

Penerapan *Natural Language Processing* pada *Chatbot* Edukasi HIV Berbasis *Support Vector Machine* dan TF-IDF untuk Remaja

^{1*}Muhamad Aldi Hawari, ²Patmi Kasih, ³Muh. Aris Saputra

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹aldihawarii@gmail.com, ²fatkasi@gmail.com, ³aris.saputra@unpkdr.ac.id

Penulis Korespondens : Muhamad Aldi Hawari

Abstrak— *Human Immunodeficiency Virus* merupakan virus yang menyerang sistem kekebalan tubuh dan menjadi permasalahan kesehatan yang signifikan di kalangan remaja. Kurangnya pemahaman tentang HIV menimbulkan stigma dan perilaku berisiko pada kelompok usia tersebut. Penelitian ini bertujuan mengembangkan sistem chatbot edukasi HIV berbasis *Natural Language Processing* (NLP) menggunakan metode *Support Vector Machine* (SVM) dan *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai perantara peer educator dan remaja. Dataset terdiri dari 300 pasangan pertanyaan dan jawaban yang dikategorikan ke dalam 10 intent. Sistem menggunakan pipeline preprocessing teks yang mencakup cleaning, case folding, tokenisasi, stopword removal, dan stemming. Cosine Similarity diterapkan sebagai typo handler, sementara forward chaining digunakan pada modul kuesioner penilaian risiko. Evaluasi menggunakan confusion matrix menghasilkan akurasi 84%, precision 0,83, recall 0,84, dan F1-score 0,82. Hasil pengujian menunjukkan sistem berjalan sesuai spesifikasi dan efektif sebagai media edukasi HIV bagi remaja..

Kata Kunci— Chatbot, Edukasi HIV, *Natural Language Processing*, *Support Vector Machine*, TF-IDF

Abstract— *Human Immunodeficiency Virus* is a virus that attacks the human immune system, representing a significant health challenge among adolescents. Limited understanding of HIV leads to stigma and risky behaviors in this age group. This research aims to develop an HIV education chatbot system based on *Natural Language Processing* (NLP) using *Support Vector Machine* (SVM) and *Term Frequency-Inverse Document Frequency* (TF-IDF) as an intermediary between peer educators and adolescents. The dataset consists of 300 question-answer pairs categorized into 10 intents. The system employs a text preprocessing pipeline including cleaning, case folding, tokenization, stopword removal, and stemming. Cosine Similarity was applied as a typo handler, while forward chaining was used in the risk assessment questionnaire module. Model evaluation using confusion matrix yielded an accuracy of 84%, precision of 0.83, recall of 0.84, and F1-score of 0.82. Testing results demonstrate that the system operates according to specifications and serves effectively as an HIV education medium for adolescents.

Keywords— Chatbot, HIV Education, *Natural Language Processing*, *Support Vector Machine*, TF-IDF

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

HIV (*Human Immunodeficiency Virus*) merupakan retrovirus yang menyerang sel CD4 sehingga melemahkan sistem imun penderitanya [1]. Pada tahun 2024, WHO mencatat 87% dari seluruh ODHIV mengetahui status mereka, namun pemahaman di kalangan remaja masih menjadi

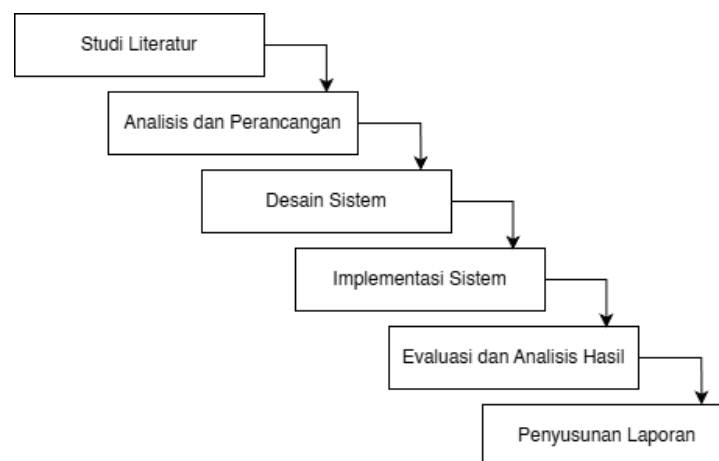
tantangan serius. Kelompok usia 20–24 tahun menyumbang 16,9% kasus HIV di Indonesia, sedangkan kelompok 15–19 tahun sebesar 3,1% [2]. Kurangnya pemahaman tentang HIV menjadi penyebab utama terbentuknya stigma buruk terhadap Orang Dengan HIV (ODHIV) [3].

Perkembangan teknologi informasi telah membuka peluang pemanfaatan chatbot sebagai media edukasi kesehatan yang interaktif dan mudah diakses. Berbagai penelitian telah mengembangkan chatbot kesehatan pada beragam domain, termasuk edukasi HIV. Salah satu penelitian mengembangkan chatbot berbasis *deep learning* menggunakan bahasa Amharik sebagai media penyampaian informasi mengenai HIV [4]. Penelitian lainnya memperkenalkan chatbot MARVIN berbasis kecerdasan buatan untuk mendukung manajemen mandiri bagi orang dengan HIV [5]. Selain itu, pemanfaatan *Large Language Model* (LLM) juga telah dikaji sebagai chatbot percakapan yang mampu mendukung upaya pencegahan HIV melalui penyampaian informasi secara lebih natural dan kontekstual [6].

Penelitian ini mengajukan chatbot edukasi HIV berbasis NLP menggunakan *Support Vector Machine* (SVM) [7] dan TF-IDF sebagai solusi yang relevan bagi remaja usia 16–25 tahun berbahasa Indonesia. Chatbot dirancang mampu mengklasifikasikan pertanyaan pengguna ke dalam intent yang sesuai, menangani kesalahan penulisan (*typo*) menggunakan *Cosine Similarity*, serta menyediakan kuesioner penilaian risiko HIV awal dengan pendekatan *Forward Chaining*. Tujuannya adalah menciptakan media edukasi yang aman, interaktif, dan mampu berfungsi sebagai perantara digital antara *Peer Educator* dan remaja/ODHIV [8].

II. METODE

Penelitian ini menggunakan pendekatan *Research and Development* (R&D) dengan model pengembangan *Waterfall* yang terdiri dari : Studi Literatur, Analisis Kebutuhan, Desain Sistem, Implementasi Sistem, Evaluasi dan Analisis Hasil, serta Penyusunan Laporan. Pendekatan *Structured Systems Analysis and Design* (SSAD) digunakan dalam perancangan sistem melalui pemodelan DFD, *Flowchart*, dan ERD.



Gambar 1. Diagram *Waterfall*

A. Dataset

Dataset yang digunakan merupakan dataset yang disusun dan dikembangkan oleh peneliti berdasarkan informasi dari berbagai sumber referensi yang kredibel, seperti WHO, UNAIDS, dan Kementerian Kesehatan Republik Indonesia. Dataset terdiri lebih dari 250 pasangan pertanyaan dan jawaban yang dikategorikan ke dalam 10 intent, yaitu: ARV_HIV, Bahaya_HIV, Definisi_HIV, Gejala_HIV, PEP_HIV, Pencegahan_HIV, Pengobatan_HIV, Penularan_HIV, PrEP_HIV, dan Testing_HIV. Dataset disimpan dalam format CSV dengan kolom: No, Intent, Pertanyaan, dan Jawaban.

B. Preprocessing Teks

Pipeline preprocessing teks mencakup lima tahapan: (1) *Cleaning*, menghapus karakter khusus, angka, dan tanda baca tidak relevan; (2) *Case Folding*, mengubah teks menjadi huruf kecil; (3) Tokenisasi, memecah kalimat menjadi token kata; (4) *Stopword Removal*, menghapus kata fungsional tidak bermakna menggunakan daftar *stopword* bahasa Indonesia; dan (5) *Stemming*, mereduksi kata turunan ke bentuk dasarnya menggunakan *library* Sastrawi [9]. Hasil preprocessing menjadi masukan bagi proses ekstraksi fitur TF-IDF.

C. Ekstraksi Fitur dan Klasifikasi

Term Frequency-Inverse Document Frequency (TF-IDF) digunakan untuk mengubah teks menjadi representasi vektor numerik [10]. Skor TF-IDF dihitung menggunakan persamaan :

$$TF - IDF(i, j) = TF(i, j) \times IDF(i) \quad (1)$$

dengan,

$$TF(i, j) = \frac{n(i, j)}{\sum_{k=1}^q n(k, j)} \quad (2)$$

$$IDF(i) = \log \frac{N}{1 + df(i)} \quad (3)$$

Vektor TF-IDF selanjutnya diklasifikasikan menggunakan *Support Vector Machine* (SVM) dengan kernel linear [7]. SVM bekerja dengan mencari *hyperplane* optimal yang memaksimalkan margin antar kelas. Dataset dibagi menggunakan *train-test split* 80:20, menghasilkan 240 data latih dan 60 data uji. Berikut adalah simulasi singkat untuk pembobotan TF-IDF dengan menggunakan 3 sample pertanyaan yaitu D1 = "Apa Gejala HIV?", D2 = "Bagaimana Gejala HIV?", D3 = "Apa cara pencegahan HIV?" :

Tabel 1. Hasil Dokumen setelah Preprocessing

No.	Dokumen	Hasil <i>Preprocessing</i>
1	Apa Gejala HIV?	gejala hiv
2	Bagaimana Gejala HIV?	penularan hiv
3	Apa cara pencegahan HIV?	pencegahan hiv

Berdasarkan Tabel 1, setiap dokumen telah direpresentasikan ke dalam kumpulan kata bermakna setelah melalui tahapan preprocessing. Selanjutnya, frekuensi kemunculan setiap kata dalam masing-masing dokumen dihitung untuk membentuk nilai *Term Frequency* (TF) dan *Document Frequency* (df). Nilai TF menunjukkan seberapa sering sebuah kata muncul dalam satu dokumen, sedangkan df mencatat pada berapa dokumen kata tersebut hadir. Hasil perhitungan TF dan df disajikan pada Tabel 2.

Tabel 2. *Term Frequency* dan *Document Frequency*

No.	Kata	TF(D1)	TF(D2)	TF(D3)	df
1.	Gejala	1	0	0	1
2.	penularan	0	1	0	1
3.	pencegahan	0	0	1	1
4.	Hiv	1	1	1	3

Nilai TF dan df pada Tabel 2 kemudian digunakan untuk menghitung bobot TF-IDF menggunakan persamaan yang telah ditetapkan. Kata yang muncul di banyak dokumen, seperti "hiv", memperoleh nilai IDF mendekati nol sehingga bobotnya menjadi 0. Sebaliknya, kata yang bersifat unik pada satu dokumen, seperti "gejala", "penularan", dan "pencegahan", memperoleh bobot TF-IDF yang lebih tinggi yakni 0,4771. Hal ini menunjukkan kemampuan TF-IDF dalam menonjolkan kata-kata yang lebih informatif dan diskriminatif. Hasil akhir pembobotan TF-IDF ditampilkan pada Tabel 3.

Tabel 3. Hasil Pembobotan TF-IDF

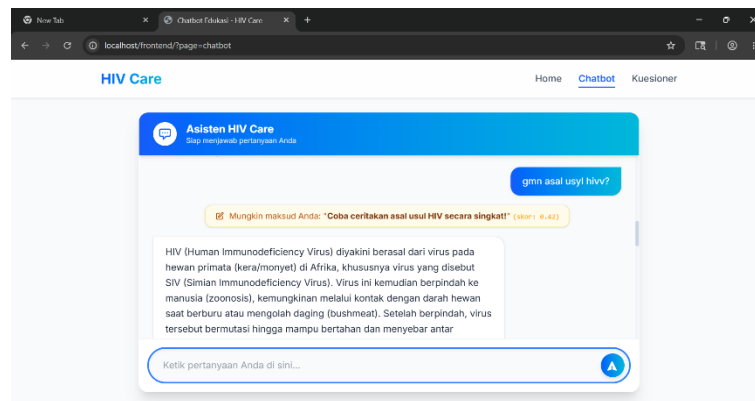
No.	Kata	(D1)	(D2)	(D3)
1.	gejala	0,4771	0	0
2.	penularan	0	0,4771	0
3.	pencegahan	0	0	0,4771
4.	hiv	0	0	0

D. Penanganan *Typo*

Cosine Similarity diintegrasikan sebagai *typo handler*. Saat input pengguna tidak dapat dikenali oleh model SVM, sistem menghitung kesamaan vektor TF-IDF antara input dan seluruh pertanyaan dalam dataset menggunakan rumus :

$$\cos \alpha = \frac{(A \cdot B)}{(|A| \times |B|)} \quad (4)$$

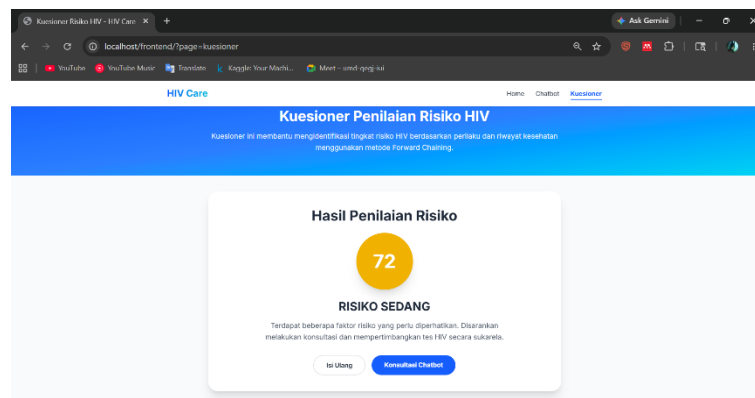
Pertanyaan dengan nilai kemiripan tertinggi digunakan sebagai referensi untuk pencocokan intent.



Gambar 2. Simulasi Penanganan *Typo* pada *Chatbot*

E. Kuesioner dan *Forward Chaining*

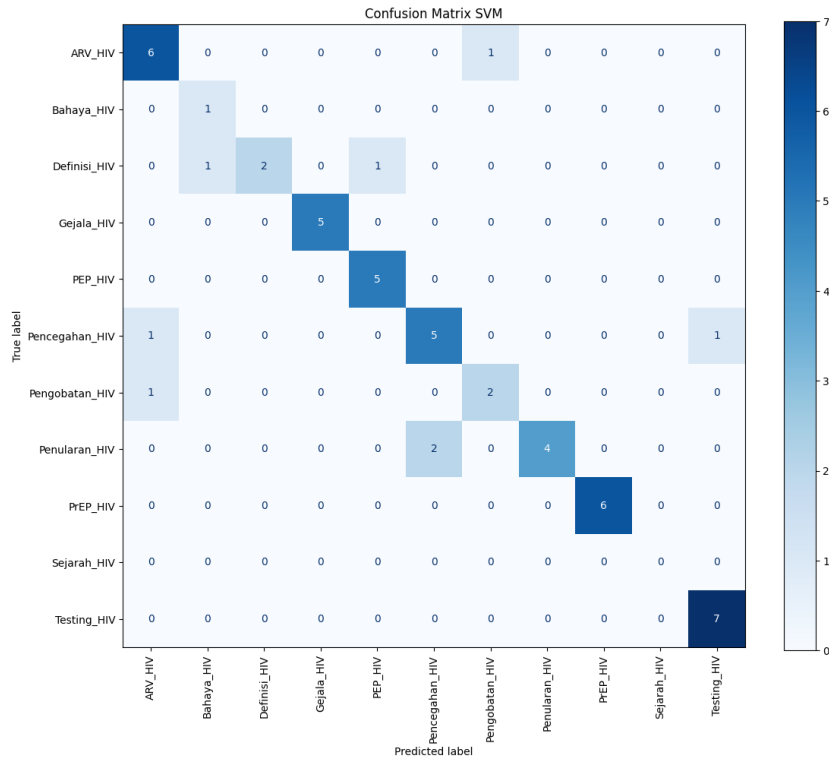
Kuesioner penilaian risiko HIV terdiri dari 20 pertanyaan dengan skala Likert berbobot 1 (Sangat Tidak Setuju) hingga 5 (Sangat Setuju). Pertanyaan dengan faktor risiko tinggi diberi bobot ganda. Mekanisme inferensi *forward chaining* diterapkan dengan aturan berbasis skor [11]. *Rule-based* dari pembobotan skro kuesioner dengan *Forward Chaining* yaitu Aturan 1 : Jika $20 \leq \text{Skor} \leq 50 \rightarrow \text{Risiko Rendah}$, Aturan 2 : Jika $51 \leq \text{Skor} \leq 80 \rightarrow \text{Risiko Sedang}$, Aturan 3 : Jika $81 \leq \text{Skor} \leq 110 \rightarrow \text{Risiko Tinggi}$, dan Aturan 4 : Jika $111 \leq \text{Skor} \leq 140 \rightarrow \text{Risiko Sangat Tinggi}$



Gambar 3. Simulasi Pembobotan Kuesioner

F. Evaluasi

Kinerja model SVM dievaluasi menggunakan *Confusion Matrix* yang menghasilkan nilai *accuracy*, *precision*, *recall*, dan *F1-score* [12]. Pengujian fungsional dilakukan terhadap seluruh fitur pada halaman utama, chatbot, dan kuesioner. Sistem diimplementasikan menggunakan *framework Flask (Python)* dengan *library* NLTK, Sastrawi, scikit-learn, Pandas, dan Numpy.



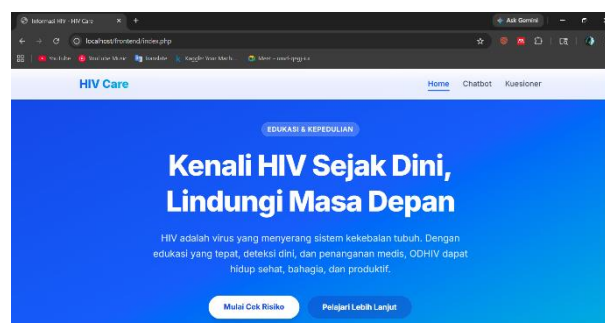
Gambar 4. Confusion Matrix

III. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

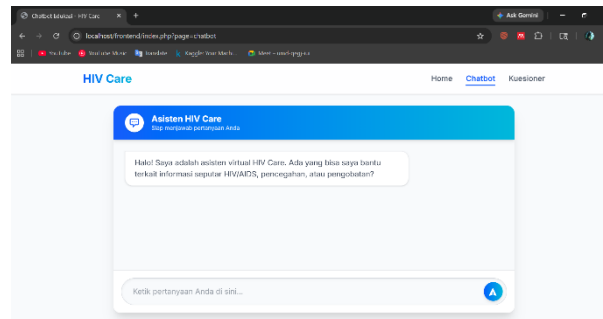
Sistem diimplementasikan sebagai aplikasi web bernama "*HIV Care*" yang dapat diakses melalui *browser*. Aplikasi terdiri dari tiga halaman utama: (1) Halaman Utama yang menampilkan banner edukasi dan navigasi; (2) Halaman Chatbot yang memuat panel percakapan interaktif dengan *header* "Asisten HIV Care"; dan (3) Halaman Kuesioner yang menyediakan form penilaian risiko HIV.

Ketiga halaman saling terintegrasi. Pengguna dapat memulai dari halaman utama, mengisi kuesioner untuk mendapat penilaian risiko, lalu melanjutkan konsultasi melalui chatbot. Setelah pengisian kuesioner selesai, sistem menampilkan skor, kategori risiko, dan rekomendasi tindak lanjut bagi pengguna.



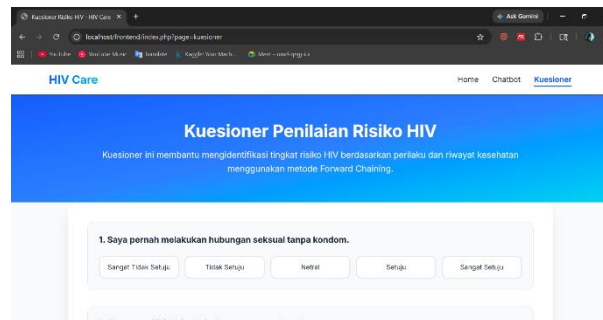
Gambar 5. Halaman *Home*

Gambar 5 menunjukkan halaman utama aplikasi *HIV Care* yang menyajikan banner edukasi dengan heading "Kenali HIV Sejak Dini, Lindungi Masa Depan" sebagai pintu masuk awal bagi pengguna. Halaman ini dilengkapi tombol navigasi "Mulai Cek Risiko" yang mengarahkan pengguna ke kuesioner, dan "Pelajari Lebih Lanjut" untuk mengakses informasi edukasi HIV. Selanjutnya, pengguna dapat mengakses fitur tanya jawab interaktif melalui halaman chatbot yang ditampilkan pada Gambar 6.



Gambar 6. Halaman Chatbot

Gambar 6 memperlihatkan antarmuka halaman chatbot dengan panel percakapan yang menampilkan header "Asisten HIV Care" beserta pesan pembuka dari sistem. Pengguna cukup mengetikkan pertanyaan pada kolom input dan menekan tombol kirim untuk memperoleh jawaban edukatif secara otomatis. Selain fitur chatbot, sistem juga menyediakan modul kuesioner penilaian risiko HIV secara mandiri sebagaimana ditampilkan pada Gambar 7.



Gambar 7. Halaman Kuesioner

B. Hasil *Preprocessing*

Tabel 1 menunjukkan hasil preprocessing untuk empat sampel teks pertanyaan :

Tabel 4. Hasil Pengujian *Preprocessing*

No	Input Teks	Tokenisasi dan <i>Stopword</i>	Output (<i>Stemming</i>)
1	Apa itu HIV/AIDS?	"hivaidis"	"hivaidis"
2	Bagaimana cara mencegah penularan HIV?	"cara", "cegah", "penularan", "hiv"	"cara", "cegah", "tular", "hiv"
3	Apa gejala2 yg ada klo kena HIV?	"gejala", "kena", "hiv"	"gejala", "kena", "hiv"
4	Bagaimana pengobatan ARV untuk penderita HIV?	"pengobatan", "arv", "penderita", "hiv"	"obat", "arv", "derita", "hiv"

Proses cleaning berhasil menghapus karakter khusus seperti "/" dan angka. *Case folding* menyeragamkan seluruh teks menjadi huruf kecil. *Stopword removal* menghapus kata fungsional seperti "apa", "itu", "bagaimana", dan "untuk". *Stemming* menggunakan *library* Sastrawi mereduksi kata turunan ke bentuk dasar, misalnya "penularan" menjadi "tular" dan "pengobatan" menjadi "obat".

C. Kinerja Model SVM

Model SVM diuji menggunakan 60 data uji (6 data per intent). *Confusion matrix* menunjukkan sebagian besar data berada pada diagonal utama, mengindikasikan klasifikasi yang tepat. Kesalahan klasifikasi umumnya terjadi antara intent yang memiliki keterkaitan semantik erat, seperti ARV_HIV dengan Pengobatan_HIV, serta Penularan_HIV dengan Pencegahan_HIV.

Tabel 5. *Classification Report* Model SVM

	Precision	Recall	F1-Score	Support
ARV_HIV	0.75	0.86	0.80	7
Bahaya_HIV	0.50	1.00	0.67	1
Definisi_HIV	1.00	0.50	0.67	4
Gejala_HIV	1.00	1.00	1.00	5
PEP_HIV	0.83	1.00	0.91	5
Pencegahan_HIV	0.71	0.71	0.71	7
Pengobatan_HIV	0.67	0.67	0.67	3
Penularan_HIV	1.00	0.67	0.80	6
PrEP_HIV	1.00	1.00	1.00	6
Testing_HIV	0.88	1.00	0.93	7
accuracy			0.84	51
macro avg	0.83	0.84	0.82	51
weighted avg	0.86	0.84	0.84	51

Berdasarkan Tabel 2, model SVM mencapai akurasi 84%, *macro precision* 0,83, *macro recall* 0,84, dan *macro F1-score* 0,82. Intent Gejala_HIV dan PrEP_HIV memperoleh nilai sempurna (F1-score 1,00), sementara Testing_HIV (0,93) dan PEP_HIV (0,91) menunjukkan performa sangat baik. Intent dengan performa lebih rendah seperti Bahaya_HIV dan Definisi_HIV (*F1-score* 0,67) mengalami kesalahan klasifikasi akibat kemiripan kosakata antar intent.

D. Pengujian Chatbot dengan Typo

Pengujian chatbot dengan pertanyaan yang mengandung kesalahan penulisan menunjukkan mekanisme *Cosine Similarity* sebagai *typo handler* bekerja efektif. Pertanyaan "Ap gejala awal HIV?" (dengan *typo* pada "ap" dan "gejala") berhasil diklasifikasikan ke intent Gejala_HIV dan menghasilkan jawaban yang relevan. Serupa, pertanyaan "Apa obt buat HIV?" (*typo*: "obt") berhasil diklasifikasikan ke Pengobatan_HIV dengan jawaban tepat mengenai terapi ARV.

E. Hasil *Forward Chaining* Kuesioner

Tabel 3 menyajikan hasil dari pengujian delapan skenario yang mewakili seluruh kategori risiko dan kondisi batas (*Boundary Condition*) :

Tabel 6. Pengujian Logika *Forward Chaining*

Skor	<i>Expected</i>	Hasil Sistem	Status
20	Risiko Rendah	Risiko Rendah	Valid
45	Risiko Rendah	Risiko Rendah	Valid
51	Risiko Sedang	Risiko Sedang	Valid
70	Risiko Sedang	Risiko Sedang	Valid
81	Risiko Tinggi	Risiko Tinggi	Valid
111	Risiko Sangat Tinggi	Risiko Sangat Tinggi	Valid
120	Risiko Sangat Tinggi	Risiko Sangat Tinggi	Valid

Seluruh skenario menghasilkan kategori risiko yang sesuai, termasuk kondisi batas (skor 51, 81, dan 111). Sistem mampu mengklasifikasikan skor minimum (20) sebagai Risiko Rendah dan skor maksimum (120) sebagai Risiko Sangat Tinggi, membuktikan mekanisme inferensi *forward chaining* bekerja konsisten.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem chatbot edukasi HIV berbasis NLP menggunakan SVM dan TF-IDF sebagai perantara *Peer Educator* dan remaja. Model SVM dengan fitur TF-IDF mencapai akurasi 84% pada 10 kelas intent, dengan *macro F1-score* 0,82. Mekanisme *Cosine Similarity* terbukti efektif menangani kesalahan penulisan pengguna. Modul kuesioner berbasis *Forward Chaining* berhasil mengklasifikasikan seluruh skenario pengujian secara konsisten sesuai kategori risiko yang ditetapkan. Pengujian fungsional menunjukkan seluruh fitur sistem berjalan valid, sementara pengujian non-fungsional membuktikan sistem responsif dan kompatibel pada berbagai perangkat.

Untuk penelitian selanjutnya, disarankan penambahan dataset yang lebih besar dan variatif, penerapan augmentasi data untuk intent dengan data terbatas, serta eksplorasi model *deep learning* berbasis *transformer* untuk meningkatkan kemampuan pemahaman konteks percakapan yang lebih kompleks. Pengembangan sistem juga dapat mencakup integrasi langsung dengan platform layanan kesehatan untuk memfasilitasi tindak lanjut *screening* kepada *Peer Educator* atau tenaga kesehatan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Nusantara PGRI Kediri, Fakultas Teknik dan Ilmu Komputer, serta Program Studi Teknik Informatika atas dukungan dan fasilitas yang diberikan selama pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada keluarga yang senantiasa memberikan doa, motivasi, dan dukungan sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] J. Ilmiah Manuntung, S. Farmasi Dan Kesehatan, S. Rachmawati, R. Fauzia, and E. Rachmawati, "PENGETAHUAN MAHASISWA UNIVERSITAS JEMBER TENTANG HIV/AIDS," *Jurnal Ilmiah Manuntung: Sains Farmasi Dan Kesehatan*, vol. 8, no. 1, pp. 106–112, May 2022, doi: 10.51352/JIM.V8I1.502.
- [2] World Health Organization (WHO), "Consolidated Guidelines on HIV Prevention, Testing, Treatment, Service Delivery And Monitoring," *Optics InfoBase Conference Papers*, no. July, p. 249, Jul. 2021, Accessed: Nov. 26, 2025. [Online]. Available: ISBN: 978-92-4-003159-3
- [3] J. Keperawatan Silampari Volume, F. Quaesita Qory Lorenz, and H. Permatasari, "Implementasi Peer Education dalam Meningkatkan Pengetahuan Remaja Mengenai Kesehatan Reproduksi," *Jurnal Keperawatan Silampari*, vol. 6, no. 2, pp. 1817–1826, May 2023, doi: 10.31539/JKS.V6I2.5867.
- [4] A. Sintayehu and E. D. Emiru, "Developing amharic text-based chatbot model for HIV/AIDS awareness and care using deep learning approaches," *BMC Artificial Intelligence 2025 1:1*, vol. 1, no. 1, pp. 2-, Jun. 2025, doi: 10.1186/S44398-025-00002-9.
- [5] Y. Ma *et al.*, "The first AI-based Chatbot to promote HIV self-management: A mixed methods usability study," *HIV Med.*, vol. 26, no. 2, pp. 184–206, Feb. 2025, doi: 10.1111/HIV.13720;PAGE:STRING:ARTICLE/CHAPTER.
- [6] A. van Heerden, S. Bosman, D. Swendeman, and W. S. Comulada, "Chatbots for HIV Prevention and Care: a Narrative Review," *Curr. HIV/AIDS Rep.*, vol. 20, no. 6, pp. 481–486, Dec. 2023, doi: 10.1007/S11904-023-00681-X.
- [7] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Information 2024, Vol. 15, Page 235*, vol. 15, no. 4, p. 235, Apr. 2024, doi: 10.3390/INFO15040235.
- [8] M. Laymouna, Y. Ma, D. Lessard, T. Schuster, K. Engler, and B. Lebouché, "Roles, Users, Benefits, and Limitations of Chatbots in Health Care: Rapid Review," *J Med Internet Res 2024;26:e56930* <https://www.jmir.org/2024/1/e56930>, vol. 26, no. 1, p. e56930, Jul. 2024, doi: 10.2196/56930.
- [9] M. Siino, I. Tinnirello, and M. La Cascia, "Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers," *Inf. Syst.*, vol. 121, pp. 306–4379, 2024, doi: 10.1016/j.is.2023.102342.
- [10] L. Zhang, "Features extraction based on Naive Bayes algorithm and TF-IDF for news classification," *PLoS One*, vol. 20, no. 7, p. e0327347, Jul. 2025, doi: 10.1371/JOURNAL.PONE.0327347.
- [11] R. Dapari *et al.*, "Developing and validating a knowledge, attitude, and practice questionnaire regarding dengue among secondary schoolchildren in Malaysia," *Discover Social Science and Health 2024 4:1*, vol. 4, no. 1, pp. 68-, Nov. 2024, doi: 10.1007/S44155-024-00130-Z.
- [12] S. Sathyanarayanan and B. R. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, vol. 27, no. 4S, pp. 4023–4031, Nov. 2024, doi: 10.53555/AJBR.V27I4S.4345.