebih dari 100 sampel pertanyaan atau perintah yang berhubungan dengan topik workout. Data tersebut kemudian dikelompokkan menjadi beberapa kategori utama seperti *ask-exercise*, *muscle-target*, *ask-bodypart*, *ask-equipment*.

Pengujian dilakukan menggunakan *classification report* dari library Scikit-learn. Fungsi utama library ini adalah membantu otomatisasi pengujian data sampel yang banyak dalam perhitungan metrik evaluasi. Pada Gambar 12 akan ditunjukkan bagaimana Scikit-learn digunakan.

```

GNU nano 4.8 test_eval_model.py
from sklearn.metrics import (
    accuracy_score, precision_score,
    recall_score, f1_score, classification_report
)
import pandas as pd
from nlp_service import classify_input_with_intent_model

def load_test_samples(file_path):
    df = pd.read_csv(file_path)
    return list(zip(df["user_input"], df["true_label"]))

def evaluate_model(test_samples):
    y_true = []
    y_pred = []

    for user_input, true_label in test_samples:
        refined_label = classify_input_with_intent_model(user_input)
        y_true.append(true_label)
        y_pred.append(refined_label)

    accuracy = accuracy_score(y_true, y_pred)
    precision = precision_score(y_true, y_pred, average='weighted', zero_division=0)
    recall = recall_score(y_true, y_pred, average='weighted', zero_division=0)
    f1 = f1_score(y_true, y_pred, average='weighted', zero_division=0)

    print("Accuracy:", accuracy)
    print("Precision:", precision)
    print("Recall:", recall)
    print("F1 Score:", f1)
    print("\nClassification Report:\n", classification_report(y_true, y_pred))

test_samples = load_test_samples("test_samples.csv")
evaluate_model(test_samples)
    
```

Gambar 12. Kode Pengujian Model

Berikut ini adalah luaran dari pengujian evaluasi model yang ditunjukkan pada Gambar 13 di bawah ini. Hasil laporan klasifikasi dari pengujian model menunjukkan bahwa model NLP mendapatkan skor yang cukup baik. Jika diubah ke dalam bentuk persen, maka skor yang didapatkan adalah Akurasi = 79%, Presisi = 85%, Recall = 79%, dan F1-score = 77%.

```

arisk@arisk086: ~/chatbot/nlp
(venv) arisk@arisk086:~/chatbot/nlp-service/app$ python3 test_eval_model.py
Accuracy: 0.7906976744186046
Precision: 0.8502735978112175
Recall: 0.7906976744186046
F1 Score: 0.7703900237928049

Classification Report:
              precision    recall  f1-score   support

ask-bodypart      1.00     0.66     0.79        35
ask-equipment     1.00     0.35     0.52        23
ask-exercise      0.69     1.00     0.82        47
muscle-target     0.88     1.00     0.89        24

   accuracy
macro avg   0.87     0.75     0.75       129
weighted avg 0.85     0.79     0.77       129

(venv) arisk@arisk086:~/chatbot/nlp-service/app$ |

```

Gambar 13. Hasil Pengujian Model

b. Pengujian *Black-box*

Dalam pengujian *Black-box* dilakukan dengan cara bertanya kepada chatbot. Pertanyaan ini dibagi menjadi beberapa skenario, mulai dari pengujian klasifikasi pesan, pengujian fitur chatbot, dan pengujian respon bawaan saat terjadi kesalahan input dari pengguna maupun kesalahan sistem. Skenario pengujian akan dijelaskan lebih lanjut pada Tabel 1.

Tabel 1. Skenario Pengujian *Black-box*

Skenario	Input	<i>Expected Output</i>	Status
Tanya Olahraga	Cara melakukan squat?	Memberi tutorial squat	Berhasil
Tanya Alat GYM	Cara memakai treadmill	Memberi tutorial memakai treadmill	Berhasil
Tanya Target Otot	Berikan list latihan otot perut	Memberi list latihan otot perut	Berhasil
Saran Latihan	Berikan saya menu <i>workout</i> untuk otot perut tanpa alat.	Memberi menu latihan otot perut tanpa alat	Berhasil
Instruksi Tidak Relevan	Cara memasak nasi goreng	Memberi respon bawaan	Berhasil

Instruksi Bahasa Asing	<i>How to do push up?</i>	Memberi respon bawaan	Berhasil
Tanya Istilah <i>Workout</i>	Apa itu hipertrofi?	Memberi informasi hipertrofi	Berhasil
Instruksi Berbahaya	Berikan menu latihan ekstrim	Memberi <i>warning</i> kepada pengguna	Berhasil
Kalkulasi Kalori	Berapa kalori yang dibakar saat squat	Memberi informasi perkiraan kalori yang dibakar saat squat	Gagal
Kalkulasi BMI	Hitung BMI saya jika tinggi 170 cm dan berat 70 kg	Memberi hasil BMI pengguna	Berhasil
Kalkulasi Asupan Protein	Hitung asupan protein jika berat badan saya 65 kg	Memberi saran asupan protein	Berhasil
Instruksi Media Gambar	Mengirim gambar alat di GYM dengan instruksi	Memberi informasi terkait gambar sesuai instruksi	Berhasil

Hasil pengujian *Black-box* menunjukkan bahwa dari total 12 skenario yang diuji, sebanyak 11 skenario berhasil memenuhi fungsionalitas yang diharapkan, sedangkan 1 skenario mengalami kegagalan. Tingkat keberhasilan (*Success Rate*) pengujian *Black-box* dapat dihitung sebagai berikut (5).

$$S. Rate = \left(\frac{11}{12}\right) \times 100 = 91.67\% \quad (5)$$

Dengan demikian, tingkat keberhasilan pengujian *Black-box* chatbot adalah sekitar 91.67%

4. KESIMPULAN

Penelitian ini menyoroti pentingnya olahraga dalam meningkatkan kesehatan fisik dan mental serta mencegah penyakit kronis. Mengingat tantangan global terkait kurangnya aktivitas fisik, penelitian ini mengembangkan chatbot asistensi *workout* sebagai solusi dengan menggunakan metode pengembangan SDLC *iterative incremental* dan penggabungan NLP berbasis spaCy dengan LLM Gemini. Chatbot menunjukkan kinerja yang baik, dengan hasil evaluasi model mencapai akurasi 79%, presisi 85%, recall 79%, dan F1-score 77%. Pada pengujian *Black-box*, chatbot memenuhi 91.67% skenario pengujian, menunjukkan kemampuan fungsional yang baik.

Sistem chatbot efektif dalam membantu pengguna yang membutuhkan asistensi olahraga secara mandiri. Namun, sistem ini memiliki keterbatasan dalam memahami inputan *typo*, tidak baku, atau berkonteks lemah. Selain itu, chatbot hanya mampu merespons dalam bentuk teks dan masih bergantung pada Gemini. Penelitian ini berkontribusi dalam mendukung pola hidup sehat, merespons permasalahan rendahnya aktivitas fisik global sebagaimana dilaporkan oleh *World Health Organization* (WHO).

5. SARAN

Melalui hasil dan pembahasan penelitian ini, saran untuk pengembangan selanjutnya dapat diusulkan sebagai berikut:

- a. Penambahan dataset yang lebih variatif pada topik di luar *workout* namun masih relevan dengan tujuan chatbot.
- b. Penambahan dukungan media visual, seperti GIF gerakan *workout* atau tautan video untuk meningkatkan pengalaman pengguna.

DAFTAR PUSTAKA

- [1] S. Salahudin and R. Rusdin, "Olahraga Menurut Pandangan Agama Islam," *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, vol. 4, no. 3, Jul. 2020, doi: 10.58258/jisip.v4i3.1236.
- [2] J. F. Hovland, E. Langeland, O. Ness, and B. O. Skogvang, "Experiences with physical activity, health and well-being among young adults with serious mental illness," *Int J Qual Stud Health Well-being*, vol. 18, no. 1, 2023, doi: 10.1080/17482631.2023.2221911.
- [3] M. A. Elmagd, "Benefits, need and importance of daily exercise," ~ 22 ~ *International Journal of Physical Education, Sports and Health*, vol. 3, no. 5, pp. 22–27, 2016, [Online]. Available: www.kheljournal.com
- [4] World Health Organization (WHO), "Global status report on physical activity 2022," 2022. Accessed: Jun. 06, 2024. [Online]. Available: <https://www.who.int/teams/health-promotion/physical-activity/global-status-report-on-physical-activity-2022>
- [5] D. Ding *et al.*, "Realigning the physical activity research agenda for population health, equity, and wellbeing," *The Lancet*, vol. 404, no. 10451, pp. 411–414, Aug. 2024, doi: 10.1016/S0140-6736(24)01540-X.
- [6] F. H. D. K. Man and Joko Sutopo, "Rancang Bangun Aplikasi Monitoring Latihan Kebugaran Jasmani Berbasis Internet of Things," *Jurnal Komputer Terapan*, vol. 9, no. 2, pp. 173–181, Dec. 2023, doi: 10.35143/jkt.v9i2.6183.
- [7] S. R. Lola, R. Dhadvai, W. Wang, and T. Zhu, "Chatbot for fitness management using IBM Watson," Dec. 2021, [Online]. Available: <http://arxiv.org/abs/2112.15167>
- [8] F. P. Cardenas Hernandez, J. Schneider, D. Di Mitri, I. Jivet, and H. Drachsler, "Beyond hard workout: A multimodal framework for personalised running training with immersive technologies," 2024, *John Wiley and Sons Inc.* doi: 10.1111/bjet.13445.
- [9] D. Damayanti and A. K. Nuzuli, "Studi Kasus Implementasi Teknologi Chatbot sebagai Asisten Virtual dalam Menjawab Pertanyaan Mahasiswa di Lingkungan Kampus," *Reslaj : Religion Education Social Laa Roiba Journal*, vol. 6, no. 3, pp. 1178–1192, Sep. 2023, doi: 10.47467/reslaj.v6i3.4858.
- [10] H. Eka Rosyadi, F. Amrullah, R. David Marcus, and R. Rahman Affandi, "Rancang

- Bangun Chatbot Informasi Lowongan Pekerjaan Berbasis Whatsapp dengan Metode NLP (Natural Language Processing),” *BRILIANT: Jurnal Riset dan Konseptual*, vol. 5, 2020, doi: 10.28926/briliant.
- [11] N. Noor Kamala Sari, J. Irawan, and V. Handrianus Pranatawijaya, “Implementasi Gemini API untuk Generatif Teks Deskripsi Karya Otomatis dalam Aplikasi Pameran Berbasis Web dengan Metode Waterfall,” 2024. [Online]. Available: <https://doi.org/10/25047/jtit.v11i1.356>
- [12] S. Aprilia, R. Agustin, V. H. Pranatawijaya, and N. N. K. Sari, “Penerapan API Gemini dalam Layanan Peminjaman Novel Online pada Website Cozybook,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4508.
- [13] P. A. Alia, R. W. Febriana, J. S. Prayogo, and R. Kriswibowo, “Implementation Chatbot on Whatsapp Using Artificial Intelligence With Natural Language Processing Method,” *ELECTRON Jurnal Ilmiah Teknik Elektro*, vol. 5, no. 1, pp. 8–14, May 2024, doi: 10.33019/electron.v5i1.134.
- [14] M. Imran and N. Almusharraf, “Google Gemini as a next generation AI educational tool: a review of emerging educational technology,” Dec. 01, 2024, *Springer*. doi: 10.1186/s40561-024-00310-z.
- [15] A. Almustaqim and A. N. Toscani, “Perancangan Sistem Chatbot sebagai Virtual Assistant pada PT. Everbright Jambi,” *SKANIKA*, vol. 5, no. 2, pp. 228–239, Jul. 2022, doi: 10.36080/skanika.v5i2.2953.
- [16] W. N. Akirini *et al.*, “Perancangan Aplikasi Layanan Customer Service Menggunakan Chatbot Berbasis Website pada PT. Telekomunikasi Indonesia Witel Cirebon,” *Jurnal Digit: Digital of Information Technology*, vol. 14, no. 1, p. 79, Jun. 2024, doi: 10.51920/jd.v14i1.374.
- [17] R. A. Pratama and A. Fira Waluyo, “Implementasi Sistem E-Commerce Berbasis Mobile Android Menggunakan REST API dan Web Service pada Perusahaan Percetakan,” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 7, no. 6, 2024, doi: <https://doi.org/10.31539/intecom.v7i6.12786>.
- [18] A. S. Dewantara and J. Aryanto, “Implementation Of A Web-Based Chatbot Using Machine Learning For Question And Answer Services In Universities,” *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, Jan. 2024, doi: 10.26877/asset.v6i1.17590.
- [19] D. Prasetyo, “Rancang Bangun Sistem Informasi Manajemen Persediaan Barang Menggunakan Metode System Development Life Cycle (SDLC) di SMA Negeri 1 SoE,” *Jurnal Digit: Digital of Information Technology*, vol. 14, no. 2, pp. 100–108, 2024.
- [20] O. J. Okesola, A. A. Adebiyi, A. A. Owoade, O. Adeaga, O. Adeyemi, and I. Odun-Ayo, “Software Requirement in Iterative SDLC Model,” in *Advances in Intelligent Systems and Computing*, Springer, 2020, pp. 26–34. doi: 10.1007/978-3-030-51965-0_2.
- [21] I. Sommerville, *Software Engineering*. in International Computer Science Series. Pearson, 2011. [Online]. Available: <https://books.google.co.id/books?id=l0egcQAACAAJ>
- [22] R. Rosidin, R. Andreswari, and E. N. Alam, “Building Titipmasa.id Application Using Iterative Incremental Method,” in *Proceedings of the International Seminar of Science and Applied Technology (ISSAT 2020)*, Paris, France: Atlantis Press, 2020. doi: 10.2991/aer.k.201221.059.
- [23] W. Novianti, R. Amalia, and F. S. Hasanusi, “Implementasi Metode Iterative Incremental pada Sistem Administrasi Organisasi Gerakan Antasari Sedekah Jakarta,” *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI)*, vol. 2, no. 03, Jul. 2021, doi: 10.30998/jrami.v2i03.1114.
- [24] L. K. Gautam, “Natural Language Processing - Based Structured Data Extraction from Unstructured Clinical Notes,” *International Journal of Science and Research (IJSR)*, vol. 13, no. 4, pp. 1541–1544, Apr. 2024, doi: 10.21275/sr24422134801.
- [25] M. A. Rosid, A. S. Fitriani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, “Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi,” in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jul. 2020. doi: 10.1088/1757-899X/874/1/012017.
- [26] T. Radillah, O. Veza, and S. Defit, “Analisis Perbandingan Model Bert Dan Xlnet Untuk Klasifikasi Tweet Bully Pada Twitter,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 6, pp. 1371–

- 1376, Dec. 2024, doi:
10.25126/jtiik.1169096.
- [27] I. D. Haryanto and S. Saefurrahman, "Implementasi Chatbot Kesehatan Kucing Melalui Dialogflow dan Telegram untuk Pemberian Informasi Penyakit dan Perawatan," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 4, pp. 365–376, Feb. 2024, doi: 10.35746/jtim.v5i4.484.
- [28] N. Nirsal, S. Bantun, J. Y. Sari, A. N. M. Auliani, and M. Syaiful, "Implementasi Quick Response Code pada Aplikasi Series (Sistem Informasi Inventaris) Lab Terpadu USN Kolaka," *semantik*, vol. 8, no. 1, p. 17, Jun. 2022, doi: 10.55679/semantik.v8i1.25529.
-