



PENGOLAHAN BAHASA ALAMI UNTUK ANAMNESIS PENYAKIT PADA ANAK

Muchtadho Dhila Arafat

21917032

Tesis diajukan sebagai syarat untuk meraih gelar Magister Komputer

Konsentrasi Informatika Medis

Program Studi Informatika Program Magister

Fakultas Teknologi Industri

Universitas Islam Indonesia

2025

Lembar Pengesahan Pembimbing

Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak

Muchtadho Dhila Arafat

21917032



Yogyakarta, 10 Juni 2025

المعتمد الاستاذ المساعد الدكتور
Pembimbing 1 Pembimbing 2

Dr. Sri Kusumadewi, S.Si., M.T

Chanifah Indah Ratnasari, S.Kom., M.Kom.

Lembar Pengesahan Penguji

Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak

Muchtadho Dhila Arafat

21917032

Irving Vitra Paputungan, S.T., M.Sc., Ph.D.

Yogyakarta, 10 Juni 2025

Tim Penguji,

Dr. Sri Kusumadewi, S.Si., M.T

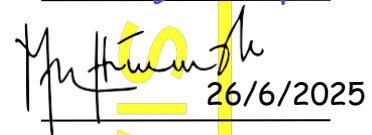
Ketua

Irving Vitra Paputungan, S.T., M.Sc., Ph.D.

Anggota I

Izzati Muhimmah, S.T., M.Sc., Ph.D.

Anggota II



26/6/2025

الجامعة الإسلامية
Mengetahui,

Ketua Program Studi Informatika Program Magister

Fakultas Teknologi Industri

Universitas Islam Indonesia



Irving Vitra Paputungan, S.T., M.Sc., Ph.D.

Abstrak

Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak

Natural Language Processing (NLP) telah banyak diterapkan dalam bidang medis, namun belum dimanfaatkan secara luas untuk proses anamnesis penyakit pada anak. Tantangan utama dalam anamnesis adalah komunikasi verbal antara pasien dan dokter yang beragam, sehingga dapat menghambat kecepatan dan ketepatan diagnosis. Kesalahan dalam proses diagnosis dapat berakibat pada penanganan yang tidak optimal dan membahayakan pasien, terutama dalam kasus penyakit dengan gejala demam seperti Demam Berdarah *Dengue* (DBD), Infeksi Saluran Pernapasan Akut (ISPA), Demam Tifoid, dan Infeksi Saluran Kemih (*Urinary Tract Infection*). Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan NLP dalam proses anamnesis penyakit pada anak dengan mengonversi keluhan suara pasien ke dalam bentuk teks dan mengidentifikasi gejala menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *Cosine Similarity*. Penelitian ini menggunakan pendekatan *Cross Industry Standard for Data Mining* (CRISP-DM) dalam memproses data anamnesis berbentuk suara (.wav) yang diperoleh dari Klinik Margorejo. Tahapan preprocessing yang diterapkan mencakup case folding, tokenizing, filtering, dan stemming, sebelum diterapkan metode TF-IDF dan *Cosine Similarity* untuk analisis kemiripan gejala. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan memungkinkan pengguna untuk melakukan konsultasi terkait gejala penyakit berbasis *website*. Sistem ini berhasil mengenali gejala dengan nilai kemiripan di atas ambang batas 0,15, serta mengelompokkan gejala yang dikenali berdasarkan “*The Sacred Seven of History Taking*” untuk memudahkan dokter dalam proses diagnosis. Penerapan NLP dalam anamnesis penyakit pada anak dapat meningkatkan keakuratan interpretasi keluhan pasien serta mendukung proses pengambilan keputusan medis. Hasil penelitian ini menunjukkan bahwa metode yang digunakan efektif dalam mengenali pola gejala dari input suara pasien. Untuk penelitian selanjutnya, pengembangan model NLP dengan pendekatan *deep learning* direkomendasikan guna meningkatkan akurasi sistem dalam mengenali gejala penyakit pada anak.

Kata Kunci

Anamnesis, *Cosine Similarity*, *Natural Language Processing*, Penyakit Anak, TF-IDF

Abstract

Natural Language Processing for Pediatric Illness Anamnesis

Natural Language Processing (NLP) has been widely applied in the medical field, yet it has not been extensively utilized for the anamnesis process of pediatric diseases. One of the main challenges in anamnesis is the variability in verbal communication between patients and doctors, which can hinder the speed and accuracy of diagnosis. Errors in the diagnostic process may lead to suboptimal treatment and pose risks to patients, especially in cases of diseases with fever symptoms such as Dengue Hemorrhagic Fever (DHF), Acute Respiratory Infections (ARI), Typhoid Fever, and Urinary Tract Infections (UTI). Therefore, this study aims to implement NLP in the anamnesis process of pediatric diseases by converting patients' voice complaints into text and identifying symptoms using the Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity methods. This research follows the Cross Industry Standard for Data Mining (CRISP-DM) framework to process anamnesis data in voice (.wav) format obtained from Klinik Margorejo. The preprocessing steps include case folding, tokenizing, filtering, and stemming, before applying TF-IDF and Cosine Similarity to analyze symptom similarity. The results indicate that the developed system enables users to consult about their symptoms through a web-based platform. The system successfully identifies symptoms with a similarity score above the threshold of 0.15 and categorizes them based on "The Sacred Seven of History Taking" to assist doctors in making diagnoses. The implementation of NLP in pediatric anamnesis enhances the accuracy of patient complaint interpretation and supports clinical decision-making. The findings demonstrate that the applied methods effectively recognize symptom patterns from patient voice inputs. Future research is recommended to develop an NLP model using deep learning approaches to improve the system's accuracy in identifying pediatric disease symptoms.

Keywords

Anamnesis, Cosine Similarity, Natural Language Processing, Pediatric Diseases, TF-IDF

Pernyataan Keaslian Tulisan

Dengan ini saya menyatakan bahwa tesis ini merupakan tulisan asli dari penulis, dan tidak berisi material yang telah diterbitkan sebelumnya atau tulisan dari penulis lain terkecuali referensi atas material tersebut telah disebutkan dalam tesis. Apabila ada kontribusi dari penulis lain dalam tesis ini, maka penulis lain tersebut secara eksplisit telah disebutkan dalam tesis ini.

Dengan ini saya juga menyatakan bahwa segala kontribusi dari pihak lain terhadap tesis ini, termasuk bantuan analisis statistik, desain survei, analisis data, prosedur teknis yang bersifat signifikan, dan segala bentuk aktivitas penelitian yang dipergunakan atau dilaporkan dalam tesis ini telah secara eksplisit disebutkan dalam tesis ini.

Segala bentuk hak cipta yang terdapat dalam material dokumen tesis ini berada dalam kepemilikan pemilik hak cipta masing-masing. Apabila dibutuhkan, penulis juga telah mendapatkan izin dari pemilik hak cipta untuk menggunakan ulang materialnya dalam tesis ini.

Yogyakarta, 07 Maret 2025

Muchtadho Dhila Arafat, S.Kom

Daftar Publikasi

Arafat, M. D., Kusumadewi, S., & Ratnasari, C. I. (2025). Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak. *JIIP-Jurnal Ilmiah Ilmu Pendidikan*, 8(3), 3263-3271.

Sebagian isi dari tesis ini telah dipublikasikan dalam bentuk artikel jurnal berjudul **"Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak"** yang diterbitkan pada **JIIP (Jurnal Ilmiah Ilmu Pendidikan)**, **Volume 8, Nomor 3, Maret 2025, halaman 3263–3271**.

Kontributor	Jenis Kontribusi
Muchtadho Dhila Arafat	Bertanggung jawab penuh atas ide penelitian, pengumpulan dan analisis data, pembuatan sistem, implementasi metode NLP berbasis TF-IDF dan <i>Cosine Similarity</i> , serta penulisan artikel jurnal.
Sri Kusumadewi	Memberikan arahan konseptual, supervisi penelitian, serta melakukan revisi substansi dalam struktur penulisan artikel.
Chanifah Indah Ratnasari	Memberikan masukan metodologis terkait penerapan NLP dalam bidang kesehatan serta melakukan review akhir naskah artikel sebelum dikirim ke jurnal.

Halaman Kontribusi

Dalam penelitian ini, kontribusi signifikan diberikan oleh Dokter Klinik Margorejo yang telah menyediakan informasi terkait gejala-gejala yang digunakan sebagai dasar dalam penelitian ini. Selain itu, Dokter Klinik Margorejo juga berperan dalam menyediakan dataset pasien yang digunakan dalam proses *voice-to-text*, sehingga memungkinkan validasi dan analisis lebih lanjut terhadap model pengolahan bahasa alami yang dikembangkan dalam penelitian ini. Kontribusi ini sangat berperan dalam memastikan bahwa penelitian memiliki dasar klinis yang kuat serta relevan dengan kondisi nyata di lapangan.

Halaman Persembahan

Dengan penuh rasa syukur dan penghormatan, karya ini saya persembahkan kepada alm. Ayah saya, dr. Cahyono Widodo, yang selalu menjadi sumber inspirasi dan semangat dalam setiap langkah hidup saya. Kepada Ibu saya, Dwikiswati Pujiastuti, yang dengan kasih sayang dan doa tak henti-hentinya mengiringi perjalanan saya.

Untuk istri tercinta, Maurizka Viera Chairunissa, yang selalu memberikan dukungan, kesabaran, dan cinta dalam setiap fase perjalanan ini. Untuk buah hati saya, Miray Alanya Ceyda Arafat, yang menjadi alasan terbesar saya untuk terus berusaha dan berkarya.

Untuk kakak saya, dr. Lutfhi Arshaldo Cakranegara, yang selalu menjadi teladan dan memberikan dukungan tanpa batas.

Ucapan terima kasih yang tulus juga saya sampaikan kepada Tim Klinik Margorejo, khususnya dr. Abid Abdurrahman dan dr. Insania Chusna, atas segala bantuan dan kontribusi dalam penelitian ini.

Tak lupa, kepada teman-teman semua yang telah menjadi bagian dari perjalanan ini, terima kasih atas kebersamaan, semangat, dan dukungan yang telah diberikan.

Saya juga menyampaikan penghargaan setinggi-tingginya kepada Ibu Dr. Sri Kusumadewi dan Ibu Chanifah Indah Ratnasari selaku dosen pembimbing, atas bimbingan, ilmu, dan arahan yang sangat berharga dalam proses penyusunan tesis ini. Terima kasih atas kesabaran dan dedikasi yang luar biasa.

Terakhir, kepada Universitas Islam Indonesia, tempat saya menimba ilmu dan berkembang, saya haturkan rasa terima kasih atas kesempatan dan pengalaman akademik yang begitu berharga.

Semoga karya ini dapat menjadi manfaat bagi banyak orang dan menjadi wujud penghormatan saya kepada semua yang telah mendukung dan mendoakan.

Kata Pengantar

Puji syukur kepada Allah SWT, Tuhan Yang Maha Esa sehingga dapat menyelesaikan thesis dengan judul “Pengolahan Bahasa Alami untuk Anamnesis Penyakit pada Anak”. Thesis ini disusun sebagai salah satu tahapan dalam menyelesaikan thesis pada Fakultas Teknologi Industri, Universitas Islam Indonesia.

Oleh karena itu, pada kesempatan ini kami ingin mengucapkan terima kasih kepada semua pihak yang telah mendukung serta membantu saya selama proses penyusunan thesis, kepada:

1. Prof. Fathul Wahid, S.T., M.Sc., Ph.D. selaku *Rector* Universitas Islam Indonesia
2. Prof. Dr. Ir. Hari Purnomo, M.T., IPU selaku Dekan Fakultas Teknologi Industri
3. Irving Vitra Paputungan, S.T., M.Sc., Ph.D. selaku Kepala Program Studi Magister Informatika
4. Dr. Sri Kusumadewi, S.Si., M.T. dan Chanifah Indah Ratnasari, S.Kom., M.Kom. selaku Dosen Pembimbing Thesis yang telah meluangkan waktu dan kesempatan untuk memberikan pengarahan dan bimbingan kepada penulis dalam penyusunan Thesis ini.
5. Orang tua dan keluarga penulis yang terus memberikan dukungan serta doa dalam penyusunan thesis kami.
6. Semua teman dan pihak-pihak yang tidak bisa disebutkan satu persatu.

Saya menyadari bahwa laporan thesis ini tidak luput dari berbagai kekurangan. Oleh karena itu, semua kritik dan saran yang bersifat memperbaiki demi kesempurnaan sangat diharapkan.

Yogyakarta,

Penulis

Muchtadho Dhila Arafat

Daftar Isi

Lembar Pengesahan Pembimbing	i
Lembar Pengesahan Penguji.....	ii
Abstrak	iii
Abstract	iv
Pernyataan Keaslian Tulisan	v
Daftar Publikasi	vi
Halaman Kontribusi.....	vii
Halaman Persembahan	viii
Kata Pengantar.....	ix
Daftar Isi.....	x
Daftar Tabel.....	xi
Daftar Gambar.....	xii
Glosarium	xv
BAB 1 Pendahuluan.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	6
1.6 Sistematika Penulisan	6
BAB 2 Kajian Pustaka.....	8
2.1 Tinjauan Pustaka	8
2.2 Landasan Teori.....	19
2.2.1 Gejala Demam.....	19
2.2.3 Metode CRISP-DM.....	21
BAB 3 Metodologi.....	30
3.1 Pendekatan Penelitian	30
3.2 Waktu dan Tempat Penelitian	30
3.3 Tahapan Penelitian	30
BAB 4 Hasil dan Pembahasan.....	54
4.1 Hasil Implementasi.....	54
4.1.1 Implementasi Program	54
4.1.2 Tampilan Aplikasi.....	55
4.1.3 Pengujian Aplikasi	69
4.2 Pembahasan.....	77
BAB 5 Kesimpulan dan Saran.....	81
5.1 Kesimpulan	81
5.2 Saran.....	81
Daftar Pustaka	83
LAMPIRAN.....	90
Lampiran A. Perancangan	90

Daftar Tabel

Tabel 2.1 Penelitian Terkait.....	11
Tabel 2.2 Bobot Jawaban.....	27
Tabel 2.3 Kategori Kelayakan	28
Tabel 3.1 Keterangan “ <i>The Sacred Seven of History Taking</i> ”	37
Tabel 3.2 Input Keluhan	40
Tabel 3.3 Perbandingan Original Dataset dan Hasil Case Folding.....	41
Tabel 3.4 Perbandingan Case Folding dan Hasil Tokenizing.....	42
Tabel 3.5 Perbandingan Tokenizing dan Hasil Stopwords.....	43
Tabel 3.6 Perbandingan Stopwords dan Hasil Stemming.....	43
Tabel 3.7 Hasil Preprocessing.....	44
Tabel 3.8 Hasil TF-IDF	44
Tabel 3.9 Hasil TF-IDF Setelah Normalisasi	45
Tabel 3.10 Hasil Perhitungan Cosine Similarity	47
Tabel 3.11 Hasil Sorting.....	48
Tabel 3.12 Hasil Top N-Recommend	49
Tabel 4.1 Pengujian <i>Black box</i> - Admin.....	67
Tabel 4.2 Pengujian <i>Black box</i> - Pengguna	69
Tabel 4.3 Pengujian <i>Black box</i> - Dokter	71
Tabel 4.4 Evaluasi <i>User Acceptance Testing</i>	72
Tabel 4.5 Pengujian NLP.....	74

Daftar Gambar

Gambar 1.1 Tren Bulanan Penyakit Anak < 5 Tahun di Klinik Margorejo Tahun 2024	2
Gambar 1.2 Tren Bulanan 4 (Empat) Penyakit Utama Anak < 5 Tahun di Klinik Margorejo Tahun 2024	2
Gambar 2.1 Tahapan CRISP-DM.	22
Gambar 3.1 Tahapan Penelitian	30
Gambar 3.2 Tahapan <i>Modeling</i>	36
Gambar 3.3 <i>Flowchart Preprocessing</i>	41
Gambar 4.1 Tampilan <i>Home</i>	54
Gambar 4.2 Tampilan <i>Login</i>	54
Gambar 4.3 Tampilan <i>Login</i> Jika <i>Username/Password</i> Tidak Sesuai	54
Gambar 4.4 Tampilan <i>Dashboard</i> – Admin	55
Gambar 4.5 Hasil Tampilan Antarmuka Modul <i>User</i>	55
Gambar 4.6 Hasil Tampilan Antarmuka Tambah <i>User</i>	55
Gambar 4.7 Hasil Tampilan Antarmuka <i>Edit User</i>	56
Gambar 4.8 Hasil Tampilan Antarmuka Hapus <i>User</i>	56
Gambar 4.9 Hasil Tampilan Antarmuka Modul Gejala.....	56
Gambar 4.10 Hasil Tampilan Antarmuka Tambah Gejala	57
Gambar 4.11 Hasil Tampilan Antarmuka <i>Edit Gejala</i>	57
Gambar 4.12 Hasil Tampilan Antarmuka Hapus Gejala	57
Gambar 4.13 Hasil Tampilan Antarmuka Laporan Hasil Konsultasi	57
Gambar 4.14 Proses NLP Pasien 1	58
Gambar 4.15 Proses NLP Pasien 2	60
Gambar 4.16 Hasil TF-IDF Salah Satu Pasien	61
Gambar 4.17 Tampilan Antarmuka Register	62
Gambar 4.18 Tampilan <i>Dashboard</i> – Pengguna	62
Gambar 4.19 Tampilan Antarmuka <i>Upload Audio</i>	63
Gambar 4.20 Tampilan Antarmuka <i>Record Audio</i>	63
Gambar 4.21 Tampilan Antarmuka Setelah Melakukan Konsultasi	63
Gambar 4.22 Tampilan Antarmuka Laporan Hasil Konsultasi	64
Gambar 4.23 Tampilan Menambahkan Keluhan Lanjutan.....	64
Gambar 4.24 Tampilan Berhasil Menambahkan Keluhan Lanjutan	65
Gambar 4.25 Tampilan <i>Dashboard</i> – Dokter	65

Gambar 4.26 Tampilan Menu Data Anamnesis.....	66
Gambar 4.27 Tampilan Antarmuka Pengelompokkan Gejala sesuai dengan “ <i>The Sacred Seven of History Taking</i> ”	66
Gambar A.1 <i>Flowchart</i> Admin.....	90
Gambar A.2 <i>Flowchart</i> Dokter.....	91
Gambar A.3 <i>Flowchart</i> Pengguna	93
Gambar A.4 <i>Use Case Diagram</i>	94
Gambar A.5 <i>Class Diagram</i>	95
Gambar A.6 Perancangan Halaman <i>Home</i>	96
Gambar A.7 Perancangan Halaman Registrasi.....	96
Gambar A.8 Perancangan Halaman <i>Login</i>	97
Gambar A.9 Perancangan Halaman <i>Dashboard</i> Admin.....	97
Gambar A.10 Perancangan Halaman Data <i>User</i>	98
Gambar A.11 Perancangan Halaman Tambah Data <i>User</i>	98
Gambar A.12 Perancangan Halaman <i>Edit</i> Data <i>User</i>	99
Gambar A.13 Perancangan Halaman Hapus Data <i>User</i>	99
Gambar A.14 Perancangan Halaman Data Gejala.....	100
Gambar A.15 Perancangan Halaman Tambah Data Gejala.....	100
Gambar A.16 Perancangan Halaman <i>Edit</i> Data Gejala	101
Gambar A.17 Perancangan Halaman Hapus Data Gejala.....	102
Gambar A.18 Perancangan Halaman Laporan Hasil Konsultasi	102
Gambar A.19 Perancangan Halaman Proses NLP	102
Gambar A.20 Perancangan Halaman <i>Dashboard</i> Dokter.....	103
Gambar A.21 Perancangan Halaman Data Anamnesis.....	104
Gambar A.22 Perancangan Halaman Pengelompokkan Gejala pada “ <i>The Sacred Seven of History Taking</i> ”	104
Gambar A.23 Perancangan Halaman <i>Dashboard</i> Pengguna	105
Gambar A.24 Perancangan Halaman Konsultasi Keluhan	105
Gambar A.25 Perancangan Halaman Berhasil Konsultasi Keluhan.....	106
Gambar A.26 Perancangan Halaman Hasil Gejala.....	106
Gambar A.27 Perancangan Halaman Tambah Keluhan	107
Gambar A.28 Hasil Pertanyaan 1	111
Gambar A.29 Hasil Pertanyaan 2	111
Gambar A.30 Hasil Pertanyaan 3	112

Gambar A.31 Hasil Pertanyaan 4	112
Gambar A.32 Hasil Pertanyaan 5	113
Gambar A.33 Hasil Pertanyaan 6	113
Gambar A.34 Hasil Pertanyaan 7	114
Gambar A.35 Hasil Pertanyaan 8	114
Gambar A.36 Hasil Pertanyaan 9	115

Glosarium

- CRISP-DM : CRISP-DM adalah metodologi yang terstruktur dan teruji waktu yang digunakan dalam proses pengembangan proyek data mining
- ISPA : Infeksi Saluran Pernafasan Akut (ISPA) adalah kondisi kesehatan yang ditandai dengan terjadinya infeksi pada saluran pernafasan bagian atas
- DBD : Demam Berdarah *Dengue* (DBD) didefinisikan sebagai suatu penyakit yang terjadi akibat gigitan nyamuk *aedes aegypti*
- NLP : NLP memberikan akses dan sarana untuk analisis dari sebuah teks atau tulisan dengan membuat sebuah struktur yang baik dan teks yang sebelumnya belum terstruktur untuk dianalisis berikutnya

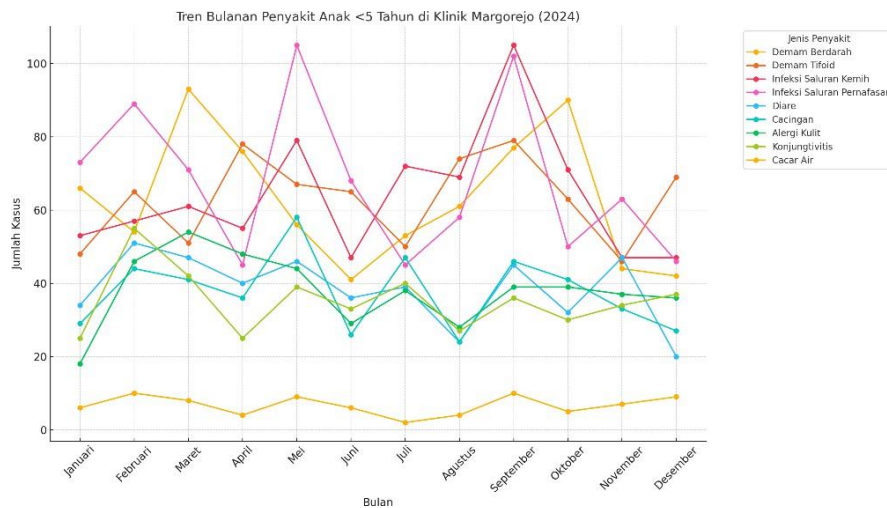
BAB 1

Pendahuluan

1.1 Latar Belakang

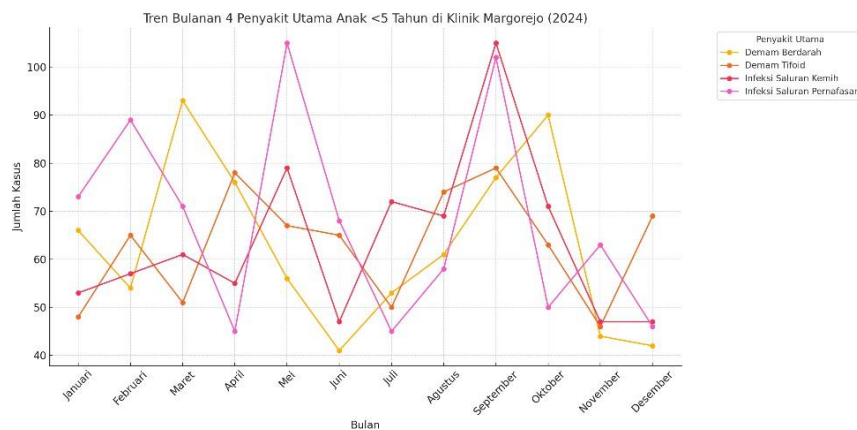
Teknologi yang semakin maju telah banyak dimanfaatkan dalam kehidupan sehari-hari salah satunya adalah pengolahan bahasa alami atau NLP. NLP merupakan salah satu kecerdasan buatan (AI) yang berfokus kepada bagaimana bahasa manusia berinteraksi dengan komputer (Toruan, 2024). Selain itu, dalam kehidupan sehari-hari NLP dapat dimanfaatkan untuk memahami konteks dari teks dengan baik dan memberikan dampak yang signifikan pada berbagai aplikasi atau sistem seperti analisis teks, penerjemah bahasa, dan sistem interaksi manusia dan komputer (Samosir, Toba, & Ayub, 2022). NLP juga menjadi salah satu alat penting dalam berbagai bidang, termasuk di bidang medis (Astuti, 2021).

Anamnesis merupakan salah satu langkah awal dalam diagnosis penyakit yang sangat bergantung pada komunikasi verbal antara pasien dan tenaga medis, dalam hal ini adalah dokter (Purba, Penmaley, & Panjaitan, 2024). Pada konteks anak usia di bawah lima tahun, komunikasi verbal biasanya disampaikan oleh orang tua atau pengasuh yang mendampingi, bukan oleh pasien itu sendiri. Hal ini menjadi tantangan tersendiri karena kemampuan orang tua dalam menggambarkan gejala dapat bervariasi, baik dari sisi detail, istilah yang digunakan, maupun urutan kronologis keluhan. Ketidaktepatan atau ketidakjelasan dalam komunikasi ini dapat memengaruhi interpretasi dokter dalam menentukan diagnosis awal. (Juwita, Hidayurrahmi, Adista, & Khairunnisa, 2023), sehingga dari permasalahan ini dapat menghambat proses diagnosa yang cepat dan tepat. Ketidaktepatan dalam mendiagnosis dapat berakibat dalam penanganan media yang kurang optimal, memperpanjang waktu diagnosis, dan pada dampak yang serius dapat membahayakan keselamatan pasien khususnya anak-anak (Nurachman & Fitrianingrum, 2022). Penyakit yang biasanya terjadi pada anak-anak, khususnya dengan gejala demam seperti *Demam Berdarah Dengue* (DBD), Infeksi Saluran Pernafasan Akut (ISPA), Demam Tifoid, dan *Urinary Tract Infection* (Andres & Purnomo, 2024), (Febriansyah, Nurtanio, & Yusuf, 2024), (Sihombing, Siska, Zebua, & Nasution, 2022). Berdasarkan tren bulanan penyakit pada anak usia kurang dari 5 (lima) tahun yang berasal dari Klinik Margorejo pada tahun 2024 dapat dilihat pada Gambar 1.1.



Gambar 1.1 Tren Bulanan Penyakit Anak < 5 Tahun di Klinik Margorejo Tahun 2024

Berdasarkan Gambar 1.1 telah divisualisasikan bahwa terdapat penyakit anak di bawah usia 5 (tahun) yang terdiri dari demam berdarah, demam tifoid, infeksi saluran kemih, infeksi saluran pernafasan, diare, cacingan, alergi kulit, konjungtivitis, dan cacar air. Dari kesembilan penyakit anak di bawah usia 5 (tahun) tersebut, tren penyakit yang tertinggi dapat dilihat pada Gambar 1.2. Dari visualisasi tersebut, diketahui bahwa penyakit demam berdarah, demam tifoid, infeksi saluran kemih, dan infeksi saluran pernafasan merupakan jenis penyakit dengan jumlah kasus tertinggi yang tercatat di Klinik Margorejo pada tahun 2024.



Gambar 1.2 Tren Bulanan 4 (Empat) Penyakit Utama Anak < 5 Tahun di Klinik Margorejo Tahun 2024

ISPA adalah kondisi kesehatan yang ditandai dengan terjadinya infeksi pada saluran pernafasan bagian atas. Penyakit tersebut disebabkan oleh virus atau bakteri yang menyerang hidung atau trakea dan mengakibatkan gangguan pada fungsi pernafasan (Andres & Purnomo, 2024). Tingkat mortalitas ISPA sangat tinggi pada bayi dan anak-anak terutama pada negara berkembang (Arafat, Simanjuntak, & Charibaldi, 2021). Sementara *Demam*

Berdarah Dengue (DBD) merupakan penyakit yang terjadi akibat gigitan nyamuk *aedes aegypti*. Penyakit tersebut mudah dijumpai di negara iklim tropis (Dewi & Sudaryanto, 2020). Selain itu, penyakit DBD di Indonesia banyak ditemukan pada kelompok usia 2 – 14 tahun (Arafat, Simanjuntak, & Charibaldi, 2021). Selanjutnya penyakit yang biasanya terjadi pada anak-anak lainnya yaitu demam tifoid, penyakit ini banyak dijumpai pada anak di negara berkembang, seperti Indonesia (Arafat, Simanjuntak, & Charibaldi, 2021). Prevalensi angka kejadian demam tifoid di Indonesia menurut data Kementerian Kesehatan RI menyebutkan sekitar 350-810 per 100.000 penduduk, yang artinya setiap tahun terdapat 600.000-1.500.000 kasus demam tifoid yang terjadi (Levani & Prastya, 2020). Sementara penyakit *Urinary Tract Infection* merupakan penyakit infeksi yang merupakan salah satu dari 10 penyakit terbanyak di Indonesia (Sihombing, Siska, Zebua, & Nasution, 2022). Penyakit *Urinary Tract Infection* pada bayi dan anak dapat menimbulkan dampak jangka panjang terhadap fungsi ginjal dan salah satunya gagal ginjal (Balqis, Anasta, & Puspasari, 2024). Untuk membantu pakar dalam mengambil tindakan diagnosa yang cepat dan tepat diperlukan solusi yang dapat menginterpretasikan komunikasi verbal antara pasien dan dokter yaitu dengan menerapkan NLP (Cannavaro, 2023). Penerapan keluhan pasien tersebut dalam penelitian ini nantinya dalam bentuk *voice to text* sehingga dapat untuk mengenali gejala yang diderita pasien. Hal ini juga didukung oleh penelitian (Yuwono, 2018) yang menerapkan data *speech to text* pada pasien. Selain itu pada penelitian yang dilakukan oleh (Ratnasari, Kusumadewi, & Rosita, 2014) juga melakukan penerapan NLP dalam data anamnesis pasien.

Penerapan teknologi NLP dalam proses anamnesis tersebut dapat berfungsi untuk menginterpretasikan pertanyaan ataupun pernyataan yang kompleks sehingga dapat memahami kalimat tersebut dan memberikan jawaban yang lebih tepat serta meningkatkan kualitas komunikasi (Cannavaro, 2023). NLP sendiri sudah banyak diterapkan dalam bidang medis. Hal ini dikarenakan NLP mampu menangkap konteks dan makna dari bahasa alami yang seringkali sulit diproses melalui metode konvensional, serta pokok utama pemanfaatan NLP di bidang medis dapat meningkatkan akurasi dalam pengenalan istilah-istilah yang penting dan mampu memberikan dukungan dalam pengambilan keputusan klinis, serta mempercepat waktu atau meningkatkan efisiensi dalam melakukan analisis data besar pada catatan medis elektronik (EHR) (Gholipour, 2023). Oleh karena itu, NLP dapat digunakan untuk menganalisis dan memproses bahasa alami pada anak-anak dengan mengidentifikasi

pola-pola tertentu dan membantu dokter dalam menginterpretasikan gejala yang disampaikan dengan cepat dan tepat (Anggoro & Akbar, 2023).

Berdasarkan uraian di atas, maka dalam penelitian ini akan melakukan penerapan *Natural Language Processing* (NLP) untuk anamnesis penyakit pada anak. Pada penelitian ini akan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), dan *Cosine Similarity* dalam mengidentifikasi data anamnesis tersebut. *Cosine similarity* adalah ukuran kesamaan yang sering digunakan untuk menghitung nilai relevansi antara kueri dan dokumen, serta nilai relevansi antara dokumen yang berbeda (Rasyid & Ningsih, 2024). Penggunaan *cosine similarity* juga didukung dari penelitian yang dilakukan oleh (Singh & Singh, 2021) pada penelitiannya membahas mengenai pengukuran *text similarity* pada artikel berita menggunakan NLP. *Similarity* yang digunakan pada penelitian tersebut yaitu seperti *Cosine similarity*, *Jaccard similarity*, dan *Euclidean similarity*. Hasil penelitian tersebut menunjukkan bahwa performa dari *cosine similarity* lebih unggul dengan akurasi sebesar 0.8215, dibandingkan kedua metode *similarity* lain seperti *Jaccard similarity* menghasilkan akurasi sebesar 0.75, dan *Euclidean similarity* dengan akurasi 0.5625.

Penggunaan teknik TF-IDF pada penelitian ini dikarenakan TF-IDF memiliki kelebihan yaitu kemampuannya untuk memberikan bobot yang tepat pada *term-term* dalam dokumen (Annisa & Kalifia, 2024). Dengan menggabungkan skor TF dan IDF, metode TF-IDF mampu mengidentifikasi *term-term* yang paling relevan dalam dokumen (Annisa & Kalifia, 2024). Penerapan TF-IDF didukung penelitian yang telah dilakukan oleh (Merryton & Augasta, 2022) yang membahas mengenai deteksi berita palsu dengan Teknik NLP dan membandingkan penggunaan *Count Vectorizer* serta TF-IDF. Hasil dari penelitian tersebut adalah TF-IDF lebih unggul dengan penerapan pada metode *passive aggressive* sebesar 93,5%, dibandingkan *Count Vectorizer* dengan penerapan metode yang sama sebesar 89,4%. Penelitian lain yang juga membandingkan TF-IDF dengan Teknik lain dilakukan oleh (Min, Lee, & Lee, 2022). Penelitian tersebut melakukan analisis karakteristik URL spam dengan menerapkan beberapa *vectorizer* seperti TF-IDF, *Count Vectorizer*, dan *Hashing Vectorizer*. Hasil penelitian tersebut adalah TF-IDF dapat menghasilkan performa terbaik yang dominan dibandingkan kedua Teknik yang ada. Seperti pada penerapan TF-IDF dengan algoritma linear SVM dihasilkan akurasi 97.99%, sementara pada *Count Vectorizer* dengan algoritma yang sama sebesar 97.75%, dan pada *Hashing Vectorizer* sebesar 97.91%. Penelitian selanjutnya yang juga menerapkan perbandingan TF-IDF dilakukan oleh (Mahendra, Murdiansyah, & Lhaksana, 2023) yang membahas analisis sentimen pada Tweet Covid-

19. Teknik yang digunakan pada penerapan metode K-Nearest Neighbors adalah TF-IDF dan *CountVectorizer*. Hasil dari penelitian tersebut adalah akurasi yang didapatkan TF-IDF sebesar 73,2%, sementara *CountVectorizer* sebesar 65,33%. Hal ini dapat dikarenakan *CountVectorizer* hanya menghitung jumlah kemunculan kata, sedangkan TF-IDF melakukan pembobotan yang skornya menunjukkan seberapa pentingnya sebuah term dalam dokumen dan seluruh korpus. Pada penelitian yang akan dilakukan akan menggunakan empat *class* yaitu DBD, ISPA, Demam Tifoid, dan *Urinary Tract Infection* dari anamnesis user. Kerangka kerja yang digunakan dalam penelitian ini adalah *Cross Industry Standard for Data Mining* (CRISP-DM). Hal ini dikarenakan CRISP-DM memiliki tahapan dan kerangka kerja yang terstruktur sehingga pengguna akan lebih terarah dan mengetahui langkah yang harus dikerjakan dalam penelitian (Dhewayani, Amelia, Alifah, Sari, & Jajuli, 2022).

1.2 Rumusan Masalah

Dari latar belakang masalah yang telah dijelaskan, maka dalam penelitian ini terdapat rumusan masalah yaitu:

1. Bagaimana mengimplementasikan *input* dari *voice to text* yang dapat mengenali gejala yang diderita pasien dari data anamnesis penyakit pada anak?
2. Bagaimana hasil anamnesis pada pengolahan bahasa alami berbasis *website* dalam mengenali penyakit dengan gejala demam pada anak berdasarkan *input voice to text* ?

1.3 Batasan Masalah

Dalam melakukan penelitian ini perlu ditetapkan batasan-batasan untuk mempertahankan fokus penelitian. Adapun batasan masalah pada penelitian ini diantaranya:

1. Data yang digunakan adalah data balita bergejala demam yang didapatkan dari Klinik Margorejo.
2. Penelitian ini melakukan tahap pengubahan input data anamnesis berupa *voice to text*, kemudian dilakukan *text preprocessing*, TF IDF - *Cosine Similarity* guna mengolah data sesuai dengan keperluan anamnesis penyakit pada anak.
3. Kasus pada penelitian ini adalah anamnesis penyakit dengan gejala demam pada anak umur dibawah 5 tahun.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dibuat, maka tujuan penelitian ini yaitu:

1. Mengimplementasikan *input* dari *voice to text* untuk dapat mengenali gejala yang diderita pasien dari data anamnesis penyakit pada anak dengan menggunakan *library SpeechRecognition*.
2. Mengetahui hasil anamnesis pada pengolahan bahasa alami berbasis *website* dalam mengenali penyakit dengan gejala demam pada anak berdasarkan *input voice to text*

1.5 Manfaat Penelitian

Dari tujuan penelitian yang telah dijelaskan, maka dalam penelitian ini terdapat manfaat penelitian sebagai berikut:

1. Dapat membantu pihak terkait seputar mengekstrak informasi penyakit dengan gejala demam pada anak melalui anamnesis berbentuk audio.
2. Dapat digunakan sebagai acuan/referensi penelitian yang akan dilakukan oleh penelitian selanjutnya dalam melakukan identifikasi ekstraksi informasi penyakit.

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini akan dibagi menjadi lima bab dengan sistematika sebagai berikut:

BAB 1 Pendahuluan

Bab ini berisi latar belakang yang menjelaskan urgensi dan permasalahan dalam proses anamnesis penyakit pada anak, khususnya tantangan dalam komunikasi verbal yang dilakukan oleh orang tua atau pengasuh. Kemudian disajikan rumusan masalah, batasan masalah, tujuan penelitian, serta manfaat dari penelitian ini. Tujuan utamanya adalah menerapkan teknologi pengolahan bahasa alami (NLP) untuk meningkatkan keakuratan identifikasi gejala pada anak berdasarkan input suara yang dikonversi menjadi teks, sehingga dapat mendukung proses diagnosis medis secara lebih cepat dan tepat.

BAB 2 Tinjauan Pustaka

Bab ini menyajikan kajian terhadap penelitian-penelitian terdahulu yang relevan sebagai landasan teoritis untuk mendukung penelitian ini. Sub-bab ini memuat konsep-konsep dasar mengenai gejala demam pada anak, Natural Language Processing (NLP), metode CRISP-DM, tahapan text processing, algoritma TF-IDF, Cosine Similarity, serta library pendukung seperti SpeechRecognition. Selain itu, disajikan pula ulasan terhadap berbagai penelitian yang menggunakan metode serupa di bidang kesehatan, serta identifikasi keunggulan dan kekurangan dari masing-masing studi untuk membedakan kontribusi penelitian ini.

BAB 3 Metode Penelitian

Bab ini menjelaskan metode penelitian yang digunakan, dimulai dari pendekatan deskriptif kuantitatif dan penerapan metode CRISP-DM sebagai kerangka kerja dalam pemrosesan data. Dijelaskan pula waktu dan tempat penelitian, tahapan penelitian secara rinci mulai dari business understanding hingga deployment. Selain itu, dijelaskan implementasi voice-to-text dengan library SpeechRecognition, tahapan preprocessing teks, perhitungan TF-IDF dan Cosine Similarity, serta pengembangan sistem berbasis website dengan pendekatan waterfall. Evaluasi sistem dilakukan melalui black box testing dan user acceptance test (UAT).

BAB 4 Hasil dan Pembahasan

Bab ini memaparkan hasil implementasi sistem NLP untuk anamnesis penyakit pada anak, mulai dari tampilan antarmuka pengguna hingga hasil pengujian sistem. Dijelaskan pula bagaimana proses NLP mengidentifikasi gejala dari input suara dan mengelompokkannya berdasarkan “The Sacred Seven of History Taking”. Pembahasan dalam bab ini mengulas efektivitas metode yang digunakan, analisis kemiripan gejala dengan cosine similarity, serta peluang pengembangan lebih lanjut, seperti perluasan cakupan penyakit dan integrasi fitur NLP lanjutan untuk mengurangi potensi bias dalam interpretasi gejala.

BAB 5 Kesimpulan dan Saran

Bab terakhir ini merangkum kesimpulan dari penelitian, menegaskan bahwa penerapan NLP dengan TF-IDF dan Cosine Similarity dapat digunakan untuk mendukung proses anamnesis penyakit pada anak secara efektif melalui input suara. Sistem berhasil mengidentifikasi gejala dengan tingkat kemiripan di atas ambang batas, memberikan struktur data yang bermanfaat untuk proses diagnosis medis. Bab ini juga memberikan beberapa saran untuk pengembangan lebih lanjut, termasuk penggunaan basis data dinamis untuk kamus gejala, perluasan domain penyakit, dan integrasi algoritma klasifikasi untuk meningkatkan akurasi diagnosis.

BAB 2

Kajian Pustaka

2.1 Tinjauan Pustaka

Pada tinjauan Pustaka menjelaskan penelitian terdahulu yang berkaitan dengan penelitian saat ini sebagai bahan referensi penelitian.

Pada penelitian yang dilakukan oleh (Athota, Shukla, Pandey, & Rana, 2020) yang membahas mengenai *chatbot* untuk sistem perawatan kesehatan menggunakan *Artificial Intelligence* (AI). Permasalahan yang melatarbelakangi penelitian ini adalah sangat sulit untuk mendapatkan konsultasi dengan dokter untuk setiap masalah kesehatan. Penelitian tersebut dilakukan untuk membuat *chatbot* medis menggunakan Pemeringkatan, perhitungan kemiripan kalimat dilakukan dengan menggunakan n-gram, TF-IDF, dan cosine similarity. Hasil dari penelitian ini adalah *chatbot* berhasil dikembangkan dan dapat memberikan kualitas jawaban dalam waktu singkat.

Pada penelitian yang dilakukan oleh (Nugraheni, Widodo, & Sari, 2021) yang membahas mengenai *case-based reasoning* dalam diagnosis penyakit pada COVID-19. Hal yang melatarbelakangi penelitian ini adalah agar dapat melakukan diagnosis penyakit yang terlihat atau dirasakan. Sehingga pada penelitian ini melakukan pengembangan *case-based reasoning* untuk penyakit COVID-19 dengan menggunakan metode *cosine similarity*. Hasil dari penelitian ini adalah dapat mendeteksi tingkat kemiripan sebesar 80%.

Pada penelitian yang dilakukan oleh (De, Goswami, Faujidar, & Singh, 2024) yang membahas mengenai sistem pakar diagnosis penyakit Ginekologi. Hal yang melatarbelakangi penelitian ini adalah di negara seperti India, hanya sebagian kecil dari masyarakat kurang mampu memiliki akses ke perawatan ginekologi berkualitas tinggi, dan beberapa wanita menolak untuk mengunjungi dokter untuk masalah tersebut. Sehingga untuk memudahkan pada penelitian tersebut dilakukan pengembangan sistem pakar diagnosis penyakit Ginekologi menggunakan *machine learning* dan NLP untuk menghitung *cosine similarity*. Hasil dari penelitian ini adalah dapat menghasilkan akurasi dan presisi sebesar 93%, *recall* dan *f1-score* sebesar 92%.

Pada penelitian yang dilakukan oleh (Manoharan & Sathish, 2020) yang membahas mengenai sistem rekomendasi diet pasien. Hal yang melatarbelakangi penelitian ini adalah sistem rekomendasi dapat diterapkan bagi diet pasien, dan dapat mempermudah pemberian

sarannya. Sehingga pada penelitian tersebut dilakukan pengembangan sistem rekomendasi diet pasien menggunakan K Clique dan perhitungan *similarity*, serta penerapan algoritma *machine learning*. Hasil dari penelitian ini adalah algoritma K-DLRS menghasilkan akurasi tertinggi sebesar 96% pada data training dan 95% pada data testing.

Pada penelitian yang dilakukan oleh (Almuhaimeed, Alhomidi, & Alenezi, 2022) yang membahas mengenai deteksi influenza. Permasalahan yang melatarbelakangi penelitian ini adalah mengekstrak informasi dari data dan kemudian, menghitung *semantic similarity* adalah tugas yang tidak mudah karena sumber data tersebut memiliki cara yang berbeda-beda untuk mendeskripsikan informasi. Sehingga pada penelitian tersebut dilakukan pendekatan *semantic similarity* untuk deteksi influenza. Hasil dari penelitian ini adalah dengan menggunakan 3 pendekatan seperti Bio_inf menghasilkan *f-score* sebesar 88%, Hussain_Inf menghasilkan *f-score* sebesar 32%, dan Cosine_Inf menghasilkan *f-score* sebesar 18%

Pada penelitian yang dilakukan oleh (Laksono, Hidayatullah, & Ratnasari, 2018) membahas mengenai keluhan pasien dalam bentuk *speech* yang akan diubah ke dalam bentuk teks untuk bahasa Indonesia. Hal yang melatarbelakangi penelitian ini adalah pengubahan *speech* (suara) menjadi teks menjadi hal yang menantang dalam bidang NLP. Sehingga pada penelitian tersebut mencoba membangun aplikasi *speech to text* dari keluhan pasien untuk bahasa Indonesia. Pada penelitian tersebut menggunakan 500 data suara yang direkam dari 10 orang, 50 data tersebut diuji dengan hasil akurasi sebesar 64%. Selain itu, pada penelitian tersebut menerapkan proses *similarity*, dan menghasilkan hanya ada 3 (tiga) kata yang dihasilkan dengan benar dan 9 (sembilan) kata dikategorikan sebagai cukup, sementara kata-kata lainnya dikategorikan sebagai buruk.

Pada penelitian yang dilakukan oleh (Ratnasari, Kusumadewi, & Rosita, 2014) yang membahas mengenai penerapan NLP untuk perumusan keluhan pasien. Permasalahan yang melatarbelakangi penelitian tersebut adalah pencatatan anamnesis dengan pencatatan yang sering terdapat kesalahan eja/salah ketik. Sehingga hal tersebut perlu ditangani dengan cara yang tepat mengingat hal tersebut dipergunakan untuk menjalankan berbagai macam pendukung keputusan klinis. Penanganan tersebut dilakukan dengan menggunakan NLP. Hasil dari penelitian ini adalah hasil pengolahan teks menjadi bentuk yang lebih terstruktur dapat dijadikan salah satu bahan yang dapat diintegrasikan dengan sistem pendukung keputusan klinis.

Pada penelitian yang dilakukan oleh (Yuwono, 2018) membahas mengenai NLP untuk penerjemahan keluhan pasien pada anamnesis dalam bentuk teks atau suara. Permasalahan yang melatarbelakangi penelitian ini adalah data keluhan pasien yang ada pada rekam medis elektronik sering tertulis dalam format penulisan bebas yang kurang baku serta tidak terstruktur sehingga menyulitkan komputer untuk melakukan pemrosesan secara otomatis terhadap data tersebut. Oleh karena itu, pada penelitian tersebut melakukan pengolahan anamnesis dalam bentuk teks atau suara. Hasil dari penelitian ini adalah pada penerapan fitur *speech recognition* setelah diuji coba masih belum cukup memuaskan karena 5 (lima) dokter yang menguji aplikasi tersebut tidak ada satupun dokter yang setuju bahwa tingkat akurasi dan kecepatan fitur *speech recognition* pada aplikasi sudah cukup baik.

Pada penelitian yang dilakukan oleh (Sebayang, 2024) membahas mengenai diagnosis penderita demam berdarah menggunakan sistem pakar. Permasalahan yang melatarbelakangi penelitian ini adalah demam berdarah merupakan penyakit yang meresahkan, terutama pada anak-anak yang lebih rentan terkena penyakit tersebut. Sehingga pada penelitian tersebut dilakukan diagnosis penderita demam berdarah dengan menggunakan algoritma naive bayes. Hasil dari penelitian ini adalah dengan menghasilkan akurasi tertinggi sebesar 97%.

Pada penelitian yang dilakukan oleh oleh (Sarwar & Mamun, 2022) yang membahas mengenai prediksi demam berdarah. Permasalahan yang melatarbelakangi penelitian ini adalah demam berdarah menjadi salah satu penyebab utama rawat inap dan kematian terutama pada anak-anak, sehingga pada penelitian tersebut dilakukan prediksi demam berdarah menggunakan algoritma *machine learning* yaitu SVM. Hasil dari penelitian tersebut yaitu menghasilkan akurasi yang tinggi sebesar 97%.

Pada penelitian yang dilakukan oleh (Yehuala, et al., 2024) yang membahas mengenai prediksi perilaku pencarian layanan kesehatan para ibu untuk ISPA dan faktor-faktor penentunya di antara anak-anak balita di Afrika. Permasalahan yang melatarbelakangi penelitian ini adalah ISPA adalah penyebab utama kematian anak di bawah usia 5 tahun. Sehingga untuk mengidentifikasi hal tersebut pada penelitian tersebut menggunakan algoritma machine learning dalam prosesnya. Hasil dari penelitian ini adalah *Random forest* (RF) menghasilkan akurasi terbaik sebesar 88.89%, precision 89.5%, f-measure 83%, AUC ROC Curve sebesar 95.8% dan *recall* 77.6%.

Pada penelitian yang dilakukan oleh (Nshimiyimana & Zhou, 2022) yang membahas mengenai analisis faktor risiko terkait ISPA pada anak di Uganda. Hal yang

melatarbelakangi penelitian ini adalah di Afrika Selatan ISPA menjadi penyebab utama dalam kematian pada anak. Sehingga pada penelitian tersebut dilakukan analisis faktor risiko gejala penyakit ISPA pada anak di bawah usia 5 tahun di Uganda menggunakan machine learning. Hasil dari penelitian ini adalah RF menghasilkan akurasi yang tinggi sebesar 88.7% dengan nilai AUC sebesar 0.951.

Pada penelitian yang dilakukan oleh (Lee, Chung, Kang, Kim, & Kwak, 2021) membahas mengenai prediksi *Serious Bacterial Infection* (SBI) pada anak-anak yang mengalami demam di Unit Gawat Darurat. SBI terdiri dari bacteremia, *Urinary Tract Infection*, lobar pneumonia, bacterial CNS infection, septic arthritis. Permasalahan yang melatarbelakangi penelitian ini adalah alat untuk mendiagnosis SBI telah dikembangkan, namun alat tersebut hanya dapat memprediksi SBI jika ada nilai dari semua faktor yang digunakan dalam alat tersebut, jika salah satu dari mereka hilang, alat tersebut menjadi tidak berguna. Sehingga pada penelitian tersebut dilakukan prediksi SBI pada anak-anak menggunakan *machine learning*. Hasil dari penelitian ini adalah RF mendapatkan hasil tertinggi sebesar 0,964 untuk nilai AUROC.

Dari uraian diatas mengenai penelitian terkait disajikan dalam Tabel 2.1

Tabel 2.1 Penelitian Terkait

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
1	<i>Chatbot for Healthcare System Using Artificial Intelligence</i>	<i>Chatbot, Healthcare, Artificial Intelligence, Virtual Assistance, TFID, N-gram</i>	Penelitian ini membahas mengenai mengenai <i>chatbot</i> untuk sistem perawatan kesehatan menggunakan <i>Artificial Intelligence</i> (AI). Hasil dari penelitian ini adalah <i>chatbot</i> berhasil dikembangkan	(Athota, Shukla, Pandey, & Rana, 2020)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
			dan dapat memberikan kualitas jawaban dalam waktu singkat	
2	<i>A Case-Based Reasoning for Detection Coronavirus (Covid-19) Using Cosine Similarity</i>	<i>Coronavirus Detection, Artificial Intelligence, Early Detection, Cosine Similarity Method</i>	Penelitian ini membahas mengenai <i>case-based reasoning</i> dalam diagnosis penyakit pada COVID-19. Hasil dari penelitian ini adalah dapat mendeteksi tingkat kemiripan sebesar 80%	(Nugraheni, Widodo, & Sari, 2021)
3	<i>Gynaecological Disease Diagnosis Expert System (GDDES) Based on Machine Learning Algorithm and Natural Language Processing</i>	<i>Gynaecological Disease Diagnosis Expert System, Machine Learning, Natural Language Processing</i>	Penelitian ini membahas mengenai sistem pakar diagnosis penyakit Ginekologi. Hasil dari penelitian ini adalah dapat menghasilkan akurasi dan presisi sebesar 93%, <i>recall</i> dan <i>f1-score</i> sebesar 92%.	(De, Goswami, Faujidar, & Singh, 2024)
4	<i>Patient Diet Recommendation</i>	<i>Patient Recommendation</i>	Penelitian ini membahas	(Manoharan & Sathish, 2020)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
	<i>System Using K Clique and Deep learning Classifiers</i>	<i>System, K-clique, Deep Learning Classifiers, Gated Recurrent Units, K-DLRS, Improved Preciseness and Accuracy</i>	mengenai sistem rekomendasi diet pasien. Hasil dari penelitian ini adalah algoritma K-DLRS menghasilkan akurasi tertinggi sebesar 96% pada data training dan 95% pada data testing	
5	<i>A modern semantic similarity method using multiple resources for enhancing influenza detection</i>	<i>Semantic web, Semantic similarity, Ontologies Infectious diseases</i>	Penelitian ini membahas mengenai deteksi influenza. Hasil dari penelitian ini adalah dengan menggunakan 3 pendekatan seperti Bio_inf menghasilkan <i>f-score</i> sebesar 88%, Hussain_Inf menghasilkan <i>f-score</i> sebesar 32%, dan Cosine_Inf menghasilkan <i>f-score</i> sebesar 18%	(Almuhaimed, Alhomidi, & Alenezi, 2022)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
6	<i>Speech to Text of Patient Complaints for Bahasa Indonesia</i>	<i>speech to text, patient complaint, anamnesis, natural language processing, mfcc, ctc</i>	<p>Penelitian ini membahas mengenai keluhan pasien dalam bentuk <i>speech</i> yang akan diubah ke dalam bentuk teks untuk bahasa Indonesia. Hasil dari penelitian ini adalah dengan 50 data tersebut diuji dengan hasil akurasi sebesar 64%. Selain itu, pada penelitian tersebut menerapkan proses <i>similarity</i>, dan menghasilkan hanya ada 3 (tiga) kata yang dihasilkan dengan benar dan 9 (Sembilan) kata dikategorikan sebagai cukup, sementara kata-kata lainnya dikategorikan sebagai buruk</p>	(Laksono, Hidayatullah, & Ratnasari, 2018)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
7	Model <i>Natural Language Processing</i> untuk Perumusan Keluhan Pasien	anamnesis, keluhan pasien, <i>natural language processing</i>	Penelitian ini membahas mengenai penerapan NLP dalam data anamnesis. Hasil dari penelitian ini adalah hasil pengolahan teks menjadi bentuk yang lebih terstruktur dapat dijadikan salah satu bahan yang dapat diintegrasikan dengan sistem pendukung keputusan klinis.	(Ratnasari, Kusumadewi, & Rosita, 2014)
8	Pengolahan Bahasa Alami untuk Penerjemahan Keluhan Pasien pada Anamnesis dalam Bentuk Teks atau Suara	pengolahan bahasa alami, keluhan pasien, anamnesis, <i>speech recognition</i>	Penelitian ini membahas mengenai NLP untuk penerjemahan keluhan pasien pada anamnesis dalam bentuk teks atau suara. Hasil dari penelitian ini adalah pada penerapan fitur <i>speech</i>	(Yuwono, 2018)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
			<p><i>recognition</i></p> <p>setelah diuji coba masih belum cukup memuaskan karena 5 (lima) dokter yang menguji aplikasi tersebut tidak ada satupun dokter yang setuju bahwa tingkat akurasi dan kecepatan fitur <i>speech recognition</i> pada aplikasi sudah cukup baik</p>	
9	<i>Designing an Expert System Application to Diagnose Dengue Fever Patients Using the Naïve Bayes Method</i>	<i>Naïve Bayes; Expert System; Dengue Fever Patients</i>	<p>Penelitian ini membahas mengenai diagnosis penderita demam berdarah. Hasil dari penelitian ini adalah dengan menghasilkan akurasi tertinggi sebesar 97%</p>	(Sebayang, 2024)
10	<i>Prediction of Dengue using Machine Learning Algorithms:</i>	<i>SVM, XGBOOST, logistic regression, machine</i>	<p>Penelitian ini membahas mengenai membahas mengenai prediksi</p>	(Sarwar & Mamun, 2022)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
	<i>Case Study Dhaka</i>	<i>learning, dengue epidemic, dengue prediction</i>	demam berdarah. Hasil dari penelitian ini adalah SVM mendapatkan akurasi yang tinggi sebesar 97%	
11	<i>Machine learning algorithms to predict healthcare-seeking behaviors of mothers for acute respiratory infections and their determinants among children under five in sub-Saharan Africa</i>	<i>Acute respiratory infections, health-seeking behaviors, prediction, machine learning algorithms, sub-Saharan Africa</i>	Penelitian ini membahas mengenai prediksi perilaku pencarian layanan kesehatan para ibu untuk ISPA dan faktor-faktor penentunya di antara anak-anak balita di Afrika. Hasil dari penelitian ini adalah <i>Random forest</i> (RF) menghasilkan akurasi terbaik sebesar 88.89%, precision 89.5%, f-measure 83%, AUC ROC Curve sebesar 95.8% dan recall 77.6%	(Yehuala, et al., 2024)
12	<i>Analysis of risk factors associated with</i>	<i>Acute respiratory infections, Risk-</i>	Penelitian ini membahas mengenai analisis	(Nshimiymana & Zhou, 2022)

No.	Sub Tema	Keywords	Ulasan Kritis	Pustaka
	<i>acute respiratory infections among under-five children in Uganda</i>	<i>factors, Under-five mortality, Uganda</i>	faktor risiko terkait ISPA pada anak di Uganda. Hasil dari penelitian ini adalah RF menghasilkan akurasi yang tinggi sebesar 88.7% dengan nilai AUC sebesar 0.951	
13	<i>Development and validation of machine learning-driven prediction model for serious bacterial infection among febrile children in emergency departments</i>	<i>Febrile children, Machine learning, serious bacterial infection,</i>	Penelitian ini membahas mengenai prediksi <i>Serious Bacterial Infection</i> (SBI) pada anak-anak yang mengalami demam di Unit Gawat Darurat. Hasil dari penelitian ini adalah RF mendapatkan hasil tertinggi sebesar 0.964 untuk nilai AUROC	(Lee, Chung, Kang, Kim, & Kwak, 2021)

Berdasarkan penjelasan penelitian terkait, maka pada penelitian ini melakukan analisis anamnesis penyakit dengan gejala demam pada anak. Perbedaan penelitian yang akan dilakukan dengan penelitian yang telah dilakukan adalah pada penelitian yang akan

dilakukan akan mengimplementasikan NLP untuk anamnesis penyakit dengan gejala demam pada anak. Sementara pada penelitian yang telah dilakukan oleh keenam peneliti (Athota, Shukla, Pandey, & Rana, 2020), (Nugraheni, Widodo, & Sari, 2021), (De, Goswami, Faujidar, & Singh, 2024), (Manoharan & Sathish, 2020), (Almuhaimed, Alhomidi, & Alenezi, 2022) melakukan penerapan *cosine similarity* pada beberapa penyakit, dan penelitian yang telah dilakukan oleh kelima peneliti (Sarwar & Mamun, 2022), (Sebayang, 2024), (Yehuala, et al., 2024), (Nshimiyimana & Zhou, 2022), (Lee, Chung, Kang, Kim, & Kwak, 2021) melakukan klasifikasi penyakit. Pada penelitian (Ratnasari, Kusumadewi, & Rosita, 2014) penerapan NLP digunakan untuk data anamnesis yang terdapat kesalahan eja, tata bahasa buruk menjadi teks yang terstruktur. Pada penelitian (Yuwono, 2018) akan melakukan pengubahan suara ke teks dari keluhan pasien dengan data yang diperoleh dari alodokter.com, dokter.id, klikdokter.com, dan tanyadok.com, pengubahan suara ke teks dari keluhan pasien juga dilakukan oleh (Laksono, Hidayatullah, & Ratnasari, 2018) dengan data yang digunakan yaitu 10 pasien. Sementara pada penelitian ini akan mengenali gejala yang diderita pasien dengan data input *voice to text* terkait anamnesis penyakit pada anak, dengan data yang berasal dari data anamnesis penyakit pada anak bersumber dari data Klinik Margorejo. Selain itu, pada penelitian ini akan menggunakan TF-IDF dan *cosine similarity* dalam mengimplementasikan NLP untuk anamnesis penyakit pada anak.

2.2 Landasan Teori

2.2.1 Gejala Demam

Demam adalah proses alami tubuh untuk melawan infeksi yang masuk ke dalam tubuh ketika suhu meningkat melebihi suhu tubuh normal ($>37,5^{\circ}\text{C}$). Demam adalah proses alami tubuh untuk melawan infeksi yang masuk ke dalam tubuh. Demam terjadi pada suhu $> 37, 2^{\circ}\text{C}$, biasanya disebabkan oleh infeksi (bakteri, virus, jamur atau parasit), penyakit autoimun, keganasan, ataupun obat – obatan (Arafat, Simanjuntak, & Charibaldi, 2021).

1. ISPA

Infeksi Saluran Pernafasan Akut (ISPA) adalah kondisi kesehatan yang ditandai dengan terjadinya infeksi pada saluran pernafasan bagian atas. Infeksi ini menyebabkan gejala awal seperti batuk, pilek, dan demam. Penyakit ISPA sangat mudah menular dari satu orang ke orang lain, terutama pada anak-anak dan orang dewasa lanjut usia (Andres & Purnomo, 2024).

ISPA didefinisikan sebagai gangguan pada saluran pernafasan yang disebabkan oleh mikroorganisme penginfeksi yang menular dari individu ke individu lainnya. Beberapa faktor yang berpotensi memicu terjadinya ISPA adalah kondisi cuaca, tingginya tingkat polusi udara akibat asap kendaraan bermotor, serta adanya zat-zat pencemar udara dari sumber-sumber tertentu. Berdasarkan gejalanya, ISPA dibedakan menjadi dua jenis yakni ISPA ringan yang gejala-gejalanya ringan dan ISPA berat yang gejalanya lebih parah. Biasanya, penyakit ini disebabkan oleh virus atau bakteri yang menyerang hidung atau trakea (saluran pernapasan), mengakibatkan gangguan pada fungsi pernapasan (Andres & Purnomo, 2024).

Gejala ISPA pada anak seperti demam, batuk-batuk, hidung tersumbat/pilek, sakit kepala/pusing, sesak nafas, nafsu makan berkurang/susah makan, tulang dan persendian anggota badan terasa sakit, bintik merah, meriang dan menggigil, nyeri (Diva, Maulindar, Sumarlinda, & Binuko, 2024).

2. Demam *Tifoid*

Demam *tifoid* merupakan penyakit infeksi sistemik akut yang mengenai sistem retikuloendotelial, kelenjar limfe saluran cerna, dan kandung empedu. Disebabkan terutama oleh *Salmonella enterica* serovar typhi (*S. typhi*) dan menular melalui jalur fekal-oral (Arafat, Simanjuntak, & Charibaldi, 2021). Gejala pada demam tifoid meliputi demam, sakit kepala, tubuh menggigil, batuk, mual, muntah, sakit perut, kehilangan selera makan, diare (Lusiyanti, Yetri, Setiawan, Affandi, & Laia, 2023).

3. Demam Berdarah

Demam Berdarah *Dengue* (DBD) didefinisikan sebagai suatu penyakit yang terjadi akibat gigitan nyamuk *aedes aegypti*. Penyakit jenis ini sangat mudah sekali dijumpai di negara yang memiliki iklim tropis. Kasus penderita DBD biasanya mengalami peningkatan pada saat musim hujan, hal ini diakibatkan oleh meningkatnya aktivitas nyamuk dalam menggigit (Dewi & Sudaryanto, 2020). Gejala demam berdarah pada anak seperti demam, tidak nafsu makan, mual, muntah, lemas/lesu, tubuh menggigil, badan lemah, nyeri otot seluruh tubuh (malgysa), sakit perut, nyeri perut (Azizah & Rohman, 2023).

4. *Urinary Tract Infection*

Infeksi saluran kemih (ISK) adalah suatu penyakit yang dimana terbentuknya mikroorganisme dalam jumlah banyak di saluran kemih. ISK merupakan kondisi dimana terdapat mikroorganisme dalam urin yang jumlahnya sangat banyak dan

mampu menimbulkan infeksi pada saluran kemih. Gejala yang dapat dirasakan pasien ISK, yaitu mual, muntah, demam, disuria, dan cairan urine berwarna keruh, perut sakit (Arafat, Simanjuntak, & Charibaldi, 2021), (Sundari, Sarmidi, & Nuraropah, 2021).

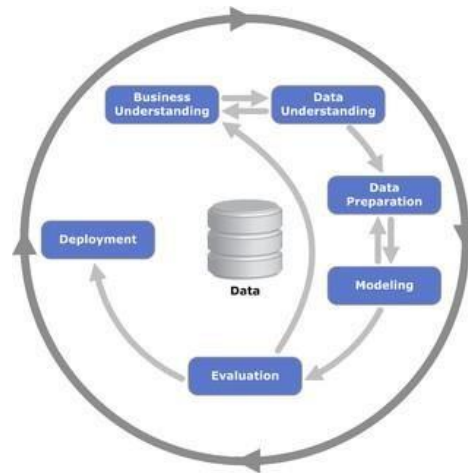
2.2.2 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) adalah sub-bidang kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. Tujuan dari teknologi ini adalah untuk membuat mesin mampu membaca dan menalar dengan bahasa manusia dan secara otomatis memprosesnya (Sihombing, 2022). Kajian NLP mencakup segmentasi tuturan (*speech segmentation*), penandaan kelas kata (*part-of-speech tagging*), segmentasi teks (*text segmentation*), penentuan makna (*word sense disambiguation*) (Rumapea, 2021).

NLP mulai diperkenalkan sekitar tahun 1970an dengan fungsi awal sebagai perangkat lunak dalam pemahaman bahasa domain block berlanjut pada decade 80an dengan tambahan pengembangan dengan berbasis simbol dan ejaan. NLP sendiri lebih dari sekedar mengidentifikasi sekaligus untuk mengekstraksi hubungan isi antar dokumen, serta mengambil dan mendapatkan data dari dokumen dan sebagainya. Sederhananya, NLP memberikan akses dan sarana untuk analisis dari sebuah teks atau tulisan dengan membuat sebuah struktur yang baik dan teks yang sebelumnya belum terstruktur untuk dianalisis berikutnya (Rahmadani & Tasrif, 2024).

2.2.3 *Metode CRISP-DM*

CRISP-DM adalah metodologi yang terstruktur dan teruji waktu yang digunakan dalam proses pengembangan proyek data mining, termasuk dalam konteks *Machine Learning*. Meskipun CRISP-DM dirancang khusus untuk data mining, namun prinsip dan tahapannya disajikan di Gambar 2.1 dapat dengan mudah diterapkan dalam konteks *Machine Learning* (Sholikhah, Harjanta, & Latifah, 2023).



Gambar 2.1 Tahapan CRISP-DM.

Dari Gambar 2.1 dapat dijelaskan tahapan-tahapan CRISP-DM adalah sebagai berikut (Dhewayani, Amelia, Alifah, Sari, & Jajuli, 2022):

1. *Business Understanding*

Business Understanding adalah proses menentukan tujuan bisnis, memahami situasi dan kondisi pada saat penelitian dan menetapkan sebuah tujuan dari penelitian yang dilakukan ke dalam permasalahan yang diselesaikan dengan data mining.

2. *Data Understanding*

Data understanding adalah tahap persiapan, melakukan pengecekan terhadap data yang digunakan, mengumpulkan data awal serta melakukan identifikasi pada kualitas data. Dalam *data understanding*, data yang digunakan akan melalui proses deskripsi dari setiap fiturnya.

3. *Data Preparation*

Data preparation merupakan proses yang dilakukan setelah data telah dikumpulkan. Pada tahap ini, data akan melalui proses identifikasi, pemilihan data, pembersihan data dan transformasi data.

4. *Modelling*

Modeling merupakan tahap implementasi algoritma yang akan digunakan untuk melakukan pencarian, identifikasi, serta menghasilkan pola yang akan digunakan pada data penelitian.

5. *Evaluation*

Evaluation adalah suatu proses untuk melakukan pengukuran hasil evaluasi dari model yang telah diimplementasikan sebelumnya di tahap modelling. Hasil evaluasi

tersebut menggambarkan proses dari data mining yang telah dilakukan dan mengukur model yang paling baik untuk digunakan.

6. *Deployment*

Deployment atau penyebaran merupakan proses menggunakan model yang dihasilkan sebelumnya. Dalam tahap ini terdapat 2 jenis kegiatan penyebaran yaitu melakukan perencanaan dan pemantauan dari penyebaran hasil yang dilakukan. Jenis kegiatan selanjutnya menyelesaikan tugas penutup dengan membuat laporan akhir dan melakukan tinjauan proyek. Kedua kegiatan ini dapat dilakukan semua atau menyelesaikan salah satunya.

2.2.4 *Text Processing*

Preprocessing text berguna untuk mengolah data yang digunakan dalam bentuk yang sesuai untuk proses identifikasi ekstraksi informasi (Haumahu, Permana, & Yaddarabullah, 2021). Berikut merupakan beberapa tahapan dalam *text processing* yang akan digunakan dalam penelitian ini (Hikmah, Ariyanti, & Pratama, 2022):

1. *Case Folding*

Case Folding merupakan tahapan yang bertujuan untuk mengubah semua respon menjadi huruf kecil, pada proses ini, karakter 'A'-'Z' yang terdapat pada data diubah menjadi karakter 'a'-'z'. Sementara itu, karakter lain yang bukan huruf dan angka, seperti tanda baca dan spasi, dianggap sebagai pembatas (Razzaq, et al., 2022).

2. *Tokenization*

Tokenizing adalah proses pengolahan data dimana tanda baca akan dihilangkan untuk menghasilkan kalimat/kata yang berdiri sendiri, entitas yang dapat disebut sebagai *token* antara lain kata, angka, simbol, tanda baca, dll. Tahapan ini bertujuan untuk memecah respon menjadi unit-unit kata (Al-Harigy, Al-Nuaim, Moradpoor, & Tan, 2022).

3. *Filtering*

Selama tahap *filtering*, ada Langkah yang disebut penghapusan *stopwords*. *Stopwords* merupakan langkah yang bertujuan untuk menghilangkan kata-kata umum yang sering muncul dalam jumlah banyak dan tidak memiliki arti dengan menggunakan algoritma *stoplist* (menghilangkan kata-kata yang kurang penting) atau *wordlist* (menyimpan kata-kata yang penting), contohnya pada penelitian ini adalah penggunaan kata penghubung seperti 'dan', 'yang', 'serta', 'setelah', dan lain-

lain. Menghapus *stopword* ini dapat mengurangi ukuran indeks dan waktu pemrosesan. Selain itu, hal ini juga dapat mengurangi tingkat kebisingan (Al-Harigy, Al-Nuaim, Moradpoor, & Tan, 2022).

4. *Stemming*

Stemming adalah tahap dimana setiap kata akan diubah dari imbuhan menjadi kata dasar, tahap ini diperlukan untuk mengurangi jumlah indeks yang berbeda dari satu data sehingga kata yang mendapat imbuhan atau awalan akan kembali ke bentuk dasarnya. Selain itu juga untuk mengelompokkan kata-kata lain yang memiliki kata dasar dan makna yang sama namun memiliki bentuk yang berbeda karena mendapatkan imbuhan yang berbeda (Al-Harigy, Al-Nuaim, Moradpoor, & Tan, 2022).

2.2.5 *Term Frequency-Inverse Document Frequency (TF-IDF)*

TF-IDF merupakan kombinasi dari *term frequency* (TF) dan *inverse document frequency* (IDF). Karena nilai asli frekuensi term dalam dokumen digunakan secara langsung, maka representasi TF merupakan salah satu TWS yang paling sederhana. TF didasarkan pada asumsi bahwa term dengan nilai frekuensi term yang lebih tinggi dianggap lebih penting dibandingkan dengan term dengan nilai frekuensi term yang lebih rendah, dan hanya bergantung pada jumlah kemunculan dari suatu dokumen tertentu. Oleh karena itu, kapasitas TF untuk membedakan semua dokumen yang relevan dengan dokumen lain yang tidak relevan sangat rendah karena ketidaktahuannya terhadap frekuensi koleksi (Jiang, et al., 2021).

Untuk mengatasi masalah ini, *inverse document frequency* (IDF) diusulkan dengan mempertimbangkan frekuensi koleksi yang meningkatkan kapasitas diskriminatif dari suatu term untuk klasifikasi teks. IDF merupakan pengembangan dari frekuensi dokumen (DF) yang berarti jumlah dokumen yang muncul. Hal ini diusulkan berdasarkan asumsi bahwa term yang muncul di beberapa dokumen dianggap lebih penting daripada term yang muncul di lebih banyak dokumen (Jiang, et al., 2021). Berikut pada persamaan (2.1) dan (2.2) merupakan rumus TF-IDF (Maulidina, 2020).

$$TF - IDF = tf_t * IDF_t \quad (2.1)$$

$$IDF = \log \left(\frac{D}{DF} \right) \quad (2.2)$$

Keterangan:

tf_t = *Term Frequency*

IDF = Inverse Document Frequency

D = Total Dokumen

DF = Total kemunculan kata pada dari seluruh dokumen

Pada Sklearn, formula IDF yang digunakan agak berbeda. Formula IDF pada sklearn dapat dilihat pada (2.3).

$$IDF = \ln \left(\frac{1 + D}{1 + DF} \right) + 1 \quad (2.3)$$

Selanjutnya pada sklearn, nilai TF-IDF juga akan dilakukan normalisasi L2 yang memiliki formula dapat dilihat pada (2.4).

$$L2 - Norm = \sqrt{\frac{TF - IDF_1^2 + TF - IDF_2^2 + \dots + TF - IDF_i^2}{2}} \quad (2.4)$$

Sehingga hasil dari TF-IDF L2 Norm adalah seperti pada (2.5).

$$TF - IDF \text{ L2 - Norm} = \frac{TF - IDF \text{ sebelum Norm}}{L2 - Norm} \quad (2.5)$$

2.2.6 Cosine Similarity

Similarity atau relevansi dihitung antara kueri dan dokumen. Skor *similarity* diperoleh dengan mengukur skor *similarity* antara dua vektor, yaitu vektor kueri dan vektor dokumen. Semakin besar nilai relevansi, semakin mirip/relevan kueri dan dokumen tersebut. Ukuran similarity digunakan untuk menentukan tingkat similarity antara titik data. Selain itu, ukuran ini digunakan untuk mengukur kesamaan antara himpunan berdasarkan irisan kedua himpunan tersebut. Selain itu, ukuran kesamaan diakui sebagai fungsi yang mengevaluasi tingkat kemiripan antara sepasang objek teks. Secara ringkas, tingkat similarity mencerminkan intensitas hubungan antara dua titik data. *Cosine similarity* adalah ukuran kesamaan yang sering digunakan. Ini digunakan sebagai dasar untuk menghitung nilai relevansi antara kueri dan dokumen, serta nilai relevansi antara dokumen yang berbeda. *Cosine Similarity* ditentukan sebagai kosinus sudut θ yang terbentuk antara vektor-vektor. Formula *Cosine Similarity* memiliki rumus seperti pada (2.6) berikut (Rasyid & Ningsih, 2024):

$$Sim(A, B) = \frac{\sum n(A) * n(B)}{\sqrt{\sum n(A)^2} * \sqrt{\sum n(B)^2}} \quad (2.6)$$

Keterangan:

A = Item A (contoh D1)

B = Item B (contoh D2)

2.2.7 *Library SpeechRecognition*

Speech recognition merupakan kemampuan mesin untuk mendengarkan kata – kata yang diucapkan dan mengidentifikasinya. Kemampuan untuk dapat mengubah suara yang masuk pada komputer ke dalam bentuk teks. Platform python dapat digunakan untuk mengubah kata – kata yang diucapkan menjadi teks, membuat kueri, atau memberikan balasan. Kata – kata yang terucap lalu diubah bentuknya ke dalam sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode – kode tertentu untuk mengidentifikasikan kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan ditampilkan dalam bentuk tulisan atau dapat dibaca oleh perangkat teknologi sebagai sebuah perintah untuk melakukan suatu pekerjaan (Adnan, Amelia, & Shiddiq, 2022).

Library SpeechRecognition digunakan untuk melakukan pengenalan suara, dengan dukungan untuk beberapa mesin dan API, daring dan luring (SpeechRecognition, 2022). Didalam *SpeechRecognition* terdapat *Recognizer class*. Tujuan utama dari *instance Recognizer* untuk mengenali ucapan. Setiap *instance* memiliki berbagai pengaturan dan fungsionalitas untuk mengenali ucapan dari sumber audio. Setiap *instance Recognizer* memiliki tujuh metode untuk mengenali ucapan dari sumber audio menggunakan berbagai API (Pujoseno, 2020). *SpeechRecognition* dikenal sebagai *SpeechRecognition* atau *speech_recognition*, adalah *library* Python yang menyediakan antarmuka yang sederhana dan mudah digunakan untuk bekerja dengan berbagai mesin ASR, seperti Google Web Speech API, CMU Sphinx, dan banyak lagi. *Library* ini mendukung berbagai sumber masukan audio dan merupakan pilihan yang sangat baik untuk menggunakannya (Pathak, 2023). *SpeechRecognition* yang digunakan dalam penelitian ini dikembangkan oleh (Zhang, 2017).

2.2.8 *Blackbox Testing*

Blackbox testing adalah tahap untuk pengujian terhadap suksesnya program yang sudah diciptakan. Tahap ini penting dilakukan supaya tidak terjadi kesalahan alur pada program. *Blackbox testing* ini pengujiannya lebih pada sisi eksternal suatu aplikasi untuk memudahkan pengguna dan juga bukan menguji program dari kode sumber. *Blackbox testing* aktivitasnya berfokus pada informasi domain. *Blackbox testing* yaitu menguji perangkat lunak dari segi perincian fungsional (Rohi, Pote, & Talakua, 2022).

Tujuan *blackbox testing* yaitu untuk mencari adanya kesalahan-kesalahan yang terjadi dalam kategori fungsi-fungsi, antarmuka sistem (*user interface*), pengaksesan basis data (*database*) eksternal sistem dan pada struktur data, kemudian kesalahan kinerja (*performa*), kesalahan pada proses inisialisasi (*initialization*) dan kesalahan pada proses keluar program (*terminasi*) (Susanti & Bismin, 2021).

2.2.9 User Acceptance Test (UAT)

UAT (*User Acceptance Test*) yaitu pengujian aplikasi terhadap pengguna yang dilakukan sehubungan dengan kebutuhan pengguna terakhir atau *end user* (Wulandari, Nofiyani, & Hasugian, 2023). UAT yang dilakukan untuk menguji sistem yang secara keseluruhan telah dibangun. UAT bertujuan untuk memastikan sistem telah memenuhi kebutuhan pengguna, tidak hanya dari sisi spesifikasi sistem. Pengujian UAT dilakukan untuk memastikan bahwa solusi yang disediakan oleh sistem dapat bekerja dengan baik bagi pengguna. Pengujian ini akan dilakukan oleh pengguna secara langsung (Tandirilambun, Rini, Filiana, & Virginia, 2024).

Saat pengujian, responden akan mengisi kuesioner bersisi daftar pertanyaan terkait sistem dengan lima opsi jawaban. Bobot jawaban yang digunakan dapat dilihat pada Tabel 2.2.

Tabel 2.2 Bobot Jawaban

Kode	Jawaban	Bobot
A	Sangat Setuju	5
B	Setuju	4
C	Netral	3
D	Tidak Setuju	2
E	Sangat Tidak Setuju	1

Sumber: (Jimmy & Suwitno, 2023)

Dalam melihat nilai UAT tersebut, lebih singkatnya dijelaskan dengan rumus berikut (Mariyani, Setiyaningsih, & Agustina, 2022):

$$\text{Nilai UAT} = \left(\frac{\text{Jumlah jawaban} \times \text{Bobot Jawaban}}{\text{Jumlah responden}} \right) \quad (2.7)$$

$$X = \text{Skor tertinggi} \times (\text{jumlah pertanyaan} \times \text{jumlah responden}) \quad (2.8)$$

$$Y = \text{Skor terendah} \times (\text{jumlah pertanyaan} \times \text{jumlah responden}) \quad (2.9)$$

Setelah mendapatkan nilai UAT, langkah berikutnya adalah menghitung nilai persentase dari manfaat sistem yang dibangun. Ini dilakukan dengan membagi nilai UAT dengan skor ideal. Selanjutnya, dikalikan dengan 100% (Mariyani, Setiyaningsih, & Agustina, 2022).

$$\text{Persentase UAT} = \frac{\text{Total skor}}{x} \times 100\% \quad (2.10)$$

Keterangan:

Total skor = Jumlah skor keseluruhan jawaban responden

x = Skor tertinggi $\times$ (jumlah pertanyaan $\times$ jumlah responden)

Dari hasil persentase UAT dapat dilihat kategori kelayakannya yang dapat dilihat pada Tabel 2.3.

Tabel 2.3 Kategori Kelayakan

No	Skor	Kategori Kelayakan
1	< 21%	Sangat Tidak Layak
2	21 – 40%	Tidak Layak
3	41 – 60%	Cukup Tidak Layak
4	60 – 80%	Layak
5	81 – 100%	Sangat Layak

Sumber: (Anam, et al., 2023)

BAB 3

Metodologi

3.1 Pendekatan Penelitian

Pendekatan penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Penelitian kuantitatif bertujuan untuk mengidentifikasi masalah yang ditetapkan sebagai tujuan penelitian. Sedangkan, pendekatan deskriptif merupakan suatu cara untuk mengetahui eksistensi dari variabel bebas (Riany & Testiana, 2023).

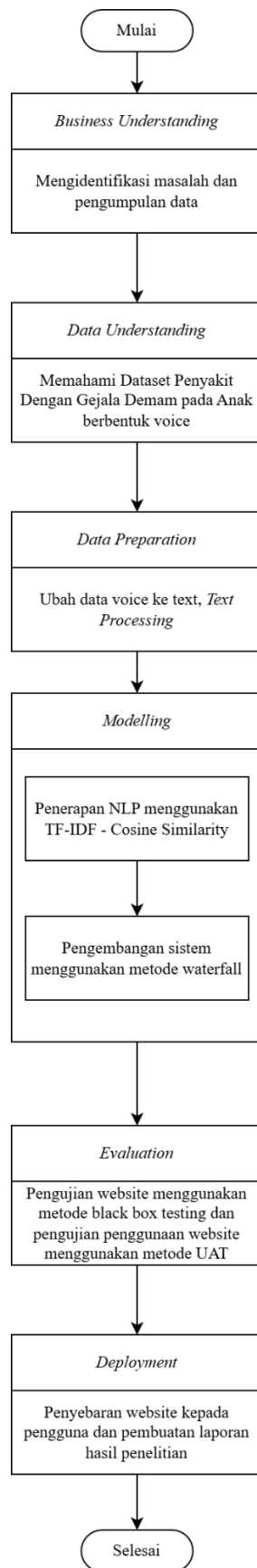
Penelitian ini menggunakan metode kuantitatif yaitu TF-IDF - *cosine similarity* untuk menghitung seberapa sering kata muncul di dalam kalimat dan menghitung kedekatan antar kalimat. Kemudian, melakukan deskripsi terhadap hasil *cosine similarity* dalam menghitung kedekatan antar kalimat.

3.2 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan selama 6 bulan dimulai pada bulan Juli hingga Desember 2024 yang bertempat pada Klinik Margorejo untuk mengumpulkan data-data terkait penyakit pada anak.

3.3 Tahapan Penelitian

Tahapan penelitian menjelaskan langkah-langkah penelitian yang dilakukan dalam identifikasi ekstraksi informasi penyakit dengan gejala demam dapat dilihat pada Gambar 3.1. Pada penelitian NLP untuk Anamnesis Penyakit pada Anak yang dilakukan, peneliti menggunakan CRISP-DM sebagai kerangka kerja yang terstruktur untuk mengembangkan, mengimplementasikan, dan mengevaluasi model NLP yang digunakan untuk menganalisis teks medis atau anamnesis pasien (Zia, et al., 2022).



Gambar 3.1 Tahapan Penelitian

3.3.1 Business Understanding

Business understanding melakukan pemahaman tujuan berdasarkan pada perspektif bisnis kemudian diubah ke dalam definisi masalah *machine learning* hingga dilakukan penentuan solusi yang akan diusulkan untuk menangani permasalahan yang ada. Dalam penelitian ini tujuan penelitian yaitu mengimplementasikan NLP untuk anamnesis penyakit dengan gejala demam pada anak menggunakan algoritma TF IDF - *cosine similarity* dalam sebuah *website*.

Selanjutnya, dilakukan pengumpulan data yang terdiri dari data primer dan data sekunder. Data sekunder merupakan data terkait penelitian terdahulu, teori-teori yang digunakan sebagai bahan referensi dalam penelitian, seperti teori gejala demam, NLP, CRISP-DM, *text processing*, TF-IDF, *cosine similarity*, *flowchart*, UML, bahasa pemrograman PHP, *framework laravel*, *database*, MySQL, *blackbox testing*, *user acceptance test* (UAT). Sedangkan, data primer merupakan dataset anamnesis yang digunakan untuk ekstraksi informasi dari penyakit dalam gejala demam yang bersumber dari Klinik Margorejo berbentuk *voice*.

3.3.2 Data Understanding

Data *understanding* merupakan tahap untuk memahami data yang digunakan. Dalam penelitian ini menggunakan dataset yang bersumber dari data Klinik Margorejo, nantinya data tersebut akan berbentuk *voice* (.wav). Dataset tersebut terdiri dari 200 data balita bergejala demam dan tidak menggunakan data uji karena pada penelitian ini akan menerapkan NLP dengan menggunakan *cosine similarity*. NLP dapat digolongkan kedalam *unsupervised learning* (Perdana, Pricillia, & Zulfachmi, 2021). Selain itu dalam *unsupervised learning* tidak diperlukan data pelatihan dan data pengujian, dan hal tersebut dapat diterapkan pada pendekatan seperti *similarity-based* (*cosine similarity*) dan *zero-shot learning* (Schopf, 2023) (Arnita, Marpaung, Aulia, Suryani, & Nabila, 2022). Data anamnesis yang ditanyakan dokter kepada pasien akan mencakup “*The Sacred Seven of history taking*”. *The Sacred Seven of history taking* merupakan teknik memberikan informasi yang paling krusial dalam menentukan diagnosis dan pengobatan pasien. Berdasarkan *The Sacred Seven of history taking*, seorang dokter dapat menggunakan beberapa parameter untuk anamnesis yang dibagi ke dalam tujuh kategori seperti *chronology*, *bodily location*, *quality*, *quantity*, *setting*, *aggravating or alleviating factors*, *associated manifestations*.

Berikut adalah penjelasan masing-masing kategori tersebut beserta bagaimana gejala-gejala yang ada dikelompokkan:

1. *Chronology*, deskripsi dari kronologis tersebut menjelaskan tentang perjalanan penyakit, dengan itu dokter harus mendapatkan laporan kronologis dengan menanyakan kapan sakit pertama kali dirasakan, dan dilanjutkan dengan pertanyaan lainnya secara spesifik. Kronologi juga mencakup durasi simptomatik, periodisitas dan apakah gejala menjadi lebih baik atau lebih buruk dari waktu ke waktu.
2. *Bodily location*, Lokasi sakitnya harus didefinisikan seakurat mungkin dengan cara menunjukkan Lokasi nyeri yang dirasakan oleh pasien dengan menggunakan Gerakan tangannya. Penting bagi dokter mengingat bahwa pasien mungkin memiliki lebih dari satu penyakit dan rasa sakit itu dapat mengindikasikan berbagai proses penyakit. Untuk itu, dokter dapat meminta pasien agar dapat membedakan dan mengkarakterisasikan masing-masing penyakitnya.
3. *Quality*, Sebagian besar pasien menggambarkan kualitas dari rasa sakitnya menggunakan analogi. Beberapa pasien juga menggunakan Bahasa yang deskriptif atau emosional seperti, “Rasanya seperti seseorang menikam saya dengan pisau”, dan pasien sulit untuk menemukan bahasa yang sedang dideskripsikannya.
4. *Quantity*, intensitas nyeri yang dirasakan pasien dapat diperkirakan dengan menggunakan skala 1 (satu) sampai 10 (sepuluh) atau dapat dibandingkan dengan nyeri yang lainnya. Selain itu, volume juga termasuk dalam salah satu contoh kuantitas. Misalnya jumlah dahak yang dikeluarkan dalam sehari.
5. *Setting*, pengaturan ini menjelaskan tentang gejala, di mana, apa, dan dengan siapa pasien pada saat merasakan sakit itu. Hal ini merupakan pertanyaan yang bagus untuk digunakan di awal wawancara.
6. *Aggravating or alleviating factors*, dokter mengumpulkan data mengenai hal-hal apa saja yang telah dilakukan oleh pasien. Misalnya, pasien telah meminum obat sebelumnya. Hasil dari data tersebut, dokter dapat menganalisis apa yang telah membuat gejala lebih buruk atau menjadi lebih baik.
7. *Associates manifestations*, gejala jarang terjadi dengan sendirinya, maka dari itu dokter harus mendengarkan gejala-gejala lainnya yang dapat memberikan informasi diagnostic tentang patologi dan organ yang terlibat.

Dengan menggunakan “*The Sacred Seven of history taking*”, dokter dapat mengklasifikasikan gejala berdasarkan tujuh kategori tersebut. Berikut adalah contoh

bagaimana teknik “*The Sacred Seven of history taking*” diterapkan dalam wawancara medis dengan pasien:

1. *Chronology*

Pertanyaan : Kapan keluhan ini pertama kali muncul, dan bagaimana perkembangannya?

Jawaban : Demam mulai sekitar 4 hari yang lalu. Awalnya demamnya tidak terlalu tinggi, tapi sekarang lebih parah. Lemas juga mulai terasa sekitar 2 hari yang lalu, dan mual serta muntah muncul sehari setelah demam dimulai. Setiap kali saya makan atau minum, saya langsung muntah.

2. *Bodily Location*

Pertanyaan : Di bagian tubuh mana Anda merasakan keluhan, atau apakah keluhan ini terasa di seluruh tubuh?

Jawaban : Demam terasa di seluruh tubuh, saya juga merasa lemas secara umum. Badan saya terasa dingin, terutama di tangan dan kaki

3. *Quality*

Pertanyaan : Bagaimana Anda menggambarkan keluhan atau gejala ini?

Jawaban : Saya merasa sangat lemas, seperti tidak ada tenaga sama sekali. Mual datang setiap kali saya makan atau minum, dan saya pasti muntah setelahnya

4. *Quantity*

Pertanyaan : Seberapa parah gejala-gejala ini? Apakah Anda bisa menilai keparahannya dari skala 1 sampai 10?

Jawaban : Saya akan memberikan nilai 8 dari 10 untuk rasa lemas, karena saya tidak bisa melakukan aktivitas apapun. Mual dan muntah sangat parah, setiap makan atau minum saya pasti muntah

5. *Setting*

Pertanyaan : Apa yang Anda lakukan ketika gejala ini pertama kali muncul?

Jawaban : Saya sedang menjalani aktivitas biasa ketika demam mulai muncul. Saya tidak tahu apa penyebabnya, tapi sejak itu saya merasa semakin lemas dan setiap kali makan atau minum, saya langsung muntah

6. *Aggravating or Alleviating factors*

Pertanyaan : Apakah ada hal-hal yang membuat gejala ini semakin parah atau lebih baik?

Jawaban : Setiap kali saya makan atau minum, mual dan muntah langsung muncul, jadi itu memperburuk kondisi saya. Istirahat sedikit membantu meringankan lemas, tapi saya tetap merasa lemah dan demam tidak hilang

7. *Associated Manifestations*

Pertanyaan : Apakah ada gejala lain yang menyertai keluhan utama, seperti batuk, sakit tenggorokan, atau gejala lainnya?

Jawaban : Selain demam, lemas, dan muntah, saya merasa badan saya menggigil, terutama di malam hari. Saya tidak batuk atau sakit tenggorokan

Berdasarkan contoh bagaimana teknik “*The Sacred Seven of history taking*” diterapkan dalam wawancara medis dengan pasien dihasilkan bahwa pada *chronology* gejala yang dialami meliputi demam, lemas, mual, muntah. *Bodily location*, demam dan lemas terasa di seluruh tubuh, *quality* yaitu lemas yang digambarkan dengan tidak ada tenaga sama sekali, serta mual terjadi setiap kali makan atau minum, dan muntah selalui menyertai setelahnya. *Quantity*, lemas dinilai 8 dari 10 karena pasien tidak beraktivitas. *Setting*, gejala pertama kali muncul saat pasien sedang menjalani aktivitas biasa dan tidak ada factor eksternal yang langsung dikenali sebagai penyebab. *Aggravating or alleviating factors*, makan atau minum memperburuk mual dan muntah. *Associated manifestations*, badan menggigil terutama di malam hari.

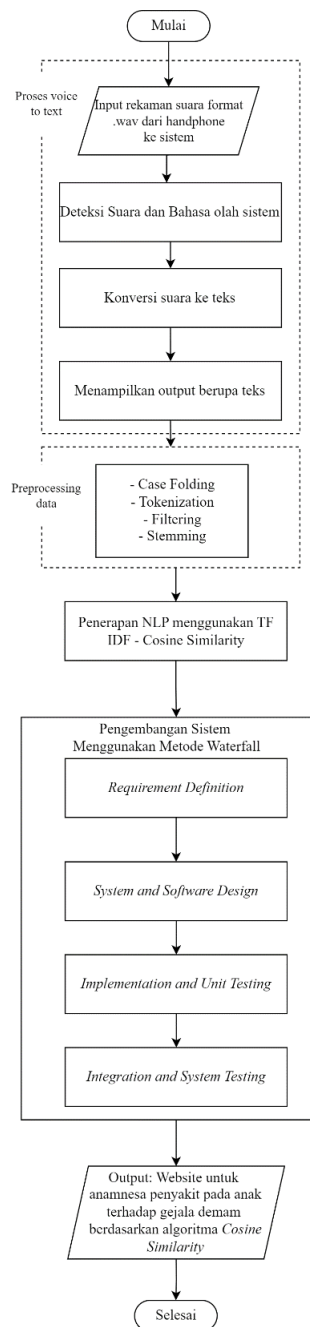
Pada penelitian ini gejala yang digunakan dalam mendiagnosis anamnesis terdiri dari Demam Intermittent, Demam Menggigil, Lama Demam, Sakit Kepala, Mual, Muntah, Nyeri Sendi, Nyeri Perut, Nyeri Ulu Hati, Bintik Merah, Lidah Kotor, Kejang, Batuk, Pilek, Sesak Nafas, Sulit BAB, BAB Cair, BAK Kemerahan, BAK Sakit, Nafsu Makan Turun, Lemas dan Mimisan. Gejala-gejala tersebut mencakup kondisi pasien yang dapat membantu dokter mengenali pola penyakit dan kemungkinan penyebabnya. Dengan mengumpulkan informasi dari berbagai gejala, dapat menjadi alat yang bermanfaat untuk mempersempit diagnosis dan menentukan langkah penanganan yang tepat. Untuk gejala yang digunakan dilakukan pengelompokkan berdasarkan “*The Sacred Seven of History Taking*” yang dapat dilihat pada Tabel 3.1.

3.3.3 Data Preparation

Data *preparation* merupakan proses melakukan persiapan data dengan menyesuaikan dataset agar dapat sesuai dengan kebutuhan pada saat tahap pemodelan. Dataset yang digunakan berbentuk *voice* (.wav) dari hasil anamnesis pasien dan dokter. Data tersebut dilakukan deteksi suara dan bahasa oleh sistem yang nantinya akan dikonversi dari suara ke teks, dan hasil akhir akan menampilkan *output* (keluaran) berupa teks. Setelah data berbentuk teks, selanjutnya dilakukan text *pre-processing*. *Preprocessing* yang digunakan dalam penelitian ini meliputi tahap *case folding* untuk mentransformasi kata yang memiliki huruf besar menjadi huruf kecil. Tahap *tokenize* untuk membagi atau memecah kata yang awalnya dalam bentuk kalimat utuh menjadi penggalan kata perkata. Tahap *filtering*, yang bertujuan untuk menghilangkan kata-kata umum yang sering muncul dalam jumlah banyak. Tahap *stemming* adalah tahap dimana setiap kata akan diubah dari imbuhan menjadi kata dasar, tahap ini diperlukan untuk mengurangi jumlah indeks yang berbeda dari satu data sehingga kata yang mendapat imbuhan atau awalan akan kembali ke bentuk dasarnya.

3.3.4 Modelling

Pada tahap *modelling* merupakan proses pembentukan model untuk anamnesis penyakit pada anak menggunakan TF-IDF dan *Cosine Similarity*. Model NLP akan diterapkan pada sistem untuk mengolah data teks yang ada dan tahap input sampai menjadi bentuk *output* seperti yang ditunjukkan oleh Gambar 3.2.



Gambar 3.2 Tahapan *Modeling*

Pada Gambar 3.2, di dalam tahapannya juga melakukan perhitungan menggunakan persamaan rumus (2.1) untuk proses TF-IDF. Setelah melakukan proses TF-IDF, dilakukan perhitungan *cosine similarity* dengan menggunakan persamaan rumus (2.2) untuk mendapatkan skor untuk setiap dokumen, yang akan digunakan untuk menentukan peringkat kesamaan dokumen dengan kueri masukan. Tujuan dari penerapan NLP pada anamnesis yaitu untuk memahami keluhan pasien dari data *voice*, sehingga dengan dilakukannya proses NLP, akan menghasilkan diagnosis yang tepat. Berikut merupakan penjabaran detail dari Gambar 3.2.

1. Input Rekaman Suara Format .wav dari Handphone ke Sistem

Tahap awal adalah *input* data yang akan dilakukan proses *similarity*. Data tersebut berupa *voice* dengan format .wav yang akan diinput dari *handphone* ke sistem. Prosesnya data tersebut akan dilakukan deteksi suara dan bahasa oleh sistem menggunakan *library speechRecognition*. *SpeechRecognition* yang digunakan dalam penelitian ini dikembangkan oleh (Zhang, 2017). Selanjutnya dilakukan pembuatan objek Recognizer dari *speechRecognition* untuk menangani pengenalan suara. Kemudian membaca file audio (.wav) dengan menentukan file .wav yang ingin di konversi, pada proses tersebut menggunakan *sr.AudioFile()* untuk membuka file audio. Kemudian proses konversi dilakukan dengan menggunakan *r.recognize_google(audio)* yang menggunakan API Google Speech Recognition untuk mengonversi suara menjadi teks. Selanjutnya sistem akan menampilkan output berupa teks yang akan digunakan untuk proses *text preprocessing*. Dari data yang telah dikonversi menjadi teks, pada Tabel 3.1 yang merupakan gejala-gejala yang dijabarkan dalam *The Sacred Seven of History Taking*. Evaluasi terhadap isian The Sacred Seven of History Taking pada Tabel 3.1 berdasarkan hasil konsensus dokter yang berpraktik di Klinik Margorejo. Hasil pengujian yang telah divalidasi oleh dokter menunjukkan bahwa dari 30 data anamnesis yang diuji, sebanyak 28 data dinyatakan benar, dengan tingkat ketepatan sebesar 93,33%.

Tabel 3.1 Keterangan “*The Sacred Seven of History Taking*”

“<i>The Sacred Seven of History Taking</i>”	Gejala
<i>Chronology (Kronologi)</i>	<ul style="list-style-type: none"> - Demam tinggi >39°C - Demam >3 hari - Demam naik turun - Demam memburuk malam hari - Demam dengan kejang - Demam sulit turun - Demam disertai rewel - Demam setelah imunisasi - Demam setelah bepergian - Demam dengan keringat malam - Kejang saat demam tinggi - Kejang berulang dalam 24 jam - Kejang berlangsung >5 menit - Demam muncul tiba-tiba (DBD, ISPA) - Demam lebih dari 5 hari (tifoid, DBD) - Demam turun sebentar lalu naik lagi (tifoid)
<i>Bodily Location (Lokasi Gejala)</i>	<ul style="list-style-type: none"> - Sakit kepala - Sakit tenggorokan

	<ul style="list-style-type: none"> - Nyeri perut - Nyeri ulu hati - Nyeri dada saat batuk - Nyeri sendi - Ruam merah kulit - Mata merah - Gusi bengkak - Nyeri saat BAK - Kejang seluruh tubuh - Kejang hanya satu sisi tubuh - Kejang dengan mata melotot - Nyeri punggung bawah (indikasi infeksi ginjal/UTI) - Nyeri di belakang mata (DBD) - Nyeri saat menelan (ISPA, radang tenggorokan)
Quality (Kualitas)	<ul style="list-style-type: none"> - Batuk kering - Batuk berdahak - Batuk menggonggong - Batuk lama >2 minggu - Hidung tersumbat - Pilek lendir kuning - Muntah berulang - Napas berbunyi (mengi) - Urine berbau tajam - BAB cair berlendir - Kejang kaku seluruh tubuh - Kejang dengan tatapan kosong - Kejang dengan tubuh tersentak-sentak - Batuk berdarah (infeksi paru berat, ISPA parah) - Tinja berdarah atau berlendir (tifoid, disentri) - Urine keruh atau berwarna merah muda (UTI, ginjal)
Quantity (Kuantitas)	<ul style="list-style-type: none"> - Muntah >3 kali sehari - Diare >3 kali sehari - Pipis sedikit-sedikit - Tidak pipis >6 jam - Batuk makin sering malam hari - Demam terus naik - Berat badan turun cepat - Anak sering bangun malam - Sering menangis rewel - Air liur berlebihan - Kejang lebih dari 1 kali dalam sehari - Kejang berlangsung >5 menit - Kejang setiap kali anak demam tinggi - Frekuensi buang air kecil meningkat (UTI) - Jumlah urine berkurang drastis (dehidrasi berat, gagal ginjal) - Diare lebih dari 5 kali dalam sehari (infeksi pencernaan, tifoid)
Setting (Kondisi yang Memicu)	<ul style="list-style-type: none"> - Batuk memburuk malam hari

	<ul style="list-style-type: none"> - Demam setelah aktivitas - Batuk setelah minum dingin - Demam setelah kontak orang sakit - Pilek makin parah di udara dingin - Sesak napas setelah asap rokok - Menangis saat makan/minum - Muntah setelah batuk - Demam setelah gigitan nyamuk - BAB berubah setelah makan tertentu - Kejang saat demam >39°C - Kejang setelah kelelahan - Kejang terjadi setelah kurang tidur - Demam muncul setelah makan makanan tertentu (tifoid, alergi) - Sesak napas memburuk saat tidur (asma, ISPA berat) - Demam setelah bermain di lingkungan banyak nyamuk (DBD)
<i>Aggravating or Alleviating Factors</i> (Faktor Pemicu atau Pereda)	<ul style="list-style-type: none"> - Batuk berkurang dengan air hangat - Demam turun sebentar lalu naik - Nyeri perut membaik setelah BAB - Pilek membaik setelah tidur - Sesak napas membaik setelah duduk - Nyeri tenggorokan makin sakit makan pedas - Rewel setelah makan asam - Nafsu makan turun saat sakit - Sakit kepala membaik setelah istirahat - Diare makin parah setelah susu - Kejang membaik setelah demam turun - Kejang makin sering jika kurang tidur - Kejang makin buruk setelah kelelahan - Batuk memburuk setelah makan (refluks, ISPA) - Nyeri sendi membaik setelah istirahat (DBD) - Sakit perut memburuk setelah makan makanan berlemak (gangguan pencernaan)
<i>Associated Manifestations</i> (Manifestasi Terkait)	<ul style="list-style-type: none"> - Bibir kering pecah-pecah - Mata cekung - Kulit pucat - Tangan kaki dingin meski demam - Mimisan saat demam - Ruam kulit tidak hilang saat ditekan - Sering kejang tanpa demam - Air mata sedikit saat menangis - Lidah merah bengkak - Kelenjar leher bengkak - Kejang dengan bibir kebiruan - Kejang dengan muntah - Kejang dengan tubuh lemas setelah sadar - Kulit kebiruan di sekitar mulut/kuku (hipoksia, ISPA berat)

	<ul style="list-style-type: none"> - Kaki dan tangan bengkak (indikasi masalah ginjal) - Napas cepat dan dangkal (infeksi paru, pneumonia)
--	--

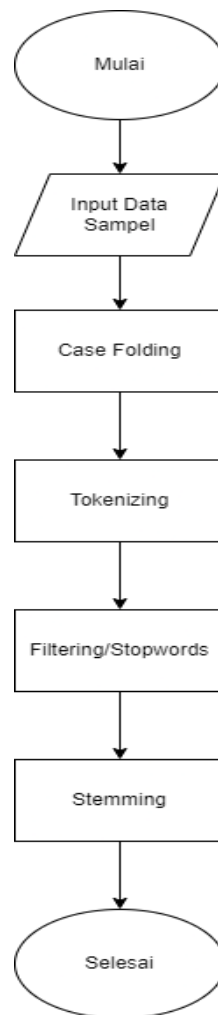
Setelah melakukan input Dataset, maka selanjutnya adalah input Dataset yang akan dijadikan Query, Untuk Dataset Query disini yang akan digunakan adalah Input keluhan yang terdapat pada Tabel 3.2 – Input Keluhan

Tabel 3.2 Input Keluhan

No.	Keluhan
1	Halo selamat pagi Dok anak saya sudah demam selama 3 hari dia sering menangis setiap buang air kecil iya dok urinenya lebih peka dan berbau tajam kadang dia juga mengeluh perutnya sakit Iya dia juga muntah beberapa kali setelah makan dia sering pipis kok tapi hanya sedikit-sedikit dong baik dong terima kasih

2. *Preprocessing* Data

Langkah *preprocessing* yang dilakukan adalah Pengolahan kata yang terdiri dari tahapan *case folding*, *tokenizing*, *filtering*, dan *stemming*, langkah tersebut dapat di lihat pada Gambar 3.3 berikut.



Gambar 3.3 Flowchart Preprocessing

a. Case Folding

Case Folding adalah tahapan yang bertujuan untuk mengubah semua respon menjadi huruf kecil, pada proses ini, karakter 'A'-'Z' yang terdapat pada data diubah menjadi karakter 'a'-'z'. Sementara itu, karakter lain yang bukan huruf dan angka, seperti tanda baca dan spasi, dianggap sebagai pembatas. Tabel 3.4 adalah perbandingan Original Dataset dan hasil *Case Folding*.

Tabel 3.3 Perbandingan Original Dataset dan Hasil Case Folding

Index	Original	Casefolding
0	Halo selamat pagi Dok anak saya sudah demam selama 3 hari dia sering menangis setiap buang air kecil iya dok urinenya lebih peka dan berbau tajam kadang dia juga mengeluh perutnya sakit Iya dia juga muntah beberapa kali setelah	[halo selamat pagi dok anak saya sudah demam selama 3 hari dia sering menangis setiap buang air kecil iya dok urinenya lebih peka dan berbau tajam kadang dia juga mengeluh perutnya sakit iya dia juga muntah beberapa kali setelah

	makan dia sering pipis kok tapi hanya sedikit-sedikit dong baik dong terima kasih	makan dia sering pipis kok tapi hanya sedikit-sedikit dong baik dong terima kasih']
--	---	---

b. Tokenizing

Tokenizing adalah proses dalam pemrosesan teks yang bertujuan untuk memecah teks menjadi unit-unit yang lebih kecil, yang disebut token. *Token* dapat berupa kata, frasa, atau bahkan karakter tergantung pada kebutuhan analisis. Proses ini penting dalam berbagai aplikasi pemrosesan bahasa alami (NLP). Dalam implementasinya, tokenizing dapat dilakukan dengan berbagai cara, seperti *word tokenization* (memecah berdasarkan kata) atau *sentence tokenization* (memecah berdasarkan kalimat). Proses ini penting dalam *Natural Language Processing* (NLP) karena membantu dalam analisis teks, ekstraksi informasi, dan pembuatan model berbasis teks. Tabel 3.5 adalah perbandingan Case Folding dan hasil Tokenizing

Tabel 3.4 Perbandingan Case Folding dan Hasil Tokenizing

Index	Casefolding	Tokenizing
0	[halo selamat pagi dok anak saya sudah demam selama 3 hari dia sering menangis setiap buang air kecil iya dok urinenya lebih peka dan berbau tajam kadang dia juga mengeluh perutnya sakit iya dia juga muntah beberapa kali setelah makan dia sering pipis kok tapi hanya sedikit-sedikit dong baik dong terima kasih]	['', 'halo', 'selamat', 'pagi', 'dok', 'anak', 'saya', 'sudah', 'demam', 'selama', '3', 'hari', 'dia', 'sering', 'menangis', 'setiap', 'buang', 'air', 'kecil', 'iya', 'dok', 'urinenya', 'lebih', 'peka', 'dan', 'berbau', 'tajam', 'kadang', 'dia', 'juga', 'mengeluh', 'perutnya', 'sakit', 'iya', 'dia', 'juga', 'muntah', 'beberapa', 'kali', 'setelah', 'makan', 'dia', 'sering', 'pipis', 'kok', 'tapi', 'hanya', 'sedikit-sedikit', 'dong', 'baik', 'dong', 'terima', 'kasih', '', '']

c. Filtering

Proses *filtering* dalam pemrosesan teks dimulai setelah tahap tokenizing, di mana teks telah dipecah menjadi unit-unit kata atau token. Langkah pertama dalam *filtering* adalah mengidentifikasi *stopwords*, yaitu kata-kata umum yang tidak memberikan makna signifikan dalam analisis. Setelah itu, sistem akan membandingkan setiap token dengan daftar *stopwords* yang telah ditentukan dan menghapus token yang termasuk dalam kategori tersebut. Selain *stopwords*,

filtering juga dapat mencakup penghapusan karakter khusus, angka, atau simbol yang tidak relevan. Tabel 3.6 adalah perbandingan Tokenizing dan Hasil Stopwords

Tabel 3.5 Perbandingan Tokenizing dan Hasil Stopwords

Index	Tokenizing	Stopwords
0	[['', "'halo', 'selamat', 'pagi', 'dok', 'anak', 'saya', 'sudah', 'demam', 'selama', '3', 'hari', 'dia', 'sering', 'menangis', 'setiap', 'buang', 'air', 'kecil', 'iya', 'dok', 'urinenya', 'lebih', 'peka', 'dan', 'berbau', 'tajam', 'kadang', 'dia', 'juga', 'mengeluh', 'perutnya', 'sakit', 'iya', 'dia', 'juga', 'muntah', 'beberapa', 'kali', 'setelah', 'makan', 'dia', 'sering', 'pipis', 'kok', 'tapi', 'hanya', 'sedikit-sedikit', 'dong', 'baik', 'dong', 'terima', 'kasih', '', ']]]	[halo' 'pagi' 'demam' '3' 'menangis' 'buang air' 'urinenya' 'peka' 'berbau' 'tajam' 'kadang' 'mengeluh' 'perutnya' 'sakit' 'muntah' 'kali' 'makan' 'pipis' 'sedikit-sedikit' 'terima' 'kasih']

d. Stemming

Proses *stemming* dalam pemrosesan teks dimulai setelah tahap *tokenizing* dan *filtering*, di mana teks telah dibersihkan dari kata-kata yang tidak relevan. *Stemming* bertujuan untuk mengubah kata-kata berimbuhan menjadi bentuk dasarnya (*root word*) dengan menghapus awalan, akhiran, atau imbuhan lainnya. Algoritma *stemming* akan membandingkan setiap token dengan aturan linguistik tertentu untuk mengidentifikasi bentuk dasarnya. Tabel 3.7 adalah perbandingan Stopwords dan Hasil Stemming

Tabel 3.6 Perbandingan Stopwords dan Hasil Stemming

Index	Stopwords	Stemming
0	[['', "'halo', 'selamat', 'pagi', 'dok', 'anak', 'saya', 'sudah', 'demam', 'selama', '3', 'hari', 'dia', 'sering', 'menangis', 'setiap', 'buang', 'air', 'kecil', 'iya', 'dok', 'urinenya', 'lebih', 'peka', 'dan', 'berbau', 'tajam', 'kadang', 'dia', 'juga', 'mengeluh', 'perutnya', 'sakit', 'iya', 'dia', 'juga', 'muntah', 'beberapa', 'kali', 'setelah', 'makan', 'dia', 'sering', 'pipis', 'kok', 'tapi', 'hanya', 'sedikit-sedikit', 'dong', 'baik', 'dong', 'terima', 'kasih', '', ']]]	halo pagi demam 3 menang buang air urine peka bau tajam kadang keluh perut sakit muntah kali makan pipis sedikit terima kasih

e. Hasil Pre-processing

Setelah dilakukan pre-processing dihasilkan Tabel 3.8 – Hasil Preprocessing

Tabel 3.7 Hasil Preprocessing

Index	Hasil Preprocessing
0	halo pagi demam 3 menang buang air urine peka bau tajam kadang keluh perut sakit muntah kali makan pipis sedikit terima kasih

3. TF-IDF

Setelah tahap *preprocessing*, melalui beberapa langkah untuk mengimplementasikan pemfilteran berbasis konten, termasuk menggunakan *cosine similarity* dan TF-IDF untuk menentukan kemiripan dan tinggi angka yang mirip (Rianti, 2024). TF-IDF adalah suatu metode yang bisa digunakan untuk pembobotan kata. *Term Weighting* atau pembobotan kata bertujuan untuk memberikan bobot nilai pada setiap kata. Perhitungan bobot ini memerlukan dua hal yaitu *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). Berdasarkan formula (2.3) dan (2.1), berikut adalah contoh perhitungan TF-IDF sebelum dilakukan normalisasi, dan pada Tabel 3.9 - Hasil TF-IDF sebelum normalisasi.

$$IDF_0(Query) = \ln\left(\frac{136 + 1}{2 + 1}\right) + 1 = 3.6376$$

$$TF - IDF_{Query, bau} = 0.148 * 3.6376 = 0.539$$

Tabel 3.8 Hasil TF-IDF

	...	bau	...	warna
Query				
Demam tinggi lebih dari 39 derajat	0.000	0.000	...	0.000
Demam lebih dari 3 hari	0.000	0.000	...	0.000
Demam naik turun	0.000	0.000	...	0.000
Demam memburuk malam hari	0.000	0.000	...	0.000
Demam dengan kejang	0.000	0.000	...	0.000
Demam sulit turun	0.000	0.000	...	0.000
Demam disertai rewel	0.000	0.000	...	0.000
Demam setelah imunisasi	0.000	0.000	...	0.000
Demam setelah bepergian	0.000	0.000	...	0.000

.....	0.000	0.000	...	0.000
Urine berbau tajam	0.000	0.2292		0.000
BAB cair berlendir	0.000	0.000	...	0.000
Kejang kaku seluruh tubuh	0.000	0.000	...	0.000
Kejang dengan tatapan kosong	0.000	0.000	...	0.000
Kejang dengan tubuh tersentak-sentak	0.000	0.000	...	0.000
Tinja berdarah atau berlendir	0.000	0.000	...	0.000
Urine keruh atau berwarna merah muda	0.000	0.000	...	0,480
Diare lebih dari 3 kali sehari	0.000	0.000	...	0.000
.....	0.000	0.000	...	0.000
Napas cepat dan dangkal				

Setelah mendapatkan hasil TF-IDF, maka selanjutnya adalah menghitung L2-Norm dengan formula (24). Misalkan kita memiliki TF-IDF dari beberapa kata dalam gejala, misalnya untuk kata-kata:

1. “air” = 0.1949
2. “bau” = 0.2292
3. “demam” = 0.1345
4. “tinggi” = 0.1789

$$L2 - Norm, Query = \sqrt{0.1949^2 + 0.2292^2 + 0.1345^2 + 0.1789^2} = 0.3751$$

Setelah mendapatkan nilai L2-Norm pada setiap dokumen, maka selanjutnya adalah dilakukan perhitungan TF-IDF L2 Norm menggunakan persamaan (2.5). Berikut adalah contoh perhitungan TF-IDF L2 Norm, dan berikut pada Tabel 3.10 merupakan hasil TF-IDF setelah normalisasi.

$$TF - IDF L2 - Norm, Query, bab = \frac{0.2292}{0,3751} = 0.6112$$

Tabel 3.9 Hasil TF-IDF Setelah Normalisasi

	...	bau	...	warna
Query				

Demam tinggi lebih dari 39 derajat	0.000	0.000	...	0.000
Demam lebih dari 3 hari	0.000	0.000	...	0.000
Demam naik turun	0.000	0.000	...	0.000
Demam memburuk malam hari	0.000	0.000	...	0.000
Demam dengan kejang	0.000	0.000	...	0.000
Demam sulit turun	0.000	0.000	...	0.000
Demam disertai rewel	0.000	0.000	...	0.000
Demam setelah imunisasi	0.000	0.000	...	0.000
Demam setelah bepergian	0.000	0.000	...	0.000
.....	0.000	0.000	...	0.000
Urine berbau tajam	0.000	0.6112		0.000
BAB cair berlendir	0.000	0.000	...	0.000
Kejang kaku seluruh tubuh	0.000	0.000	...	0.000
Kejang dengan tatapan kosong	0.000	0.000	...	0.000
Kejang dengan tubuh tersentak-sentak	0.000	0.000	...	0.000
Tinja berdarah atau berlendir	0.000	0.000	...	0.000
Urine keruh atau berwarna merah muda	0.000	0.000	...	0,480
Diare lebih dari 3 kali sehari	0.000	0.000	...	0.000
.....	0.000	0.000	...	0.000
Napas cepat dan dangkal				

4. Cosine Similarity

Pada tahapan ini akan menghitung similaritas dengan jenis *Cosine Similarity*. Penggunaan *cosine similarity* dikarenakan dari hasil TF-IDF yang digunakan dimana formula *Cosine Similarity* tidak memanfaatkan nilai rata-rata suatu item. Berdasarkan formula (2.6), berikut adalah contoh perhitungan similarity A (Query) terhadap B (Gejala Demam Intermittent)

$$Sim \left(\begin{smallmatrix} A \\ B \end{smallmatrix} \right) = \frac{\Sigma(0 * 0) + (0 * 0) + .. + (0 * 0)}{\sqrt{\Sigma 0^2 + 0^2 + .. + 0^2} * \sqrt{\Sigma 0^2 + 0^2 + .. + 0^2}} = 0$$

Perhitungan Similarity dilakukan pada Query terhadap Semua Gejala, dan didapatkan hasil perhitungan Similarity Query terhadap 22 gejala pada Tabel 3.11 – Hasil Perhitungan Cosine Similarity

Tabel 3.10 Hasil Perhitungan Cosine Similarity

Gejala	Similaritas
Demam tinggi lebih dari 39 derajat	0.039
Demam lebih dari 3 hari	0.117
Demam naik turun	0.063
Demam memburuk malam hari	0.046
Demam dengan kejang	0.079
Demam sulit turun	0.041
Demam disertai rewel	0.039
Demam setelah imunisasi	0.049
Demam setelah bepergian	0.049
.....	
Urine berbau tajam	0.382
BAB cair berlendir	0
Kejang kaku seluruh tubuh	0
Kejang dengan tatapan kosong	0
Kejang dengan tubuh tersentak-sentak	0
Tinja berdarah atau berlendir	0
Urine keruh atau berwarna merah muda	0.080
Diare lebih dari 3 kali sehari	0.100
.....	0
Napas cepat dan dangkal	0.065

Berdasarkan Tabel 3.11 – Hasil Perhitungan *Cosine Similarity*, sebagai fokus terhadap dokumen Query didapatkan hal sebagai berikut.

- **Urine Berbau Tajam (0.382):** pada Gejala Demam Intermittent ini memiliki kemiripan 38.2% dengan Keluhan Query, Hasil kemiripan tersebut didapatkan dari Hasil Cosine Similarity yang di ubah dalam bentuk persen, hal ini dapat

disimpulkan Gejala Demam Intermittent tidak memiliki kemiripan pada Input Query

- **Pipis Sedikit (0,314):** pada Gejala Mual ini memiliki kemiripan 31.4% dengan Keluhan Query, Hasil kemiripan tersebut didapatkan dari Hasil Cosine Similarity yang di ubah dalam bentuk persen, hal ini dapat disimpulkan bahwa Gejala Pipis Sedikit hanya memiliki sedikit kemiripan seperti terhadap input query
- **Sakit Perut Memburuk (0.273):** pada Gejala Nyeri Perut memiliki kemiripan 27.3% dengan Keluhan Query, Hasil kemiripan tersebut didapatkan dari Hasil Cosine Similarity yang di ubah dalam bentuk persen.

Berdasarkan Tabel 3.12 – Hasil Sorting, dapat disimpulkan bahwa terdapat 13 gejala yang memiliki kemiripan dengan query (input user).

Tabel 3.11 Hasil Sorting

Gejala	Similaritas
Urine berbau tajam	0.3829
Pipis sedikit-sedikit	0.3142
Sakit perut memburuk setelah makan	0.2737
Air mata sedikit saat menangis	0.232
Muntah lebih dari 3 kali sehari	0.2158
Menangis saat makan atau minum	0.2007
Frekuensi buang air kecil meningkat	0.1953
Kejang setiap kali anak demam tinggi	0.1917
Demam muncul setelah makan makanan tertentu	0.1673
Kejang dengan muntah	0.1639
Nyeri perut	0.1601
Muntah setelah batuk	0.1519
Nafsu makan turun saat sakit	0.1512
....	0
Napas Cepat dan Dangkal	0

Pada penelitian yang dilakukan (Jepriana, 2020) penentuan objektifitas dilakukan berdasarkan indeks prestasi (yang diprediksi) dengan ambang batas rekomendasi 2.75, hal tersebut dijelaskan bahwasanya 2.75 sudah mencakup keberhasilan mahasiswa dalam memahami matakuliah yang didapatkan. Pada manualisasi ini ditetapkan untuk ambang batas yaitu > 0.15 , hal ini dikarenakan pada manualisasi ini maksimal nilai Cosine adalah 1 dan apabila ambang batas ditetapkan berdasar penelitian (Jepriana, 2020) yang memiliki ambang batas 2.75 maka apabila nilai batas ambang ditetapkan seperti penelitian (Jepriana, 2020) tidak akan mendapatkan gejala similaritas, dan menghasilkan Gejala similar seperti pada Tabel 3.13 – Hasil Top N-Recommend

Tabel 3.12 Hasil Top N-Recommend

Gejala	Similaritas
Urine berbau tajam	0.3829
Pipis sedikit-sedikit	0.3142
Sakit perut memburuk setelah makan	0.2737
Air mata sedikit saat menangis	0.232
Muntah lebih dari 3 kali sehari	0.2158
Menangis saat makan atau minum	0.2007
Frekuensi buang air kecil meningkat	0.1953
Kejang setiap kali anak demam tinggi	0.1917
Demam muncul setelah makan makanan tertentu	0.1673
Kejang dengan muntah	0.1639
Nyeri perut	0.1601
Muntah setelah batuk	0.1519
Nafsu makan turun saat sakit	0.1512
....	0
Napas Cepat dan Dangkal	0

5. *The Sacred Seven of History Taking*

Setelah mendapatkan hasil cosine similarity, langkah selanjutnya adalah melakukan pembagian berdasarkan "*The Sacred Seven of History Taking*". *The Sacred Seven*

merupakan metode standar dalam anamnesis medis yang membantu dalam mengidentifikasi gejala pasien secara sistematis. Dengan menggunakan hasil cosine similarity, sistem akan mengklasifikasikan informasi yang telah diolah ke dalam tujuh aspek anamnesis ini untuk memastikan kelengkapan dan keakuratan data. Berikut adalah pembagian sesuai dengan *The Sacred Seven of History Taking*:

Tabel 3.14 Hasil *The Sacred Seven of History Taking*

<i>“The Sacred Seven of History Taking”</i>	Gejala
<i>Chronology (Kronologi)</i>	- <i>(Tidak Ada)</i>
<i>Bodily Location (Lokasi Gejala)</i>	- Nyeri Perut
<i>Quality (Kualitas)</i>	- Urine berbau tajam
<i>Quantity (Kuantitas)</i>	- Muntah lebih dari 3 kali sehari - Pipis sedikit-sedikit - Kejang setiap kali anak demam tinggi - Frekuensi buang air kecil meningkat
<i>Setting (Kondisi yang Memicu)</i>	- Menangis saat makan atau minum - Muntah dengan Batuk
<i>Aggravating or Alleviating Factors (Faktor Pemicu atau Pereda)</i>	- Demam muncul setelah makan makanan tertentu
<i>Associated Manifestations (Manifestasi Terkait)</i>	- Air mata sedikit saat menangis - Kejang dengan muntah - Nafsu makan turun saat sakit - Sakit perut memburuk setelah makan

6. Pengembangan Sistem menggunakan Metode Waterfall

Setelah dilakukan *modelling*, maka dilanjutkan pengembangan sistem menggunakan metode *waterfall*. Dalam metode *waterfall* terdiri dari tahap *requirement identification* untuk mendefinisikan kebutuhan perangkat lunak yang terdiri dari kebutuhan fungsional dan kebutuhan non-fungsional. Tahap kedua yaitu *system and software design* untuk melakukan pembuatan rancangan sistem dengan membuat *flowchart* sistem, UML terdiri dari *use case diagram* dan *class diagram* serta desain *user interface website*. Tahap ketiga yaitu *implementation and unit testing* untuk mengimplementasikan sistem kedalam bahasa pemrograman *Python* dengan *framework flask* dan *database MySQL* yang menghasilkan *website*. Tahap terakhir yaitu *integration and system testing* untuk melakukan pengujian sistem menggunakan *black box testing* dan UAT untuk menguji *website* yang telah dibuat. Penjelasan tahapan tersebut dapat dilihat pada Lampiran A.

3.3.5 Evaluation

Pada tahap *evaluation* dilakukan evaluasi terhadap sistem yang dibuat, dalam hal ini berupa *website* dengan melakukan pengujian menggunakan *black box testing*. *Black box testing* merupakan pengujian sistem dengan memeriksa fungsionalitas sistem apakah dapat berjalan dengan semestinya atau tidak. Evaluasi berikutnya yaitu *User Acceptance Test* (UAT) yang dilakukan dengan memberikan *survey* kepada pengguna sistem untuk mengetahui tanggapan penggunaan aplikasi tersebut apakah sudah sesuai dengan kebutuhan pengguna.

3.3.6 Deployment

Tahap terakhir yaitu tahap *deployment*, melakukan penyebaran hasil penelitian dengan menyebarkan *website* untuk digunakan pengguna dan laporan hasil penelitian serta artikel jurnal hasil penelitian yang berisi hasil ekstraksi informasi dari data anamnesis tersebut dalam bentuk ilmu pengetahuan.

BAB 4

Hasil dan Pembahasan

4.1 Hasil Implementasi

Berikut merupakan penjabaran terkait implementasi yang dilakukan pada penelitian ini.

4.1.1 Implementasi Program

Berikut penjabaran terkait implementasi program pada *data voice* yang digunakan..

1. Business Understanding

Pada tahap ini dilakukan pemahaman pemahaman tujuan bisnis. Tujuan penelitian ini adalah mengimplementasikan NLP untuk anamnesis penyakit dengan gejala demam pada anak menggunakan algoritma TF IDF - *cosine similarity* dalam sebuah *website*.

2. Data Understanding

Pada tahap *data understanding*, dilakukan *Exploratory Data Analysis* (EDA) pada *data voice* yang ada. *Data voice* yang digunakan merupakan hasil rekaman anamnesis pasien terkait gejala atau kondisi yang dirasakan. Selanjutnya, dari *data voice* tersebut digunakan untuk proses NLP.

3. Data Preparation

Selanjutnya, dilakukan proses pengubahan *voice to text* yang prosesnya dapat dilihat Code Snippet A. 1. Dari Code Snippet A. 1 terdapat beberapa fungsi yang diimplementasikan seperti *index* dan *process_voice*. Fungsi *index* untuk merender halaman HTML dengan nama “pages/konsultasiKeluhan/index.html” yang menggunakan *flask*. Sementara pada fungsi *process_voice* digunakan untuk menangani proses pengenalan suara dari audio yang dikirimkan pengguna. Fungsi tersebut dimulai dengan pengambilan data audio dalam bentuk blob *audioBlob* atau file unggahan *audiofile* dari permintaan HTTP. Setelah data audio tersedia, file tersebut dikonversi ke format .WAV menggunakan *library pydub* dan *SpeechRecognition*. File audio tersebut kemudian diolah menggunakan *Recognizer* dari *SpeechRecognition*, yang merekam audio dan mengenali teks menggunakan API *Google Speech Recognition*. Selanjutnya dilakukan proses pengambilan data gejala dari *database* melalui fungsi *Gejala.get().serialize()*. Teks hasil pengenalan suara kemudian digabungkan dengan daftar gejala untuk diproses lebih lanjut seperti menggunakan Code Snippet A. 2.

Dari Code Snippet A. 2, digunakan untuk *preprocessing* untuk teks yang dihasilkan pada tahap sebelumnya. *Preprocessing* dilakukan dengan proses *case folding* untuk mengubah semua huruf menjadi huruf kecil. *Tokenizing* merupakan proses pemecahan kata menjadi beberapa bagian. Hasil kata yang telah dipecah disebut dengan token. *Tokenizing*

dilakukan dengan memanfaatkan *library nlt.tokenize*. Pada tahap *filtering* yaitu dilakukan proses penghapusan kata yang tidak dibutuhkan berupa kata keterangan dan kata sambung dari hasil *token*. Tahap *stemming* merupakan proses mengubah kata dengan mengembalikan bentuknya ke dalam bentuk aslinya atau kata dasar.

4. Modelling

Setelah melakukan *preprocessing*, dilanjutkan dengan proses TF-IDF dan *cosine similarity* yang prosesnya dapat dilihat pada Code Snippet A. 3. Berdasarkan Code Snippet A. 3, hasil dari tahap sebelumnya diproses menggunakan TF-IDF untuk membuat representasi bobot. Kemudian, dilakukan perhitungan menggunakan *cosine similarity* untuk menghasilkan nilai kesamaan. Hasil kesamaan tersebut diurutkan berdasarkan nilai tertinggi, dan gejala dengan nilai kesamaan lebih dari 0,15 difilter sebagai gejala yang relevan. Jenis gejala yang lolos filter kemudian digunakan untuk mencocokkan ID gejala dengan *database*, sehingga data keluhan dapat disimpan.

5. Evaluation

Tahap ini dilakukan evaluasi dengan melakukan pengujian menggunakan *black box testing*, serta UAT. Pengujian *black box* dilakukan untuk pengujian fungsional untuk mengetahui sistem berfungsi sesuai dengan skenario yang ditentukan. Sementara UAT digunakan untuk mengidentifikasi apakah sistem memenuhi kebutuhan operasional pengguna.

6. Deployment

Tahap terakhir yaitu tahap *deployment*, melakukan penyebaran hasil penelitian dengan membuat laporan hasil penelitian yang berisi hasil penerapan NLP untuk anamnesis penyakit dengan gejala demam pada anak menggunakan algoritma TF IDF - *cosine similarity* dalam bentuk ilmu pengetahuan.

4.1.2 Tampilan Aplikasi

Berikut merupakan tampilan aplikasi yang dihasilkan dengan menggunakan *framework flask*, dengan menggunakan bahasa pemrograman *python*. Pada aplikasi ini terdiri dari beberapa aktor seperti admin, dokter, dan pengguna. Masing-masing aktor tersebut memiliki hak akses yang berbeda-beda, yang dapat dilihat sebagai berikut:

1. Tampilan Home

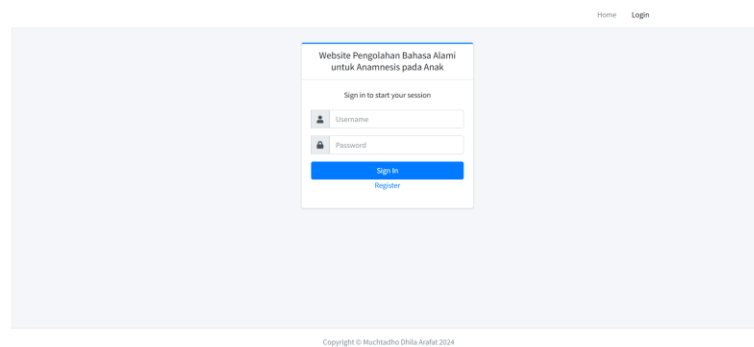
Tampilan ini digunakan oleh pengguna ketika pertama kali mengakses *website*. Rancangan antarmuka home dapat dilihat pada Gambar 4.1.



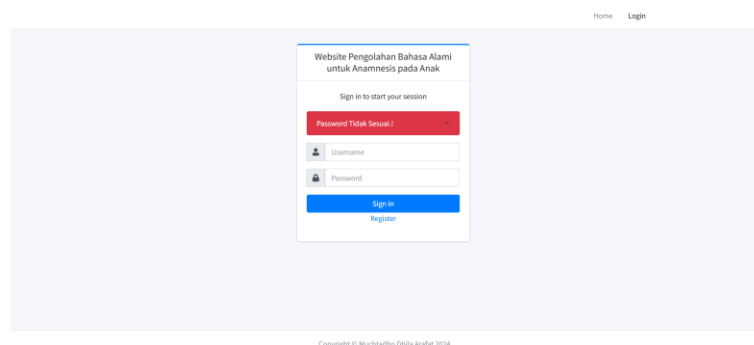
Gambar 4.1 Tampilan *Home*

2. Tampilan *Login*

Tampilan pada Gambar 4.2 ini digunakan oleh pengguna untuk melakukan *login* agar dapat mengakses *website* dengan *input field username* dan *password*. Jika *input field username* dan *password* benar maka akan menampilkan halaman *dashboard*, jika *input field username* dan *password* salah maka akan tampil seperti pada Gambar 4.3 berikut.



Gambar 4.2 Tampilan *Login*

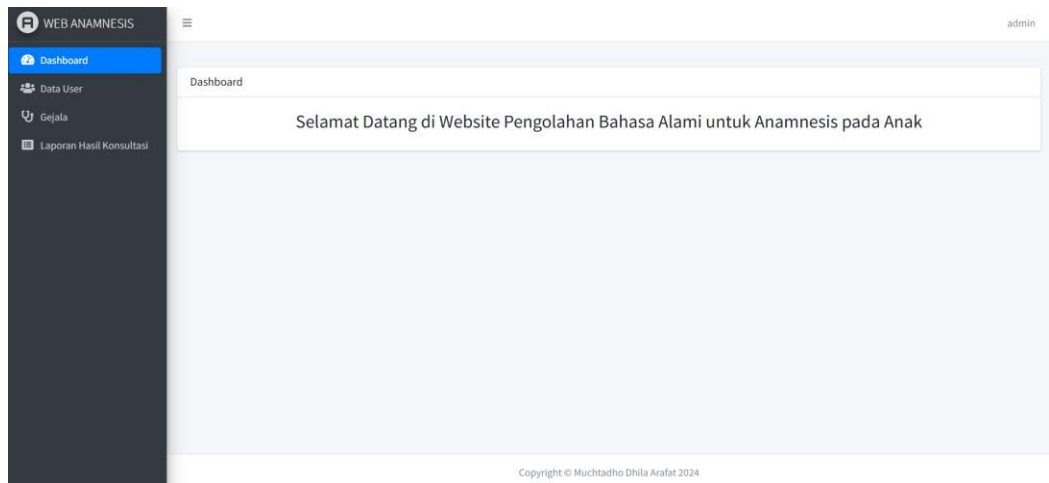


Gambar 4.3 Tampilan *Login* Jika *Username/Password* Tidak Sesuai

3. Aktor Admin

a. Tampilan Menu *Dashboard*

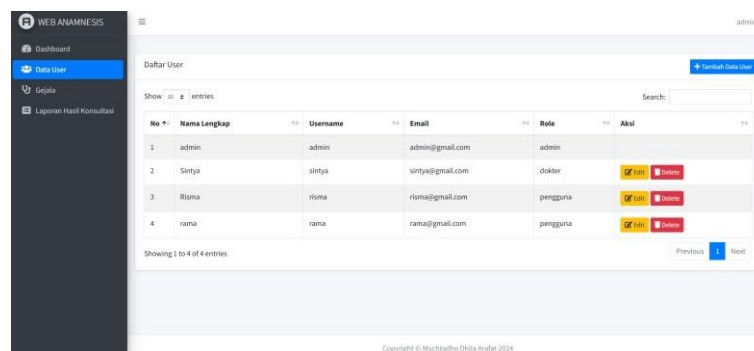
Gambar 4.4 merupakan halaman *dashboard* digunakan untuk menampilkan halaman pertama yang akan ditampilkan ketika admin berhasil melakukan *login*.



Gambar 4.4 Tampilan *Dashboard* – Admin

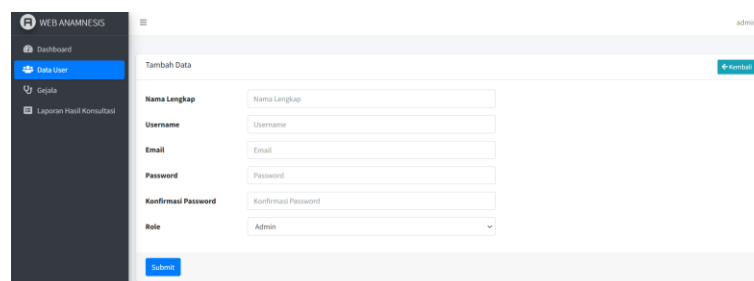
b. Tampilan Menu *Data User*

Menu *data user* digunakan oleh admin dalam mengelola data *user* yang dapat mengakses sistem. Berikut pada Gambar 4.5 merupakan tampilan menu *user* pada halaman admin.



Gambar 4.5 Hasil Tampilan Antarmuka Modul *User*

Jika akan melakukan tambah *user*, maka admin harus mengisi *form* tambah *user* berupa nama lengkap, *email*, *username*, *role*, *password*, konfirmasi *password*. Berikut pada Gambar 4.6 merupakan tampilan tambah *user*.



Gambar 4.6 Hasil Tampilan Antarmuka Tambah *User*

Selain itu juga admin dapat mengedit data user dengan cara klik tombol “Edit”, secara otomatis program akan memunculkan tampilan form untuk mengedit data user, terdapat Gambar 4.7 merupakan tampilan Edit user.

WEB ANAMNESIS

admin

Edit Data

Nama Lengkap: Sintya

Username: sintya

Email: sintya@gmail.com

Password: Password

Role: Dokter

Update

Gambar 4.7 Hasil Tampilan Antarmuka *Edit User*

Selain itu juga admin dapat menghapus data user dengan klik *button* hapus. Berikut pada Gambar 4.8 merupakan tampilan hapus user.

127.0.0.1:5000 says
Apakah anda yakin?

OK Cancel

Daftar User

Show 4 entries

No	Nama Lengkap	Username	Email	Role	Aksi
1	admin	admin	admin@gmail.com	admin	
2	Sintya	sintya	sintya@gmail.com	dokter	Edit Hapus
3	Rima	rima	rima@gmail.com	pengguna	Edit Hapus
4	rama	rama	rama@gmail.com	pengguna	Edit Hapus

Showing 1 to 4 of 4 entries

Previous 1 Next

Copyright © Muchlisno Dikla Arafat 2024

Gambar 4.8 Hasil Tampilan Antarmuka Hapus User

c. Tampilan Menu Gejala

Menu gejala digunakan oleh admin dalam mengelola data gejala. Berikut pada Gambar 4.9 merupakan tampilan menu gejala.

WEB ANAMNESIS

admin

Daftar Gejala

Show 10 entries

No	Kode Gejala	Nama Gejala	Aksi
1	G01	Demam Intermittent	Edit Hapus
2	G02	Demam Menggigil	Edit Hapus
3	G03	Lena Demam	Edit Hapus
4	G04	Sakit Kepala	Edit Hapus
5	G05	Mual	Edit Hapus
6	G06	Muntah	Edit Hapus
7	G07	Iyeri Tulang atau Sendi	Edit Hapus
8	G08	Iyeri Perut	Edit Hapus
9	G09	Iyeri Ura Hati	Edit Hapus
10	G10	Bintik Merah	Edit Hapus

Showing 1 to 10 of 22 entries

Previous 1 2 3 Next

Copyright © Muchlisno Dikla Arafat 2024

Gambar 4.9 Hasil Tampilan Antarmuka Modul Gejala

Jika akan melakukan tambah gejala, maka admin harus mengisi *form* tambah gejala berupa kode gejala, nama gejala. Berikut pada Gambar 4.10 merupakan tampilan tambah gejala.

Gambar 4.10 Hasil Tampilan Antarmuka Tambah Gejala

Selain itu juga admin dapat mengedit data gejala dengan cara klik tombol “Edit”, secara otomatis program akan memunculkan tampilan form untuk mengedit data gejala, terdapat Gambar 4.11 merupakan tampilan Edit gejala.

Gambar 4.11 Hasil Tampilan Antarmuka *Edit* Gejala

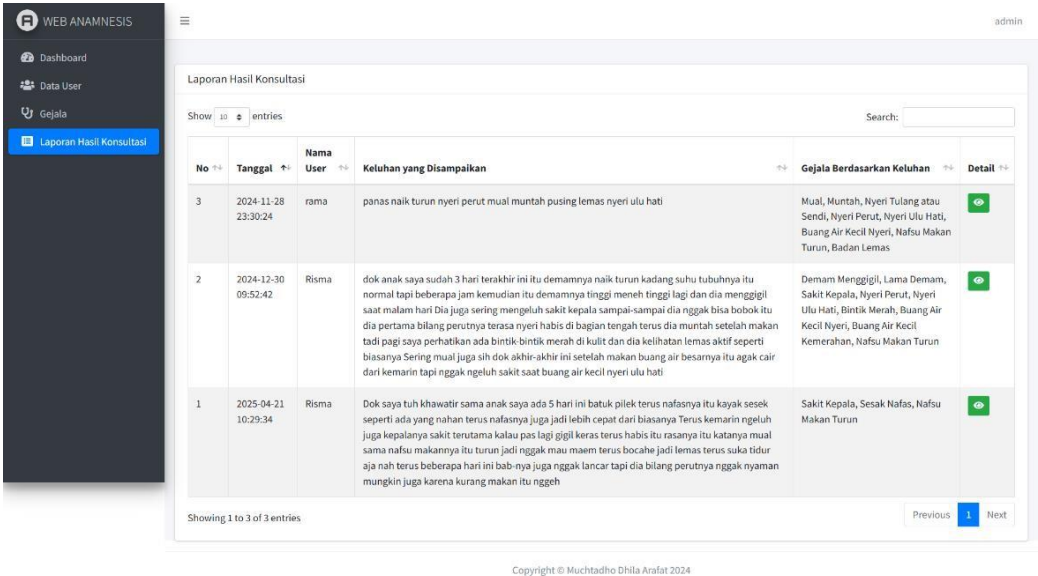
Selain itu juga admin dapat menghapus data gejala dengan klik *button* hapus. Berikut pada Gambar 4.12 merupakan tampilan hapus gejala.

No	Kode Gejala	Nama Gejala	Aksi
1	G01	Demam Intermittent	Edit Hapus
2	G02	Demam Menggigil	Edit Hapus
3	G03	Lama Demam	Edit Hapus
4	G04	Sakit Kepala	Edit Hapus
5	G05	Mual	Edit Hapus
6	G06	Muntah	Edit Hapus
7	G07	Nyeri Tulang atau Sendi	Edit Hapus
8	G08	Nyeri Perut	Edit Hapus

Gambar 4.12 Hasil Tampilan Antarmuka Hapus Gejala

d. Tampilan Menu Laporan Hasil Konsultasi

Pada menu ini admin dapat melihat data berupa keluhan yang disampaikan oleh pengguna, beserta gejala dari keluhan tersebut. Berikut pada Gambar 4.13 merupakan tampilan antarmuka laporan hasil konsultasi.



Gambar 4.13 Hasil Tampilan Antarmuka Laporan Hasil Konsultasi

Selain itu, admin dapat melakukan klik pada *icon eye* untuk melihat detail dari proses NLP sehingga dihasilkan gejala berdasarkan keluhan pasien, dan prosesnya dapat di lihat pada Gambar 4.14 dan Gambar 4.15.

Hasil Preprocessing:

Show 10 entries

No	Casefolding	Tokenizing	Filtering	Stemming
1	['halo selamat pagi dok anak saya sudah demam selama 3 hari dia sering menangis setiap buang air kecil iya dok urinenya lebih peka dan berbau tajam kadang dia juga mengeluh perutnya sakit iya dia juga muntah beberapa kali setelah makan dia sering pipis kok tapi hanya sedikit-sedikit dong baik dong terima kasih']	[['', 'halo', 'selamat', 'pagi', 'dok', 'anak', 'saya', 'sudah', 'demam', 'selama', '3', 'hari', 'dia', 'sering', 'menangis', 'setiap', 'buang', 'air', 'kecil', 'iya', 'dok', 'urinenya', 'lebih', 'peka', 'dan', 'berbau', 'tajam', 'kadang', 'dia', 'juga', 'mengeluh', 'perutnya', 'sakit', 'iya', 'dia', 'juga', 'muntah', 'beberapa', 'kali', 'setelah', 'makan', 'dia', 'sering', 'pipis', 'kok', 'tapi', 'hanya', 'sedikit-sedikit', 'dong', 'baik', 'dong', 'terima', 'kasih', '', '']]	['halo pagi demam 3 menangis buang air urinenya peka berbau tajam kadang mengeluh perutnya sakit muntah kali makan pipis sedikit-sedikit terima kasih ']	halo pagi demam 3 menang buang air urine peka bau tajam kadang keluh perut sakit muntah kali makan pipis sedikit terima kasih

Cosine Similarity antara keluhan yang disampaikan dengan gejala pada database:

Show 10 entries

No ↑↓	Gejala ↑↓	Cosine Similarity ↑↓
1	Demam tinggi lebih dari 39 derajat	0.03976903368840338
2	Demam lebih dari 3 hari	0.11706648982446094
3	Demam naik turun	0.06364889268367768
4	Demam memburuk malam hari	0.04652024208455639
5	Demam dengan kejang	0.07954043630029864
6	Demam sulit turun	0.04161121446502996
7	Demam disertai rewel	0.0392322919620579
8	Demam setelah imunisasi	0.049772620586850315
9	Demam setelah bepergian	0.049772620586850315
10	Demam dengan keringat malam	0.04062322797726553

Mengambil gejala dengan nilai cosine similarity diatas 0.15:

Show 10 entries

No ↑↓	Gejala ↑↓	Cosine Similarity ↑↓
1	Urine berbau tajam	0.38293250590787375
2	Pipis sedikit-sedikit	0.3142180977032224
3	Sakit perut memburuk setelah makan	0.27373569895900735
4	Air mata sedikit saat menangis	0.23209560359667541
5	Muntah lebih dari 3 kali sehari	0.21581258922262966
6	Menangis saat makan atau minum	0.20075962627753974
7	Frekuensi buang air kecil meningkat	0.1953389977068135
8	Kejang setiap kali anak demam tinggi	0.19170489759467982
9	Demam muncul setelah makan makanan tertentu	0.16738117695460691
10	Kejang dengan muntah	0.1635434596517789

Gambar 4.14 Proses NLP Pasien 1

Hasil Preprocessing:

Show 10 entries

Search:

No	Casefolding	Tokenizing	Filtering	Stemming
1	['jadi gini dong anaknya tadi malam itu mulai demam tapi pagi ini udah turun soalnya saya udah kasih penurun panas tapi ya masih anget sama lengan gitu panas dingin rasanya badannya terus tadi malam itu sempat mual tapi nggak sampai muntah mungkin karena itu juga dia jadi nafsu makannya berkurang minumannya juga nggak mau banyak soalnya rasa mual itu kalau buang air kecil nya lancar terus bab-nya juga lancar nggak ada diare tak tanya perutnya sakit apa enggak katanya enggak katanya tenggorokannya sakit tapi perutnya nggak sakit']	['!', '"', 'jadi', 'gini', 'dong', 'anaknya', 'tadi', 'malam', 'itu', 'mulai', 'demam', 'tapi', 'pagi', 'ini', 'udah', 'turun', 'soalnya', 'saya', 'udah', 'kasih', 'penurun', 'panas', 'tapi', 'ya', 'masih', 'anget', 'sama', 'lengan', 'gitu', 'panas', 'dingin', 'rasanya', 'badannya', 'terus', 'tadi', 'malam', 'itu', 'sempat', 'mual', 'tapi', 'nggak', 'sampai', 'muntah', 'mungkin', 'karena', 'itu', 'juga', 'dia', 'jadi', 'nafsu', 'makannya', 'berkurang', 'minumannya', 'juga', 'nggak', 'mau', 'banyak', 'soalnya', 'rasa', 'mual', 'itu', 'kalau', 'buang', 'air', 'kecil', 'nya', 'lancar', 'terus', 'bab-nya', 'juga', 'lancar', 'nggak', 'ada', 'diare', 'tak', 'tanya', 'perutnya', 'sakit', 'apa', 'enggak', 'katanya', 'enggak', 'katanya', 'tenggorokannya', 'sakit', 'tapi', 'perutnya', 'nggak', 'sakit', '"', '']	['jadi gini anaknya malam demam pagi udah turun udah kasih penurun panas anget lengan gitu panas dingin badannya malam mual nggak muntah nafsu makannya berkurang minumannya nggak mual buang air nya lancar bab-nya lancar nggak diare perutnya sakit tenggorokannya sakit perutnya nggak sakit ']	jadi gin anak malam demam pagi udah turun udah kasih turun panas anget lengan gitu panas dingin badan malam mual nggak muntah nafsu makan kurang minum nggak mual buang air nya lancar bab lancar nggak diare perut sakit tenggorok sakit perut nggak sakit

Cosine Similarity antara keluhan yang disampaikan dengan gejala pada database:

Show 10 entries

Search:

No	Gejala	Cosine Similarity
1	Demam tinggi lebih dari 39 derajat	0.020427269657105923
2	Demam lebih dari 3 hari	0.06013092433152143
3	Demam naik turun	0.1827497986964085
4	Demam memburuk malam hari	0.14875713974013627
5	Demam dengan kejang	0.04085575610613455
6	Demam sulit turun	0.11790772871472337
7	Demam disertai rewel	0.02015157354474409
8	Demam setelah imunisasi	0.025565588297532298
9	Demam setelah bepergian	0.025565588297532298
10	Demam dengan keringat malam	0.12936820038232108

Mengambil gejala dengan nilai cosine similarity diatas 0.15:

Show 10 entries

Search:

No	Gejala	Cosine Similarity
1	Nafsu makan turun saat sakit	0.31874946719369973
2	Sakit perut memburuk setelah makan	0.28370573730846516
3	Sakit tenggorokan	0.2557263385846367
4	Nyeri tenggorokan makin sakit makan pedas	0.19471643293536237
5	Demam naik turun	0.1827497986964085
6	Nyeri perut membaik setelah BAB	0.16883936140893874
7	Nyeri perut	0.16456459155573877
8	Sakit kepala	0.16348798025472347

Gambar 4.15 Proses NLP Pasien 2

	24	39	air	aktivitas	asam	asap	bab	badan	baik	bak	bangun	batuk	bau	bengkal	berat	bibir	biru	buang	bunyi
0	0	0	0,1943	0	0	0	0	0	0	0	0	0	0,2232	0	0	0	0	0,2232	0
1	0	0,6651	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0,5824	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0,4673	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0,6333	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0,9405	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0,5645	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0,5323	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0,5323	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0,5323	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,7877	0
41	0	0	0	0	0	0	0	0	0	0	0	0,5385	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0,5468	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0,5645	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Gambar 4.16 Hasil TF-IDF Salah Satu Pasien

Berdasarkan Gambar 4.14 dan Gambar 4.15 terlihat bagaimana proses NLP untuk memproses keluhan pasien. Dari keluhan yang disampaikan, akan diproses dengan menggabungkan *text* hasil *voice* dengan gejala pada *database*. Selanjutnya ditampilkan hasil *preprocessing* yang meliputi *casefolding*, *tokenizing*, *filtering*, dan *stemming*. Kemudian diproses menggunakan TF-IDF yang hasilnya dapat diunduh pada *file* berformat *excel* yang dapat dilihat pada Gambar 4.16. Selanjutnya dihasilkan nilai *cosine similarity* pada setiap gejala yang ada, dan dilakukan pengambilan gejala berdasarkan kronologi dari cerita pasien.

4. Aktor Pengguna

a. Tampilan Register

Ketika pengguna belum memiliki akun untuk mengakses sistem, maka pengguna terlebih dahulu melakukan pendaftaran akun. Pengguna dapat klik tombol “Register” pada form *login*, kemudian pengguna akan diarahkan untuk mengisi form register dengan mengisi *field* berupa nama lengkap, email, *username*, *password*, dan konfirmasi *password*. Halaman register dapat terlihat seperti pada Gambar 4.17.

Gambar 4.17 Tampilan Antarmuka Register

b. Tampilan Menu *Dashboard*

Gambar 4.18 merupakan halaman *dashboard* digunakan untuk menampilkan halaman pertama yang akan ditampilkan ketika pengguna berhasil melakukan *login*.

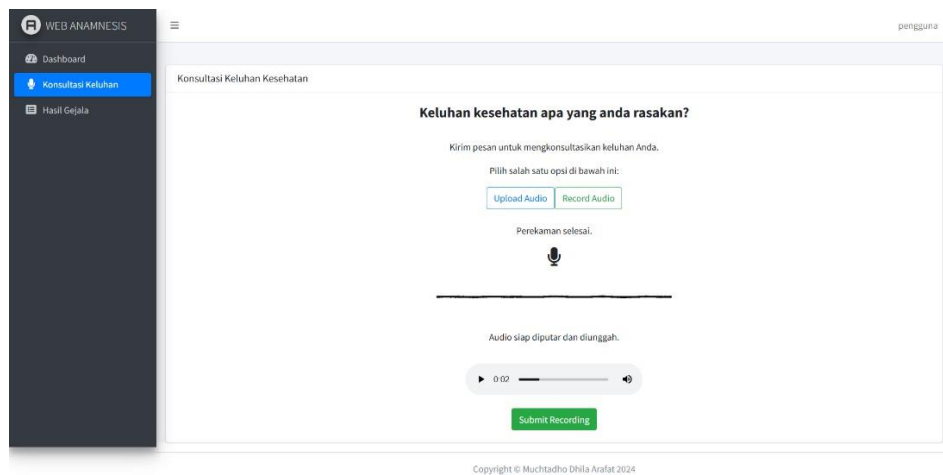


Gambar 4.18 Tampilan *Dashboard* – Pengguna

c. Tampilan Menu Konsultasi Keluhan

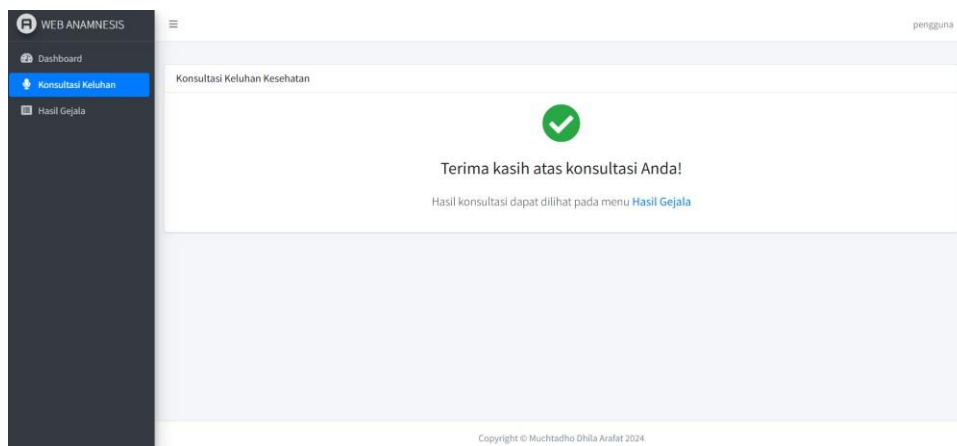
Pada menu konsultasi keluhan, pengguna dapat melakukan konsultasi keluhan dengan cara *upload audio*, atau *record audio*. Berikut pada Gambar 4.19 merupakan tampilan untuk melakukan *upload audio*, sementara pada Gambar 4.20 tampilan untuk melakukan *record audio*.

Gambar 4.19 Tampilan Antarmuka *Upload Audio*



Gambar 4.20 Tampilan Antarmuka *Record Audio*

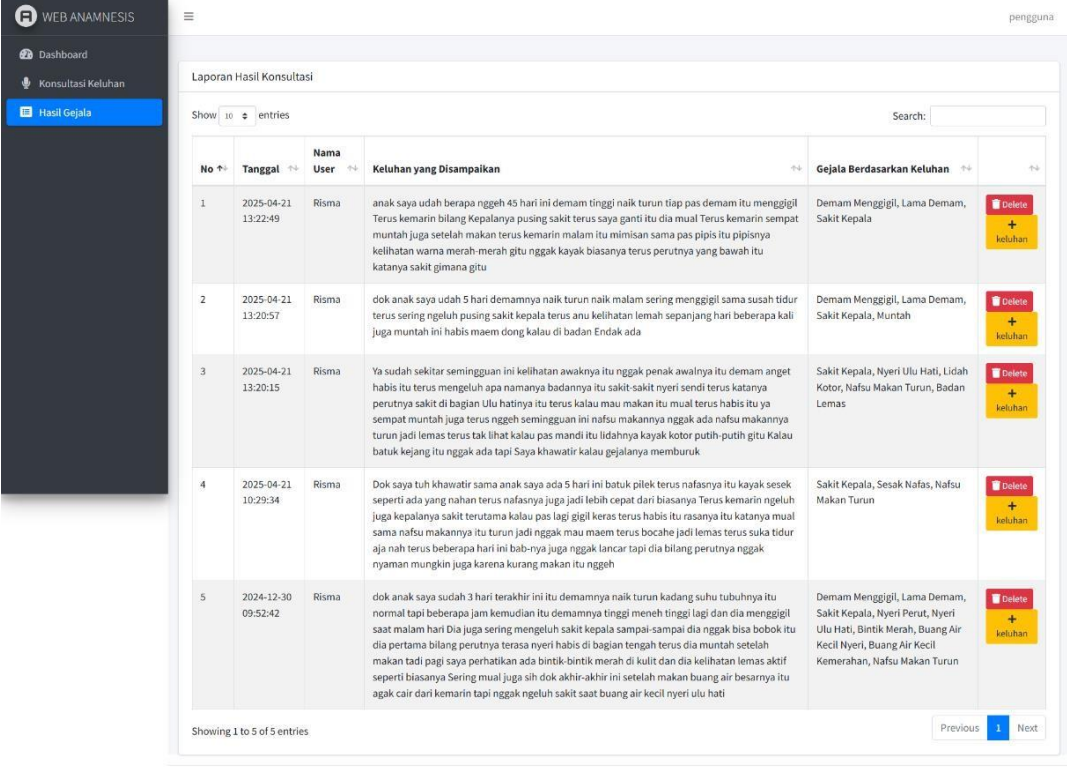
Berikut pada Gambar 4.21 merupakan tampilan ketika pengguna telah melakukan *upload/submit audio*. Hasil konsultasi tersebut dapat dilihat pada menu hasil gejala.



Gambar 4.21 Tampilan Antarmuka Setelah Melakukan Konsultasi

d. Tampilan Menu Hasil Gejala

Berikut pada Gambar 4.22 merupakan tampilan antarmuka laporan konsultasi pengguna yang berisi nama dari pengguna, keluhan yang disampaikan, gejala berdasarkan keluhan, dan hasil anamnesis dari dokter. Pengguna juga dapat melakukan hapus data tersebut dengan melakukan klik pada tombol “Hapus”.



WEB ANAMNESIS

Dashboard

Konsultasi Keluhan

Hasil Gejala

pengguna

Laporan Hasil Konsultasi

Show 10 entries

Search:

No	Tanggal	Nama User	Keluhan yang Disampaikan	Gejala Berdasarkan Keluhan	
1	2025-04-21 13:22:49	Risma	anak saya udah berapa nggeh 45 hari ini demam tinggi naik turun tiap pas demam itu menggigil Terus kemarin bilang Kepalanya pusing sakit terus saya ganti itu dia mual Terus kemarin sempat muntah juga setelah makan terus kemarin malam itu mimisan sama pas pipis itu pipisnya kelihatan warna merah-merah gitu nggak kayak biasanya terus perutnya yang bawah itu katanya sakit gimana gitu	Demam Menggigil, Lama Demam, Sakit Kepala	Delete + keluhan
2	2025-04-21 13:20:57	Risma	dok anak saya udah 5 hari demamnya naik turun naik malam sering menggigil sama susah tidur terus sering ngeluh pusing sakit kepala terus anu kelihatan lemah sepanjang hari beberapa kali juga muntah ini habis maem dong kalau di badan Endak ada	Demam Menggigil, Lama Demam, Sakit Kepala, Muntah	Delete + keluhan
3	2025-04-21 13:20:15	Risma	Ya sudah sekitar seminggu ini kelihatan awaknya itu nggak penak awalnya itu demam anget habis itu terus mengeluh apa namanya badannya itu sakit-sakit nyeri sendi terus katanya perutnya sakit di bagian Ulu hatinya itu terus kalau mau makan itu mual terus habis itu ya sempat muntah juga terus nggeh seminggu ini nafsu makannya nggak ada nafsu makannya turun jadi lemas terus tak lihat kalau pas mandi itu lidahnya kayak kotor putih-putih gitu Kalau batuk kejang itu nggak ada tapi Saya khawatir kalau gejalanya memburuk	Sakit Kepala, Nyeri Ulu Hati, Lidah Kotor, Nafsu Makan Turun, Badan Lemas	Delete + keluhan
4	2025-04-21 10:29:34	Risma	Dok saya tuh khawatir sama anak saya ada 5 hari ini batuk pilek terus nafasnya itu kayak sesek seperti ada yang nahan terus nafasnya juga jadi lebih cepat dari biasanya Terus kemarin ngeluh juga kepalanya sakit terutama kalau pas lagi gigil keras terus habis itu rasanya itu katanya mual sama nafsu makannya itu turun jadi nggak mau maem terus bocah jadi lemas terus suka tidur aja nah terus beberapa hari ini bab-nya juga nggak lancar tapi dia bilang perutnya nggak nyaman mungkin juga karena kurang makan itu nggeh	Sakit Kepala, Sesak Nafas, Nafsu Makan Turun	Delete + keluhan
5	2024-12-30 09:52:42	Risma	dok anak saya sudah 3 hari terakhir ini itu demamnya naik turun kadang suhu tubuhnya itu normal tapi beberapa jam kemudian itu demamnya tinggi meneh tinggi lagi dan dia menggigil saat malam hari Dia juga sering mengeluh sakit kepala sampai-sampai dia nggak bisa bobok itu dia pertama bilang perutnya terasa nyeri habis di bagian tengah terus dia muntah setelah makan tadi pagi saya perhatikan ada bintik-bintik merah di kulit dan dia kelihatan lemas aktif seperti biasanya Sering mual juga sih dok akhir-akhir ini setelah makan buang air besarnya itu agak cair dari kemarin tapi nggak ngeluh sakit saat buang air kecil nyeri ulu hati	Demam Menggigil, Lama Demam, Sakit Kepala, Nyeri Perut, Nyeri Ulu Hati, Bintik Merah, Buang Air Kecil Nyeri, Buang Air Kecil Kemerahan, Nafsu Makan Turun	Delete + keluhan

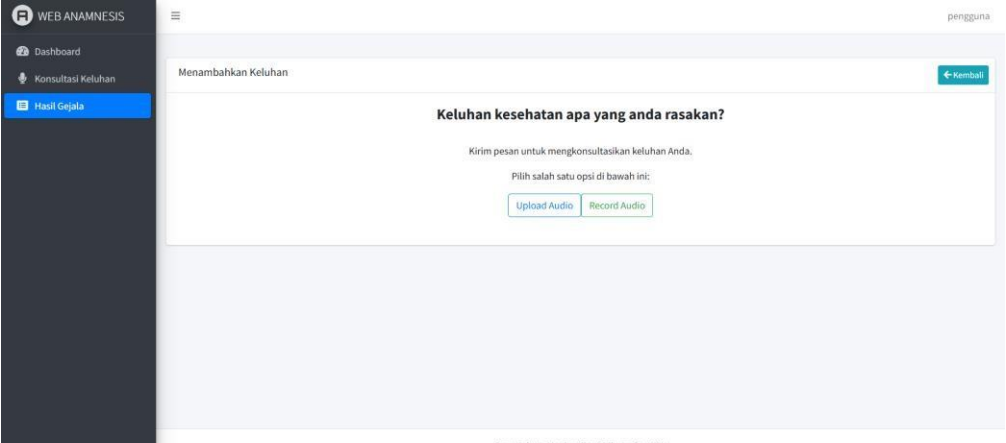
Showing 1 to 5 of 5 entries

Previous 1 Next

Copyright © Muchtadho Dhila Arifat 2024

Gambar 4.22 Tampilan Antarmuka Laporan Hasil Konsultasi

Selain itu, apabila terdapat keluhan yang belum tersampaikan, pengguna dapat menambah keluhan tersebut dengan melakukan klik pada *button* keluhan. Tampilan tersebut dapat dilihat pada Gambar 4.23.



WEB ANAMNESIS

Dashboard

Konsultasi Keluhan

Hasil Gejala

pengguna

Menambahkan Keluhan

Kembali

Keluhan kesehatan apa yang anda rasakan?

Kirim pesan untuk mengkonsultasikan keluhan Anda.

Pilih salah satu opsi di bawah ini:

Upload Audio Record Audio

Copyright © Muchtadho Dhila Arifat 2024

Gambar 4.23 Tampilan Menambahkan Keluhan Lanjutan

Gambar 4.24 merupakan hasil penambahan keluhan lanjutan tersebut.

No ↑↓	Nama User ↑↓	Keluhan yang Disampaikan ↑↓	Gejala Berdasarkan Keluhan ↑↓	↑↓
1	rama	dok anak saya sudah 3 hari terakhir ini itu demamnya naik turun kadang suhu tubuhnya itu normal tapi beberapa jam kemudian itu demamnya tinggi meneng tinggi lagi dan dia menggigil saat malam hari Dia juga sering mengeluh sakit kepala sampai-sampai dia nggak bisa bobok itu dia pertama bilang perutnya terasa nyeri habis di bagian tengah terus dia muntah setelah makan tadi pagi saya perhatikan ada bintik-bintik merah di kulit dan dia kelihatan lemas aktif seperti biasanya Sering mual juga sih akhir-akhir ini setelah makan buang air besarnya itu agak cair dari kemarin tapi nggak ngeluh sakit saat buang air kecil anak-anak saya tambah sakit perutnya tapi rasanya	Demam Menggigil, Lama Demam, Sakit Kepala, Nyeri Perut, Bintik Merah, Buang Air Kecil Nyeri, Buang Air Kecil Kemerahan, Nafsu Makan Turun	<div>Delete</div> <div>+ keluhan</div>

Gambar 4.24 Tampilan Berhasil Menambahkan Keluhan Lanjutan

5. Aktor Dokter

a. Tampilan Menu *Dashboard*

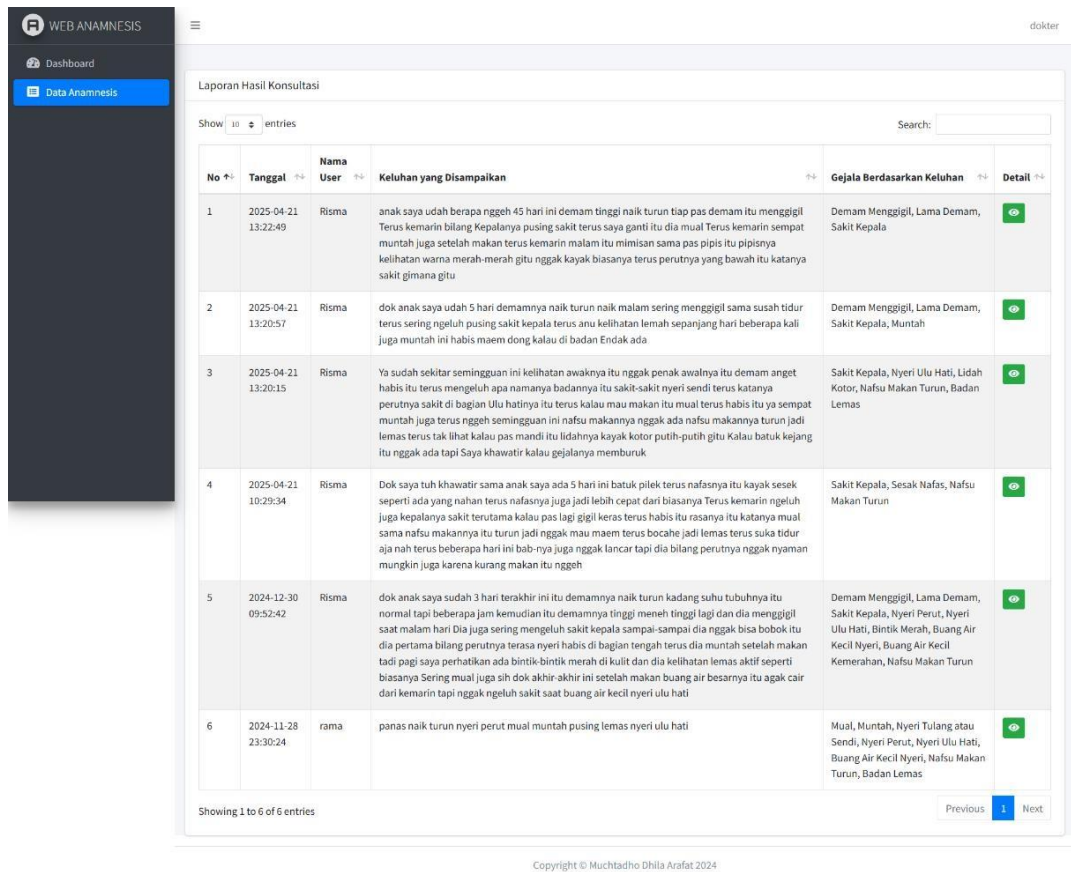
Gambar 4.25 merupakan halaman *dashboard* digunakan untuk menampilkan halaman pertama yang akan ditampilkan ketika dokter berhasil melakukan *login*.



Gambar 4.25 Tampilan *Dashboard* – Dokter

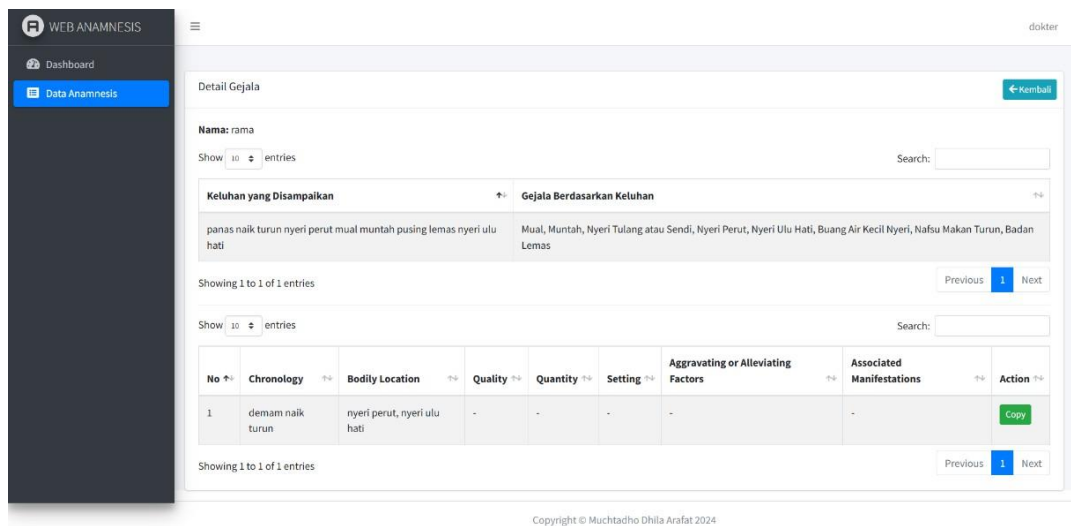
b. Tampilan Menu Data Anamnesis

Berikut pada Gambar 4.26 merupakan tampilan menu data anamnesis agar dokter dapat melihat gejala berdasarkan keluhan tersebut.



Gambar 4.26 Tampilan Menu Data Anamnesis

Berikut pada Gambar 4.27 merupakan tampilan ketika dokter melihat pengelompokan gejala pada “*The Sacred Seven of History Taking*”.



Gambar 4.27 Tampilan Antarmuka Pengelompokkan Gejala sesuai dengan “*The Sacred Seven of History Taking*”

4.1.3 Pengujian Aplikasi

Berikut merupakan penjabaran terkait pengujian aplikasi yang dilakukan dengan teknik *black box testing* dan UAT.

1. *Black box Testing*

Skenario pengujian *black box* merupakan rancangan pengujian sistem yang dilakukan pada sistem yang dibuat. Dalam penelitian ini menggunakan jenis pengujian *black box* yang menguji setiap fungsionalitas sistem. Skenario pengujian *black box* dapat dilihat pada Tabel 4.1, Tabel 4.2, dan Tabel 4.3.

Tabel 4.1 Pengujian *Black box* - Admin

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
1.	<i>Login</i>	Melakukan <i>login</i> dengan admin mengakses URL aplikasi, kemudian mengisi <i>username</i> dan <i>password</i> . Selanjutnya, menekan tombol “ <i>Sign in</i> ”.	Sistem menampilkan halaman <i>login</i> dan menampilkan halaman <i>dashbaord</i> .	Sesuai	Gambar 4.4
2.	Mengelola data user	Melakukan proses tambah data user dengan mengisi data secara lengkap dan sesuai dan menekan tombol “ <i>Submit</i> ”.	Data user baru berhasil ditambahkan pada sistem.	Sesuai	Gambar 4.6
		Melakukan ubah data user dengan memilih salah satu data dan klik tombol “ <i>Edit</i> ”. Kemudian,	Sistem berhasil melakukan perubahan data user	Sesuai	Gambar 4.7

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
		mengubah data pada halaman <i>form Edit</i> data user dan klik tombol “ <i>Update</i> ”			
		Melakukan hapus data user dengan memilih salah satu data dan klik tombol “Apakah Anda Yakin?”	Sistem berhasil menghapus data user	Sesuai	Gambar 4.8
3.	Mengelola Data gejala	Melakukan proses tambah data gejala dengan mengisi data secara lengkap dan sesuai dan menekan tombol “Submit”.	Data gejala baru berhasil ditambahkan pada sistem.	Sesuai	Gambar 4.10
		Melakukan ubah data gejala dengan memilih salah satu data dan klik tombol “Edit”. Kemudian, mengubah data pada halaman <i>form Edit</i> data gejala dan klik tombol “ <i>Update</i> ”	Sistem berhasil melakukan perubahan data gejala	Sesuai	Gambar 4.11

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
		Melakukan hapus data gejala dengan memilih salah satu data dan klik tombol “Apakah anda yakin?”	Sistem berhasil menghapus data gejala	Sesuai	Gambar 4.12
4.	Melihat laporan hasil konsultasi	Klik pada menu “Laporan Hasil Konsultasi”	Sistem menampilkan laporan hasil konsultasi	Sesuai	Gambar 4.13
		Melihat hasil NLP, admin dapat melakukan klik pada <i>icon eye</i>	Sistem menampilkan proses NLP pada keluhan pasien beserta hasil <i>cosine similarity</i>	Sesuai	Gambar 4.14 hingga Gambar 4.16

Tabel 4.2 Pengujian *Black box* - Pengguna

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
1.	<i>Register</i>	Melakukan registrasi dengan mengisi <i>field</i> registrasi. Selanjutnya menekan tombol “Register”	Sistem menyimpan data, dan pengguna dapat melakukan <i>login</i>	Sesuai	Gambar 4.17

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
2.	<i>Login</i>	Melakukan <i>login</i> dengan pengguna mengakses URL aplikasi, kemudian mengisi <i>username</i> dan <i>password</i> . Selanjutnya, menekan tombol “ <i>Sign in</i> ”.	Sistem menampilkan halaman <i>login</i> dan menampilkan halaman <i>dashbaord</i> .	Sesuai	Gambar 4.18
3.	Melakukan konsultasi keluhan	Melakukan proses konsultasi melalui upload audio dari penyimpanan local. Selanjutnya klik “Upload Audio”	Data konsultasi berhasil diproses dan ditampilkan di menu hasil gejala	Sesuai	Gambar 4.21
		Melakukan proses konsultasi melalui <i>record</i> audio. Selanjutnya klik “ <i>Submit Recording</i> ”	Data konsultasi berhasil diproses dan ditampilkan di menu hasil gejala	Sesuai	Gambar 4.21
4.	Melihat laporan hasil konsultasi	Klik pada menu “Laporan Hasil Konsultasi”	Sistem menampilkan laporan hasil konsultasi	Sesuai	Gambar 4.22

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
		Ketika pengguna akan melakukan penambahan keluhan lanjutan, pengguna dapat menambah keluhan tersebut dengan melakukan klik pada <i>button</i> keluhan	Sistem menampilkan halaman untuk mengupload keluhan, dan keluhan berhasil ditambahkan	Sesuai	Gambar 4.23 dan Gambar 4.24

Tabel 4.3 Pengujian *Black box* - Dokter

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
1.	<i>Login</i>	Melakukan <i>login</i> dengan dokter mengakses URL aplikasi, kemudian mengisi <i>username</i> dan <i>password</i> . Selanjutnya, menekan tombol " <i>Sign in</i> ".	Sistem menampilkan halaman <i>login</i> dan menampilkan halaman <i>dashbaord</i> .	Sesuai	Gambar 4.25

No	Butir Uji	Skenario Pengujian	Hasil yang diharapkan	Hasil Pengujian	Reference
2.	Melihat data gejala keluhan pasien	Melihat data gejala keluhan pasien dengan cara klik menu “Data Anamnesis” dan klik <i>icon eye</i>	Sistem menampilkan detail gejala yang telah dikelompokkan pada “ <i>The Sacred Seven of History Taking</i> ”.	Sesuai	Gambar 4.26 dan Gambar 4.27

2. User Acceptance Test

Evaluasi ini dilakukan dengan menyebarkan kuesioner dan diisi oleh pengguna aplikasi. Jumlah responden dari UAT ini adalah 10 (sepuluh) responden yang dipilih berdasarkan peran pengguna pada sistem, terdiri dari 6 (enam) responden yang berperan sebagai admin sistem dan perawat dan 4 (empat) responden yaitu dua dokter bertindak sebagai pakar. Tabel 4.4 merupakan pertanyaan dari UAT (Abraham & Ismail, 2021) (Bastari, Darmansah, & Rakhmadani, 2022) (Alfian, Putra, Rafsanjani, & Witjaksono, 2022).

Proses *User Acceptance Testing* (UAT) di Klinik Margorejo dilakukan oleh berbagai pihak yang terlibat langsung dalam layanan kesehatan, yaitu dokter, perawat, dan beberapa pasien. Pengujian ini bertujuan untuk memastikan bahwa sistem yang dikembangkan dapat berfungsi dengan baik sesuai kebutuhan klinis dan pengalaman pengguna.

Tabel 4.4 Evaluasi *User Acceptance Testing*

No	Pertanyaan	Jawaban Pertanyaan				
		SS	S	N	TS	STS
1.	Apakah informasi yang ditampilkan pada sistem NLP untuk anamnesis penyakit pada anak sudah sesuai dengan kebutuhan?	5	5	0	0	0

No	Pertanyaan	Jawaban Pertanyaan				
		SS	S	N	TS	STS
2.	Apakah penyusunan menu serta isi tiap-tiap menu pada pada sistem NLP untuk anamnesis penyakit pada anak sudah rapih?	5	5	0	0	0
3.	Apakah pemilihan warna pada pada sistem NLP untuk anamnesis penyakit pada anak sudah terlihat nyaman oleh pengguna?	4	4	2	0	0
4.	Apakah dari sisi operasional pada sistem NLP untuk anamnesis penyakit pada anak dapat memberikan kemudahan bagi pengguna?	7	3	0	0	0
5.	Apakah sistem NLP untuk anamnesis penyakit pada anak dapat diakses sesuai dengan hak akses pengguna?	10	0	0	0	0
6.	Apakah proses memuat informasi pada pada sistem NLP untuk anamnesis penyakit pada anak tidak membutuhkan waktu yang lama?	6	2	2	0	0
7.	Apakah pada sistem NLP untuk anamnesis penyakit pada anak sudah berjalan dengan baik?	5	5	0	0	0
8.	Apakah data yang dibutuhkan dalam pada sistem NLP untuk anamnesis penyakit pada anak sudah sesuai?	6	4	0	0	0
9.	Apakah pada sistem NLP tersebut sudah layak dijadikan sistem NLP untuk anamnesis penyakit pada anak?	5	5	0	0	0
Total Bobot		53	33	4	0	0
Skor		265	132	12	0	0
Jumlah Seluruh Skor		409				

Kemudian dilakukan perhitungan nilai x (skor tertinggi) menggunakan persamaan rumus pada (2.6), yang dapat dijabarkan sebagai berikut:

$$\begin{aligned}
 N &= \text{Skor tertinggi} \times (\text{jumlah pertanyaan} \times \text{jumlah responden}) \\
 &= 5 \times (9 \times 10)
 \end{aligned}$$

$$= 450$$

Setelah didapatkan nilai tertinggi, maka dilakukan perhitungan persentase UAT menggunakan persamaan rumus pada (2.6) sebagai berikut:

$$\text{Persentase UAT} = \frac{\text{Total skor}}{x} \times 100 \%$$

$$\text{Persentase UAT} = \frac{409}{450} \times 100 \%$$

$$\text{Persentase UAT} = 90.88 \%$$

Maka didapatkan kesimpulan, total skor yang diperoleh adalah 409 (90.88%) dari skor tertinggi 450 (100%). Hal ini menunjukkan bahwa NLP untuk anamnesis penyakit pada anak termasuk dalam kategori yang “Sangat Layak” karena termasuk kedalam range 81 – 100%.

3. Pengujian NLP

Tabel 4.5 merupakan penjabaran pengujian NLP dengan menggunakan 2 (dua) kasus.

Tabel 4.5 Pengujian NLP

No	Keluhan	Gejala dengan nilai cosine similarity diatas 0,15	Cosine Similarity	Persentase
1	Jadi gini dong anaknya tadi malam itu mulai demam tapi pagi ini udah turun soalnya saya udah kasih penurun panas tapi ya masih anget sama lengan gitu panas dingin rasanya badannya Terus tadi malam itu sempat mual Tapi nggak sampai muntah mungkin karena itu juga dia jadi nafsu makannya berkurang minumnya juga nggak mau banyak soalnya	Nafsu makan turun saat sakit	0,318749467	21,6%
		Sakit perut memburuk setelah makan	0,283705737	
		Sakit tenggorokan	0,255726339	
		Nyeri tenggorokan makin sakit makan pedas	0,194716433	
		Demam naik turun	0,182749799	
		Nyeri perut membaik setelah BAB	0,168839361	
		Nyeri perut	0,164564592	

	rasa mual itu kalau buang air kecil nya lancar terus bab-nya juga lancar nggak ada diare tak tanya	Sakit kepala	0,16348798	
No	Keluhan	Gejala dengan nilai <i>cosine similarity</i> diatas 0,15	<i>Cosine Similarity</i>	Persentase
	perutnya sakit apa enggak Katanya enggak Katanya tenggorokannya sakit tapi perutnya nggak sakit			

Berdasarkan nilai persentase rata-rata *cosine similarity* yang disajikan pada Tabel 4.5, dapat disimpulkan bahwa gejala-gejala yang dihasilkan memiliki kemiripan dengan data *voice* keluhan yang digunakan. Hal ini ditunjukkan oleh hasil evaluasi, dimana keluhan pertama memiliki nilai persentase rata-rata 21.6%, hasil tersebut berada di atas nilai ambang batas yaitu 0,15, yang menunjukkan bahwa gejala yang dianalisis memiliki tingkat kemiripan yang cukup signifikan.

4.2 Pembahasan

Berdasarkan hasil pengujian sistem NLP untuk anamnesis penyakit pada anak, diperoleh bahwa sistem mampu mengenali gejala penyakit berbasis input *voice-to-text* dengan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan Cosine Similarity. Dalam dua kasus uji yang dilakukan, sistem menghasilkan nilai *cosine similarity* di atas ambang batas 0,15, dengan rata-rata tingkat kesamaan sebesar 21,6%. Hal ini menunjukkan bahwa pendekatan yang digunakan memiliki kemampuan awal dalam menemukan pola kesamaan antara input keluhan pasien dan kamus gejala yang telah disusun. Meskipun nilai similarity yang dihasilkan tergolong rendah secara absolut, hal ini tetap merepresentasikan keberhasilan sistem dalam mengekstraksi gejala dari input bahasa alami yang tidak terstruktur. Nilai ini juga menunjukkan efektivitas metode vektorisasi teks dalam mengenali istilah medis secara parsial, sekaligus membuka ruang pengembangan untuk meningkatkan akurasi.

Implikasi dari hasil ini terhadap proses anamnesis adalah bahwa sistem dapat membantu

mempercepat identifikasi gejala awal secara otomatis, mengurangi ketergantungan pada pencatatan manual, serta memberikan struktur data anamnesis yang lebih terstandarisasi. Hal

ini tentunya berkontribusi pada proses diagnosis yang lebih cepat dan akurat, serta mendukung integrasi dengan *Electronic Medical Record* (EMR).

Namun demikian, untuk meningkatkan keakuratan dan keluasan deteksi gejala, diperlukan beberapa langkah perbaikan. Pertama, memperbanyak kamus gejala untuk menangkap variasi ekspresi pengguna, termasuk sinonim dan frasa medis yang sering digunakan dalam praktik klinis. Kedua, mengoptimalkan metode pembentukan kamus secara otomatis menggunakan teknik seperti *clustering* atau *word embedding*. Ketiga, memperluas cakupan penyakit yang ditangani agar tidak hanya terbatas pada penyakit bergejala demam.

Selanjutnya, disarankan agar pengelolaan kamus gejala dilakukan melalui *basis data dinamis* yang terstruktur, menggantikan pendekatan *hardcoded dictionary*. Hal ini akan memudahkan pembaruan data secara periodik, memungkinkan integrasi sistem dengan database klinis, dan mendukung kolaborasi langsung dengan tenaga medis dalam proses validasi gejala.

Dalam pengujian sistem, juga ditemukan adanya potensi bias NLP, terutama ketika input berupa kalimat keluhan yang memiliki struktur tidak umum atau ambigu. Bias ini ditunjukkan melalui ketidaksesuaian hasil ekstraksi dengan gejala yang sebenarnya dimaksud. Hal ini disebabkan karena sistem saat ini hanya menggunakan pendekatan berbasis kemiripan vektor, tanpa mempertimbangkan struktur kalimat atau hubungan sintaktik antar kata.

Sebagai solusi, pengembangan selanjutnya perlu mengintegrasikan teknik parsing NLP lanjutan seperti *Part-of-Speech (POS) Tagging*, *Named Entity Recognition (NER)*, dan *Dependency Parsing*. Pendekatan ini akan memungkinkan sistem untuk memahami konteks dan fungsi kata dalam kalimat, sehingga dapat meningkatkan akurasi pemetaan gejala serta mengurangi bias interpretasi.

Meskipun seluruh kasus dalam penelitian ini mengarah pada gejala umum berupa demam, keunggulan sistem justru terletak pada kemampuannya dalam mengidentifikasi gejala pendukung spesifik yang menyertai demam tersebut. Contohnya, gejala seperti “urine berbau tajam”, “muntah setelah makan”, atau “frekuensi buang air kecil meningkat” berhasil dikenali dan dikelompokkan ke dalam kemungkinan penyakit yang berbeda. Hal ini membuktikan bahwa sistem tidak hanya mengenali demam sebagai gejala utama, tetapi juga mampu mengklasifikasikan demam berdasarkan kombinasi gejala yang menyertainya, seperti ISPA, DBD, demam tifoid, atau infeksi saluran kemih (ISK), dengan menggunakan struktur yang mengacu pada *The Sacred Seven of History Taking*.

Terakhir, untuk memperoleh sistem yang lebih adaptif, pengujian lebih luas terhadap variasi input suara dari berbagai latar belakang bahasa daerah juga diperlukan. Hal ini penting agar sistem dapat diterapkan secara lebih inklusif di berbagai wilayah layanan kesehatan.

BAB 5

Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan proses implementasi dan pembahasan yang telah dilakukan mengenai pengolahan Bahasa alami untuk penyakit pada anak, dapat disimpulkan sebagai berikut:

1. Dalam mengimplementasikan *input* dari *voice to text* yang dapat mengenali gejala yang diderita pasien dari data anamnesis penyakit pada anak dapat dilakukan dengan menggunakan *library SpeechRecognition*. Proses ini diimplementasikan menggunakan bahasa pemrograman *python*, dengan menggunakan *framework flask*, serta MySQL sebagai manajemen databasenya. *User* dapat mengunggah atau merekam audio melalui *website* yang telah dibangun. Kemudian *file audio* tersebut diolah menggunakan *Recognizer* dari *SpeechRecognition* untuk merekam audio dan mengenali teks melalui API *Google Speech Recognition*. Setelah itu, data gejala diambil dari *database* dan diproses lebih lanjut menggunakan teknik *preprocessing*, TF-IDF, serta *cosine similarity*.
2. Hasil anamnesis pada pengolahan Bahasa alami berbasis *website* dalam mengenali penyakit dengan gejala demam pada anak berdasarkan *input voice to text*, user dapat melakukan konsultasi terkait penyakit yang dialami tersebut. Sistem akan memproses data tersebut menggunakan TF-IDF serta *cosine similarity*. Dalam pengujian NLP menggunakan percobaan 2 (dua) kasus menghasilkan gejala-gejala yang dihasilkan memiliki kemiripan dengan data *voice* keluhan yang digunakan karena nilai yang dihasilkan diatas ambang batas 0,15 yaitu dengan nilai persentase rata-rata sebesar 21.6% pada kasus pertama Hasil gejala-gejala tersebut juga dikelompokkan pada “*The Sacred Seven of History Taking*” yang dapat dilihat oleh dokter.

5.2 Saran

Dari hasil penelitian yang telah dilakukan, dapat diberikan saran untuk pengembangan penelitian berikutnya:

1. Dapat memperluas cakupan penyakit yang ditangani, tidak hanya terbatas pada penyakit anak, tetapi juga mencakup rentang usia yang lebih luas. Selain itu, gejala yang digunakan dalam sistem dapat diperluas tidak hanya untuk penyakit demam,

tetapi juga mencakup beragam penyakit lainnya agar sistem lebih komprehensif dalam mendeteksi kondisi pasien.

2. Untuk pengembangan lebih lanjut, disarankan agar proses pengisian kamus gejala dan pengelompokan gejala berdasarkan *The Sacred Seven of History Taking* dilakukan secara dinamis melalui basis data, bukan lagi menggunakan metode *hardcoded*. Pendekatan ini akan meningkatkan fleksibilitas sistem dalam melakukan pembaruan serta mempermudah penambahan atau penghapusan entri sesuai dengan kebutuhan klinis terkini. Dengan basis data yang terstruktur, kolaborasi antara tim pengembang dan tenaga medis dalam memperbarui kamus gejala dapat dilakukan secara terpusat dan lebih efisien, serta memungkinkan integrasi dengan sistem pendukung keputusan klinis secara lebih optimal.
3. Untuk mendukung implementasi *voice-to-text* dalam *Rekam Medis Elektronik* (RME) khususnya di lingkungan klinik pratama, disarankan agar sistem dilengkapi dengan *proof of concept* yang lebih komprehensif. Hal ini mencakup simulasi skenario penggunaan nyata, pengujian dengan berbagai aksen dan variasi bahasa pasien, serta integrasi langsung dengan format standar RME. Dengan adanya *proof of concept* yang lebih lengkap, maka validitas dan kesiapan sistem untuk digunakan di lapangan dapat lebih terukur, sekaligus membuka peluang adopsi yang lebih luas di fasilitas layanan kesehatan tingkat pertama.
4. Dapat dilakukan penerapan algoritma seperti *Naïve Bayes* agar dapat melakukan klasifikasikan gejala-gejala yang diperoleh dari *input voice*. Penerapan tersebut akan membantu dalam memetakan gejala yang dengan penyakit yang digunakan.

Daftar Pustaka

- Abraham, J., & Ismail, I. E. (2021). Unit Testing dan User Acceptance Testing pada Sistem Informasi Pelayan Kategorial Pelayanan Anak. *Repository PNJ*, 1-7.
- Adnan, F., Amelia, I., & Shiddiq, S. (2022). Implementasi Voice Recognition Berbasis Machine Learning. *Edu ElektriKA Journal*, 24-29.
- Alfian, A. N., Putra, M. Y., Rafsanjani, R., & Witjaksono, A. P. (2022). USER ACCEPTANCE TEST TERHADAP APLIKASI AUGMENTED REALITY QUIVERVISION 3D SEBAGAI MEDIA PEMBELAJARAN MEWARNAI. *INFORMATICS FOR EDUCATORS AND PROFESSIONALS*, 6(2), 197-206.
- Al-Harigy, L. M., Al-Nuaim, H. A., Moradpoor, N., & Tan, Z. (2022). Building towards Automated Cyberbullying Detection: A Comparative Analysis. *Computational Intelligence and Neuroscience*, 1-20.
- Almuhaimeed, A., Alhomidi, M. A., & Alenezi, M. N. (2022). A modern semantic similarity method using multiple resources for enhancing influenza detection. *Expert Systems With Applications*, 1-8.
- Anam, M. H., Tamara, D., Widari, W. S., Farida, A., Saori, S., & Eikasari, B. (2023). Pengembangan Sistem Informasi Tenaga Kesehatan Hewan Kabupaten Bondowoso Berbasis Web. *JACIS : Journal Automation Computer Information System*, 3(1), 10-18.
- Andres, & Purnomo, A. S. (2024). Rancang Bangun Sistem Pakar Diagnosa Ispa Pada Apotek Adifarma Metode Certainty Factor. *Jurnal Teknik Informatika*, 1-11.
- Anggoro, N. C., & Akbar, M. (2023). Chatbot Pemilihan Makanan dan Minuman Berdasarkan Kalori menggunakan Natural Language Processing. *Jurnal ForAI*, 29-38.
- Annisa, L., & Kalifia, A. (2024). Analisis Teknik TF-IDF Dalam Identifikasi Faktor-Faktor Penyebab Depresi Pada Individu Depresi Pada Individu. *Gudang Jurnal Multidisiplin Ilmu*, 2, 302-307.
- Arafat, M. D., Simanjuntak, O. S., & Charibaldi, N. (2021). Classification Of Diseases With Fever Symptoms In Children Using K-Nearest Neighbor Genetic Modified Algorithm. *Telematika: Jurnal Informatika dan Teknologi Informasi*, 1-10.

- Arnita, Marpaung, F., Aulia, F., Suryani, N., & Nabila, R. c. (2022). *Computer Vision Dan Pengolahan Citra Digital*. Medan: PUSTAKA AKSARA. Retrieved from <https://digilib.unimed.ac.id/id/eprint/53012/1/Book.pdf>
- Astuti, F. A. (2021). Pemanfaatan Teknologi Artificial Intelligence untuk Penguatan Kesehatan dan Pemulihan Ekonomi Nasional. *Jurnal Sistem Cerdas*, 25-34.
- Athota, L., Shukla, V. K., Pandey, N., & Rana, A. (2020). Chatbot for Healthcare System Using Artificial Intelligence. *International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 1-4.
- Azizah, A., & Rohman, A. (2023). Certainty Factor Pada Sistem Pakar Diagnosis Penyakit Anak. *JUTIS (Jurnal Teknik Informatika Unis)*, 1-12.
- Balqis, F., Anasta, I., & Puspasari, A. (2024). Hubungan Pemakaian Pampers terhadap Kejadian Infeksi Saluran Kemih (ISK) pada Anak di Bawah 2 Tahun di RSUD H. Abdul Manap Kota Jambi. *JOMS*, 1-9.
- Bastari, M. A., Darmansah, & Rakhmadani, D. P. (2022). Sistem Informasi Jasa Cuci Interior Rumah dan Mobil Menggunakan Metode User Acceptance Test. *JURIKOM (Jurnal Riset Komputer)*, 9(2), 305-315.
- Cannavaro, N. (2023). Aplikasi Chatbot untuk Layanan Akademik Menggunakan Platform RASA Open Source dengan Fitur Two Stage Fallback. *Jurnal Ilmu Komputer dan Informatika*, 53-64.
- De, S., Goswami, P., Faujidar, N., & Singh, G. (2024). Gynaecological Disease Diagnosis Expert System (GDDES) Based on Machine Learning Algorithm and Natural Language Processing. *IEEE Access*, 1-12.
- Dewi, S. K., & Sudaryanto, A. (2020). Validitas dan Reliabilitas Kuis Pengetahuan, Sikap dan Perilaku Pencegahan Demam Berdarah. *Prosiding Seminar Nasional Keperawatan Universitas Muhammadiyah Surakarta*, 70-77.
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi dan Informasi (JATI)*, 12, 64-77.
- Diva, T. M., Maulindar, J., Sumarlinda, S., & Binuko, K. P. (2024). Sistem Pakar Diagnosis Penyakit ISPA pada Anak. *Jurnal Teknologi Informasi Komunikasi (e-Journal)*, 1-6.

- Febriansyah, M. A., Nurtanio, I., & Yusuf, M. (2024). Sistem Pakar Berbasis Pohon Keputusan untuk Diagnosa Awal Penyakit Bergejala Demam. *Jurnal Minfo Polgan*, 1-8.
- Gholipour, M. K. (2023). Extracting cancer concepts from clinical notes using natural language processing: a systematic review. *BMC Bioinformatics*, 1-16.
- Haumahu, J. P., Permana, S. D., & Yaddarabullah, Y. (2021). Fake news classification for Indonesian news using Extreme Gradient Boosting (XGBoost). *IOP Conf. Series: Materials Science and Engineering*, 1-9.
- Hikmah, N., Ariyanti, D., & Pratama, F. A. (2022). Implementasi Chatbot Sebagai Virtual Assistant di Universitas Panca Marga Probolinggo menggunakan Metode TF-IDF. *JTIM : Jurnal Teknologi Informasi dan Multimedia*, 133-148.
- Jepriana, I. W. (2020). METODE ITEM-BASED COLLABORATIVE FILTERING UNTUK MODEL SISTEM REKOMENDASI KONSENTRASI PENJURUSAN DI STMIK STIKOM BALI. *Jurnal Teknologi Informasi dan Komputer*.
- Jiang, Z., Gao, B., He, Y., Han, Y., Doyle, P., & Zhu, Q. (2021). Text classification using novel term weighting scheme-based improved tf-idf for internet media reports. *Hindawi: Mathematical Problems in Engineering*, 1-30.
- Jimmy, & Suwitno. (2023). Analisis Dan Perancangan Sistem Informasi Penjualan Online Berbasis Web Menggunakan Pengujian User Acceptance Test (UAT) Pada PT. Putra Jarum Mas Mandiri. *AKSELERATOR*, 4, 11-20.
- Juwita, Hidayurrahmi, Adista, A., & Khairunnisa. (2023). Tinjauan Biomolekular Patofisiologi Dan Penegakan Diagnosis. *Jurnal Sinaps*, 1-16.
- Laksono, T. P., Hidayatullah, A. F., & Ratnasari, C. I. (2018). Speech to Text of Patient Complaints for Bahasa Indonesia . *IEEE*, 79-85.
- Lee, B., Chung, H. J., Kang, H. M., Kim, D. K., & Kwak, Y. H. (2021). Development and validation of machine learning-driven prediction model for serious bacterial infection among febrile children in emergency departments. *PLOS ONE*, 1-12.
- Levani, Y., & Prasty, A. D. (2020). Demam Tifoid : Manifestasi Klinis, Pilihan Terapi Dan Pandangan Dalam Islam. *AL-IQRA MEDICAL JOURNAL : JURNAL BERKALA ILMIAH KEDOKTERAN*, 10-16.
- Lusiyanti, Yetri, M., Setiawan, F., Affandi, E., & Laia, J. (2023). Implementasi Expert System Menggunakan Metode Case Based Reasoning Mendiagnosa Penyakit

- Typhoid Fever Pada Anak. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, 517-527.
- Mahendra, M. H., Murdiansyah, D. T., & Lhaksana, K. M. (2023). Analisis Sentimen Tweet COVID-19 Menggunakan Metode K-Nearest Neighbors dengan Ekstraksi Fitur TF-IDF dan CountVectorizer. *Dike : Jurnal Ilmu Multidisiplin*, 37-43.
- Manoharan, S., & Sathish. (2020). Patient Diet Recommendation System Using K Clique and Deep learning Classifiers. *Journal of Artificial Intelligence and Capsule Networks*, 1-10.
- Mariyani, U. D., Setyaningsih, W., & Agustina, R. (2022). Pengembangan Sistem Koreksi Jawaban Esai Otomatis Menggunakan Naive Bayes Dan Pengujian Menggunakan User Acceptance Test (UAT). *Jurnal Terapan Sains & Teknologi*, 61-73.
- Maulidina, M. K. (2020). ANALISIS SENTIMEN KOMENTAR WARGANET TERHADAP POSTINGAN INSTAGRAM MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN TF-IDF.
- Merryton, A. R., & Augasta, M. G. (2022). AN EMPIRICAL ANALYSIS OF FAKE NEWS DETECTION WITH NLP AND MACHINE LEARNING TECHNIQUES . *Advanced Engineering Science*, 1351-1365.
- Min, M., Lee, J. J., & Lee, K. (2022). Detecting Illegal Online Gambling (IOG) Services in the Mobile Environment. *Security and Communication Networks*, 1-12.
- Nshimiyimana, Y., & Zhou, Y. (2022). Analysis of risk factors associated with acute respiratory infections among under-five children in Uganda. *BMC Public Health*, 1-10.
- Nugraheni, M., Widodo, & Sari, I. P. (2021). A Case-Based Reasoning for Detection Coronavirus (Covid-19) Using Cosine Similarity. *Conference on Broad Exposure to Science and Technology*, 1-6.
- Nurachman, M. T., & Fitrianingrum, I. (2022). Pengaruh Komunikasi Efektif Terhadap Kejadian Tidak Diharapkan (KTD). *Jurnal Cerebellum*, 12-15.
- Pathak, J. (2023, September 20). *Speech Recognition in Python*. Retrieved from Medium: <https://medium.com/naukri-engineering/speech-recognition-in-python-fcda027a97a1>
- Perdana, K., Pricillia, T., & Zulfachmi. (2021). Optimasi Text Blob Menggunakan Support Vector Machine untuk Analisis Sentimen (Studi Kasus Layanan Telkomsel). *Bangkit Indonesia*, 1-3.

- Pujoseno, J. (2020, September 17). *Speech to Text pada Python*. Retrieved from Thinkstudio21: <https://thinkstudoo.blogspot.com/2020/09/speech-to-text-pada-python.html>
- Purba, M., Penmaley, F., & Panjaitan, J. (2024). Dugaan Pelanggaran Disiplin Terbanyak Akibat Kurangnya Komunikasi Dokter dan Pasien. *Jurnal Global Ilmiah*, 337-340.
- Rahmadani, S., & Tasrif, E. (2024). Implementasi Natural Language Processing (NLP) pada Layanan Pengelolaan Surat Kantor Camat Bukit Barisan. *JURNAL TEKNIK KOMPUTER DAN INFORMATIKA*, 19-24.
- Rasyid, R. A., & Ningsih, D. H. (2024). Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata. *Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)*, 1-9.
- Ratnasari, C. I., Kusumadewi, S., & Rosita, L. (2014). Model Natural Language Processing untuk Perumusan Keluhan Pasien. *Seminar Nasional Informatika Medis (SNIMed)*.
- Razzaq, A., Abbas, T., Hashim, S., Qadri, S., Mumtaz, I., Saher, N., . . . Nawaz, S. A. (2022). Extraction of psychological effects of COVID-19 pandemic through topic-level sentiment dynamics. *Complexity*, 1-10.
- Rianti, A. (2024). MACHINE LEARNING JOURNAL ARTICLE RECOMMENDATION SYSTEM USING CONTENT BASED FILTERING. *JUTI*.
- Riany, A., & Testiana, G. (2023). Penerapan Data Mining untuk Klasifikasi Penyakit Stroke Menggunakan Algoritma Naïve Bayes. *Jurnal Saintekom : Sains, Teknologi, Komputer dan Manajemen*, 42-54.
- Rohi, R. A., Pote, J. Y., & Talakua, A. C. (2022). Perancangan Dan Implementasi Sistem Informasi Perpustakaan Berbasis Website Menggunakan Metode Waterfall Di Sd Masehi Kambaniru 2. *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, 10, 63-70.
- Rumapea, H. (2021). Deteksi Kemiripan Artikel Melalui Keywords dengan Metode Fuzzy String Matching dalam Natural Language Processing. *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntans*, 60-67.
- Samosir, F. V., Toba, H., & Ayub, M. (2022). BESklus : BERT Extractive Summarization with K-Means Clustering in Scientific Paper. *Jurnal Teknik Informatika dan Sistem Informasi*.

- Sarwar, T., & Mamun, A. (2022). Prediction of Dengue using Machine Learning Algorithms: Case Study Dhaka. *International Conference on Electrical, Computer Telecommunication Engineering (ICECTE)*, 1-6.
- Schopf, T. (2023). Zero-Shot vs. Similarity-Based Text Classification. *Medium*. Retrieved from <https://towardsdatascience.com/zero-shot-vs-similarity-based-text-classification-83115d9879f5>
- Sebayang, F. A. (2024). Designing an Expert System Application to Diagnose Dengue Fever Patients Using the Naïve Bayes Method. *Bulletin of Information Technology (BIT)*, 31-36.
- Sholikhah, I., Harjanta, A., & Latifah, K. (2023). Machine Learning Untuk Deteksi Berita Hoax Menggunakan BERT. *Seminar Nasional Informatika – FTI UPGRIS*.
- Sihombing, Y. R., Siska, F., Zebua, K. C., & Nasution, D. W. (2022). Seminar Evaluasi Penggunaan Antibiotik Ciprofloxacin pada Pasien Infeksi Saluran Kemih (ISK) Di Instalasi Rawat Inap Rumah Sakit Grandmed Lubuk Pakam. *Jurnal Pengabdian Kepada Masyarakat*, 1-6.
- Sihombong, D. O. (2022). Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis. *Jurnal Sistem Komputer dan Informatika (JSON)*, 396-406.
- Singh, R., & Singh, S. (2021). Text Similarity Measures in News Articles by Vector Space Model Using NLP. *J. Inst. Eng. India Ser. B*, 329-338.
- SpeechRecognition. (2022). *SpeechRecognition 3.11.0*. Retrieved from Python Software Foundation: <https://pypi.org/project/SpeechRecognition/>
- Sundari, S. S., Sarmidi, & Nuraropah, A. F. (2021). Sistem Pakar Diagnosis Penyakit Gangguan Kehamilan Menggunakan Metode Forward Chaining Dan Certainty Factor Berbasis Web. *IT Journal*, 1-12.
- Susanti, E., & Bismin. (2021). Pengujian Web E-Katalog PT Pilar Cipta Solusi Integratika (PISCI) Menggunakan Metode Black-Box. *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2021*, 162-168.
- Tandirilambun, Rini, M., Filiana, A., & Virginia, G. (2024). Pengembangan Sistem Manajemen Rencana Strategis Studi Kasus Yayasan SATUNAMA Yogyakarta. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 183-188.
- Toruan, A. M. (2024). Penggunaan NLP dalam Analisis Sentimen untuk Meningkatkan Kepuasan Pelanggan pada Pengguna E-commerce: Lazada.

- <https://jurnal.larisma.or.id/index.php/SAINTEK/article/view/452%0Ahttps://jurnal.larisma.or.id/index.php/SAINTEK/article/download/452/310>, 18-20.
- Wulandari, Nofiyani, & Hasugian, H. (2023). User Acceptance Testing (UAT) Pada Electronic Data Preprocessing Guna Mengetahui Kualitas Sistem. *JMIK (JURNAL MAHASISWA ILMU KOMPUTER)*, 20-27.
- Yehuala, T. Z., Agimas, M. C., Derseh, N. M., Wubante, S. M., Fente, B. M., Yismaw, G. A., & Tesfie, T. K. (2024). Machine learning algorithms to predict healthcare-seeking behaviors of mothers for acute respiratory infections and their determinants among children under five in sub-Saharan Africa. *frontiers*, 1-12.
- Yuwono, P. (2018). Pengolahan Bahasa Alami untuk Penerjemahan Keluhan Pasien pada Anamnesis dalam Bentuk Teks atau Suara. Retrieved from <https://dspace.uui.ac.id/handle/123456789/7594>
- Zhang, A. (2017). *SpeechRecognition* 3.14.2. Retrieved from pypi: <https://pypi.org/project/SpeechRecognition/>
- Zia, A., Aziz, M., Popa, I., Khan, S. A., Hamedani, A. F., & Asif, A. R. (2022). Artificial Intelligence-Based Medical Data Mining. *Journal of Personalized Medicine (MDPI)*.

LAMPIRAN

Lampiran A. Perancangan

Berikut penjelasan metode pengembangan sistem dari *waterfall* yang digunakan dalam penelitian ini:

1. *Requirement Definition*

Requirement definition merupakan tahap mendefinisikan analisis kebutuhan yang dibutuhkan dalam pembuatan sistem. Dalam penelitian ini melakukan analisis kebutuhan fungsional dan non fungsional pada *website* anamnesis penyakit pada anak terhadap gejala batuk berdasarkan *input* dari *voice to text*. Analisis kebutuhan fungsional menghasilkan fitur yang ada didalam sistem yang nantinya digunakan untuk perancangan desain sistem. Sedangkan, analisis kebutuhan non-fungsional menghasilkan spesifikasi perangkat lunak dan perangkat keras untuk membangun aplikasi.

A. Kebutuhan Fungsional

Kebutuhan fungsional merupakan fungsi yang akan ada pada sistem yang dibangun. Kebutuhan fungsional dalam penelitian ini sebagai berikut:

1) Admin

Pada aktor admin dapat melakukan beberapa aktivitas sebagai berikut:

- a) Melakukan *login*
- b) Mengelola data *user*
- c) Mengelola data gejala
- d) Melihat laporan hasil konsultasi

2) Dokter

Pada aktor dokter dapat melakukan aktivitas sebagai berikut:

- a) Melakukan *login*
- b) Melihat data anamnesis

3) Pengguna

Pada aktor pengguna dapat melakukan aktivitas sebagai berikut:

- a) Melakukan registrasi
- b) Melakukan *login*
- c) Melakukan konsultasi terhadap keluhan
- d) Melihat hasil gejala

B. Kebutuhan Non-Fungsional

Analisis kebutuhan non-fungsional merupakan tahap mendefinisikan kebutuhan perangkat dalam pembuatan *website* anamnesis penyakit pada anak terhadap gejala batuk berdasarkan *input* dari *voice to text*, kebutuhan non-fungsional terdiri dari kebutuhan perangkat lunak dan kebutuhan perangkat keras.

Berikut ini merupakan kebutuhan perangkat keras dan perangkat lunak yang digunakan dalam pembuatan *website* anamnesis penyakit pada anak terhadap gejala batuk berdasarkan *input* dari *voice to text*.

1) Kebutuhan Perangkat Keras

Perangkat keras adalah komponen fisik yang memiliki spesifikasi untuk menjalankan sistem. Berikut ini perangkat keras yang digunakan untuk membuat *website* anamnesis penyakit pada anak terhadap gejala batuk berdasarkan *input* dari *voice to text* sebagai berikut:

Perangkat keras dalam pembuatan sistem yang digunakan yaitu:

- a) Komputer atau Laptop
- b) *Processor Intel Core i3*
- c) *RAM 4 GB*
- d) *Harddisk 500 GB*
- e) *Mouse dan Keyboard*
- f) *Internet Connection*

2) Kebutuhan Perangkat Lunak

Kebutuhan perangkat lunak (*Software*) merupakan aplikasi yang dibutuhkan untuk dapat membangun sistem yang akan dibuat. Berikut ini perangkat lunak dalam pembuatan sistem yang digunakan sebagai berikut:

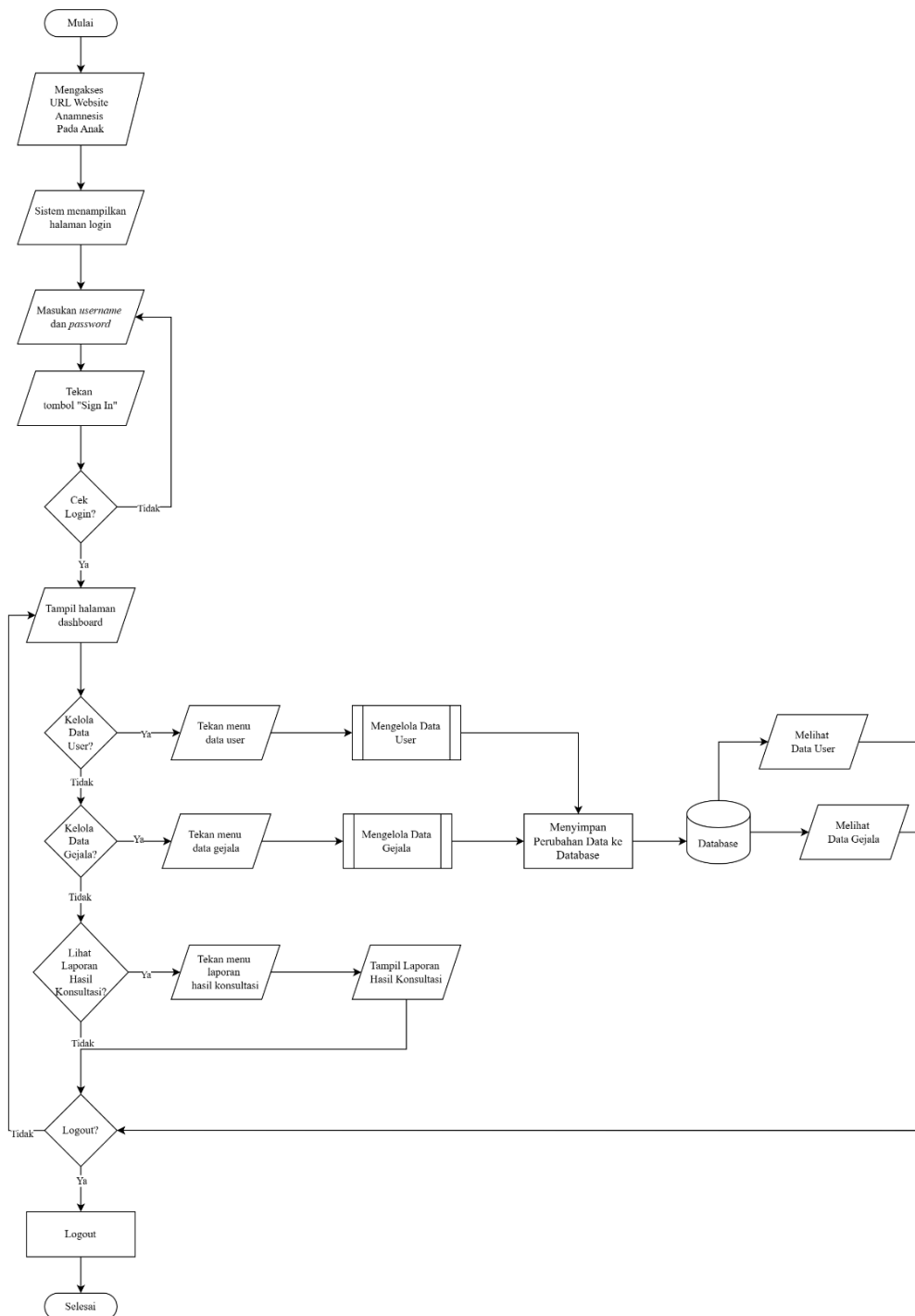
- a) *Operating System: Windows 10*
- b) *Database Management System: MySQL*
- c) *Web Server: XAMPP*
- d) *Web Browser: Google Chrome*
- e) *Text Editor: Visual Studio Code*
- f) *Draw.io*
- g) *Balsamic Mockup*
- h) *Microsoft Word*
- i) *Microsoft Excel*

2. *System and Software Design*

Pada tahap ini akan dilakukan pembuatan perancangan sistem dari hasil analisa kebutuhan yang telah dilakukan bertujuan untuk mempermudah proses pembuatan sistem. Perancangan sistem yang dibuat menggunakan *flowchart*, UML seperti diagram *use case diagram* dan *class diagram* serta perancangan *user interface*.

A. *Flowchart*

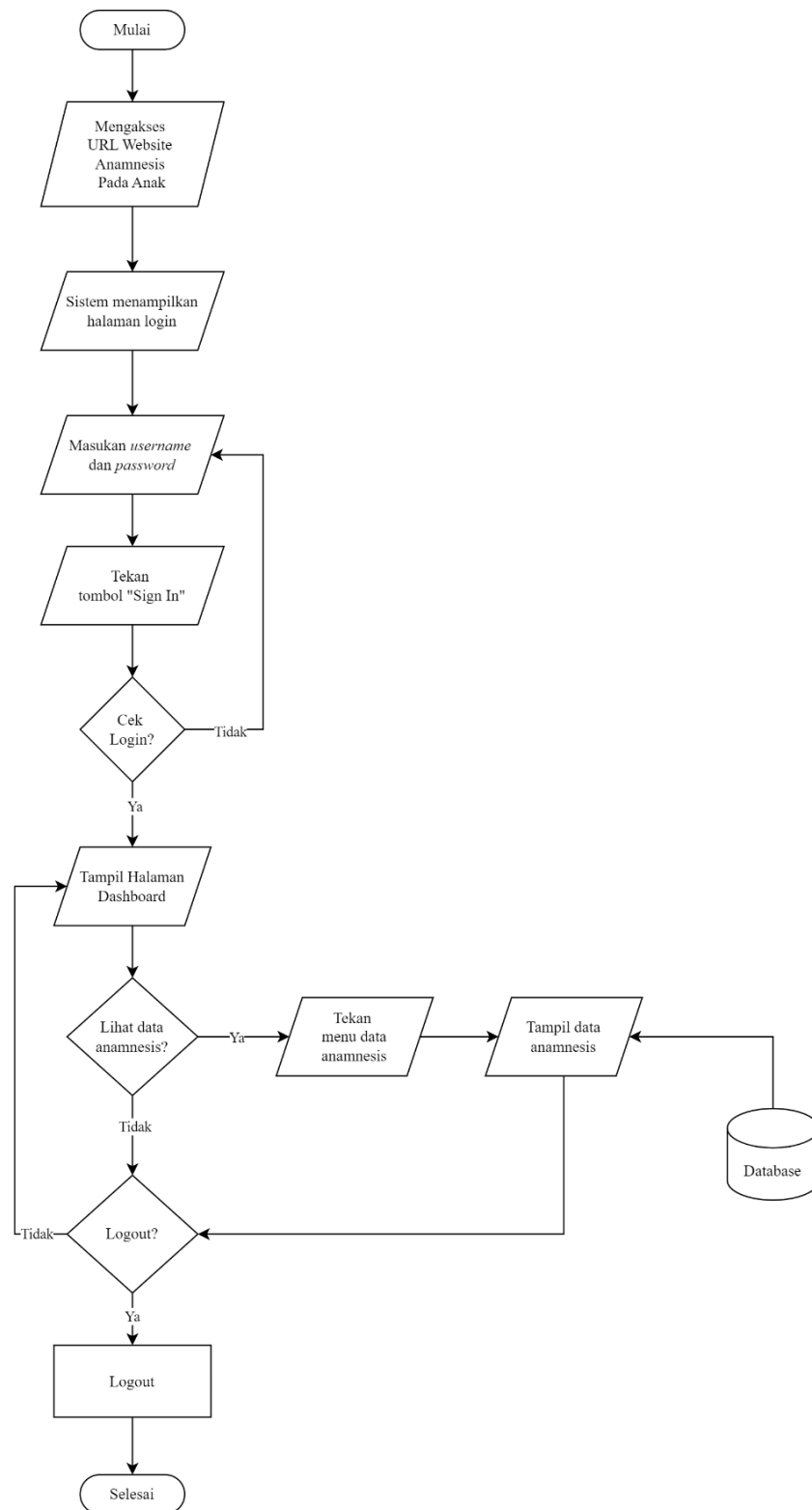
Flowchart yang digunakan dalam pembuatan *website* pengolahan bahasa alami untuk anamnesis pada anak terdiri dari *flowchart* admin, dokter dan pengguna. *Flowchart* admin dimulai dengan admin mengakses URL *website*, maka sistem menampilkan halaman *login* berupa *form* yang terdiri dari kolom *username* dan *password*. Setelah mengisi data *login*, sistem mengecek apakah data tersebut terdaftar atau tidak. Jika data terdaftar akan menampilkan halaman *dashboard*, namun jika tidak terdaftar dapat melakukan *login* kembali. Pada halaman *dashboard*, admin dapat mengelola data *user*, data gejala dan melihat laporan hasil gejala. Semua perubahan data disimpan kedalam *database*. Setelah melakukan aktivitas, admin dapat *logout* dengan menekan tombol *logout*. *Flowchart* admin dapat dilihat pada Gambar A.1.



Gambar A.1 *Flowchart Admin*

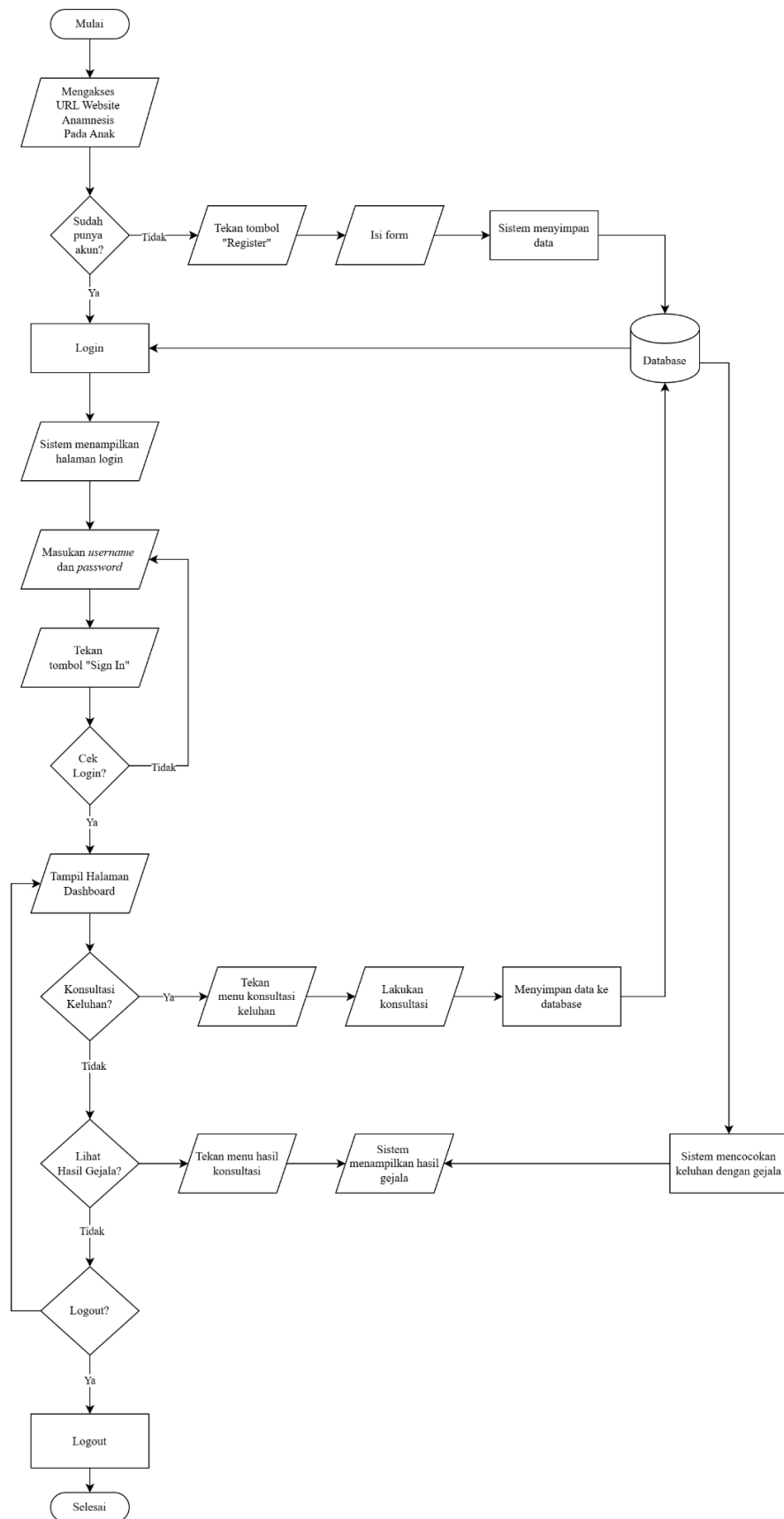
Flowchart dokter dimulai dengan dokter mengakses URL *website*, maka sistem menampilkan halaman *login* berupa *form* yang terdiri dari kolom *username* dan *password*. Setelah mengisi data *login*, sistem mengecek apakah data tersebut terdaftar atau tidak. Jika data terdaftar akan menampilkan halaman *dashboard* dokter, namun jika tidak terdaftar dapat melakukan *login* kembali. Pada halaman *dashboard*, dokter dapat melihat data anamnesis yang dirasakan pengguna/pasien.

Semua perubahan data disimpan kedalam *database*. Setelah melakukan aktivitas, dokter dapat *logout* dengan menekan tombol *logout*. *Flowchart* dokter dapat dilihat pada Gambar A.2.



Gambar A.2 *Flowchart* Dokter

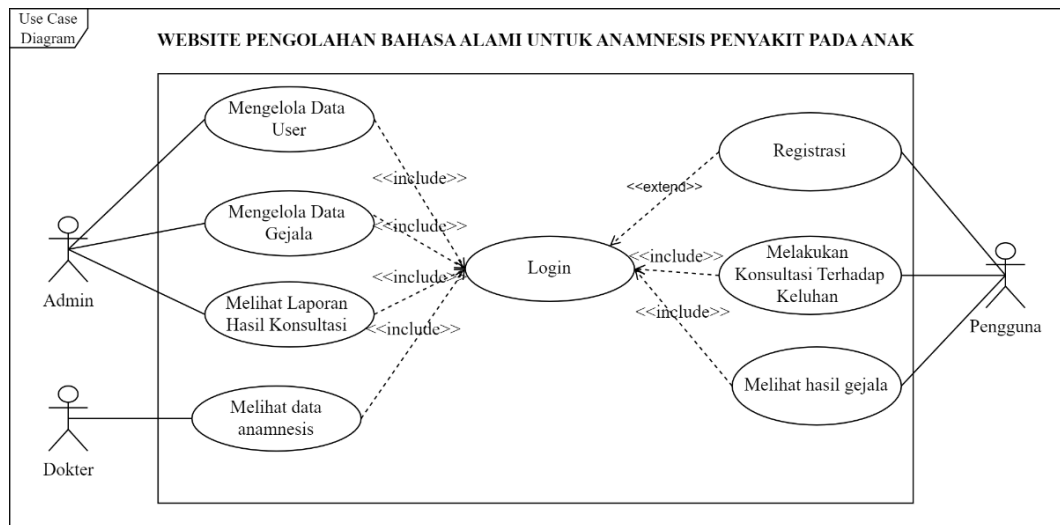
Flowchart pengguna dimulai dengan pengguna mengakses URL *website*, maka sistem menampilkan halaman utama dan untuk melakukan *login* pilih menu *login* yang menampilkan *form* yang terdiri dari kolom *username* dan *password*. Namun, ketika belum mempunyai akun dapat menekan tombol “Registrasi Akun”. Pada halaman registrasi akun terdapat *form* pengisian daftar akun yang harus diisi dengan benar dan lengkap, setelah selesai tekan tombol “Register”. Maka, sistem mengecek data dan menyimpannya pada *database*. Pengguna yang telah memiliki akun dapat langsung *login*. Setelah mengisi data *login*, sistem mengecek apakah data tersebut terdaftar atau tidak. Jika data terdaftar akan menampilkan halaman *dashboard* pengguna, namun jika tidak terdaftar dapat melakukan *login* kembali. Pada halaman *dashboard*, pengguna dapat melakukan konsultasi terhadap keluhan yang dialaminya dengan memilih menu “Konsultasi Keluhan”. Maka, sistem menampilkan halaman konsultasi keluhan yang dapat digunakan pengguna dengan menekan tombol “Voice”. Semua perubahan data disimpan kedalam *database*. Hasil gejala dapat dilihat pada menu “Hasil Gejala”. Setelah melakukan aktivitas, pengguna dapat *logout* dengan menekan tombol *logout*. *Flowchart* pengguna dapat dilihat pada Gambar A.3.



Gambar A.3 Flowchart Pengguna

B. Use Case Diagram

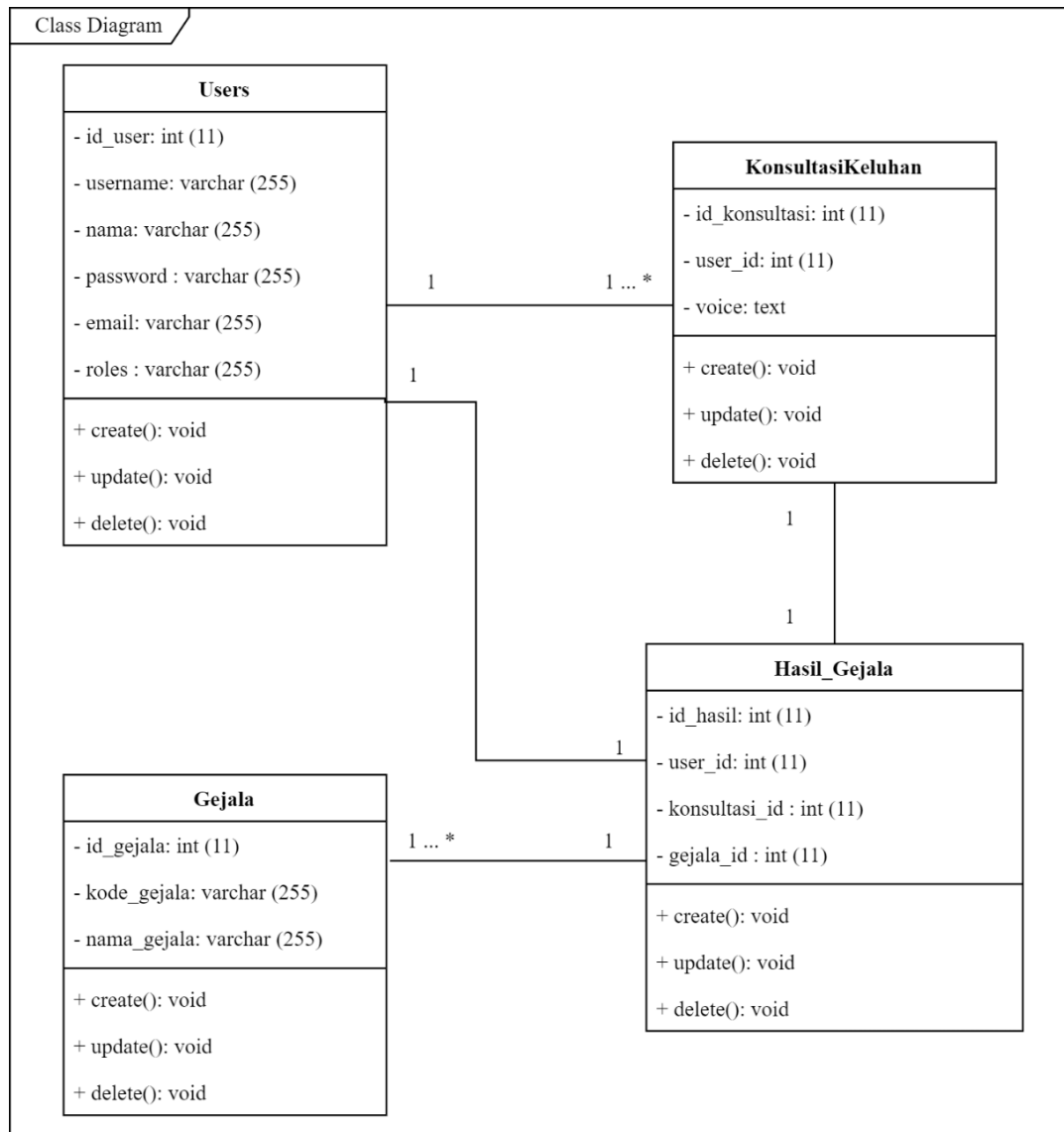
Use case diagram yang digunakan terdiri dari 3 aktor yaitu admin, dokter dan pengguna yang dapat dilihat pada Gambar A.4.



Gambar A.4 Use Case Diagram

C. Class Diagram

Class diagram yang digunakan terdiri dari *class* Users, KonsultasiKeluhan, Gejala dan HasilGejala. Setiap *class* memiliki atribut dan *method* serta memiliki relasi dengan *class* lainnya. Relasi yang digunakan yaitu *One to One (1:1)* yang berarti terdapat satu data dari entitas A dihubungkan dengan satu data entitas B. Relasi berikutnya yaitu *One to Many* yang berarti terdapat satu data pada entitas A memiliki banyak data pada entitas B. *Class diagram* yang digunakan dapat dilihat pada Gambar A.5.



Gambar A.5 Class Diagram

D. Perancangan User Interface

Perancangan *user interface* terdiri dari halaman untuk aktor admin, dokter dan pengguna. Pada perancangan *user interface website* pengolahan bahasa alami untuk anamnesis pada anak terdapat halaman *home* yang dapat dilihat pada Gambar A.6. Pada halaman *home* menampilkan informasi tentang *website* dan terdapat menu *login*.



Gambar A.6 Perancangan Halaman *Home*

Ketika pengguna mengakses tombol “Registrasi Akun”, maka sistem menampilkan halaman registrasi akun yang berisi form untuk mendaftarkan akun dengan mengisi nama lengkap, *username*, *email* dan *password*. Setelah diisi semua dapat menekan tombol “Register”. Perancangan halaman registrasi dapat dilihat pada Gambar A.7.

Gambar A.7 Perancangan Halaman Registrasi

Pada perancangan halaman *login* digunakan untuk semua aktor melakukan *login* dengan mengisi data *login* yaitu *username* dan *password*. Kemudian, sistem mengecek data *login*. Jika terdaftar akan diarahkan ke halaman *dashboard* masing-masing. Namun, jika gagal *login* diminta untuk melakukan *login* kembali. Perancangan halaman *login* dapat dilihat pada Gambar A.8.

Gambar A.8 Perancangan Halaman *Login*

1) Perancangan *User Interface* Admin

Perancangan *user interface* admin terdiri dari menu *dashboard*, data *user*, data gejala, dan laporan hasil konsultasi.

a) Perancangan Halaman *Dashboard* Admin

Perancangan halaman *dashboard* merupakan halaman yang ditampilkan ketika admin berhasil melakukan *login*. Perancangan halaman *dashboard* dapat dilihat pada Gambar A.9.

Gambar A.9 Perancangan Halaman *Dashboard* Admin

b) Perancangan Halaman Data User

Perancangan halaman *user* merupakan halaman yang menampilkan daftar *user* yang ada dalam sistem. Pada halaman *user* dapat melakukan tambah, *edit* dan hapus data *user*. Perancangan halaman *user* dapat dilihat pada Gambar A.10.

No.	Nama Lengkap	Username	Email	Role	Aksi
1	Admin	admin	admin@gmail.com	Admin	☒ Edit ☐ Delete
2	Sintiya	dokter	dokter@gmail.com	Dokter	☒ Edit ☐ Delete
3	Risma	risma	risma@gmail.com	Pengguna	☒ Edit ☐ Delete

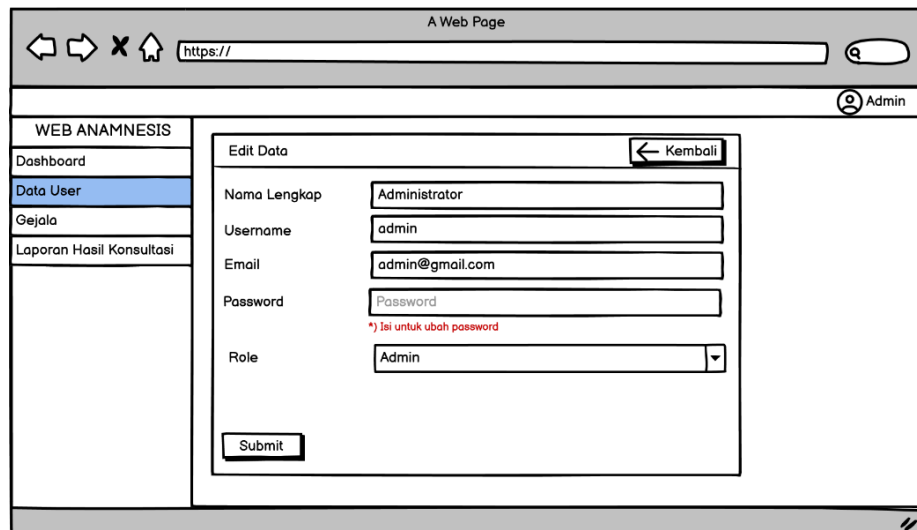
Gambar A.10 Perancangan Halaman Data User

Perancangan halaman tambah *user* merupakan halaman yang digunakan admin untuk menambahkan data *user*. Perancangan halaman tambah *user* dapat dilihat pada Gambar A.11.

No.	Nama Lengkap	Username	Email	Role	Aksi
1	Admin	admin	admin@gmail.com	Admin	☒ Edit ☐ Delete
2	Sintiya	dokter	dokter@gmail.com	Dokter	☒ Edit ☐ Delete
3	Risma	risma	risma@gmail.com	Pengguna	☒ Edit ☐ Delete

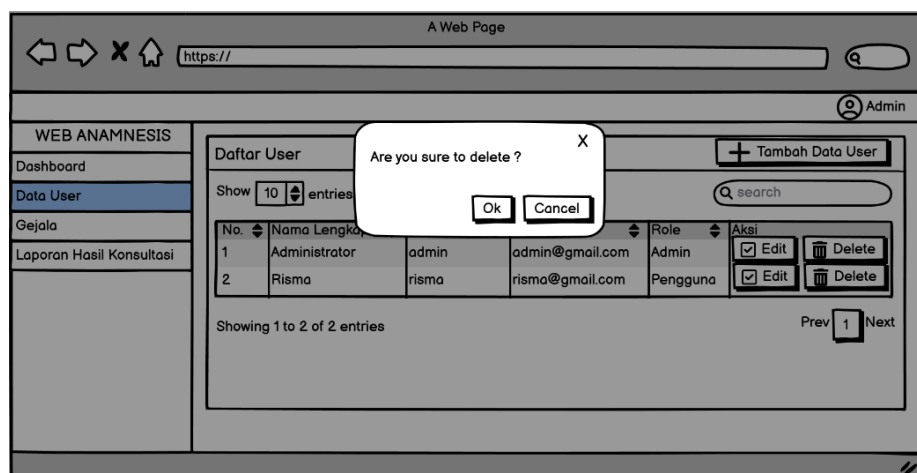
Gambar A.11 Perancangan Halaman Tambah Data User

Perancangan halaman *edit user* merupakan halaman yang digunakan admin untuk mengubah data *user*. Perancangan halaman *edit user* dapat dilihat pada Gambar A.12.



Gambar A.12 Perancangan Halaman *Edit Dat User*

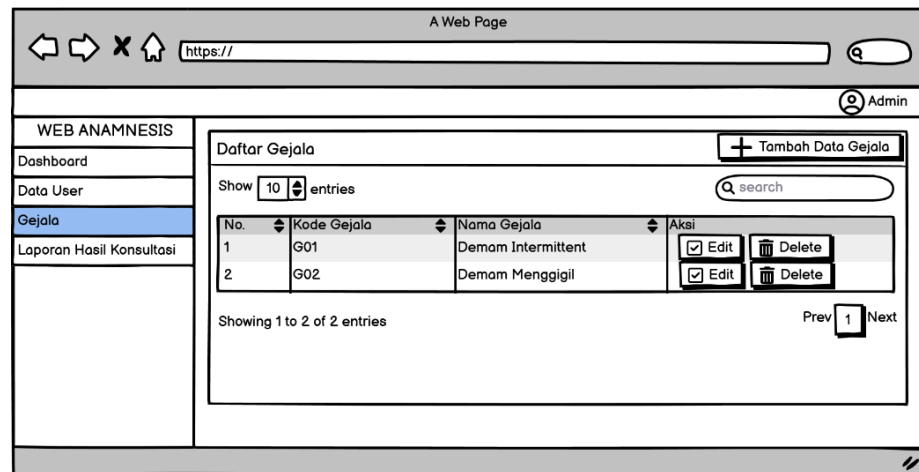
Perancangan halaman hapus *user* merupakan halaman yang digunakan admin untuk menghapus data *user*. Perancangan halaman hapus *user* dapat dilihat pada Gambar A.13.



Gambar A.13 Perancangan Halaman Hapus Data *User*

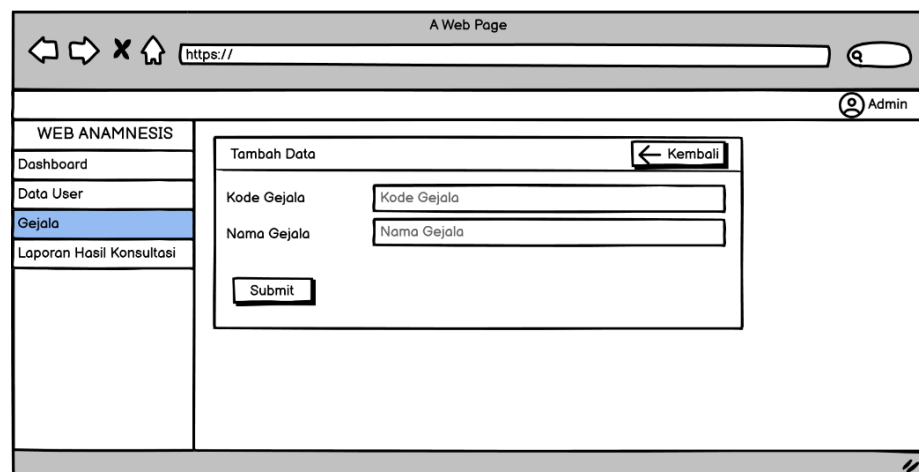
c) Perancangan Halaman Data Gejala

Perancangan halaman gejala merupakan halaman yang menampilkan daftar gejala yang ada dalam sistem. Pada halaman gejala dapat melakukan tambah, *edit* dan hapus data gejala. Perancangan halaman gejala dapat dilihat pada Gambar A.14.



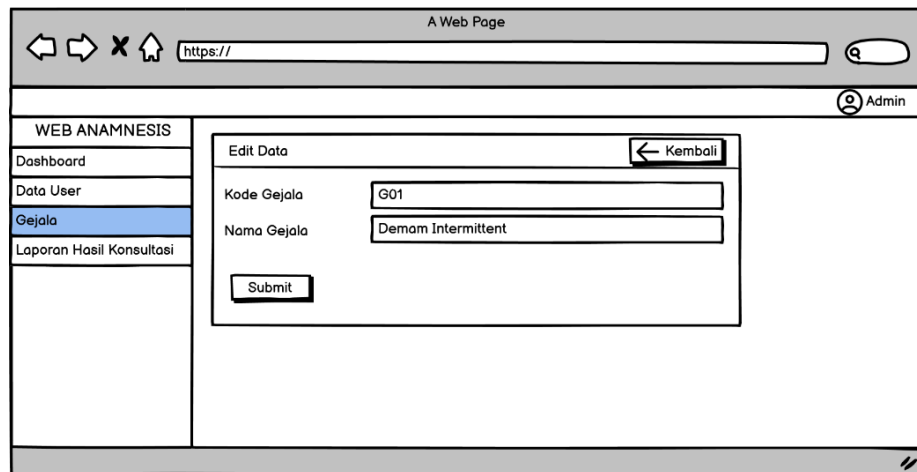
Gambar A.14 Perancangan Halaman Data Gejala

Perancangan halaman tambah gejala merupakan halaman yang digunakan admin untuk menambahkan data gejala. Perancangan halaman tambah gejala dapat dilihat pada Gambar A.15.



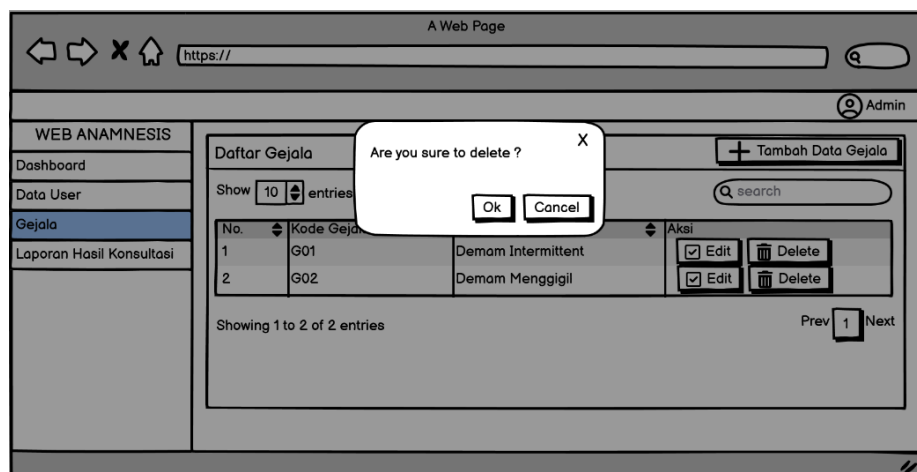
Gambar A.15 Perancangan Halaman Tambah Data Gejala

Perancangan halaman *edit* gejala merupakan halaman yang digunakan admin untuk mengubah data gejala. Perancangan halaman *edit* gejala dapat dilihat pada Gambar A.16.



Gambar A.16 Perancangan Halaman *Edit Data Gejala*

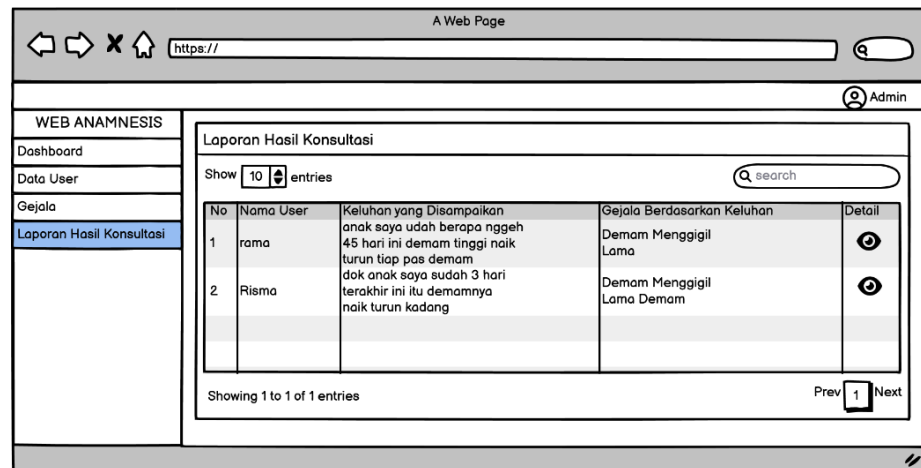
Perancangan halaman hapus gejala merupakan halaman yang digunakan admin untuk menghapus data gejala. Perancangan halaman hapus gejala dapat dilihat pada Gambar A.17.



Gambar A.17 Perancangan Halaman Hapus Data Gejala

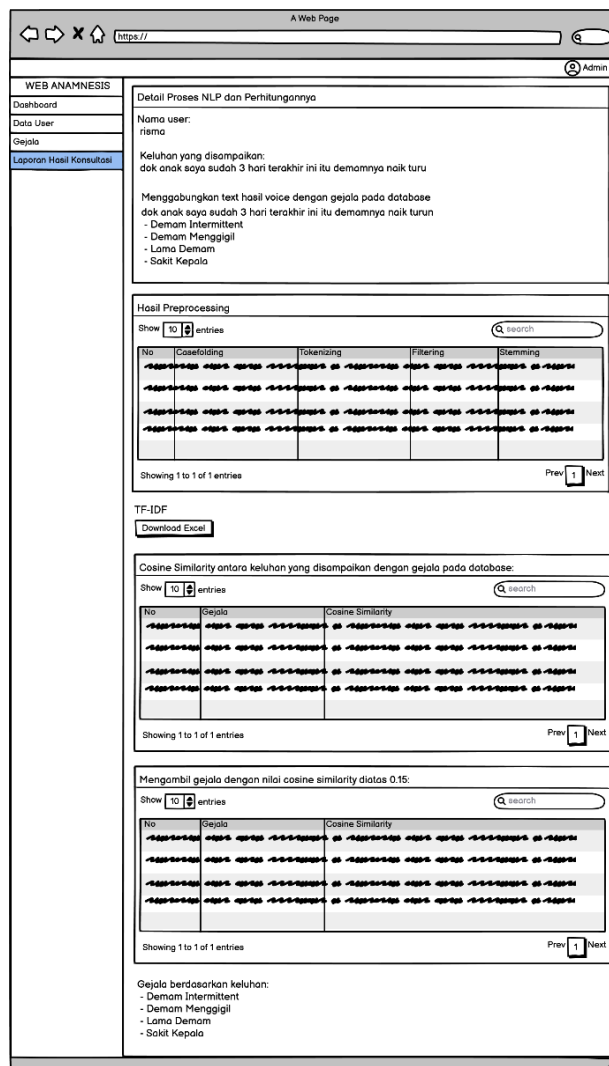
d) Perancangan Halaman Laporan Hasil Konsultasi

Perancangan halaman laporan hasil konsultasi digunakan admin untuk melihat informasi hasil konsultasi yang telah dilakukan oleh pengguna. Perancangan halaman laporan hasil konsultasi dapat dilihat pada Gambar A.18.



Gambar A.18 Perancangan Halaman Laporan Hasil Konsultasi

Selain itu, admin dapat melihat proses NLP dengan melakukan klik pada tombol icon eye, dan hasil tampilan tersebut dapat dilihat pada Gambar A.19.



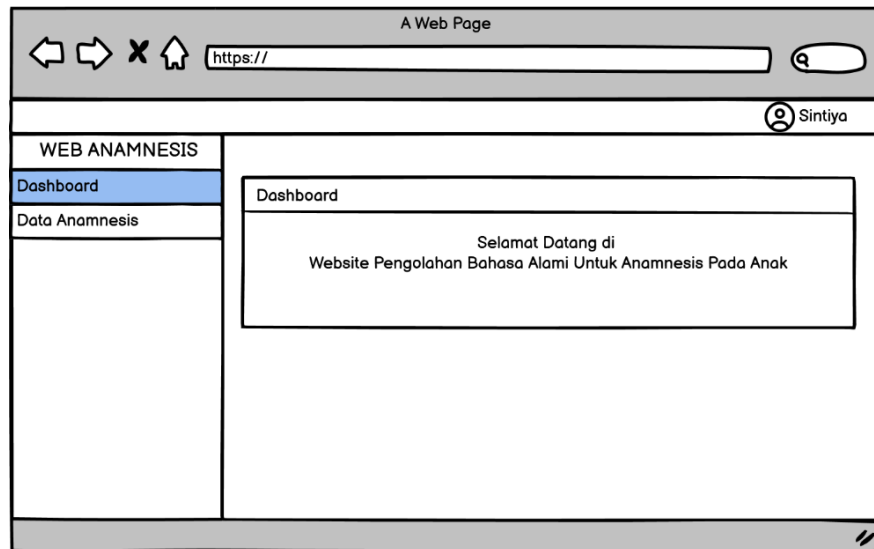
Gambar A.19 Perancangan Halaman Proses NLP

2) Perancangan *User Interface* Dokter

Perancangan *user interface* dokter terdiri dari menu *dashboard* dan data anamnesis.

a) Perancangan Halaman *Dashboard* Dokter

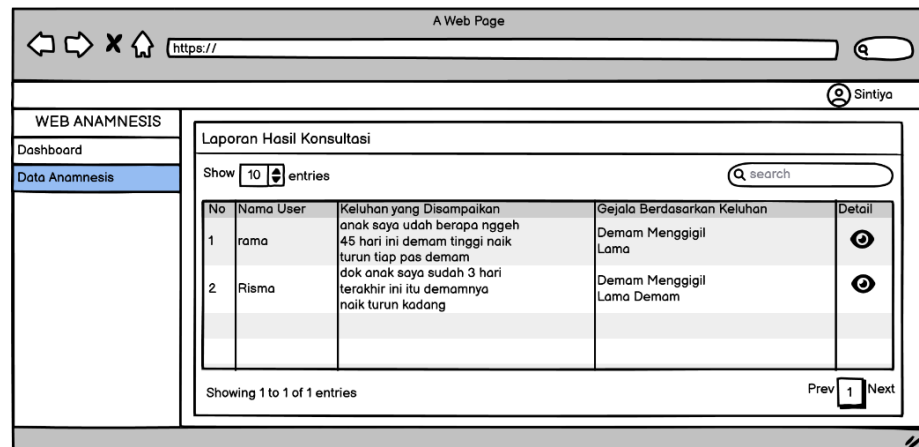
Berikut pada Gambar A.20 merupakan rancangan halaman *dashboard* dokter.



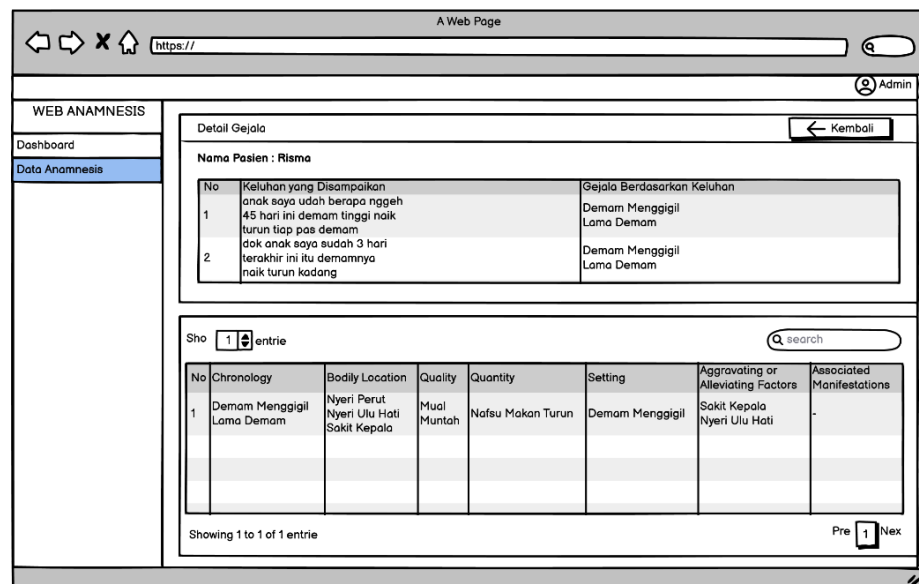
Gambar A.20 Perancangan Halaman *Dashboard* Dokter

b) Perancangan Halaman Data Anamnesis

Perancangan halaman data anamnesis merupakan halaman yang menampilkan data gejala dari keluhan hasil dari algoritma *TF-IDF-cosine similarity*, dari hasil tersebut dokter dapat melihat hasil anamnesis dengan pengelompokkan pada “*The Sacred Seven of History Taking*”. Perancangan halaman data anamnesis dapat dilihat pada Gambar A.21 dan Gambar A.22 merupakan halaman untuk melihat pengelompokkan gejala pada “*The Sacred Seven of History Taking*”.



Gambar A.21 Perancangan Halaman Data Anamnesis



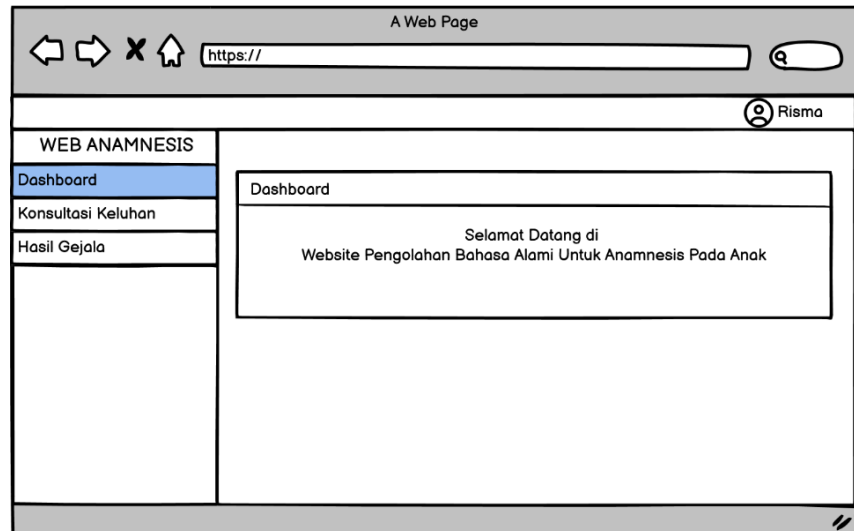
Gambar A.22 Perancangan Halaman Pengelompokkan Gejala pada “The Sacred Seven of History Taking”

3) Perancangan *User Interface* Pengguna

Perancangan *user interface* pengguna terdiri dari menu *dashboard*, konsultasi keluhan dan hasil gejala.

a) Perancangan Halaman *Dashboard* Pengguna

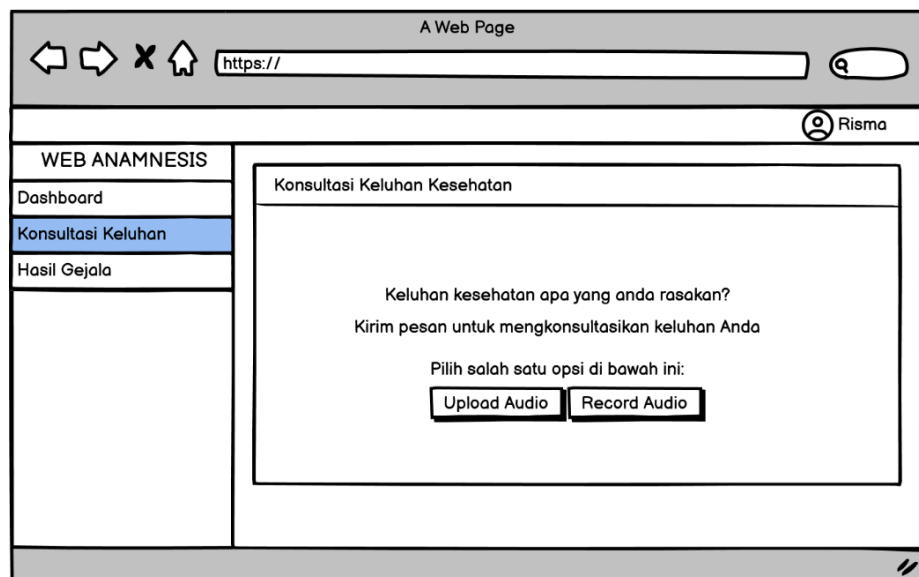
Perancangan halaman *dashboard* pengguna merupakan halaman yang ditampilkan ketika berhasil melakukan *login*. Perancangan halaman *dashboard* pengguna dapat dilihat pada Gambar A.23.



Gambar A.23 Perancangan Halaman *Dashboard* Pengguna

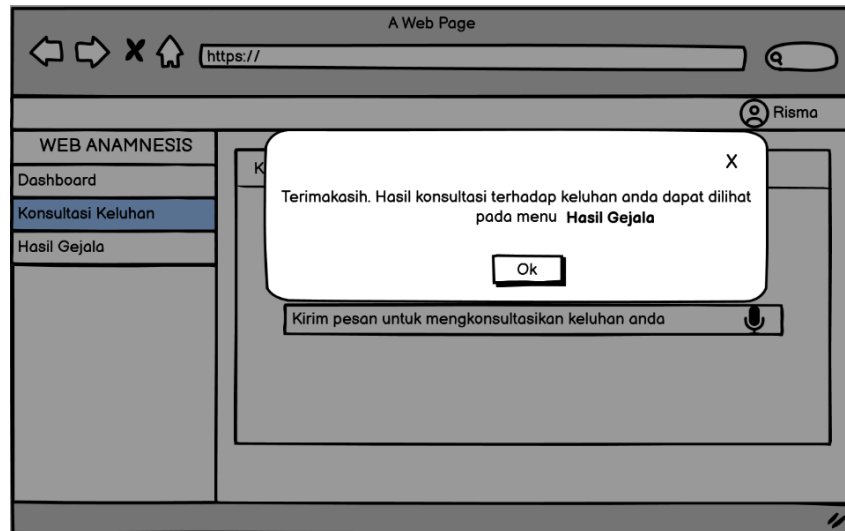
b) Perancangan Halaman Konsultasi Keluhan

Perancangan halaman konsultasi keluhan merupakan halaman yang digunakan pengguna untuk melakukan konsultasi keluhan dengan menekan tombol “Voice” ketika sistem menanyakan terkait keluhan kesehatan. Perancangan halaman konsultasi keluhan dapat dilihat pada Gambar A.24.



Gambar A.24 Perancangan Halaman Konsultasi Keluhan

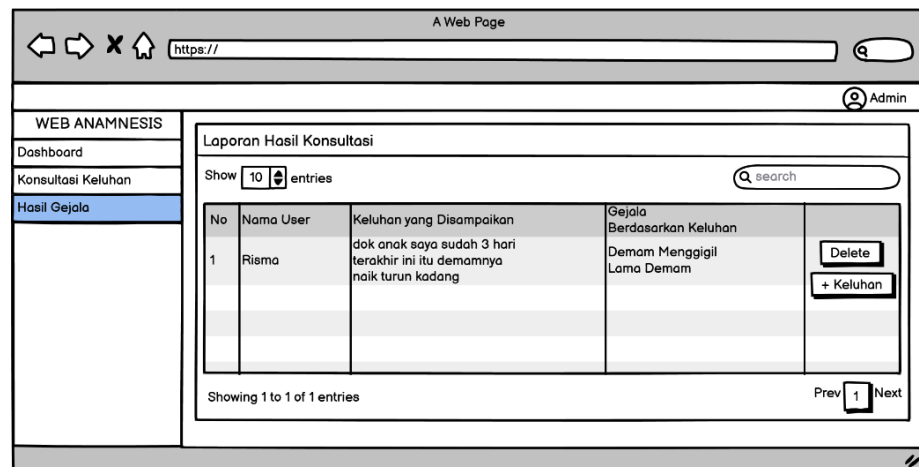
Perancangan halaman berhasil melakukan konsultasi ditampilkan ketika selesai melakukan konsultasi yang dapat dilihat pada Gambar A.25.



Gambar A.25 Perancangan Halaman Berhasil Konsultasi Keluhan

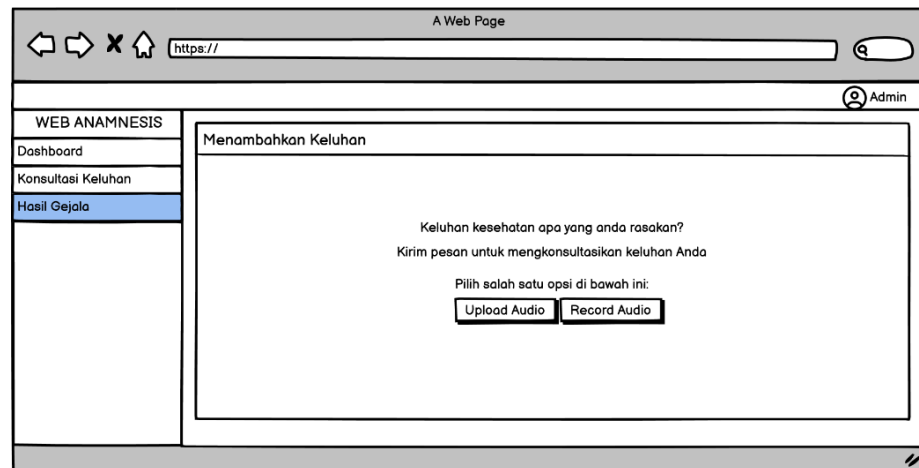
c) Perancangan Halaman Hasil Gejala

Perancangan halaman hasil gejala merupakan halaman yang digunakan pengguna untuk melihat hasil gejala berdasarkan algoritma *TF-IDF-cosine similarity*. Perancangan halaman hasil gejala dapat dilihat pada Gambar A.26.



Gambar A.26 Perancangan Halaman Hasil Gejala

Perancangan halaman hasil gejala, pasien juga dapat melakukan tambah keluhan dengan cara melakukan klik Keluhan. Perancangan halaman tambah keluhan dapat dilihat pada Gambar A.27.



Gambar A.27 Perancangan Halaman Tambah Keluhan

3. *Implementation and Unit Testing*

Pada tahap ini akan dilakukan pembuatan sistem sesuai dengan hasil analisis kebutuhan dan perancangan sistem. Pembuatan sistem menggunakan bahasa pemrograman *Python* dengan *framework flask*, *MySQL* sebagai *database* sehingga dibutuhkan sebuah perangkat untuk menjalankan aplikasi yang dibuat. Berikut merupakan implementasi coding penelitian ini.

Code Snippet A. 1 Proses *Voice to Text*

```
def index():
    return render_template('pages/konsultasiKeluhan/index.html', segment='keluhan')

def process_voice():
    audio_blob = request.form.get('audioBlob') # Untuk mendapatkan hasil perekaman suara
    filename = request.form.get('filename', 'recording.webm')

    uploaded_file = request.files.get('audioFile') # Untuk file yang diunggah oleh pengguna

    # Jika kedua input kosong, kembalikan error
    if not audio_blob and not uploaded_file:
        flash('No audio data received!', 'danger')
        return redirect(url_for('keluhan_index'))

    # Tentukan jalur sementara untuk file audio
    temp_audio_path = tempfile.NamedTemporaryFile(delete=False, suffix=".webm").name
    wav_audio_path = tempfile.NamedTemporaryFile(delete=False, suffix=".wav").name
```

```

# Jika menerima audioBlob (rekaman langsung)
if audio_blob:
    # Decode base64 dan simpan ke file sementara
    audio_data = base64.b64decode(audio_blob)
    with open(temp_audio_path, 'wb') as f:
        f.write(audio_data)

# Jika menerima file unggahan
elif uploaded_file:
    temp_audio_path = tempfile.NamedTemporaryFile(delete=False, suffix=".webm").name
    uploaded_file.save(temp_audio_path) # Simpan file unggahan sementara

# Konversi audio ke format WAV menggunakan pydub
audio = AudioSegment.from_file(temp_audio_path)
audio.export(wav_audio_path, format="wav")

# Gunakan SpeechRecognition untuk mengenali teks dari audio
recognizer = sr.Recognizer()
with sr.AudioFile(wav_audio_path) as source:
    audio_data = recognizer.record(source)

# Deteksi teks menggunakan Google Speech Recognition
text = recognizer.recognize_google(audio_data, language='id-ID') # Bahasa Indonesia
print(f"Recognized Text: {text}")

# Mendapatkan list gejala
gejala = Gejala.get().serialize()
list_gejala = [item['nama_gejala'] for item in gejala]
print(list_gejala)

# Menggabungkan text hasil voice dan gejala
corpus = [text] + list_gejala

```

Code Snippet A. 2 *Preprocessing*

```

hasil_preprocessing = []
for txt in corpus:
    _casefolding = case_folding(txt)
    _tokenizing = tokenizing(_casefolding)
    _filtering = filtering(_tokenizing)

```

```

    _stemming = stemming(_filtering)
    hasil_preprocessing.append(_stemming)
def case_folding(text):
    return text.casefold()
def tokenizing(text):
    text = nltk.tokenize.word_tokenize(text)
    return text
def filtering(token):
    stopword = set(stopwords.words('indonesian'))
    filter_ = [t for t in token if t not in stopword]
    return " ".join(filter_)
def stemming(filtered_token):
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    words = filtered_token.split()
    backstr = []
    for w in words:
        stemmed = stemmer.stem(w)
        backstr.append(stemmed)

    return ' '.join(backstr)

```

Code Snippet A. 3 TF-IDF dan *Cosine Similarity*

```

# Membuat TF-IDF vektor
vectorizer = TfidfVectorizer()
tfidf_matrix = vectorizer.fit_transform(hasil_preprocessing)
tfidf = pd.DataFrame(tfidf_matrix.toarray(), columns = vectorizer.get_feature_names_out())
print(tfidf)

# Menghitung cosine similarity
similarity = cosine_similarity(tfidf_matrix[0:1], tfidf_matrix[1:])
print(similarity)

# Menampilkan hasilnya
similarities = list(zip(list_gejala, similarity[0]))
print(similarities)

# Mengurutkan dari nilai similarities terbesar
sorted_similarities = sorted(similarities, key=lambda x: x[1], reverse=True)

```

```

# Filter Gejala dengan similarities > 0.15
filtered_gejala = [item for item in sorted_similarities if item[1] > 0.15]
print(filtered_gejala)

# Mengambil nama gejalanya saja
filtered_names = [item[0] for item in filtered_gejala]
print(filtered_names)

# Mendapatkan id gejala yang terpilih
list_id = []
for g in range(len(gejala)):
    if gejala[g]['nama_gejala'] in filtered_names:
        list_id.append(gejala[g]['id'])

# Menyimpan data ke database
keluhan = konsultasiKeluhan()
keluhan.user_id = session['user']['id']
keluhan.voice = text
keluhan.save()

hasil_gejala = hasilGejala()
hasil_gejala.user_id = session['user']['id']
hasil_gejala.konsultasi_id = keluhan.serialize()['id']
hasil_gejala.gejala_id = json.dumps(list_id)
hasil_gejala.save()

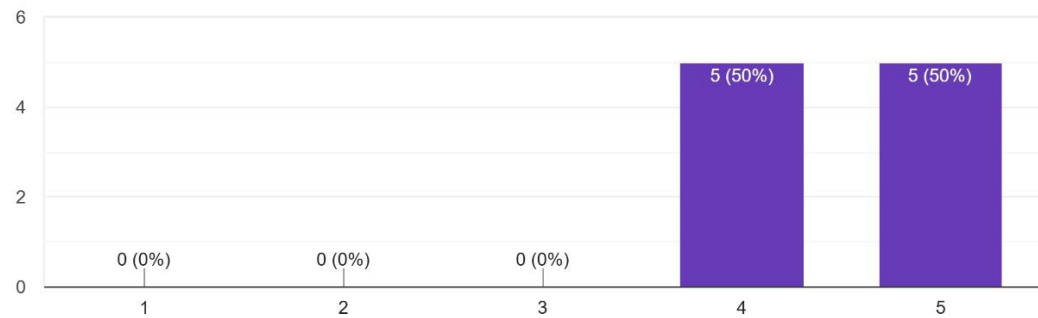
```

4. *Integration and System Testing*

Setelah pembuatan sistem dilakukan maka akan dilakukan pengujian sistem untuk memastikan sistem dapat berjalan dengan baik. Pengujian yang dilakukan seperti pengujian sistem menggunakan metode *black box testing* untuk menguji fungsionalitas sistem dan pengujian penggunaan *website* kepada pengguna menggunakan metode *user acceptance test* (UAT) dengan melakukan *survey*. Berikut pada Gambar A.28 hingga Gambar A.36 merupakan penjabaran terkait hasil penyebaran UAT.

Apakah informasi yang ditampilkan pada sistem NLP untuk anamnesis penyakit pada anak sudah sesuai dengan kebutuhan?

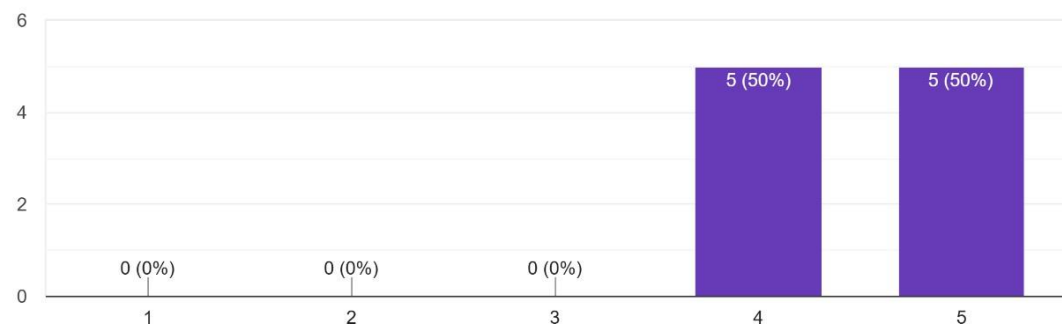
10 jawaban



Gambar A.28 Hasil Pertanyaan 1

Apakah penyusunan menu serta isi tiap-tiap menu pada pada sistem NLP untuk anamnesis penyakit pada anak sudah rapih?

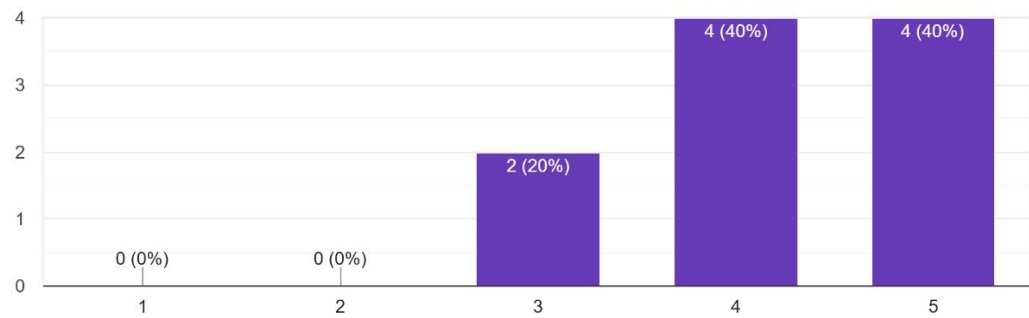
10 jawaban



Gambar A.29 Hasil Pertanyaan 2

Apakah pemilihan warna pada pada sistem NLP untuk anamnesis penyakit pada anak sudah terlihat nyaman oleh pengguna?

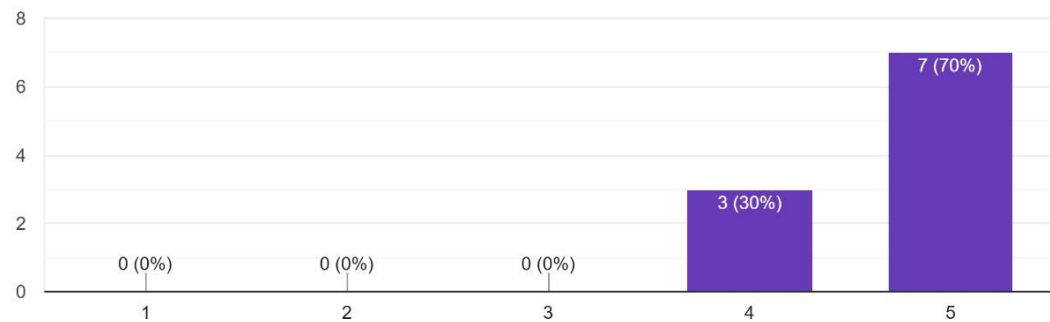
10 jawaban



Gambar A.30 Hasil Pertanyaan 3

Apakah dari sisi operasional pada sistem NLP untuk anamnesis penyakit pada anak dapat memberikan kemudahan bagi pengguna?

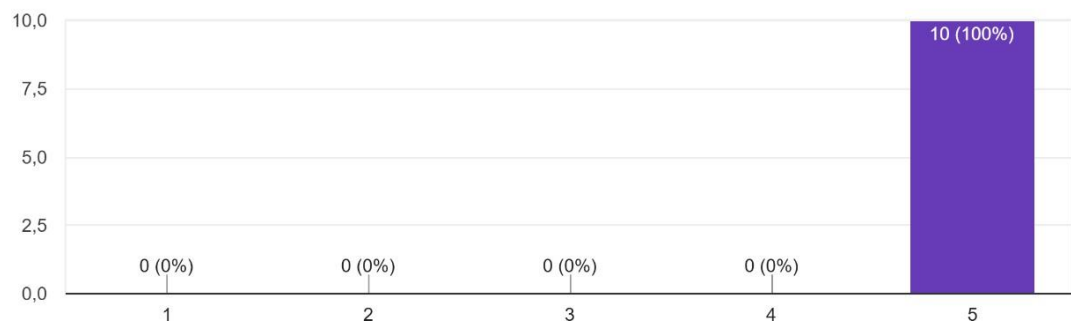
10 jawaban



Gambar A.31 Hasil Pertanyaan 4

Apakah sistem NLP untuk anamnesis penyakit pada anak dapat diakses sesuai dengan hak akses pengguna?

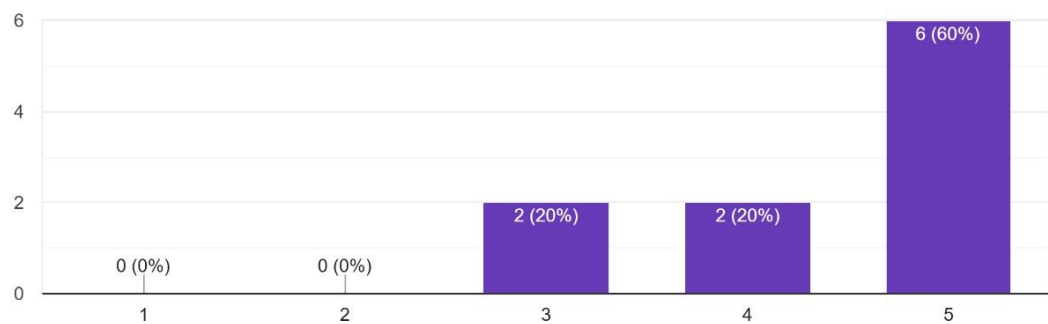
10 jawaban



Gambar A.32 Hasil Pertanyaan 5

Apakah proses memuat informasi pada pada sistem NLP untuk anamnesis penyakit pada anak tidak membutuhkan waktu yang lama?

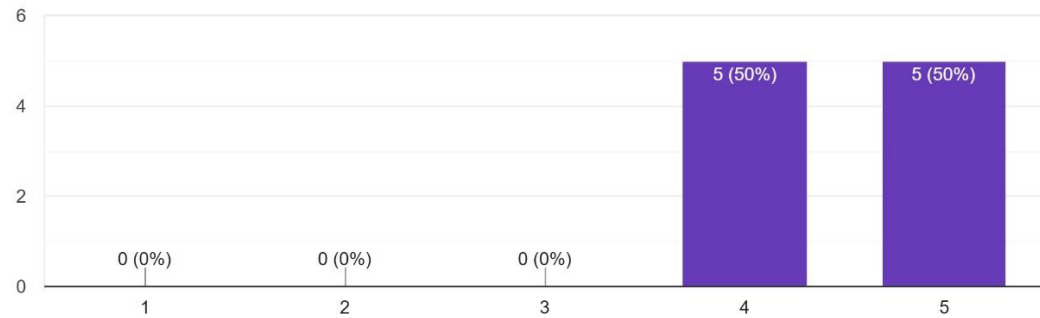
10 jawaban



Gambar A.33 Hasil Pertanyaan 6

Apakah pada sistem NLP untuk anamnesis penyakit pada anak sudah berjalan dengan baik?

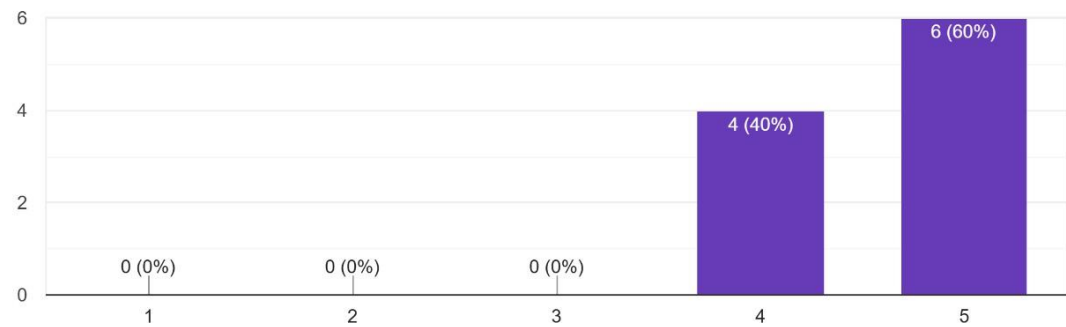
10 jawaban



Gambar A.34 Hasil Pertanyaan 7

Apakah data yang dibutuhkan dalam pada sistem NLP untuk anamnesis penyakit pada anak sudah sesuai?

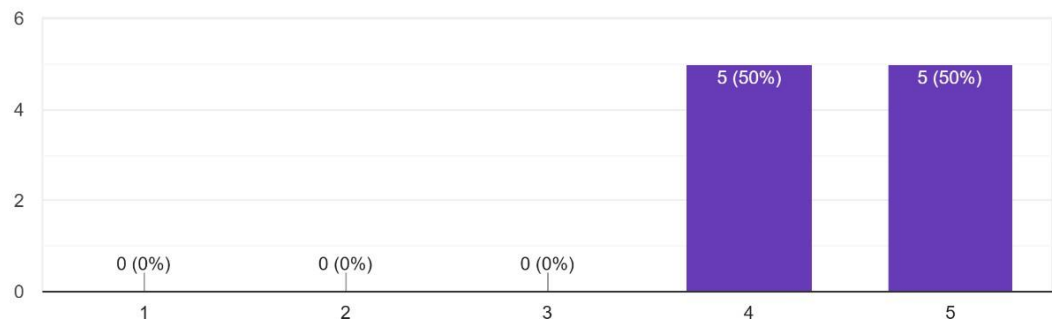
10 jawaban



Gambar A.35 Hasil Pertanyaan 8

Apakah pada sistem NLP tersebut sudah layak dijadikan sistem NLP untuk anamnesis penyakit pada anak?

10 jawaban



Gambar A.36 Hasil Pertanyaan 9

No	Keluhan yang Disampaikan	Gejala Sistem (Hasil NLP)	Gejala Berdasarkan Keluhan	Valid (Benar/Tidak)
1	Anak saya sudah tiga hari demam, muntah setiap habis makan, dan sekarang jarang buang air kecil.	Demam, Muntah setelah makan, Frekuensi BAK menurun	Demam, Muntah setelah makan, Frekuensi BAK menurun	Ya
2	Setiap malam anak saya batuk keras, demam tinggi, dan mengeluh sakit tenggorokan.	Batuk malam hari, Demam tinggi, Nyeri tenggorokan	Batuk malam hari, Demam tinggi, Nyeri tenggorokan	Ya
3	Anak sering menangis saat pipis, pipisnya sedikit dan berbau tajam.	Menangis saat BAK, Pipis sedikit, Urine berbau	Menangis saat BAK, Pipis sedikit, Urine berbau	Ya
4	Demam naik turun selama dua hari terakhir dan anak mengeluh pusing dan nyeri kepala.	Demam naik turun, Nyeri kepala, Pusing	Demam naik turun, Nyeri kepala, Pusing	Ya
5	Anak muntah lebih dari tiga kali sehari dan tidak mau makan, terlihat lemas.	Muntah >3 kali, Tidak nafsu makan, Lemas	Muntah >3 kali, Tidak nafsu makan, Lemas	Ya
6	Anak demam dan setiap makan langsung muntah, perut juga sering dipegangi.	Demam, Muntah setiap makan, Nyeri perut	Demam, Muntah setiap makan, Nyeri perut	Ya
7	Keluhan anak meliputi demam, batuk, dan napas berbunyi.	Demam, Batuk, Wheezing	Demam, Batuk, Wheezing	Ya
8	Anak demam, tidak mau makan, dan buang air kecil menjadi sangat jarang.	Demam, Anoreksia, Frekuensi BAK menurun	Demam, Anoreksia, Frekuensi BAK menurun	Ya
9	Demam lebih dari tiga hari, muntah, dan muncul ruam merah di kulit.	Demam lama, Muntah, Ruam kulit	Demam lama, Muntah, Ruam kulit	Ya
10	Anak mengalami nyeri tenggorokan, demam, dan suara serak.	Nyeri tenggorokan, Demam, Suara serak	Nyeri tenggorokan, Demam, Suara serak	Ya
11	Anak saya demam sejak semalam, tampak lemas, dan keluar bercak merah di kulit.	Demam, Lemas, Ruam kulit	Demam, Lemas, Ruam kulit	Ya
12	Demam tinggi disertai batuk kering dan napas anak berbunyi saat tidur.	Demam tinggi, Batuk kering, Napas berbunyi	Demam tinggi, Batuk kering, Napas berbunyi	Ya
13	Anak tidak mau makan sejak kemarin, rewel terus dan muntah setelah minum susu.	Tidak nafsu makan, Muntah, Rewel	Tidak nafsu makan, Muntah, Rewel	Ya
14	Demam anak tidak turun selama dua hari, dia juga mengeluh perutnya sakit.	Demam, Nyeri perut	Demam, Nyeri perut	Ya
15	Setiap kali makan, anak saya muntah dan mengatakan perutnya sakit.	Muntah setelah makan, Nyeri perut	Muntah setelah makan, Nyeri perut	Ya
16	Anak sering buang air kecil tapi sedikit dan menangis setiap kali pipis.	Menangis saat pipis, Pipis sedikit	Menangis saat pipis, Pipis sedikit	Ya
17	Malam hari batuknya parah, disertai demam dan tidak bisa tidur nyenyak.	Batuk malam hari, Demam, Sulit tidur	Batuk malam hari, Demam, Sulit tidur	Ya
18	Demam, batuk, dan keluar cairan dari hidung sudah dua hari ini.	Demam, Batuk, Pilek	Demam, Batuk, Pilek	Ya

19	Demam selama dua hari dan sekarang anak tampak sangat lemas dan tidak aktif.	Demam, Lemas, Tidak aktif	Demam, Lemas, Tidak aktif	Ya
20	Demam disertai mual dan ruam kecil muncul di tangan dan kaki.	Demam, Mual, Ruam tangan dan kaki	Demam, Mual, Ruam tangan dan kaki	Ya
21	Anak saya demam tinggi, muntah, dan terlihat gelisah sejak pagi.	Demam tinggi, Muntah, Gelisah	Demam tinggi, Muntah, Gelisah	Ya
22	Anak terlihat tidak nyaman saat buang air kecil dan mengeluh perih.	Nyeri saat BAK, Frekuensi BAK meningkat	Nyeri saat BAK, Frekuensi BAK meningkat	Ya
23	Sudah tiga hari anak demam dan tidak ada nafsu makan sama sekali.	Demam, Anoreksia	Demam, Anoreksia	Ya
24	Anak rewel terus, demam, dan muntah beberapa kali dalam sehari.	Demam, Muntah, Rewel	Demam, Muntah, Rewel	Ya
25	Anak susah tidur karena batuk dan demam di malam hari.	Batuk, Demam malam hari	Batuk, Demam malam hari	Ya
26	Anak mengeluh pusing, lemas, dan tubuhnya hangat sejak kemarin.	Demam, Pusing, Lemas	Demam, Pusing, Lemas	Ya
27	Pipis anak jadi jarang, berbau tajam, dan dia terus menangis.	Frekuensi BAK menurun, Urine berbau, Rewel	Frekuensi BAK menurun, Urine berbau, Rewel	Ya
28	Anak mengalami demam, nyeri sendi, dan bintik merah muncul di lengan.	Demam, Nyeri sendi, Bintik merah	Demam, Nyeri sendi, Bintik merah	Ya
29	Anak saya muntah, tidak mau makan, dan demam sejak kemarin.	Demam, Muntah, Tidak nafsu makan	Gejala tidak relevan	Tidak
30	Anak batuk terus, demam, dan seperti sesak saat bernapas.	Batuk, Demam, Sesak napas	Gejala tidak relevan	Tidak