

SKRIPSI

**SISTEM DETEKSI DINI KARIES GIGI MENGGUNAKAN
ALGORITMA *K-NEAREST NEIGHBOR* BERBASIS WEB**

DISUSUN OLEH

AGUNG KENCANA SUGIHARTO

2009010102



UMSU

Unggul | Cerdas | Terpercaya

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA

MEDAN

2025

**SISTEM DETEKSI DINI PADA KARIES GIGI MENGGUNAKAN
ALGORITMA K-NEAREST NEIGHBOR BERBASIS WEB**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom) dalam Program Studi Sistem Informasi pada Fakultas
Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah
Sumatera Utara**

**AGUNG KENCANA SUGIHARTO
NPM. 2009010102**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

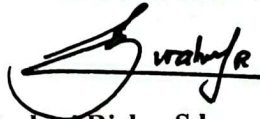
MEDAN

2025

LEMBAR PENGESAHAN

Judul Skripsi : SISTEM DETEKSI DINI KARIES GIGI
MENGUNAKAN ALGORITMA K-NEAREST
NEIGHBOR BERBASIS WEB
Nama Mahasiswa : AGUNG KENCANA SUGIHARTO
NPM : 2009010102
Program Studi : SISTEM INFORMASI

Menyetujui
Dosen Pembimbing



(Dr. Firahmi Rizky, S.kom., M.Kom.)
NIDN.0128029302

Ketua Program Studi



(Dr. Firahmi Rizky, S.kom., M.Kom.)
NIDN.0128029302

Dekan



(Dr. Al-Khowarizmi, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

SISTEM DETEKSI DINI PADA KARIES GIGI MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR BERBASIS WEB

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 08 Agustus 2025

Yang membuat pernyataan



AGUNG KENCANA SUGIHARTO

PERNYATAAN PERSETUJUAN PUBLIKASI

KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama	: Agung Kencana Sugiharto
NPM	: 2009010102
Program Studi	: Sistem Informasi
Karya Ilmiah	: Skripsi

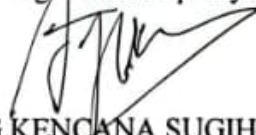
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

SISTEM DETEKSI DINI PADA KARIES GIGI MENGUNAKAN ALGORITMA K-NEAREST NEIGHBOR BERBASIS WEB

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 08 Agustus 2025
Yang membuat pernyataan



AGUNG KENCANA SUGIHARTO
NPM. 2009010102

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Agung Kencana Sugiharto
Tempat dan Tanggal Lahir : Tanjung Mulyo, 31 maret 2001
Alamat Rumah : Dusun XIII Tanjung Mulyo
Telepon/Faks/HP : 083183101521
E-mail : agungkencana2303@gmail.com
Instansi Tempat Kerja : Belum Bekerja
Alamat Kantor : -

DATA PENDIDIKAN

SD	: SD NEGERI 058295	TAMAT: 2013
SMP	: SMP SWASTA AMPERA	TAMAT: 2016
SMA	: SMK SPP PUTRA JAYA STABAT	TAMAT: 2019

KATA PENGANTAR



Puji syukur Alhamdulillah, Penulis panjatkan atas kehadiran Allah SWT., yang telah melimpahkan rahmat dan karunia-Nya yang telah memberi kekuatan kepada Penulis untuk yang menuntaskan tugas akhir dalam meraih Strata 1 ini. Skripsi ini Penulis sajikan dalam bentuk buku yang sederhana. Judul Skripsi pada penelitian ini adalah sebagai berikut. **“SISTEM DETEKSI DINI PADA KARIES GIGI MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR BERBASIS WEB”**.

Tujuan dari penulisan skripsi ini adalah sebagai salah satu syarat kelulusan Program Strata Satu (S1) Sistem Informasi Universitas Muhammadiyah Sumatera Utara. Penulisan ini didasarkan pada hasil penelitian, observasi dan beberapa sumber referensi yang mendukung penulisan ini. Penulis menyadari bahwa tanpa dukungan dan bimbingan dari semua pihak, maka Penulis skripsi ini tidak akan bisa berjalan lancar. Oleh karena itu pada kesempatan ini penulis ingin mengucapkan terima kasih kepada semua pihak yang telah membantu, terimakasih kepada :

1. Bapak Prof. Dr. Agussani, M.AP., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom sebagai Dekan Fakultas Ilmu Komputer dan Teknologi Informasi.
3. Bapak Martiano, S.Pd., S.Kom., M.Kom sebagai Ketua Prodi Sistem Informasi yang selalu memberikan dukungan.

4. Ibu Dr. Firahmi Rizky, S.Kom., M.Kom selaku Dosen Pembimbing saya yang telah sabar membimbing saya untuk menyelesaikan skripsi ini.
5. Teristimewa kepada Papa Ir.Sukirman dan Mama Alm. Ngatini sebagai Orang Tua Penulis yang tak pernah lelah mendoakan, mengusahakan dan memberikan fasilitas serta dukungan terbaik secara material maupun non-material untuk kesuksesan penulis.
6. Tersayang kepada Angga Pratama Wijanarko dan Elinta Tri Agustine yang merupakan saudara kandung penulis yang selalu menjadi motivasi penulis dan selalu mengingatkan untuk selalu rajin, tekun untuk menyelesaikan pendidikan.
7. Yang terhormat kepada Pakde Sunarto yang telah banyak mendukung saya baik material maupun moral yang sangat berarti bagi saya.
8. Untuk kakek, nenek dan seluruh anggota keluarga terimakasih atas do'a serta dukungan dan semuanya yang telah kalian berikan kepada penulis.
9. Seluruh Staff dan Keanggotaan Biro Kemahasiswaan yang mendukung dalam proses pengerjaan Penelitian ini.
10. Untuk teman seperjuangan, terima kasih atas segala bantuan, waktu, support dan kebaikan yang diberikan kepada penulis disaat masa sulit mengerjakan skripsi ini.
11. Kepada Muhammad Ismail Alwi, muhammad Kevin Masyaid Siregar, Fawwaz Belva Calosa Soesilo, sahabat penulis, terimakasih banyak atas dukungan dan dorongan yang diberikan selama proses pengerjaan skripsi ini berlangsung. Untuk selalu mengatakan bahwa Penulis pasti mampu untuk melewati segala hal berat pada semester akhir ini, selalu mengingatkan untuk terus memprioritaskan diri sendiri dahulu dari pada orang lain.

12. Kepada Agung Nurizki, Prayoga Aditya Pratama, Romi Wijaya sahabat penulis yang selalu menemani, memberikan motivasi dan semangat yang luar biasa kepada penulis dari awal masuk perkuliahan hingga saat ini. Terima kasih sudah menjadi sahabat yang sangat baik bahkan seperti saudara. Terima kasih karena tidak pernah meninggalkan penulis sendirian dan selalu menjadi garda terdepan saat penulis membutuhkan bantuan serta selalu mendengarkan keluhan kesah penulis selama berada di perantauan ini.
 13. Seluruh teman-teman Angkatan Sistem Informasi 2020 yang telah sama-sama berjuang, yang namanya tidak dapat penulis sebutkan satu-persatu.
 14. Kepada seseorang sangat penting kehadirannya, manusia spesial penulis Vinka Varera, terimakasih karena selalu menghibur penulis dalam kesedihan, mendengar keluhan kesah dan selalu memberikan motivasi bahwa penulis mampu melewati segala hal berat pada semester akhir ini.
 15. Dan yang terakhir, kepada diri saya sendiri. Agung Kencana Sugiharto. Terima kasih sudah berjuang sejauh ini, terima kasih tetap masih berdiri dan bertahan, walau sering kali merasa putus asa atas apa yang diusahakan dan belum berhasil, namun terima kasih tetap menjadi manusia yang selalu mau berusaha dan tidak lelah mencoba. Terima kasih karena memutuskan untuk tidak menyerah sesulit apapun proses penyusunan skripsi ini dan telah menyelesaikannya sebaik dan semaksimal mungkin, ini merupakan pencapaian yang patut dibanggakan untuk diri sendiri. Berbahagialah selalu dimanapun berada, terima kasih Agung telah berusaha menjadi manusia yang berguna. Dimasa depan kamu harus sukses.
- Serta semua pihak terlibat yang namanya tidak bisa disebutkan satu persatu

sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini jauh dari kata sempurna, untuk itu Penulis memohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan ini di mendatang. Akhir kata, semoga Penelitian ini dapat berguna bagi para pembaca yang berminat pada topik penelitian ini.

ABSTRAK

Karies gigi merupakan masalah kesehatan mulut yang umum di Indonesia dengan prevalensi yang tinggi. Proses deteksi konvensional yang mengandalkan pemeriksaan visual dan radiografi memiliki keterbatasan seperti subjektivitas, biaya, dan paparan radiasi. Penelitian ini bertujuan untuk mengembangkan dan menganalisis akurasi sistem pakar berbasis web menggunakan algoritma *K-Nearest Neighbor* (KNN) untuk deteksi dini karies gigi. Metode pengembangan sistem yang digunakan adalah model *Waterfall*, yang mencakup tahapan analisis kebutuhan, pengumpulan data rekam medis dari Klinik Gigi Rahma Dewi, perancangan sistem dengan UML, implementasi, dan pengujian. Data pasien yang mencakup gejala dan atribut diolah menggunakan normalisasi *Z-Score* untuk menyeragamkan skala fitur. Model KNN kemudian dilatih untuk mengklasifikasikan tingkat risiko karies ke dalam tiga kategori: Karies Awal, Karies Sedang, dan Karies Lanjut. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan berhasil diimplementasikan dengan antarmuka yang fungsional. Berdasarkan pengujian menggunakan *Confusion Matrix*, model klasifikasi KNN mencapai akurasi sebesar 91.67%. Kinerja ini didukung oleh nilai *precision*, *recall*, dan *F1-score* yang seimbang, menunjukkan bahwa sistem ini berpotensi menjadi alat bantu yang efektif dan efisien bagi dokter gigi dalam mempercepat proses diagnosis dan menyediakan deteksi dini yang mudah diakses.

ABSTRACT

Dental caries is a common oral health problem in Indonesia with a high prevalence. Conventional detection processes that rely on visual examinations and radiography have limitations such as subjectivity, cost, and radiation exposure. This research aims to develop and analyze the accuracy of a web-based expert system using the K-Nearest Neighbor (KNN) algorithm for the early detection of dental caries. The system development method used is the Waterfall model, which includes stages of requirements analysis, medical record data collection from the Rahma Dewi Dental Clinic, system design with UML, implementation, and testing. Patient data, including symptoms and attributes, were processed using Z-Score normalization to standardize feature scales. The KNN model was then trained to classify the level of caries risk into three categories: Early Caries, Moderate Caries, and Advanced Caries. The results show that the developed system was successfully implemented with a functional interface. Based on testing using a Confusion Matrix, the KNN classification model achieved an accuracy of 91.67%. This performance is supported by balanced precision, recall, and F1-score values, indicating that this system has the potential to be an effective and efficient auxiliary tool for dentists in accelerating the diagnostic process and providing accessible early detection.

DAFTAR ISI

DAFTAR ISI	i
DAFTAR GAMBAR	v
DAFTAR TABEL	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II LANDASAN TEORI	5
2.1 Karies Gigi	5
2.2 Sistem Pakar	5
2.3 Algoritma K-Nearest Neighbor	7
2.4 Confussion Matrix	9
2.5 Pra-Pemrosesan Data	10
2.6 Normalisasi Z-Score	11
2.7 Hypertext Markup Language (HTML)	12
2.8 Hypertext Preprocessor (PHP)	13
2.9 MySQL	13
2.10 Python	13

2.11	Pandas	14
2.12	Scikit-Learn	14
2.13	Unified Modelling Language (UML)	14
BAB III METODE PENELITIAN		22
3.1	Metode Perancangan Sistem	22
3.2	Perancangan Sistem	25
3.3	Simulasi Perhitungan K-Nearest Neighbor	42
3.4	Perancangan Basis Data	44
3.5	Desain <i>Interface</i>	48
BAB IV HASIL DAN PEMBAHASAN		57
4.1.	Perhitungan Manual Metode <i>K-Nearest Neighbor</i>	57
4.2.	Penerapan dalam Sistem Berbasis Web	65
4.3.	Kelebihan dan Kelemahan Program	78
4.3.1.	Kelebihan Program	78
4.3.2.	Kelemahan Program	79
4.4.	Pengujian Program	80
BAB V KESIMPULAN DAN SARAN		82
5.1	Kesimpulan	82
5.2	Saran	83
DAFTAR PUSTAKA		84

DAFTAR GAMBAR

Gambar 3. 1 Metode Waterfall	23
Gambar 3. 2 Use Case Diagram	27
Gambar 3. 3 Activity Diagram Login	28
Gambar 3. 4 Activity Diagram Kelola Data Master	29
Gambar 3. 5 Activity Diagram Pengguna	30
Gambar 3. 6 Activity Diagram Latih Model	32
Gambar 3. 7 Activity Diagram Lihat Evaluasi Model	33
Gambar 3. 8 Activity Diagram Riwayat Diagnosa	34
Gambar 3. 9 Activity Diagram Logout Admin	35
Gambar 3. 10 Activity Diagram Input dan Gejala	36
Gambar 3. 11 Activity Diagram Lakukan Diagnosa	37
Gambar 3. 12 Activity Diagram Lihat Evaluasi Model	38
Gambar 3. 13 Activity Diagram Cetak Laporan	39
Gambar 3. 14 Activity Diagram Logout Dokter	40
Gambar 3. 15 Sequence Diagram Admin	40
Gambar 3. 16 Sequence Diagram Dokter	41
Gambar 3. 17 Class Diagram	42
Gambar 3. 18 Rancangan Menu Login	49
Gambar 3. 19 Rancangan Menu Dashboard	50
Gambar 3. 20 Rancangan Menu Kelola Gejala	51
Gambar 3. 21 Kelola Data Atribut	52
Gambar 3. 22 Rancangan Menu Kelola Penyakit	53
Gambar 3. 23 Rancangan Kelola Dataset Latih	54
Gambar 3. 24 Rancangan Menu Latih Model KNN	55
Gambar 3. 25 Rancangan Menu Evaluasi Model	56
Gambar 3. 26 Rancangan Menu Riwayat Diagnosa	57
Gambar 3. 27 Form Diagnosa Pasien Baru	58
Gambar 4. 1 Tampilan Halaman Login	65
Gambar 4. 2 Tampilan Halaman Dashboard	66
Gambar 4. 3 Tampilan Kelola Data Pasien	67

Gambar 4. 4 Tampilan Tambah Data Pasien	68
Gambar 4. 5 Tampilan Kelola Data Dokter	69
Gambar 4. 6 Tampilan Tambah Data Dokter	69
Gambar 4. 7 Tampilan Kelola Dataset Latih	70
Gambar 4. 8 Tampilan Riwayat Konsultasi	71
Gambar 4. 9 Tampilan Awal Kelola Model KNN	72
Gambar 4. 10 Bagian EDA dalam Sistem	73
Gambar 4. 11 Bagian Pra-Pemrosesan & Optimasi	74
Gambar 4. 12 Gambar Evaluasi dan Performa Model	75
Gambar 4. 13 Tampilan Klasifikasi Data	78

DAFTAR TABEL

Tabel 2. 1 Simbol Dalam Use Case	16
Tabel 2. 2 Simbol Activity Diagram	17
Tabel 2. 3 Tabel Simbol Sequence Diagram	18
Tabel 3. 1 Contoh Data Latih	43
Tabel 3. 2 Contoh Data yang ingin diprediksi	44
Tabel 3. 3 Contoh Urutan Data dari Terdekat	45
Tabel 3. 4 Rancangan Tabel user	46
Tabel 3. 5 Rancangan Tabel Pasien	46
Tabel 3. 6 Rancangan Tabel Gejala	47
Tabel 3. 7 Rancangan Tabel Atribut	47
Tabel 3. 8 Rancangan Tabel Penyakit	47
Tabel 3. 9 Rancangan Tabel Diagnosa	48
Tabel 3. 10 Rancangan Tabel Dataset Latih	48
Tabel 4. 1 Sampel Data Latih	58
Tabel 4. 2 Sampel Data Uji	59
Tabel 4. 3 Hasil normalisasi Z-score untuk Data Latih	61
Tabel 4. 4 Hasil Normalisasi Z-Score untuk Data Uji	61
Tabel 4. 5 Hasil Perhitungan Jarak Euclidean dan Pengurutan untuk Data Uji U01	63
Tabel 4. 6 Penentuan Kelas Berdasarkan 3 Tetangga Terdekat untuk U01	63
Tabel 4. 7 Ringkasan Hasil Prediksi Model KNN untuk Seluruh Data Uji	64
Tabel 4. 8 Confusion Matrix Hasil Klasifikasi	64
Tabel 4. 9 Tabel Confusion Matrix	76

Tabel 4. 10 Tabel Hasil Klasifikasi	77
Tabel 4. 11 Tabel Pengujian Program	80

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Karies gigi merupakan infeksi yang paling umum terjadi dalam rongga mulut yang dapat menyerang individu dari masa bayi hingga usia lanjut (Diastuti et al., 2024) . Penyakit ini ditandai dengan kerusakan jaringan keras gigi yang disebabkan oleh asam yang dihasilkan oleh bakteri plak. Prevalensi karies gigi di Indonesia masih tergolong tinggi. Riset Kesehatan Dasar 2018 menunjukkan prevalensi karies gigi pada penduduk usia 12 tahun mencapai 57,6%, untuk Remaja 89,5%, Dewasa 88,1% dan Lansia 98,4%, menunjukkan bahwa lebih dari separuh penduduk Indonesia memiliki masalah karies gigi (Nur Rahma & Mulyanti, 2021) . Di Klinik Gigi Rahma, yang berlokasi di Medan, kasus karies gigi menjadi salah satu keluhan utama yang sering dialami oleh pasien, menunjukkan bahwa masalah ini membutuhkan perhatian lebih dalam upaya pencegahan dan deteksi dini.

Dampak karies gigi tidak hanya terbatas pada kerusakan gigi, tetapi juga dapat menyebabkan rasa sakit, infeksi, gangguan pengunyahan, dan penurunan kualitas hidup (Wicaksono et al., 2024). Pada kasus yang parah, karies gigi dapat menyebabkan komplikasi serius seperti abses gigi, selulitis, dan bahkan sepsis. Selain dampak Kesehatan, karies gigi juga menimbulkan beban ekonomi yang signifikan. Biaya perawatan karies gigi, terutama untuk stadium lanjut, cukup mahal.

Metode konvensional yang umum digunakan untuk mendeteksi karies gigi adalah pemeriksaan visual dan radiografi (rontgen gigi). Pemeriksaan visual dilakukan oleh dokter gigi dengan mengamati kondisi gigi secara langsung menggunakan alat bantu cermin mulut dan *explorer*. Metode ini memiliki keterbatasan, diantaranya subjektif dalam penilaian dan ketergantungan pada keahlian dokter gigi. Radiografi disisi lain dapat memberikan Gambaran struktur gigi yang lebih detail, namun, metode ini memiliki beberapa kelemahan, seperti paparan radiasi, biaya yang relative mahal, dan keterbatasan aksesibilitas, terutama di daerah yang minim Kesehatan (Wicaksono et al., 2024). Oleh karena itu, diperlukan solusi untuk mendeteksi dini karies gigi yang lebih efektif dan mudah diakses oleh Masyarakat.

Perkembangan teknologi telah membuka peluang baru dalam bidang kedokteran gigi, termasuk dalam deteksi karies gigi. Salah satu cabang teknologi computer yang banyak digunakan oleh manusia untuk membantu pekerjaannya adalah pembuatan sistem pakar. Sistem pakar adalah sistem yang memindahkan pengetahuan manusia atau pakar ke dalam komputer untuk memecahkan masalah tertentu seperti seorang pakar (Firdaus & Yanti, 2022a) . Implementasi sistem pakar dalam bidang kedokteran atau kesehatan berupa diagnosa penyakit, konsultasi penjagaan Kesehatan, serta memberikan saran penentuan solusi berdasarkan hasil yang ada.

Agar sistem pakar dapat berfungsi secara optimal, diperlukan algoritma yang mampu membantu dalam menghasilkan diagnosis dan rekomendasi yang akurat, salah satu algoritma yang bisa digunakan adalah algoritma *K-Nearest Neighbor*. Algoritma *K-Nearest Neighbor* adalah salah satu algoritma klasifikasi

objek berdasarkan atribut kelas dan data latih, konsep dasar dari KNN adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga terdekatnya dalam data pelatihan (Supriyanto et al., 2023).

Dengan menggabungkan sistem pakar dan algoritma *K-Nearest Neighbor* diharapkan dapat dikembangkan model deteksi dini karies gigi yang lebih cepat dan mudah diakses dan juga dapat mengurangi ketergantungan pada metode konvensional yang memiliki keterbatasan, serta memberikan serta memberikan rekomendasi perawatan yang lebih tepat.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang dirincikan sebelumnya, rumusan masalah yang penulis ambil adalah :

1. Bagaimana mengembangkan sistem pakar untuk mendeteksi karies gigi menggunakan algoritma *K-Nearest Neighbor* (KNN)?
2. Bagaimana akurasi algoritma *K-Nearest Neighbor* dalam mendiagnosis karies gigi?

1.3 Batasan Masalah

Dalam penelitian ini, penulis memberikan Batasan masalah sebagai berikut :

1. Penelitian hanya berfokus pada Klinik Gigi Rahma Dewi
2. Data yang digunakan dalam penelitian adalah data kunjungan dan rekam medis pasien yang telah tersedia di Klinik Gigi Rahma Dewi
3. Algoritma yang digunakan untuk diagnosa awal adalah algoritma *K-Nearest Neighbor* (KNN).

4. Sistem pakar hanya berfokus pada deteksi karies gigi berdasarkan gejala dan karakteristik yang teridentifikasi, tidak mencakup diagnosis penyakit gigi lainnya.
5. Sistem yang dikembangkan merupakan *prototype* yang diajukan dalam skenario simulasi dan belum digunakan dalam praktik klinis secara langsung.

1.4 Tujuan Penelitian

Adapun yang menjadi tujuan bagi penulis dalam melaksanakan penelitian ini adalah sebagai berikut :

1. Mengembangkan sistem pakar berbasis algoritma *K-Nearest Neighbor* untuk mendeteksi karies gigi secara otomatis
2. Menganalisis performa dan akurasi sistem pakar berbasis *K-Nearest Neighbor* dalam diagnosa karies gigi.
3. Menyediakan sistem deteksi dini karies gigi lebih cepat dan akurat serta mudah diakses.

1.5 Manfaat Penelitian

Setelah penelitian ini berlangsung, diharapkan dapat memberikan manfaat sebagai berikut

1. Menambah wawasan mengenai penerapan sistem pakar dalam mendeteksi karies gigi
2. Memberikan alternatif dalam mendeteksi dini karies gigi tanpa harus selalu bergantung pada pemeriksaan dengan metode konvensional.
3. Membantu dokter gigi dalam mendeteksi karies gigi dengan sistem yang lebih cepat, sehingga dapat meningkatkan efisiensi dalam pemeriksaan pasien.

BAB II

LANDASAN TEORI

2.1 Karies Gigi

Karies gigi adalah penyakit infeksi kronis yang merusak struktur gigi. Penyakit ini ditandai dengan demineralisasi jaringan keras gigi yang disebabkan asam organik yang dihasilkan dari fermentasi karbohidrat oleh bakteri di dalam plak gigi. Proses ini menyebabkan terbentuknya kavitas atau lubang pada gigi (Wicaksono et al., 2024).

2.2 Sistem Pakar

Sistem pakar merupakan sistem berbasis komputer yang menggunakan pengetahuan, fakta dan teknik penalaran dalam pemecahan masalah yang biasa hanya dapat dipecahkan oleh pakar dalam bidang tertentu (Firahmi et al., 2020). Istilah sistem pakar sendiri berasal dari gabungan “pakar” dan “sistem”. “pakar” merujuk kepada seorang ahli yang memiliki kompetensi dan pengalaman di bidang tertentu, sedangkan “sistem” merujuk kepada sekumpulan komponen yang saling berinteraksi untuk mencapai sesuatu. Sistem pakar tidak hanya meniru kemampuan pengambilan keputusan manusia tetapi juga membagikan pengetahuan dan kemampuan tersebut ke dalam sistem agar dapat dimanfaatkan oleh pengguna lain (sulistyono et al., 2024). Maka, sistem pakar dapat menjadi sarana untuk mendistribusikan keahlian pakar secara lebih luas tanpa harus bergantung pada kehadiran pakar secara langsung.

Ruang lingkup sistem pakar dalam bidang kesehatan memiliki peranan yang sangat penting sekali. Salah satu penerapannya yaitu dalam membantu mendiagnosa penyakit. Sistem pakar dapat digunakan untuk mengidentifikasi gejala-gejala yang dialami pasien, kemudian mencocokkan dengan basis pengetahuan yang sudah dimiliki untuk menghasilkan kemungkinan diagnosa penyakit beserta rekomendasi

penanganannya. Hal ini sangat berguna terutama untuk daerah yang kekurangan tenaga medis dan spesialis. Sistem pakar juga dapat digunakan sebagai alat bantu konsultasi medis. Contohnya, pasien dapat menginput gejala yang dialami ke dalam sistem, lalu sistem mengirimkan informasi awal mengenai kemungkinan penyakit dan juga tindakan yang sebaiknya dilakukan, seperti, istirahat, mengkonsumsi obat, atau segera berkonsultasi dengan dokter. Dengan demikian sistem pakar dapat menjadi *first opinion* yang membantu pasien sebelum mendapatkan perawatan medis secara langsung. Pada bidang farmasi sistem pakar dapat digunakan untuk merekomendasikan penggunaan obat. Sistem dapat memberikan informasi mengenai dosis pemberian obat, penggunaan obat, serta efek samping yang dapat terjadi. Hal ini dapat membantu Dokter maupun pasien dalam meningkatkan keamanan dalam penggunaan obat. Lebih lanjut lagi, sistem pakar dalam bidang kesehatan berfungsi dalam pendukung pengambilan keputusan klinis (clinical decision support system). Dengan adanya sistem ini, dokter dapat memperoleh kemungkinan tambahan terkait kemungkinan penyakit yang jarang ditemui, penentuan metode perawatan yang lebih tepat, hingga prediksi perkembangan penyakit berdasarkan riwayat perkembangan penyakit pasien.

Keterbatasan terhadap sistem pakar itu tetap ada, salah satunya sistem hanya bekerja sesuai basis pengetahuan yang dimiliki. Jika data atau aturan yang dimasukkan kurang lengkap maka hasil diagnosa yang dimasukkan juga tidak akurat. Maka itu, pengembangan sistem pakar memerlukan tenaga medis yang kompeten dan pembaruan basis pengetahuan secara berkala.

Tujuan utama dari sistem pakar adalah untuk memberikan solusi atau jawaban yang akurat terhadap permasalahan, berdasarkan pengetahuan dan pengalaman para pakar (Firdaus & Yanti, 2022b). Dengan mentransfer pengetahuan manusia ke dalam komputer, sistem pakar dapat mencapai beberapa tujuan, antara lain

1. Membantu diagnosis: Sistem pakar dapat digunakan untuk mendiagnosis berbagai kondisi, seperti penyakit mata , tingkat risiko penyakit jantung (Mahdalena Simanjorang et al., 2024) dan tingkat kecanduan media sosial. Sistem ini mengumpulkan data dan menganalisisnya untuk memberikan diagnosis awal atau informasi yang relevan.

2. Membantu pengambilan keputusan: Sistem pakar dapat membantu dalam pengambilan keputusan di berbagai bidang.

3. Meningkatkan efisiensi: Sistem pakar dapat mengotomatiskan proses dan memberikan informasi dengan cepat dan mudah.

4. Meningkatkan aksesibilitas: Sistem pakar dapat membuat informasi dan keahlian lebih mudah diakses oleh orang awam.

2.3 Algoritma K-Nearest Neighbor

K-Nearest Neighbor adalah suatu metode yang biasa digunakan dalam proyek *Data Mining*. Metode ini menggunakan algoritma *Supervised Learning*. *Supervised Learning* melibatkan penggunaan data yang telah diberi tanda dari hasil sebelumnya. Tujuan dari *Supervised Learning* adalah untuk melatih model komputer yang dapat mempelajari pola-pola dalam data dan melakukan prediksi akurat terhadap data yang belum diketahui (Cahya Mestika et al., 2022).

K-Nearest Neighbor adalah suatu metode yang menggunakan algoritma *Supervised Learning* dimana hasil dari permintaan data yang baru diklasifikasikan berdasarkan data mayoritas dari *label class* dalam KNN. Tujuan dari algoritma ini adalah mengklasifikasikan objek baru berdasarkan atribut dan data latih. Algoritma ini bekerja berdasarkan jarak terpendek dari permintaan data ke data latih untuk menentukan KNN nya. Salah satu cara untuk menghitung jarak dekat

atau jauhnya setiap data atau tetangga dalam data adalah menggunakan metode *Eucledian Distance* (Choirun Anisa, 2020).

Euclidean Distance adalah metode yang sering digunakan untuk menghitung jarak antar tunggal. Jarak ini digunakan untuk menguji interpretasi perkiraan jarak antara dua objek (Rozi et al., 2023). Rumus untuk menghitung jarak *Eucledian Distance* adalah sebagai berikut :

$$D(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \dots \dots \dots (2.1)$$

Dimana,

$D(x,y)$: Jarak *Eucledian* antara dua titik x dan y.

n : Dimensi ruang eucledian

X_i, Y_i : Koordinat titik x dan y di dimensi ke-i

Dalam melakukan klasifikasi menggunakan algoritma KNN, kita harus menentukan nilai parameter k , nilai k pada KNN merupakan jumlah tetangga terdekat, jika k bernilai 1, maka kelas dari satu data latih yang terdekat akan menjadi kelas bagi data uji yang baru, jika k bernilai 3 akan diambil tiga data latih terdekat menjadi kelas untuk data uji yang baru (Mulyati et al., 2020).

Pemilihan nilai k sendiri akan mempengaruhi hasil prediksi. Nilai k yang terlalu besar dapat mengakibatkan distorsi data yang besar pula, hal ini dikarenakan setiap tetangga mempunyai bobot yang sama terhadap data uji, sedangkan k yang terlalu kecil bisa menyebabkan algoritma terlalu sensitif terhadap *noise* pada data .

Berikut merupakan Langkah-langkah untuk memprediksi menggunakan algoritma *K-Nearest Neighbor* :

1. Menentukan jumlah tetangga terdekat (parameter k)

2. Membagi data menjadi data latih dan data uji
3. Menghitung jarak *Eucledian Distance*.
4. Melakukan pengurutan data dari data yang memiliki jarak *Eucledian* terkecil ke terbesar.
5. Melakukan klasifikasi data berdasarkan kelas terbanyak berdasarkan nilai k yang telah ditentukan.

2.4 Confussion Matrix

Confusion Matrix adalah tabel yang menyatakan klasifikasi jumlah uji yang benar dan jumlah uji yang salah (Normawati & Prayogi, 2021). *Confusion Matrix* sendiri memiliki empat istilah :

1. *True Negative* (TN) yang berarti model klasifikasi memprediksi data ada di kelas *negative* namun yang sebenarnya data memang ada di kelas *negative*.
2. *True Positive* (TP) yang berarti model klasifikasi memprediksi data ada di kelas positif namun sebenarnya memang data berada di kelas positif.
3. *False Negative* (FN) yang berarti model klasifikasi memprediksi data ada di kelas *negative* namun yang sebenarnya data ada di kelas positif.
4. *False Positive* (FP) yang berarti model klasifikasi memprediksi data ada di kelas positif namun yang sebenarnya data ada di kelas *negative*.

Dari defenisi *confusion matrix*, beberapa poin poin dalam *confusion matrix* digunakan untuk menghitung *precision*, *recall*, dan *f1 score*. *Precision* adalah perbandingan antara *True Positive* (TP) dengan banyaknya data yang diprediksi positif, secara matematis dapat dilihat dibawah ini

$$Precision = \frac{TP}{TP+FP} \dots \dots \dots (2.2)$$

Untuk *recall* sendiri adalah perbandingan antara *true positive* (TP) dengan banyaknya data yang sebenarnya positif. Dapat dinyatakan secara matematis seperti dibawah ini

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (2.3)$$

Sedangkan *F1 Score* adalah nilai tengah dari *precision* dan *recall*. Nilai terbaik dari *F1 Score* adalah 1 dan nilai terburuknya adalah 0, secara matematis dapat dituliskan seperti dibawah ini

$$\frac{1}{F1} = \frac{1}{2} \left(\frac{1}{precision} + \frac{1}{recall} \right) \dots\dots\dots (2.4)$$

Nilai *F1 Score* yang baik menandakan bahwa model klasifikasi kita punya *precision* dan *recall* yang baik (S. Setiawan, 2020).

2.5 Pra-Pemrosesan Data

Pra-pemrosesan data adalah tahapan untuk mengolah data mentah dengan menghilangkan beberapa permasalahan yang mengganggu saat pemrosesan data. Hal tersebut dikarenakan oleh data yang formatnya tidak konsisten. Melalui proses ini, pemodelan algoritma KNN akan berjalan lebih efektif dan efisien. Adapun tahapan tahapan dalam pra-pemrosesan data adalah :

1. Pertama-tama dalam tahap awal *preprocessing* data, langkah yang esensial adalah melakukan data cleaning. Proses ini melibatkan pemilihan kembali data mentah untuk menghilangkan entri yang tidak lengkap, tidak relevan, atau tidak akurat. Dengan melakukan ini, kita dapat menghindari kesalahpahaman saat melakukan analisis terhadap data tersebut.
2. Langkah berikutnya adalah *data integration*, yang diperlukan karena preprocessing data melibatkan penggabungan data dari berbagai sumber ke

dalam satu dataset. Penting untuk memastikan bahwa data dari berbagai sumber memiliki format yang seragam.

3. Setelah itu, kita melanjutkan ke tahap transform
4. asi data. Sebagaimana dijelaskan sebelumnya, data yang berasal dari sumber yang berbeda mungkin memiliki format yang beragam. Oleh karena itu, perlu dilakukan penyesuaian format agar seluruh data yang terkumpul memiliki struktur yang seragam, memudahkan proses analisis data.
5. Tahap terakhir dalam *preprocessing* data adalah mengurangi jumlah data, yang dikenal sebagai *data reduction*. Tujuan utamanya adalah mengurangi sampel data tanpa mengubah hasil analisis. Terdapat tiga teknik yang dapat diterapkan pada tahap ini, yaitu pengurangan dimensi (*dimensionality reduction*), pengurangan jumlah (*numerosity reduction*), dan kompresi data (Binus Student Corner, 2022).

2.6 Normalisasi Z-Score

Normalisasi data adalah bagian dari praproses data di mana nilai-nilai dalam dataset disesuaikan ulang untuk memudahkan proses pengolahan. Proses ini penting karena dataset seringkali memiliki rentang nilai yang berbeda untuk setiap atributnya. Perbedaan signifikan dalam rentang nilai antar atribut dapat mengganggu kinerja optimal atribut dalam dataset. Oleh karena itu, normalisasi dilakukan untuk menyamakan skala nilai atribut agar proses analisis data menjadi lebih efisien (Whendasmoro & Joseph, 2022).

Normalisasi *Z-score* merupakan teknik normalisasi di mana nilai-nilai data disesuaikan berdasarkan nilai rata-rata dan deviasi standar dari data tersebut

(Emalia Saqila et al., 2023) . Dalam *Z-Score*, data mengalami transformasi atau perubahan untuk menciptakan rentang nilai yang baru, berdasarkan rentang nilai yang telah ada sebelumnya dalam dataset. Adapun rumus yang digunakan dalam *Z-score* adalah sebagai berikut :

$$Z = \frac{X - \mu}{\sigma} \dots\dots\dots (2.5)$$

Dimana,

Z : *Z-Score*

X : Nilai data

μ : Rata-rata data

σ : Standar deviasi

Adapun rumus untuk mencari Standar Deviasi adalah sebagai berikut :

$$\sigma = \sqrt{\sum \frac{(X_i - \mu)^2}{n}} \dots\dots\dots (2.6)$$

Dimana,

X_i : Nilai data

μ : Rata-rata data

σ : Standar deviasi

n : Jumlah data

2.7 Hypertext Markup Language (HTML)

HTML atau *HyperText Markup Language* adalah Bahasa markup standar yang digunakan untuk membuat halaman web. HTML berfungsi untuk membuat struktur dari sebuah website. HTML digunakan untuk menandai bagian mana yang akan menjadi judul artikel, bagian mana yang berfungsi sebagai isi artikel, atau bagian mana yang disajikan dalam bentuk tabel (Andre Pratama, 2020).

2.8 Hypertext Preprocessor (PHP)

PHP adalah Bahasa *scripting* yang menyatu dengan HTML yang dijalankan pada server sedang yang dikirim ke browser hanya hasilnya saja. Ketika pengguna internet membuka situs yang menggunakan fasilitas *serverside scripting* PHP, maka terlebih dahulu *server* yang bersangkutan akan memproses semua perintah PHP lalu mengirimkan hasilnya dalam format HTML ke *web server* pengguna internet tadi. Sehingga kode asli yang ditulis dengan PHP tidak terlihat di *browser* pengguna. PHP memiliki fungsi untuk mengambil informasi dari *form* berbasis *web* dan menggunakannya untuk berbagai macam fungsi (Aniqsa, 2019).

2.9 MySQL

MySQL merupakan perangkat lunak basis data relasi atau *Relational Database Management System* (RDBMS) yang diterapkan menggunakan lisensi GPL (*General Public License*). MySQL memiliki kinerja, kecepatan proses yang tidak kalah dengan database-database besar lainnya (Aniqsa, 2019). Beberapa pertimbangan dalam memilih MySQL antara lain :

1. Kecepatan
2. Mudah Digunakan
3. Terbuka
4. Kapabilitas
5. Konektifitas dan Keamanan

2.10 Python

Python adalah salah satu Bahasa pemrograman yang populer, yang dibuat pertama kali pada tahun 1991 oleh Guido van Rossum. Bahasa Pemrograman

Python sendiri dapat digunakan dalam pengembangan aplikasi web, pengembangan software, hingga menyelesaikan permasalahan matematika dari yang dasar dan kompleks dan bisa digunakan untuk pengembangan algoritma prediksi.

Keunggulan menggunakan Python adalah, dapat digunakan di berbagai platform, memiliki *syntax* yang mudah dan *simple*, dan dapat digunakan dalam pemrograman procedural, pemrograman berorientasi objek, maupun pemrograman fungsional.

2.11 *Pandas*

Pandas adalah *library* pada *Python* yang paling sering dipakai untuk menganalisis data. Cara kerja *pandas* cukup menarik, Data berupa format CSV atau SQL diubah menjadi objek dalam *Python* dengan bentuk baris dan kolom yang disebut sebagai *Dataframe*. Objek *Dataframe* ini akan terlihat sangat mirip dengan tabel yang terdapat dalam aplikasi pengolahan statistika pada umumnya.

Penggunaan *Library Pandas* juga dapat membantu dalam verifikasi data, pengolahan data, normalisasi data, penggabungan dan penyatuan data, inspeksi data, serta membuat dan menyimpan data .

2.12 *Scikit-Learn*

Scikit-Learn adalah *library* dalam bahasa pemrograman *python* yang membantu melakukan proses pada data ataupun melakukan pelatihan pada data untuk kebutuhan *data science* seperti klasifikasi, *clustering*, dan lain-lain. *Library* ini sangat populer dikarenakan banyak sekali model pembelajaran mesin yang dapat dipanggil menggunakan *scikit-learn* (Bisa AI Academy, 2023).

2.13 *Unified Modelling Language (UML)*

Unified Modelling Language (UML) adalah bahasa pemodelan visual yang digunakan untuk menentukan, memvisualisasikan, membangun berbagai aspek dari sebuah sistem perangkat lunak. UML sendiri memiliki fungsi sebagai alat untuk menangkap pemahaman mengenai sistem yang perlu dibangun.

Bayangkan sistem sebagai kelompok objek yang berbeda yang bekerja sama untuk melakukan pekerjaan yang berguna untuk pengguna. Bagian struktur statis mendefinisikan jenis objek yang penting untuk sistem dan bagaimana mereka terkait satu sama lain. Bagian perilaku dinamis menggambarkan bagaimana objek berubah seiring waktu dan berkomunikasi satu sama lain untuk mencapai tujuan tertentu. Dengan memodelkan sistem dari berbagai sudut pandang yang terkait, kita dapat memahami sistem tersebut untuk berbagai keperluan (Rumbaugh et al., 2021).

Alat bantu yang digunakan dalam perancangan sistem menggunakan Unified Modelling Language (UML) adalah sebagai berikut :

1. *Use Case Diagram*

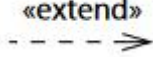
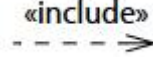
Use Case Diagram menunjukkan bagaimana perilaku suatu sistem ketika digunakan oleh orang lain, *Use Case Diagram* memecah fungsionalitas sistem menjadi tindakan yang bermakna bagi pengguna (*Actor*) yang menggunakan sistem. Pengguna (*Actor*) mencakup manusia, serta sistem dan proses komputer lainnya).

Sedangkan *Actor* adalah representasi ideal dari orang, proses, atau objek eksternal yang berinteraksi dengan suatu sistem, subsistem, atau kelas. Untuk identifikasi *Actor*, harus ditentukan tugas tugas yang berkaitan

dengan peran pada konteks sistem. Orang atau sistem bisa muncul dalam beberapa peran, perlu diketahui jika *Actor* berinteraksi dengan *use case*.

Selain Actor, terdapat simbol simbol lain yang bisa digunakan didalam *Use Case* untuk memodelkan fungsi apa saja dari sistem yang dibangun beserta hubungan antar fungsi nya. Adapun simbol-simbol dalam *Use Case* antara lain sebagai berikut :

Tabel 2. 1 Simbol Dalam Use Case

Gambar	Function	Keterangan
	Asosiasi	Asosiasi antara <i>actor</i> dan <i>use case</i> , digambarkan dengan garis tanpa panah yang menunjukkan siapa atau apa yang meminta interaksi secara langsung
	Extend	Perluasan dari <i>use case</i> lain ketika kondisi atau syarat terpenuhi
	Include	Merujuk pada penambahan perilaku tambahan dalam suatu <i>use case</i> yang secara eksplisit menjelaskan penambahan tersebut
	Use case generalization	Hubungan antara suatu <i>use case</i> umum dengan <i>use case</i>

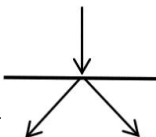
		yang lebih spesifik yang mewarisi dan menambahkan fitur fitur kepadanya
--	--	---

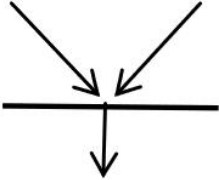
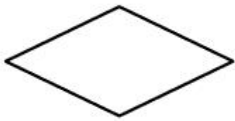
(Sumber : Rumbaugh et al., 2021).

2. Diagram Aktivitas (*Activity Diagram*)

Diagram aktivitas (*Activity Diagram*) menggambarkan aliran kerja dari sebuah sistem. Diagram aktivitas dapat mencakup cabang dan bercabangnya control dalam sebuah sistem yang berjalan secara bersamaan. Cabang cabang ini mewakili aktivitas yang dapat dilakukan secara bersamaan (Rumbaugh et al., 2021) . Simbol-simbol yang digunakan dalam *Activity Diagram* adalah sebagai berikut :

Tabel 2. 2 Simbol Activity Diagram

Gambar	Function	Keterangan
	<i>Start Point</i>	Diletakkan pada pojok kiri atas dan merupakan awal aktifitas
	<i>End Point</i>	Akhir dari aktifitas
	<i>Activities</i>	Menggambarkan suatu proses / kegiatan dalam aplikasi
	<i>Fork</i>	Digunakan untuk menunjukkan kegiatan yang dilakukan secara

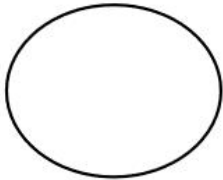
		parallel untuk menggabungkan dua kegiatan parallel menjadi satu
	<i>Join</i>	Digunakan untuk menunjukkan adanya dekomposisi
	<i>Decisions Point</i>	Menggambarkan pilihan untuk pengambilan keputusan, <i>true</i> , <i>false</i>

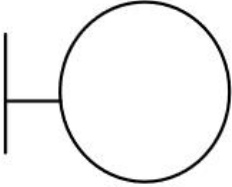
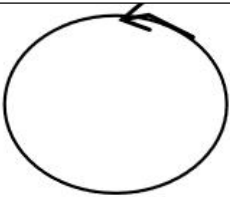
(Sumber : Rumbaugh et al., 2021).

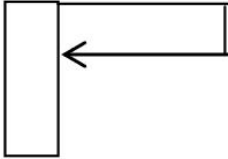
3. *Sequence Diagram*

Sequence Diagram menggambarkan kelakuan objek pada *use case* dengan mendeskripsikan waktu hidup objek dan pesan yang dikirimkan dan diterima antara objek. Simbol-simbol yang digunakan dalam *sequence diagram* yaitu :

Tabel 2. 3 Tabel Simbol *Sequence Diagram*

Gambar	Function	Keterangan
	<i>Entity Class</i>	Merupakan bagian dari system yang berisi kumpulan kelas berupa entitas-entitas yang membentuk gambaran awal sistem dan menjadi landasan untuk menyusun

		basis data
	<i>Boundary Class</i>	Berisi kumpulan kelas yang menjadi <i>interface</i> atau interaksi antara satu atau lebih <i>actor</i> dengan system, seperti tampilan form entry dan <i>form</i> cetak
	<i>Control Class</i>	Suatu objek yang berisi logika aplikasi yang tidak memiliki tanggung jawab kepada entitas, contohnya adalah kalkulasi dan aturan bisnis yang melibatkan berbagai objek
	<i>Message</i>	Simbol mengirim pesan antar <i>class</i>

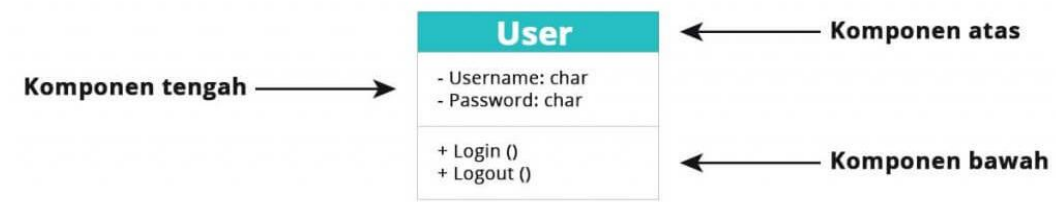
	<i>Recursive</i>	Menggambarkan pengiriman pesan yang dikirim untuk dirinya sendiri
	<i>Activation</i>	<i>Activation</i> mewakili sebuah eksekusi operasi dari objek, Panjang kotak ini berbanding lurus dengan durasi aktivitas sebuah operasi
	<i>Lifeline</i>	garis titik-titik yang terhubung dengan objek, sepanjang <i>lifeline</i> terdapat <i>activation</i>

(Sumber : Rumbaugh et al., 2021).

4. *Class Diagram* (Diagram Kelas)

Class diagram adalah salah satu jenis diagram struktur pada UML yang menggambarkan dengan jelas struktur serta deksripsi *class*, atribut, metode, dan hubungan dari setiap objek. *Class Diagram* bersifat statis, dalam artian diagram ini tidak menjelaskan apa yang terjadi jika kelas-kelasnya berhubungan, melainkan menjelaskan hubungan apa yang terjadi.

Class Diagram memiliki tiga komponen penyusun. Dapat dilihat dalam gambar dibawah ini



Gambar 2. 1 Komponen dalam Class Diagram

(Sumber : R. Setiawan, 2021)

Berikut merupakan penjelasan komponen-komponen diatas :

1. Komponen Atas

Komponen ini berisikan nama *class*. Setiap class memiliki nama yang berbeda-beda

2. Komponen Tengah

Komponen ini berisikan atribut dari *class*, komponen ini digunakan untuk menjelaskan kualitas dari suatu kelas. Atribut ini dapat ditulis dengan detail, dengan cara memasukkan tipe nilai

3. Komponen Bawah

Komponen ini menyertakan operasi yang ditampilkan dalam bentuk daftar. Operasi ini dapat menggambarkan bagaimana suatu *class* dapat berinteraksi dengan data.

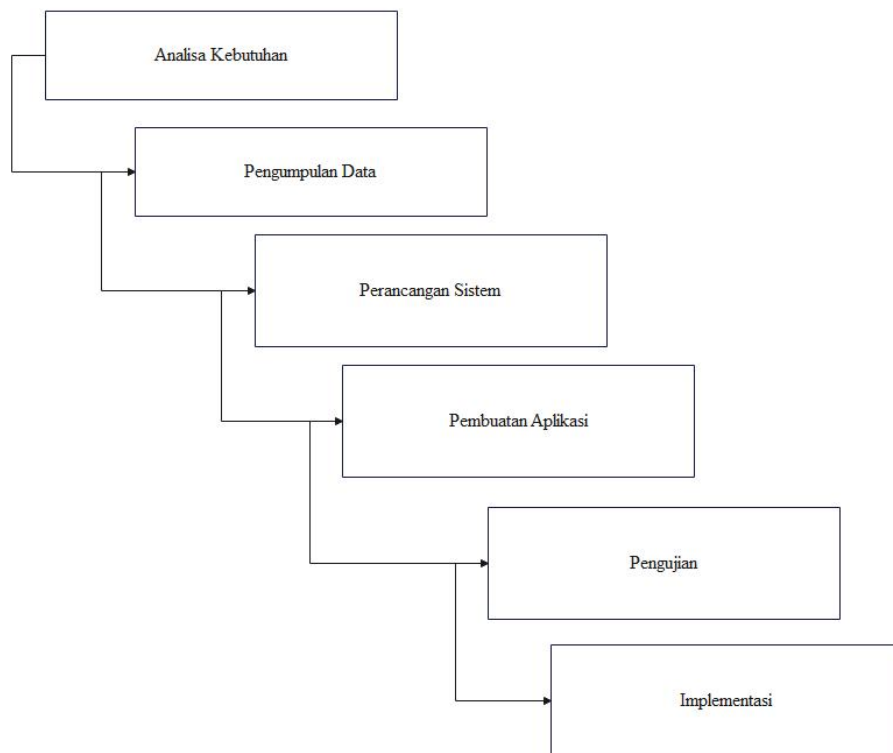
BAB III

METODE PENELITIAN

3.1 Metode Perancangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah metode Waterfall. Metode ini merupakan model pengembangan perangkat lunak yang paling sederhana dan bersifat sekuensial atau berurutan. Setiap tahapan dalam metode ini harus diselesaikan terlebih dahulu sebelum melanjutkan ke tahapan berikutnya, sehingga proses pengembangan sistem dilakukan secara sistematis dan bertahap.

Metode *Waterfall* adalah model SDLC (*System Development Life Cycle*) yang cocok digunakan untuk pengembangan perangkat lunak dengan spesifikasi yang tidak berubah-ubah. Proses pengerjaannya dilakukan secara berurutan, mulai dari analisis kebutuhan hingga tahap pengujian, sehingga sistem yang dihasilkan memiliki kualitas baik karena pelaksanaannya tidak melompat-lompat antar tahap (Hidayat et al., 2022). Berikut merupakan beberapa tahapan untuk penelitian ini.



Gambar 3. 1 Metode Waterfall

1. Analisa Kebutuhan

Tahapan ini dilakukan untuk mengidentifikasi kebutuhan pengguna terhadap sistem yang akan dikembangkan. Dalam konteks penelitian ini, analisa kebutuhan meliputi pemahaman terhadap masalah karies gigi, proses diagnosis yang berjalan di Klinik Gigi Rahma Dewi, serta kebutuhan pengguna terhadap sistem pakar berbasis KNN untuk membantu deteksi dini karies gigi.

2. Pengumpulan Data

Pada tahap ini, data yang dibutuhkan untuk proses klasifikasi dikumpulkan, seperti data rekam medis pasien, gejala-gejala karies, dan diagnosis dokter. Data ini menjadi dasar bagi algoritma K-Nearest Neighbor dalam melakukan pembelajaran dan pengambilan keputusan.

3. Perancangan Sistem

Perancangan sistem dilakukan untuk membuat blueprint sistem yang akan dibangun. Ini mencakup desain antarmuka pengguna, perancangan basis data, serta pembuatan diagram UML seperti use case, activity diagram, sequence diagram, dan class diagram agar sistem dapat dikembangkan secara terstruktur.

4. Pembuatan Aplikasi

Setelah desain selesai, tahap ini adalah proses pengkodean atau implementasi sistem berdasarkan perancangan sebelumnya. Aplikasi dikembangkan menggunakan bahasa pemrograman dan tools yang telah ditentukan, seperti PHP untuk backend, HTML untuk frontend, dan Python untuk pengolahan data klasifikasi menggunakan KNN.

5. Pengujian

Tahap ini bertujuan untuk memastikan bahwa sistem berjalan sesuai dengan kebutuhan dan bebas dari kesalahan. Pengujian dilakukan dengan memasukkan data uji dan mengevaluasi performa sistem menggunakan Confusion Matrix, dengan metrik seperti akurasi, precision, recall, dan F1-score.

6. Implementasi

Setelah sistem dinyatakan lulus tahap pengujian, sistem siap untuk diimplementasikan. Pada tahap ini, sistem dapat mulai digunakan dalam

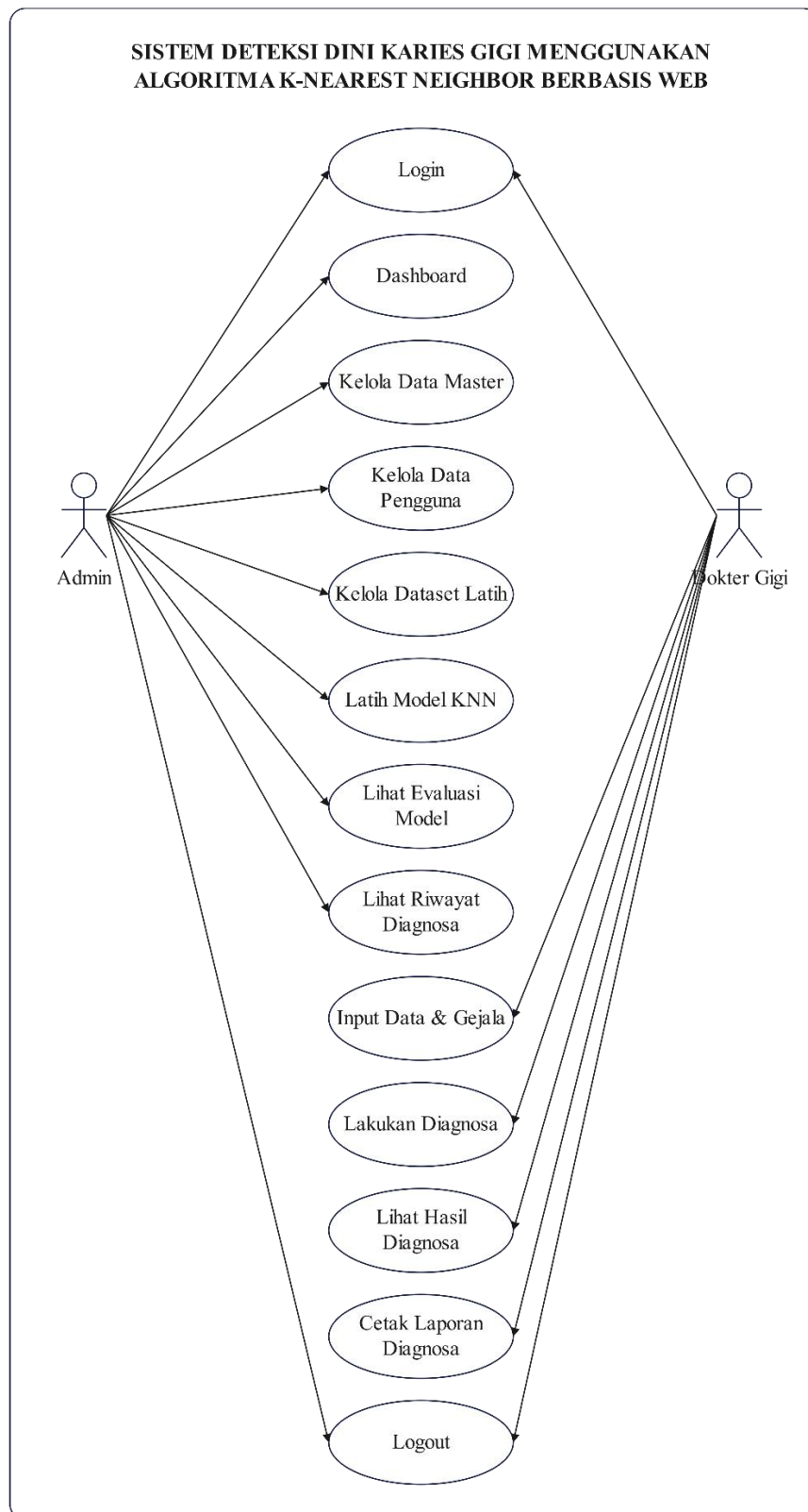
lingkungan uji coba atau simulasi sebagai prototype, serta dievaluasi kembali berdasarkan masukan dari pengguna untuk pengembangan lebih lanjut.

3.2 Perancangan Sistem

Proses ini berfokus pada perancangan seperti struktur data dan arsitektur perangkat lunak yang dibuat dengan pemodelan UML seperti *use case diagram*, *activity diagram*, *sequence diagram*, dan *class diagram*.

3.2.1. Use Case Diagram

Use Case Diagram ini menggambarkan interaksi antara dua aktor utama, yaitu Admin dan Dokter Gigi, terhadap sistem deteksi dini karies gigi berbasis web. Admin memiliki hak akses penuh terhadap sistem, termasuk melakukan login, mengakses dashboard, mengelola data master (gejala, atribut, penyakit), mengelola data pengguna (dokter), mengelola dataset latih, melatih model KNN, melihat hasil evaluasi model, dan memantau riwayat diagnosa pasien. Sementara itu, Dokter Gigi sebagai pengguna terbatas dapat melakukan login, menginput data dan gejala pasien, melakukan diagnosa berbasis algoritma K-Nearest Neighbor, melihat hasil diagnosa, mencetak laporan, serta melakukan logout dari sistem. Diagram ini menunjukkan pembagian peran dan fungsi masing-masing aktor secara efisien dan jelas untuk mendukung implementasi sistem yang terstruktur dan terkontrol.



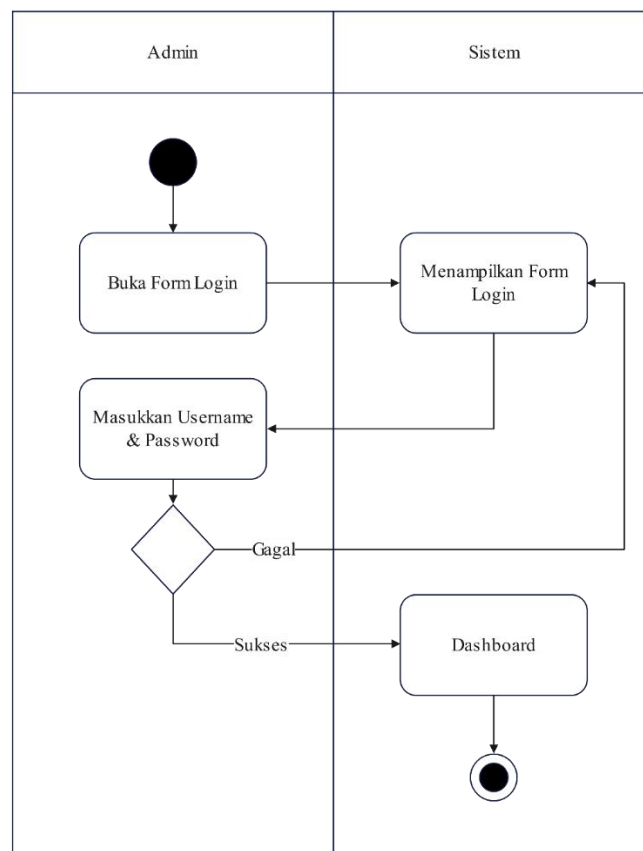
Gambar 3. 2 Use Case Diagram

3.2.2. Activity Diagram

3.2.2.1. Activity Diagram Admin

1. Activity Diagram Login

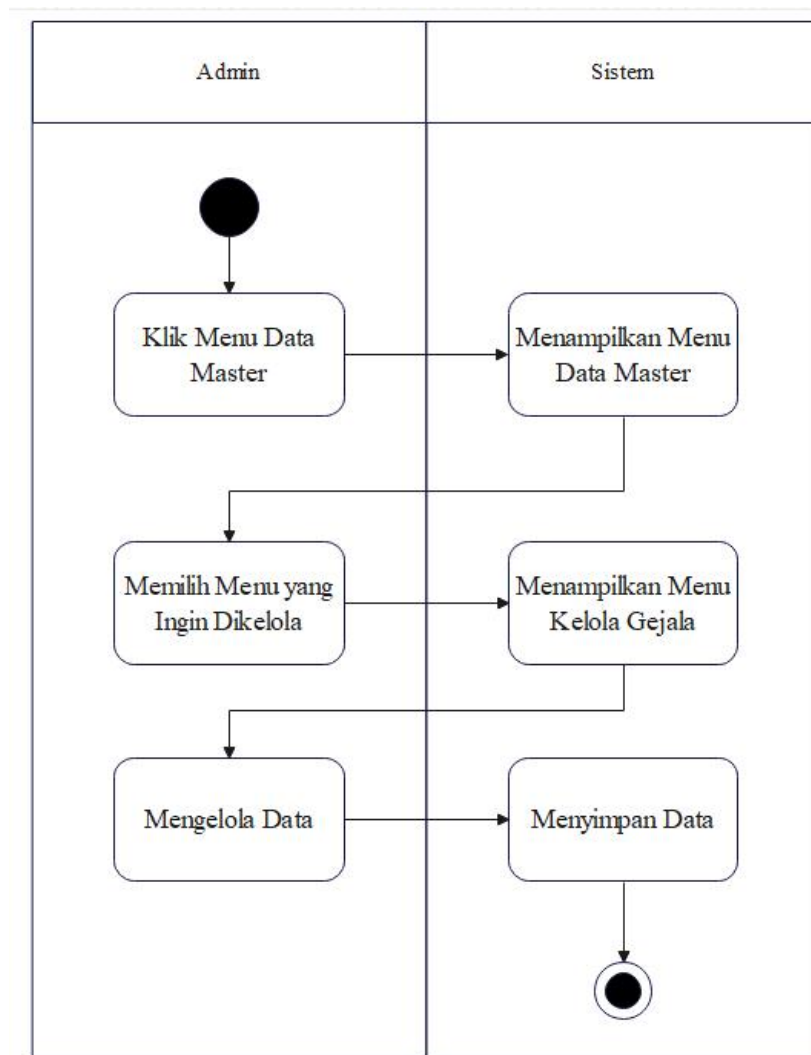
Proses login dimulai ketika admin membuka halaman form login. Sistem kemudian menampilkan form login kepada admin. Setelah itu, admin memasukkan username dan password ke dalam form yang tersedia. Sistem akan melakukan validasi terhadap kredensial yang dimasukkan. Jika data tidak valid, maka sistem akan menolak login dan kembali menampilkan form login dengan notifikasi kesalahan. Namun, jika validasi berhasil, sistem akan mengarahkan admin ke halaman dashboard untuk melanjutkan aktivitas lainnya. Dapat dilihat pada gambar dibawah ini



Gambar 3. 3 Activity Diagram Login

2. Activity Diagram Kelola Data Master

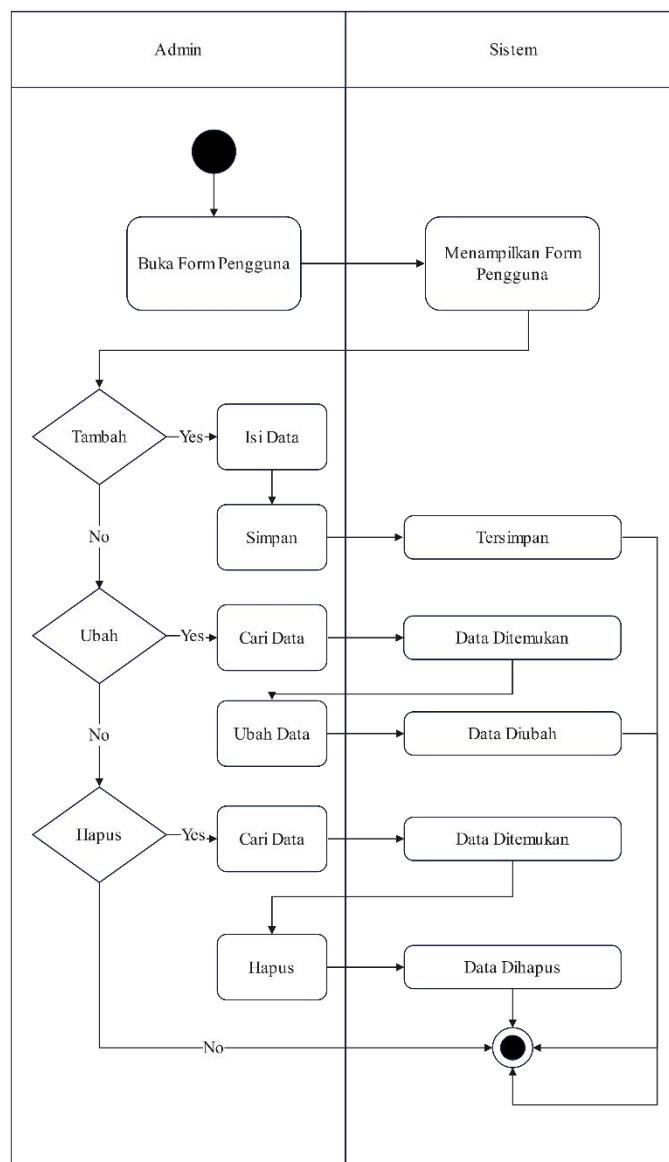
Dalam aktivitas ini, Admin mengakses menu data master untuk mengelola data gejala, atribut, dan penyakit. Admin dapat menambah, mengubah, atau menghapus data. Setelah aksi dilakukan, sistem menyimpan perubahan dan menampilkan notifikasi berhasil. Dapat dilihat pada gambar dibawah ini



Gambar 3. 4 Activity Diagram Kelola Data Master

3. Activity Diagram Kelola Data Pengguna

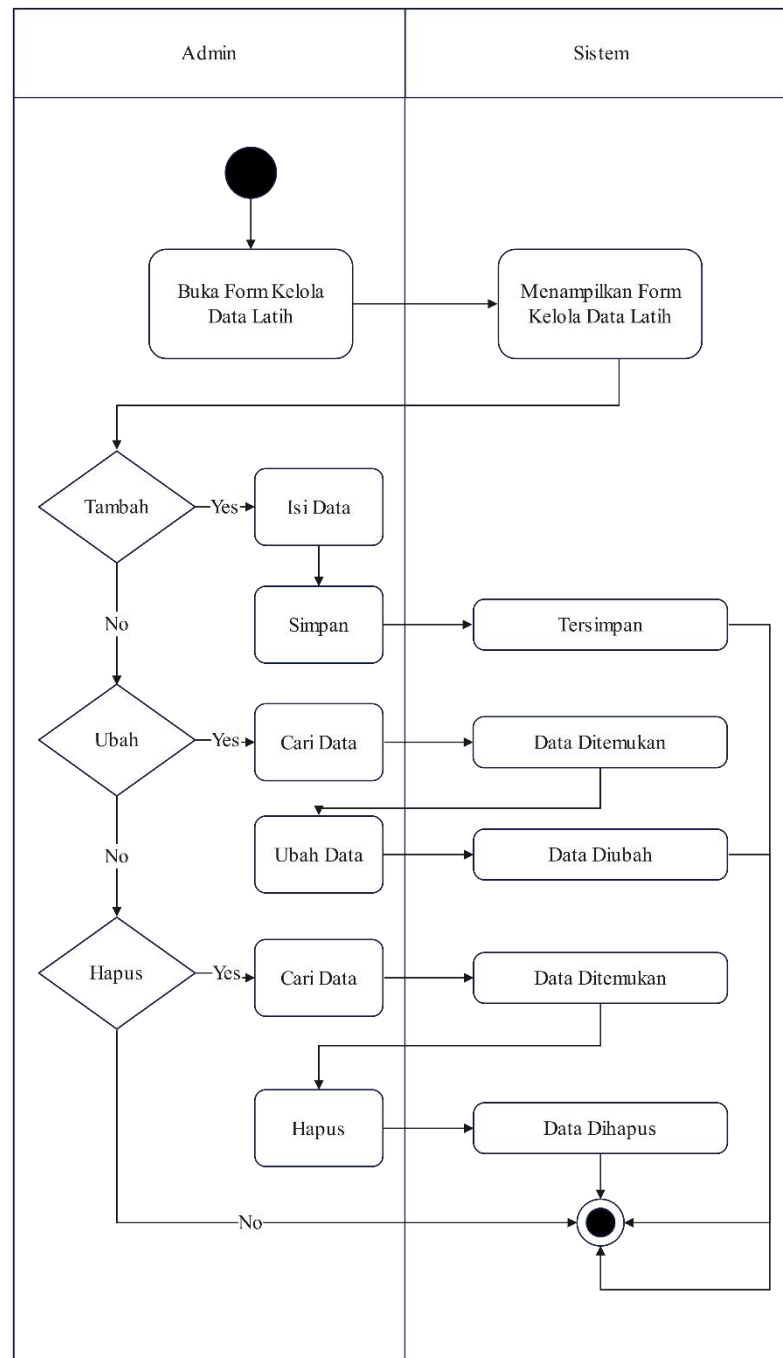
Dalam aktivitas ini, Admin mengakses menu *Kelola Data Pengguna*, lalu sistem menampilkan daftar pengguna. Admin dapat menambah, mengedit, atau menghapus data pengguna. Setelah aksi dilakukan, sistem menyimpan perubahan. Dapat dilihat pada gambar dibawah ini.



Gambar 3. 5 Activity Diagram Pengguna

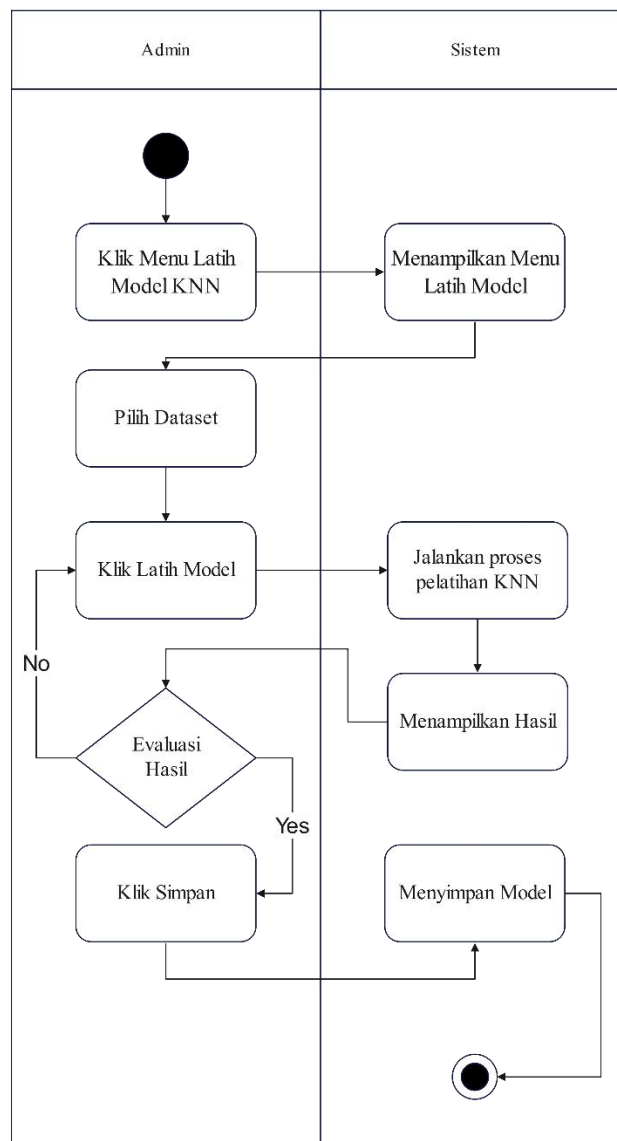
4. Activity Diagram Kelola Dataset Latih

Dalam aktivitas ini Admin membuka menu *Kelola Data Latih*, sistem menampilkan daftar dataset. Admin dapat menambahkan data latih baru berupa gejala, atribut, dan label penyakit, lalu menyimpan ke database. Dapat dilihat pada gambar dibawah ini.



5. Activity Diagram Latih Model KNN

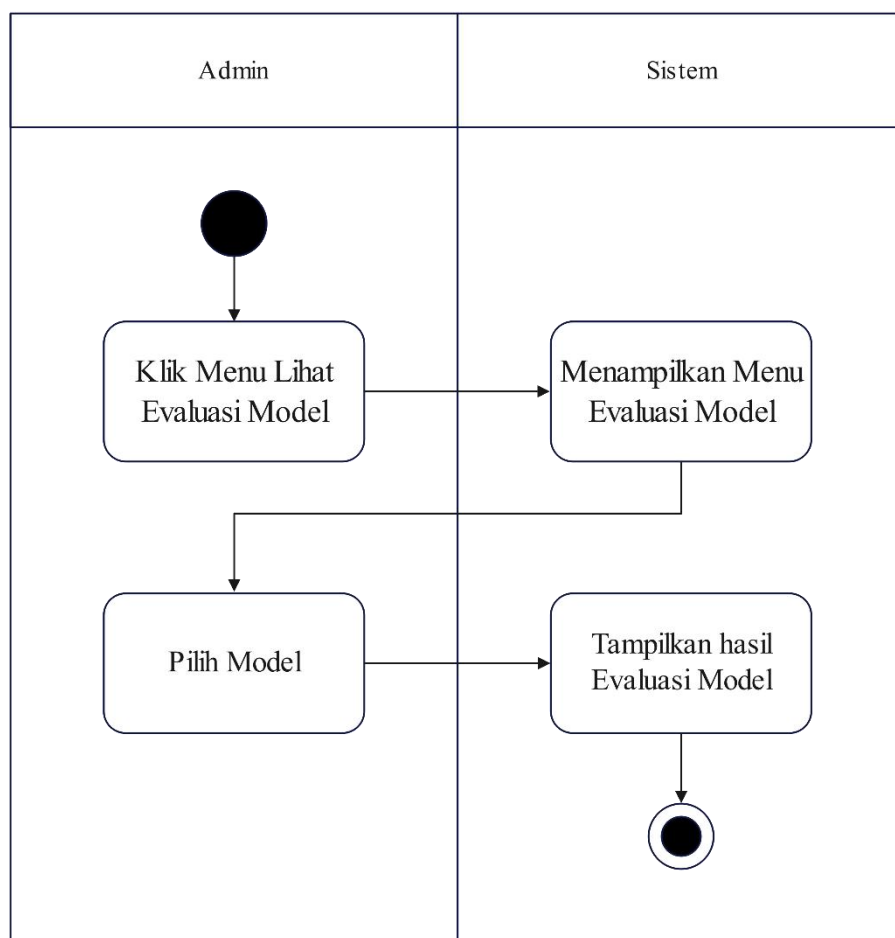
Dalam aktivitas ini, Admin membuka menu *Latih Model KNN*, lalu menekan tombol untuk memulai pelatihan. Sistem memuat data latih yang telah dimasukkan sebelumnya, memproses pelatihan menggunakan algoritma K-Nearest Neighbor, menyimpan model yang dihasilkan. Dapat dilihat pada gambar dibawah ini.



Gambar 3. 6 Activity Diagram Latih Model

6. Activity Diagram Lihat Evaluasi Model

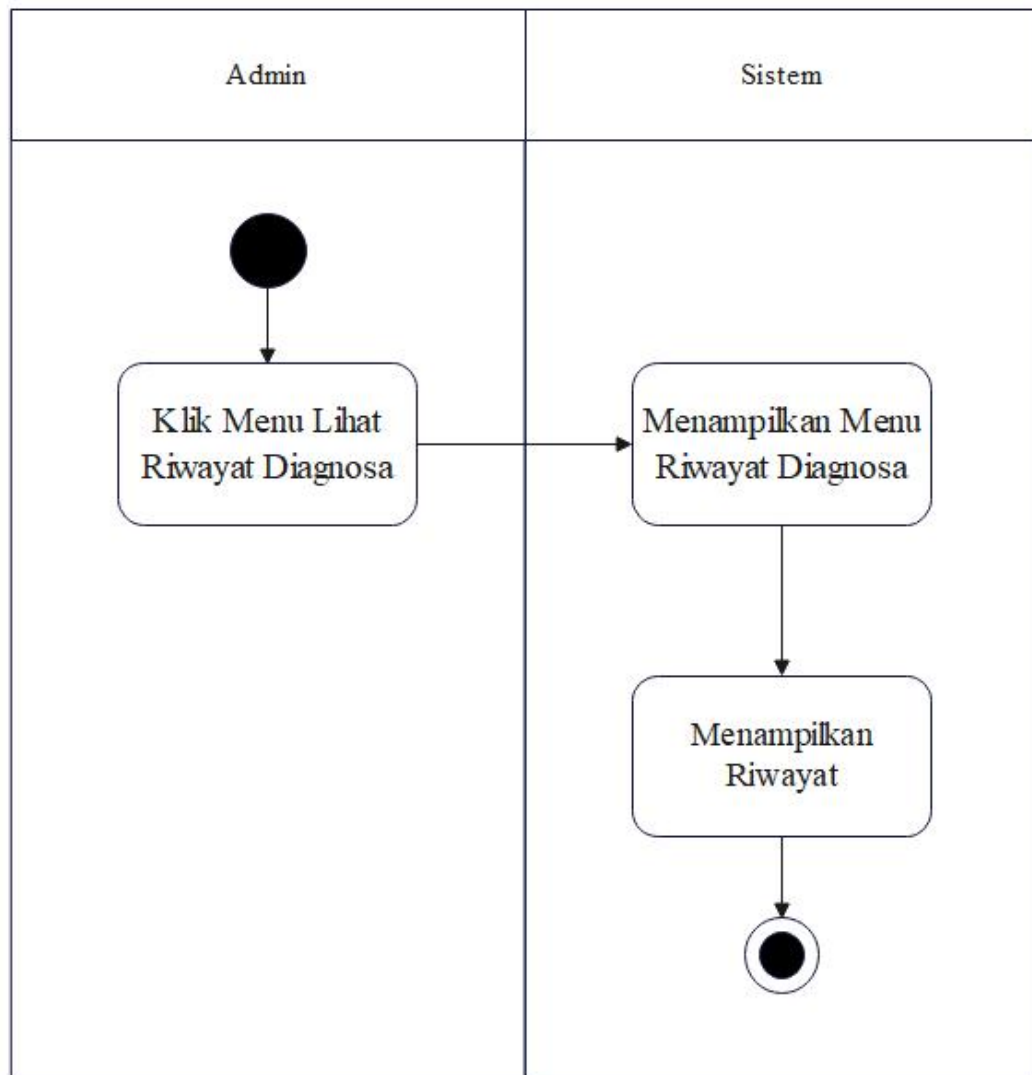
Dalam aktivitas ini, Admin mengakses menu Lihat Evaluasi Model. Sistem mengambil hasil evaluasi dari model KNN yang telah dilatih sebelumnya dan menampilkan metrik evaluasi seperti akurasi, precision, recall, dan F1-score kepada admin. Dapat dilihat pada gambar dibawah ini



Gambar 3. 7 Activity Diagram Lihat Evaluasi Model

7. Activity Diagram Lihat Riwayat Diagnosa

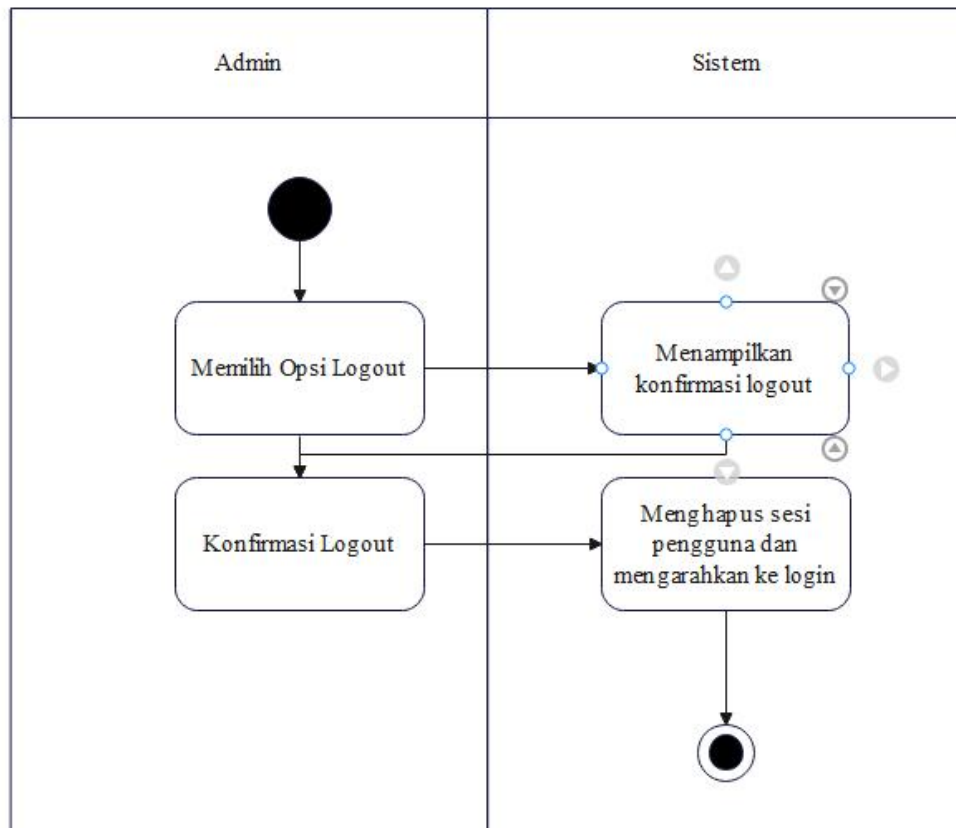
Dalam aktivitas ini, Admin membuka menu Lihat Riwayat Diagnosa. Sistem mengambil seluruh data hasil diagnosa yang pernah dilakukan oleh dokter gigi, lalu menampilkannya dalam bentuk tabel atau daftar lengkap.



Gambar 3. 8 Activity Diagram Riwayat Diagnosa

8. Activity Diagram Logout

Aktivitas yang terjadi saat pengguna keluar dari sistem dapat digambarkan pada diagram dibawah ini

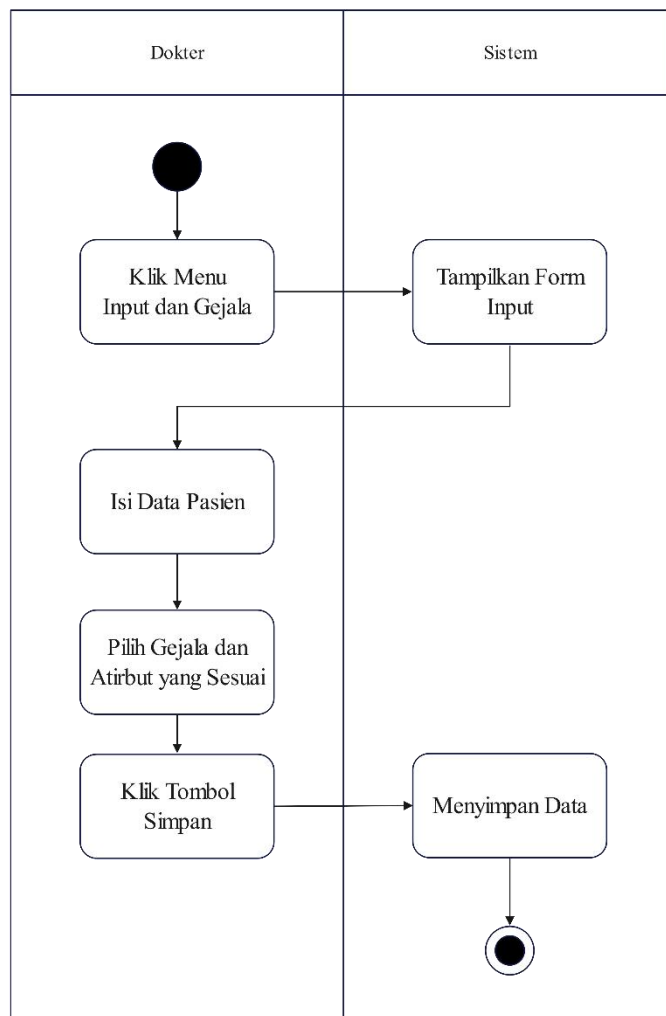


Gambar 3. 9 Activity Diagram Logout Admin

3.2.2.2. Activity Diagram Dokter

1. Activity Diagram Input Data dan Gejala

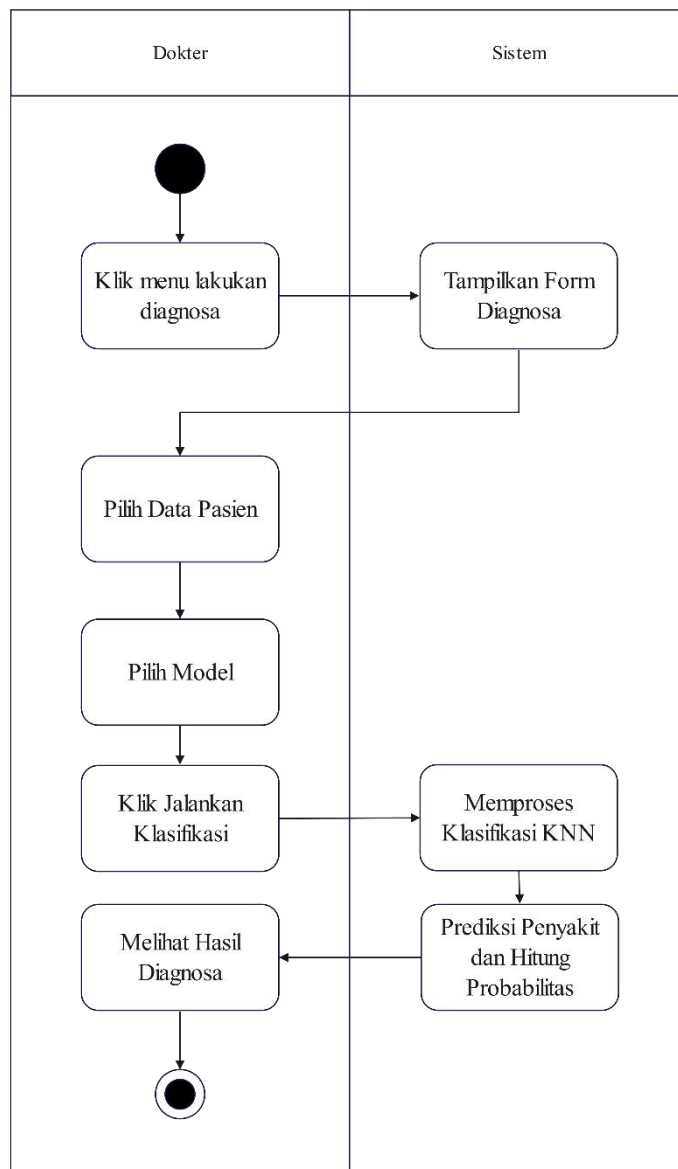
Dalam aktivitas ini, Dokter gigi mengakses menu *Input Data & Gejala*, lalu mengisi data pasien seperti nama, umur, dan memilih gejala serta atribut yang dialami pasien. Setelah seluruh data diisi, dokter menyimpan data tersebut dan sistem menyimpannya ke dalam database untuk proses diagnosa. Dapat dilihat pada gambar dibawah ini.



Gambar 3. 10 Activity Diagram Input dan Gejala

2. Activity Diagram Lakukan Diagnosa

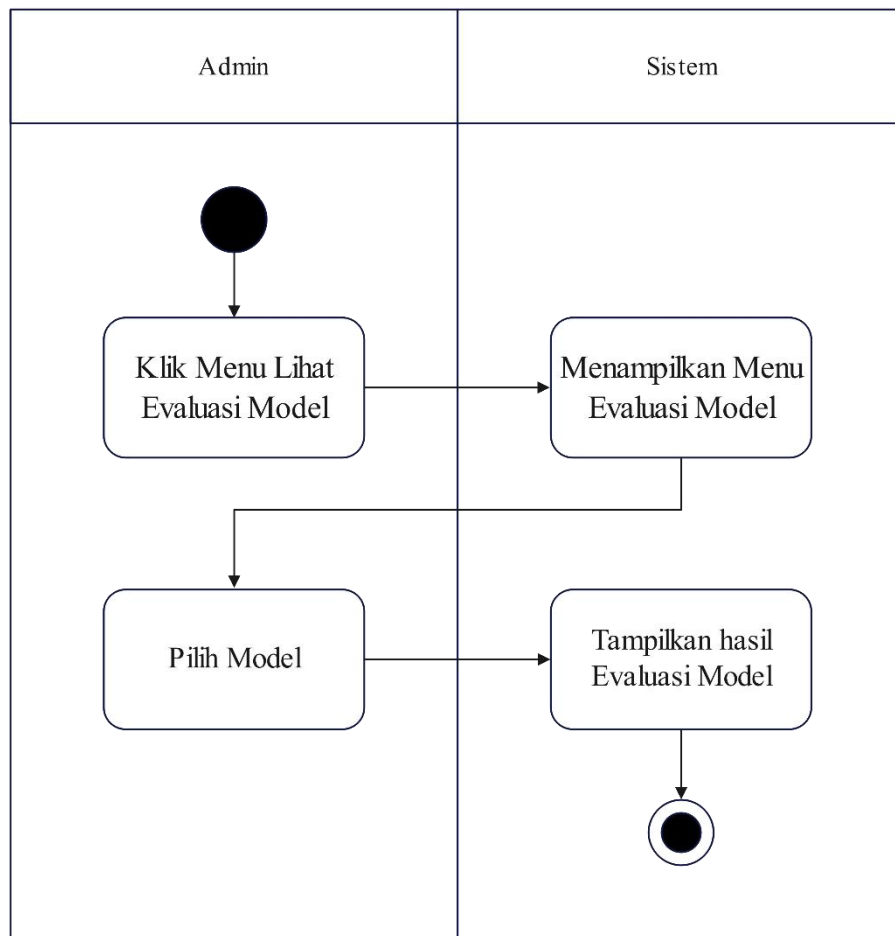
Dalam aktivitas ini, Setelah menginput data pasien, dokter gigi mengakses menu *Lakukan Diagnosa*. Sistem mengambil data pasien terbaru, kemudian menjalankan proses klasifikasi menggunakan model K-Nearest Neighbor. Hasil prediksi penyakit ditampilkan kepada dokter bersamaan dengan tingkat probabilitas.



Gambar 3. 11 Activity Diagram Lakukan Diagnosa

3. *Activity Diagram* Lihat Hasil Diagnosa

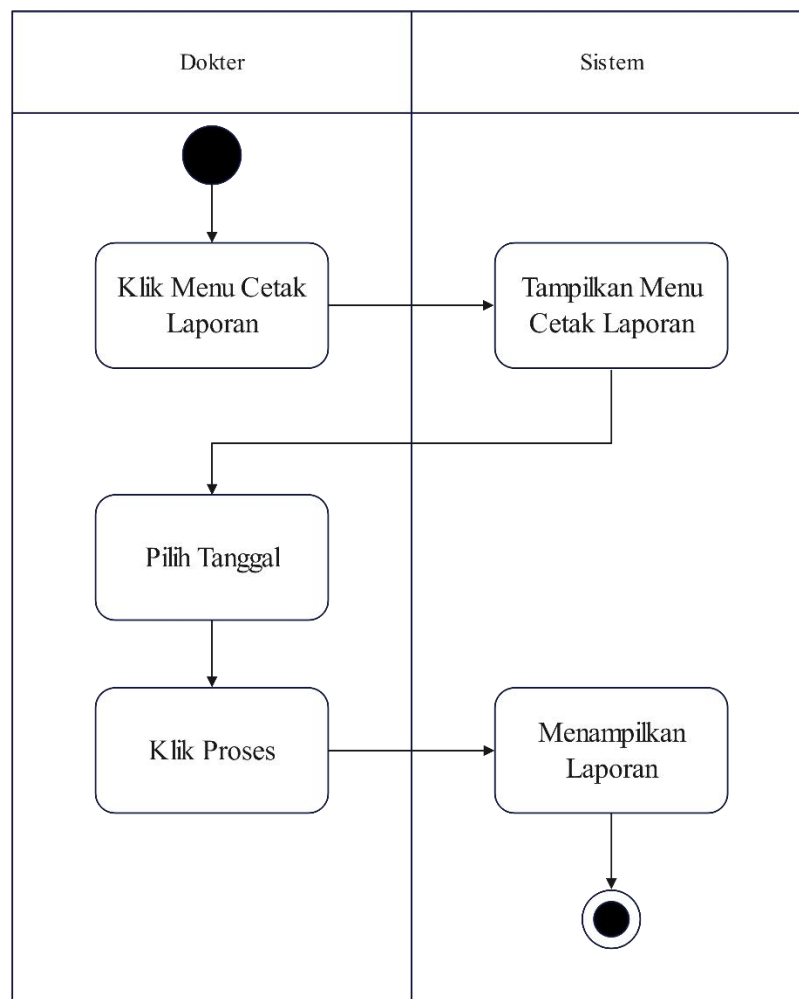
Dalam aktivitas ini, Setelah proses diagnosa selesai, dokter gigi mengakses halaman *Lihat Hasil Diagnosa* untuk meninjau hasil prediksi dari sistem. Sistem menampilkan jenis penyakit yang telah terdeteksi beserta tingkat keyakinan (probabilitas) dari model KNN.



Gambar 3. 12 Activity Diagram Lihat Evaluasi Model

4. *Activity Diagram* Cetak Laporan Diagnosa

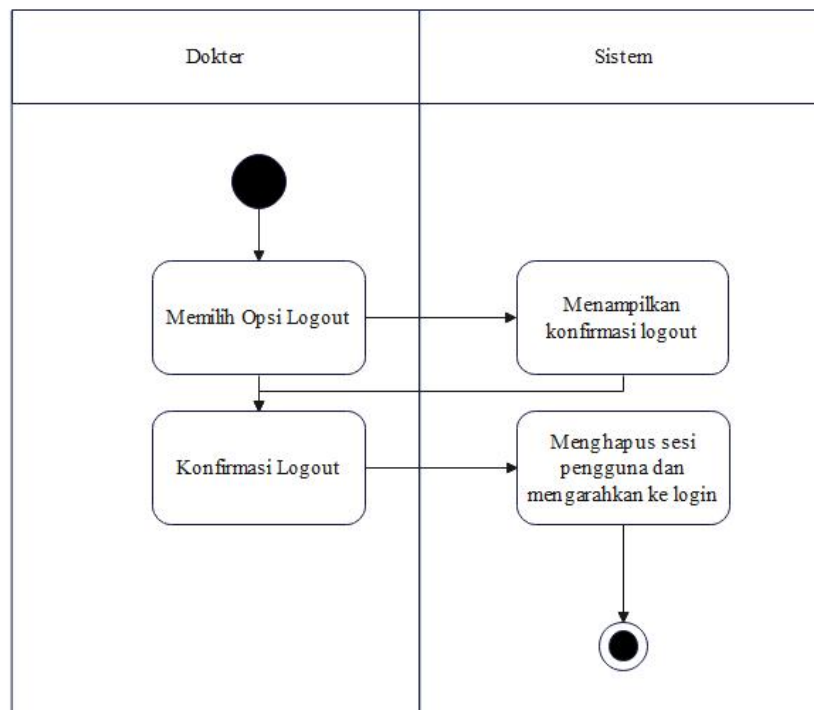
Dalam aktivitas ini dokter gigi dapat mencetak laporan untuk pasien. Dokter mengklik tombol Cetak Laporan, lalu sistem memproses data hasil diagnosa dan menghasilkan file PDF yang berisi identitas pasien, gejala yang diinput, hasil diagnosa, dan rekomendasi. Laporan dapat langsung diunduh atau dicetak. Dapat dilihat pada gambar dibawah ini.



Gambar 3. 13 Activity Diagram Cetak Laporan

5. *Activity Diagram Logout*

Aktivitas saat dokter keluar dari sistem dapat dilihat pada gambar dibawah ini

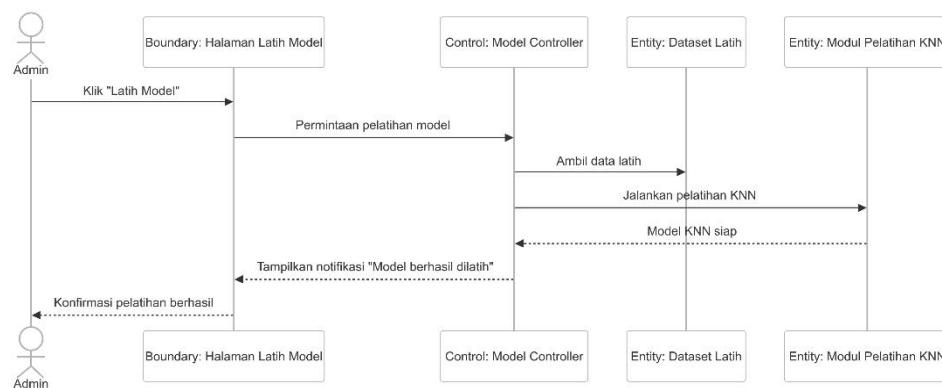


Gambar 3. 14 Activity Diagram Logout Dokter

3.2.3. Sequence Diagram

1. Sequence Diagram Admin

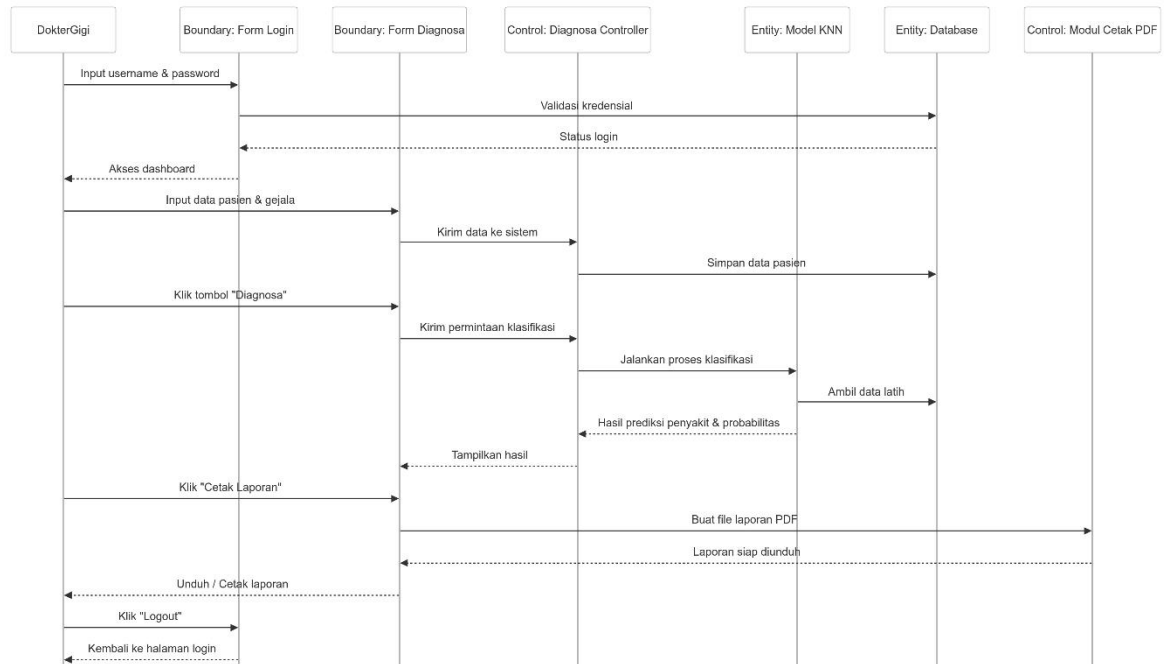
Diagram ini menggambarkan proses saat admin melatih model K-Nearest Neighbor. Admin mengakses menu pelatihan model, sistem mengambil dataset latih dari database, lalu memproses pelatihan dan menyimpan model hasil klasifikasi ke dalam sistem. Dapat dilihat pada gambar dibawah ini



Gambar 3. 15 Sequence Diagram Admin

2. Sequence Diagram Dokter Gigi

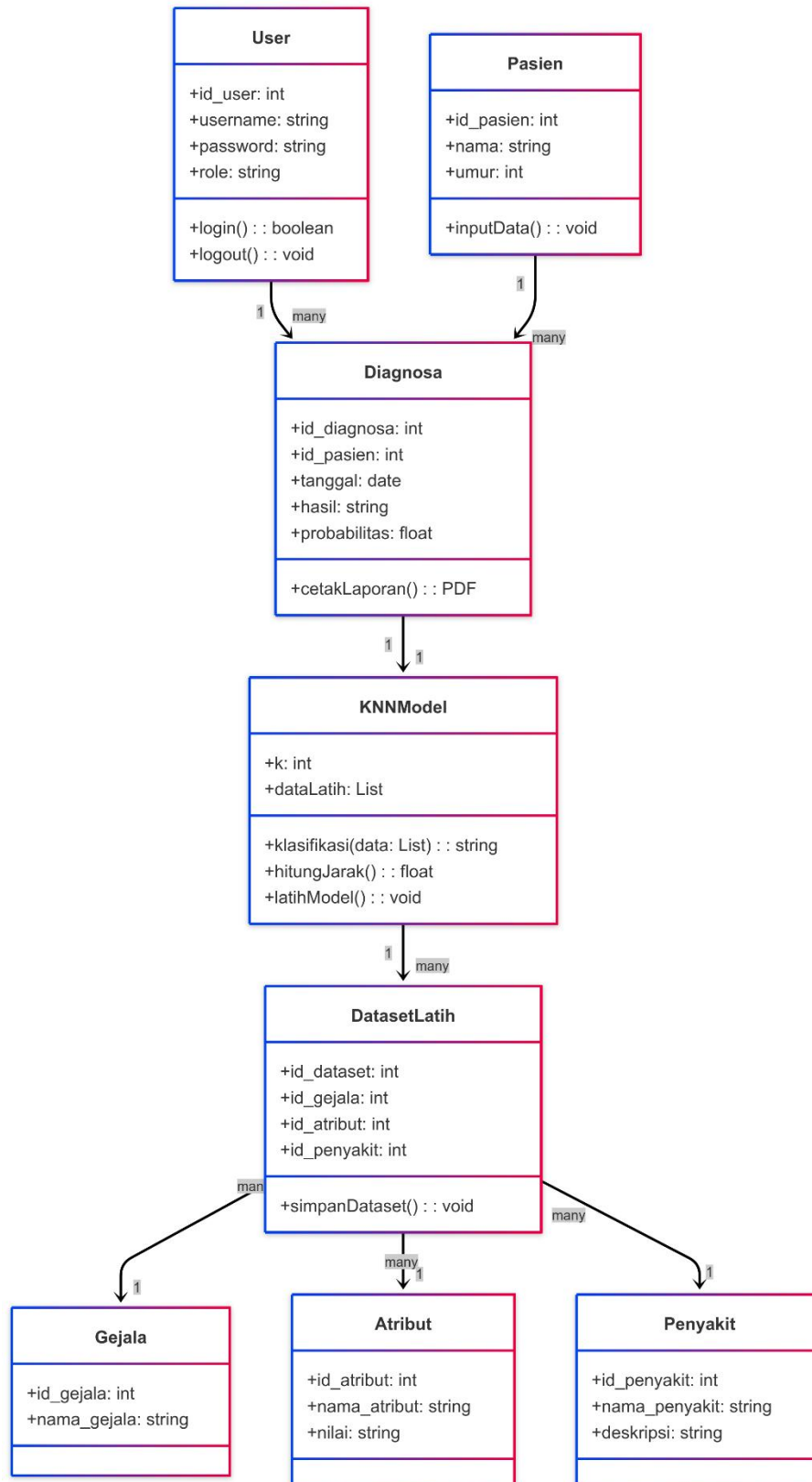
Dalam diagram ini, Sequence diagram ini menggambarkan alur interaksi dokter gigi saat menggunakan sistem, dimulai dari login, input data pasien, melakukan diagnosa menggunakan algoritma KNN, melihat hasil prediksi, hingga mencetak laporan diagnosa dalam format PDF.



Gambar 3. 16 Sequence Diagram Dokter

3.2.4. Class Diagram

Class diagram ini menggambarkan struktur objek dalam sistem. Terdapat beberapa kelas utama: User, Pasien, Gejala, Atribut, Penyakit, DatasetLatih, HasilDiagnosa, dan ModelKNN. Relasi antar kelas mencerminkan proses bisnis, seperti dokter melakukan input data pasien, memilih gejala, dan mendapatkan hasil klasifikasi dari model KNN. Admin dapat mengelola data master dan melakukan pelatihan model. Dapat dilihat pada gambar dibawah ini



Gambar 3. 17 Class Diagram

3.3 Simulasi Perhitungan K-Nearest Neighbor

Simulasi ini bertujuan untuk menggambarkan bagaimana algoritma K-Nearest Neighbor (KNN) digunakan untuk menentukan jenis penyakit karies gigi berdasarkan kombinasi gejala dan atribut yang dimiliki pasien. Dalam sistem ini, satu pasien dapat memiliki lebih dari satu gejala, dan setiap gejala memiliki atribut tertentu seperti intensitas atau tingkat keparahan. Jarak antar pasien dihitung menggunakan Euclidean Distance, dan klasifikasi dilakukan berdasarkan mayoritas tetangga terdekat (jumlah k).

Adapun Langkah-langkah dalam proses KNN adalah sebagai berikut :

1. Menentukan nilai parameter k (misalnya $k = 3$)
2. Mengambil data latih yang telah ditandai label penyakit
3. Melakukan normalisasi (jika diperlukan)
4. Menghitung jarak menggunakan *Euclidean Distance* untuk masing-masing pasien latih
5. Mengurutkan berdasarkan jarak terkecil
6. Mengambil k tetangga terdekat
7. Menentukan kelas mayoritas sebagai hasil klasifikasi

Berikut contoh data latih yang dapat digunakan sebagai data untuk memprediksi karies gigi :

Tabel 3. 1 Contoh Data Latih

ID	Gigi ngilu	Bercak Putih	Tingkat Nyeri	Lama keluhan	Penyakit
D1	1	0	2	1	Karies Awal
D2	1	1	3	2	Karies Sedang
D3	0	1	1	2	Karies Lanjut
D4	1	1	2	3	Karies Sedang

Keterangan:

Gejala 1 = "Gigi terasa ngilu"

Gejala 2 = "Terdapat bercak putih"

Atribut 1 = Tingkat nyeri (1=rendah, 2=sedang, 3=tinggi)

Atribut 2 = Lama keluhan (1=sebentar, 2=sedang, 3=lama)

Berikut merupakan contoh data uji

Tabel 3. 2 Contoh Data yang ingin diprediksi

Gigi ngilu	Bercak Putih	Tingkat Nyeri	Lama keluhan
1	1	2	2

Maka perhitungan *Euclidean Distance* adalah sebagai berikut

Jarak ke Pasien D1

$$D1 = \sqrt{(1 - 1)^2 + (1 - 0)^2 + (2 - 2)^2 + (2 - 1)^2}$$

$$D1 = \sqrt{0 + 1 + 0 + 1}$$

$$D1 = \sqrt{2}$$

$$D1 = 1.41$$

Jarak ke Pasien D2

$$D2 = \sqrt{(1 - 1)^2 + (1 - 1)^2 + (2 - 3)^2 + (2 - 2)^2}$$

$$D2 = \sqrt{0 + 0 + 1 + 1}$$

$$D2 = \sqrt{1}$$

$$D2 = 1$$

Jarak ke Pasien D3

$$D3 = \sqrt{(1-0)^2 + (1-1)^2 + (2-1)^2 + (2-2)^2}$$

$$D3 = \sqrt{1+0+1+0}$$

$$D3 = \sqrt{2}$$

$$D3 = 1.41$$

Jarak ke Pasien D4

$$D4 = \sqrt{(1-1)^2 + (1-1)^2 + (2-2)^2 + (2-3)^2}$$

$$D4 = \sqrt{0+0+0+1}$$

$$D4 = \sqrt{1}$$

$$D4 = 1$$

Berikut hasil urutan data dari terdekat ke terjauh

Tabel 3. 3 Contoh Urutan Data dari Terdekat

ID	Penyakit	Jarak
D2	Karies Sedang	1.00
D4	Karies Sedang	1.00
D1	Karies Awal	1.41
D3	Karies Lanjut	1.41

Jika mengambil K=3 maka diambil 3 tetangga terdekat, yaitu D2, D4, dan D1, jika dilihat, mayoritas kelas dari 3 tetangga adalah Karies Sedang, maka pasien terkait diprediksi mengalami Karies Sedang.

3.4 Perancangan Basis Data

Basis data sistem ini dirancang untuk menyimpan informasi terkait pengguna (admin dan dokter gigi), pasien, gejala, atribut, penyakit, dataset latih, serta hasil diagnosa. Struktur basis data disusun secara relasional, dengan relasi antar tabel yang mencerminkan hubungan satu ke banyak (one-to-many) maupun banyak ke satu (many-to-one). Setiap entitas utama diimplementasikan dalam

bentuk tabel terpisah, dengan foreign key untuk menghubungkan antar entitas. Perancangan ini memudahkan pengelolaan data dan integrasi dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor.

1. Rancangan Tabel *users*

Tabel ini menyimpan informasi akun pengguna sistem, yang terdiri dari admin dan dokter gigi. Setiap pengguna memiliki username, password, dan peran.

Tabel 3. 4 Rancangan Tabel user

Nama Field	Tipe Data	Keterangan
id_user	INT (PK)	ID unik pengguna
username	VARCHAR(50)	Username akun login
password	VARCHAR(255)	Password (dienkripsi)
role	ENUM	Peran: 'admin' atau 'dokter'

2. Rancangan Tabel Pasien

Tabel ini Menyimpan informasi dasar pasien, seperti nama, umur, dan tanggal kunjungan. Data ini menjadi acuan dalam proses diagnosa.

Tabel 3. 5 Rancangan Tabel Pasien

Nama Field	Tipe Data	Keterangan
id_pasien	INT (PK)	ID unik pasien
nama_pasien	VARCHAR(100)	Nama lengkap pasien
umur	INT	Umur pasien

3. Rancangan Tabel Gejala

Tabel ini Berisi daftar gejala yang digunakan dalam input diagnosa dan dataset latih.

Tabel 3. 6 Rancangan Tabel Gejala

Nama Field	 Tipe Data	Keterangan
id_gejala	INT (PK)	ID unik gejala
nama_gejala	VARCHAR(100)	Nama atau deskripsi gejala

4. Rancangan Tabel Atribut

Tabel ini Digunakan untuk menggambarkan karakteristik atau kondisi dari gejala yang dialami pasien. Contoh: intensitas, frekuensi, atau tingkat keparahan.

Tabel 3. 7 Rancangan Tabel Atribut

Nama Field	 Tipe Data	Keterangan
id_atribut	INT (PK)	ID unik atribut
nama_atribut	VARCHAR(100)	Nama atribut
nilai	VARCHAR(50)	Nilai dari atribut (misalnya: berat, sedang, ringan)

5. Rancangan Tabel Penyakit

Tabel ini menyimpan label target klasifikasi, yaitu jenis penyakit karies gigi. Data penyakit dikaitkan ke dataset latih sebagai output dari proses pembelajaran mesin.

Tabel 3. 8 Rancangan Tabel Penyakit

Nama Field	 Tipe Data	Keterangan
id_penyakit	INT (PK)	ID unik penyakit
nama_penyakit	VARCHAR(100)	Nama penyakit
deskripsi	TEXT	Penjelasan penyakit karies gigi

6. Rancangan Tabel Diagnosa

Tabel ini Menyimpan hasil diagnosa yang dilakukan oleh dokter terhadap pasien. Data ini berisi hasil klasifikasi penyakit, nilai probabilitas, dan tanggal pemeriksaan.

Tabel 3. 9 Rancangan Tabel Diagnosa

Nama Field	Tipe Data	Keterangan
id_diagnosa	INT (PK)	ID unik hasil diagnosa
id_user	INT (FK)	Relasi ke user (dokter yang memeriksa)
id_pasien	INT (FK)	Relasi ke pasien
tanggal	DATE	Tanggal diagnosa
hasil	VARCHAR(100)	Prediksi penyakit dari klasifikasi
probabilitas	FLOAT	Tingkat keyakinan terhadap hasil klasifikasi

7. Rancangan Tabel Dataset Latih

Tabel ini Menyimpan kombinasi data yang digunakan untuk melatih model KNN. Setiap entri merepresentasikan hubungan antara satu gejala, atribut, dan label penyakit.

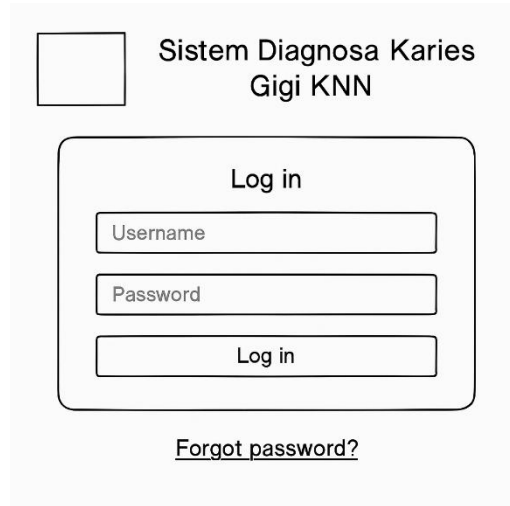
Tabel 3. 10 Rancangan Tabel Dataset Latih

Nama Field	Tipe Data	Keterangan
id_dataset	INT (PK)	ID unik dataset
id_gejala	INT (FK)	Relasi ke gejala
id_atribut	INT (FK)	Relasi ke atribut
id_penyakit	INT (FK)	Relasi ke penyakit sebagai label

3.5 Desain *Interface*

1. Rancangan menu *Login*

Rancangan menu login dapat dilihat pada gambar dibawah ini

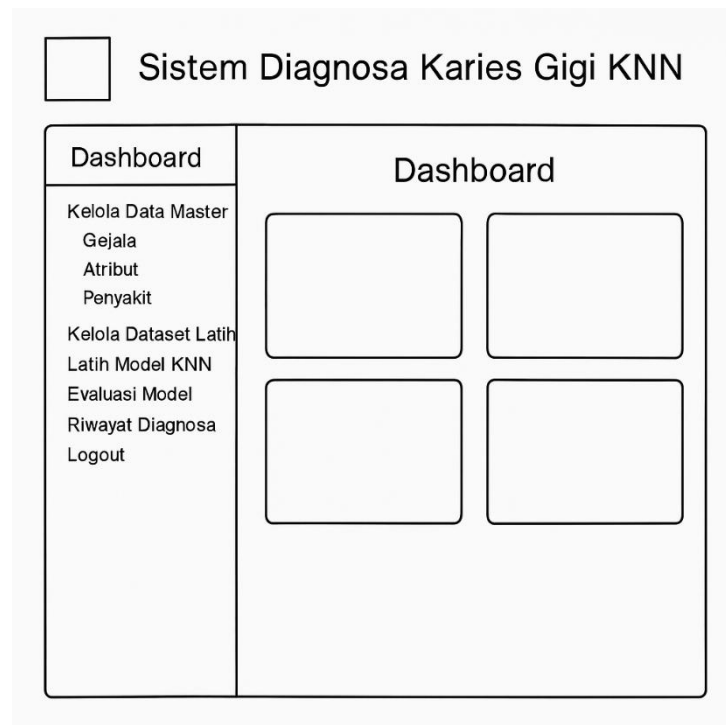


The login form is titled "Sistem Diagnosa Karies Gigi KNN". It features a "Log in" header, followed by input fields for "Username" and "Password", and a "Log in" button. Below the button is a link for "[Forgot password?](#)".

Gambar 3. 18 Rancangan Menu Login

2. Rancangan menu *Dashboard*

Rancangan menu *Dashboard* dapat dilihat pada gambar dibawah ini



The dashboard layout is titled "Sistem Diagnosa Karies Gigi KNN". It consists of a sidebar menu on the left and a main content area on the right. The sidebar menu includes the following items: "Dashboard", "Kelola Data Master" (with sub-items "Gejala", "Atribut", "Penyakit"), "Kelola Dataset Latih", "Latih Model KNN", "Evaluasi Model", "Riwayat Diagnosa", and "Logout". The main content area is titled "Dashboard" and contains four empty rectangular boxes arranged in a 2x2 grid.

Gambar 3. 19 Rancangan Menu Dashboard

3. Rancangan menu Kelola Gejala

Rancangan menu Kelola Gejala dapat dilihat pada gambar dibawah ini



Gambar 3. 20 Rancangan Menu Kelola Gejala

4. Rancangan Menu Kelola Atribut

Rancangan menu Kelola atribut dapat dilihat pada gambar dibawah ini

Logo

Sistem Diagnosa Karies Gigi KNN

Dashboard

Kelola Data Master

Gejala

Atribut

Penyakit

Kelola Dataset Latih

Latih Model KNN

Evaluasi Model

Riwayat Diagnosa

Logout

Kelola Atribut

Tambah

No	Nama Atribut	Aksi
1		Edit Hapus
2		Edit Hapus
3		Edit Hapus

Gambar 3. 21 Kelola Data Atribut

5. Rancangan Menu Kelola Penyakit

Rancangan menu Kelola penyakit dapat dilihat pada gambar dibawah ini

Logo

Sistem Diagnosa Karies Gigi KNN

Dashboard

Kelola Data Master
Gejala
Atribut
Penyakit
Kelola Dataset Latih
Latih Model KNN
Riwayat **Penyakit**
Riwayat Diagnosa
Logout

Kelola Penyakit

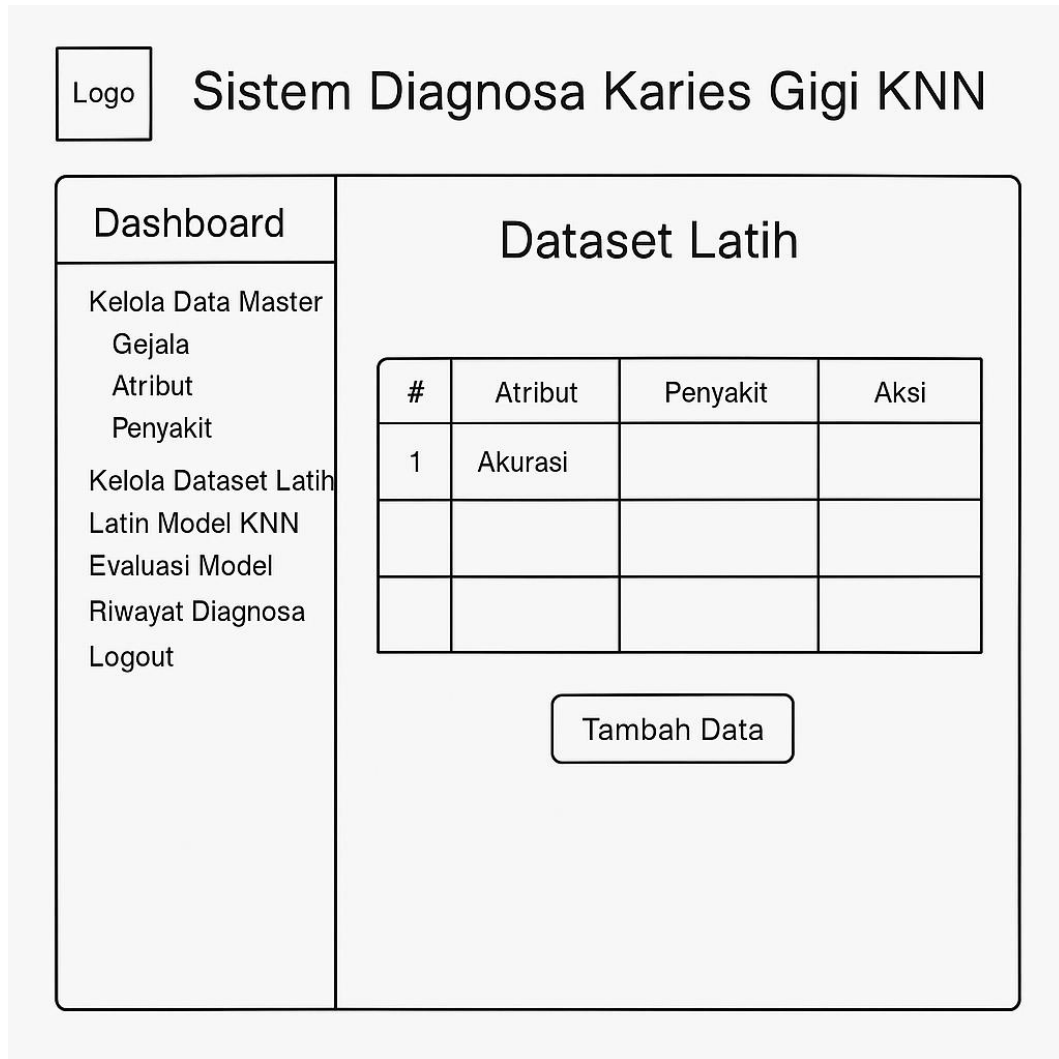
Tambah

No	Nama Penyakit	Aksi
1		Edit Hapus
2		Edit Hapus
3		Edit Hapus

Gambar 3. 22 Rancangan Menu Kelola Penyakit

6. Rancangan Menu Kelola Dataset Latih

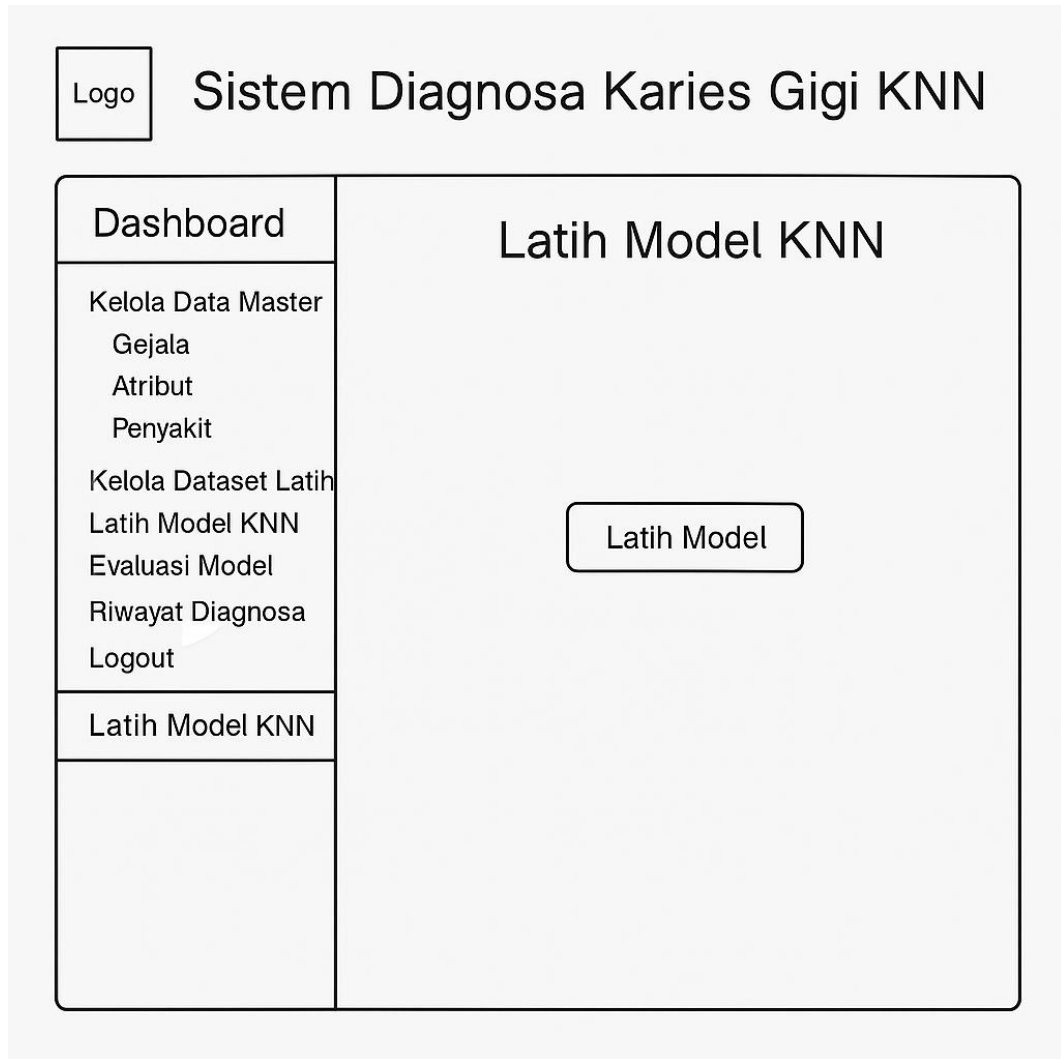
Rancangan menu Kelola Dataset Latih dapat dilihat pada gambar dibawah ini



Gambar 3. 23 Rancangan Kelola Dataset Latih

7. Rancangan Menu Latih Model KNN

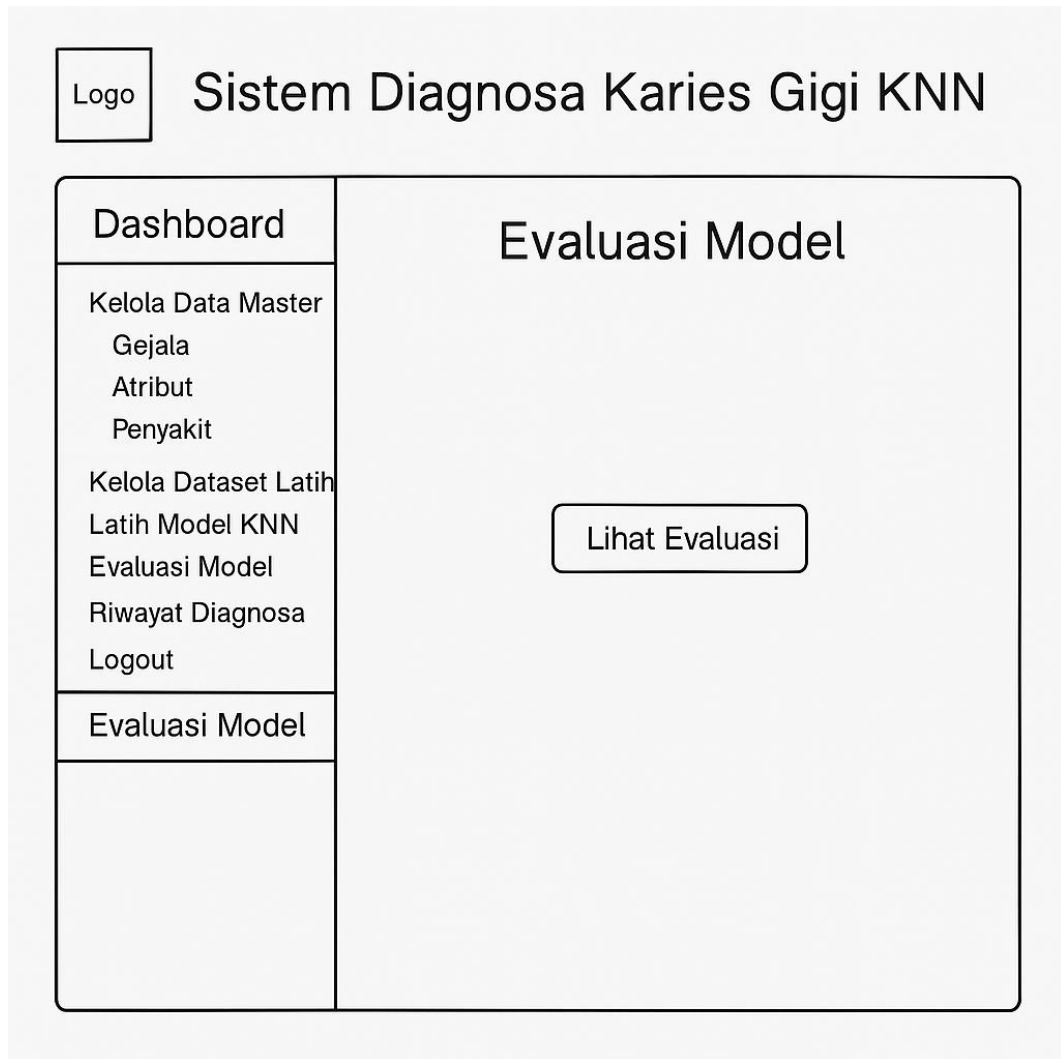
Rancangan menu Kelola Dataset Latih dapat dilihat pada gambar dibawah ini



Gambar 3. 24 Rancangan Menu Latih Model KNN

8. Rancangan Menu Evaluasi Model

Rancangan menu Evaluasi Model dapat dilihat pada gambar dibawah ini



Gambar 3. 25 Rancangan Menu Evaluasi Model

9. Rancangan Menu Riwayat Diagnosa

Rancangan menu Riwayat Diagnosa dapat dilihat pada gambar dibawah ini

Dashboard		Riwayat Diagnosa			
Kelola Data Master		ID	Pasien	Tanggal	Diagnosa
Gejala					
Atribut					
Penyakit					
Kelola Dataset Latih					
Latih Model KNN					
Riwayat Diagnosa					
Logout					
Riwayat Diagnosa					

Gambar 3. 26 Rancangan Menu Riwayat Diagnosa

10. Rancangan Form Diagnosa Pasien Baru

Rancangan Form Diagnosa Pasien baru dapat dilihat pada gambar dibawah ini

Form Diagnosa Pasien Baru

Data Pasien

Nama

Usia

Gejala

☐ Gejala 1

☐ Gejala 2

☐ Gejala 3

☐ Gejala 4

Atribut

Diagnosa

Gambar 3. 27 Form Diagnosa Pasien Baru

BAB IV

HASIL DAN PEMBAHASAN

4.1. Perhitungan Manual Metode *K-Nearest Neighbor*

Tujuan dari bagian ini adalah untuk menyajikan demonstrasi perhitungan manual yang terperinci dari algoritma *K-Nearest Neighbor*. Simulasi ini bertujuan untuk memberikan gambaran yang jelas dan transparan mengenai setiap langkah yang dilakukan oleh algoritma, mulai dari persiapan data, perhitungan jarak, hingga proses klasifikasi akhir. Dengan menyajikan perhitungan secara manual, diharapkan pembaca dapat memahami mekanisme internal KNN dalam konteks deteksi karies gigi.

Metodologi yang digunakan dalam simulasi ini mengacu pada prinsip-prinsip dasar algoritma KNN yang telah diuraikan pada Bab II, seperti penggunaan metrik jarak Euclidean dan klasifikasi berdasarkan suara mayoritas (majority voting). Untuk memastikan kedalaman dan kejelasan analisis, simulasi ini akan menggunakan sampel data yang representatif, sebuah praktik yang umum dilakukan dalam penulisan akademis untuk menyeimbangkan antara detail perhitungan dan keterbacaan. Struktur penyajian dan tingkat kerincian dalam perhitungan ini juga merujuk pada praktik terbaik yang ditunjukkan dalam penelitian sejenis untuk menjaga rigor akademis.

Sebagai dasar perhitungan, digunakan sampel data latih dan data uji yang merupakan representasi dari data rekam medis pasien. Data ini mencakup fitur-fitur yang relevan untuk diagnosis karies gigi, seperti ada atau tidaknya gejala spesifik dan tingkat keparahan yang menyertainya.

4.1.1. Sampel Data Latih untuk Perhitungan Manual

Tabel berikut menyajikan 15 sampel data rekam medis pasien yang akan digunakan sebagai data latih (*training data*). Setiap baris merepresentasikan satu kasus pasien dengan empat fitur prediktor dan satu kelas target (diagnosa aktual).

Tabel 4. 1 Sampel Data Latih

ID Pasien	Gigi Ngilu (G1)	Bercak Putih (G2)	Tingkat Nyeri (A1)	Lama Keluhan (hari) (A2)	Penyakit (Kelas Aktual)
L01	1	1	1	3	Karies Awal
L02	1	0	2	2	Karies Awal
L03	0	1	1	5	Karies Awal
L04	1	1	2	1	Karies Awal
L05	0	1	1	4	Karies Awal
L06	1	1	3	7	Karies Sedang
L07	1	0	4	10	Karies Sedang
L08	1	1	4	5	Karies Sedang
L09	0	1	3	8	Karies Sedang
L10	1	0	5	6	Karies Sedang
L11	1	1	5	14	Karies Lanjut
L12	1	0	4	12	Karies Lanjut
L13	1	1	5	9	Karies Lanjut
L14	0	1	4	11	Karies Lanjut
L15	1	0	5	13	Karies Lanjut

4.1.2. Sampel Data Uji untuk Perhitungan Manual

Tabel berikut menyajikan 3 sampel data pasien baru yang akan diklasifikasikan oleh model. Kolom Penyakit sengaja dikosongkan karena nilai inilah yang akan diprediksi oleh algoritma KNN.

Tabel 4. 2 Sampel Data Uji

ID Pasien	Gigi Ngilu	Bercak Putih (G2)	Tingkat Nyeri (A1)	Lama Keluhan	Penyakit (Kelas
-----------	------------	-------------------	--------------------	--------------	-----------------

	(G1)			(hari) (A2)	Prediksi
U01	1	1	2	4	?
U02	1	0	4	8	?
U03	0	1	5	10	?

4.1.3. Pra-Pemrosesan Data : Standardisasi Fitur

Sebelum menerapkan rumus jarak Euclidean, tahap pra-pemrosesan data menjadi langkah yang fundamental dan tidak dapat diabaikan. Algoritma KNN sangat sensitif terhadap skala atau rentang nilai dari fitur-fitur yang digunakan. Dalam dataset yang dimiliki, terdapat perbedaan skala yang signifikan antara fitur gejala (Gigi Ngilu dan Bercak Putih dengan rentang biner 0-1) dan fitur atribut (Tingkat Nyeri dengan rentang 1-5 dan Lama Keluhan dengan rentang 1-14). Jika normalisasi tidak dilakukan, fitur dengan rentang nilai yang lebih besar akan secara tidak proporsional mendominasi perhitungan jarak.

Untuk mengatasi masalah perbedaan skala, metode normalisasi Z-Score diterapkan pada fitur-fitur numerik, yaitu Tingkat Nyeri (A1) dan Lama Keluhan (A2). Normalisasi Z-Score menstandarkan setiap nilai dalam suatu fitur dengan mengubahnya menjadi representasi dari seberapa banyak standar deviasi nilai tersebut dari rata-rata. Proses ini mengubah distribusi data sehingga memiliki rata-rata.

$$Z = \frac{X - \mu}{\sigma}$$

Parameter statistik (rata-rata dan standar deviasi) harus dihitung hanya dari data latih (Tabel 4.1) untuk mencegah kebocoran informasi (*data leakage*) dari data uji ke dalam proses pelatihan.

- a) Langkah 1 : Menghitung Rata-Rata dari Data Latih dan Standar Deviasi

Untuk Fitur Tingkat Nyeri (A1), Data: {1, 2, 1, 2, 1, 3, 4, 4, 3, 5, 5, 4, 5, 4, 5}

Rata-Ratanya adalah :

$$\mu = \frac{1 + 2 + 1 + 2 + 1 + 3 + 4 + 4 + 3 + 5 + 5 + 4 + 5 + 4 + 5}{15}$$

$$\mu = 3,27$$

Standar deviasinya adalah

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2}$$

$$\sigma = \sqrt{\frac{(1 - 3,27)^2 + (2 - 3,27)^2 + \dots + (5 - 3,27)^2}{15}}$$

$$\sigma = \sqrt{\frac{31.73}{15}}$$

$$\sigma = \sqrt{2.115}$$

$$\sigma = 1.45$$

Dengan perhitungan yang sama, dilakukan perhitungan untuk lama keluhan dengan rata-rata 7.33 dan standar deviasi 3.55

b) Menerapkan normalisasi Z-score

Dengan parameter rata-rata dan standar deviasi sudah dihitung, setiap nilai pada fitur tingkat nyeri dan lama keluhan di data latih dan data uji ditransformasikan menggunakan rumus z-score. Berikut merupakan contoh perhitungan untuk data L01 dengan nilai A1 = 1 dan A2 = 3.

$$Z(A1) = \frac{X - \mu}{\sigma}$$

$$Z(A1) = \frac{1 - 3.27}{1.45} = -1.57$$

$$Z(A2) = \frac{X - \mu}{\sigma}$$

$$Z(A2) = \frac{3 - 7.33}{3.55} = -1.22$$

Perhitungan dilakukan ke seluruh titik data, sehingga menghasilkan data sebagai berikut :

Tabel 4. 3 Hasil normalisasi Z-score untuk Data Latih

ID Pasien	Gigi Ngilu (G1)	Bercak Putih (G2)	Tingkat Nyeri (Z-A1)	Lama Keluhan (Z-A2)	Penyakit (Kelas Aktual)
L01	1	1	-1.57	-1.22	Karies Awal
L02	1	0	-0.88	-1.50	Karies Awal
L03	0	1	-1.57	-0.66	Karies Awal
L04	1	1	-0.88	-1.78	Karies Awal
L05	0	1	-1.57	-0.94	Karies Awal
L06	1	1	-0.19	-0.09	Karies Sedang
L07	1	0	0.50	0.75	Karies Sedang
L08	1	1	0.50	-0.66	Karies Sedang
L09	0	1	-0.19	0.19	Karies Sedang
L10	1	0	1.20	-0.37	Karies Sedang
L11	1	1	1.20	1.88	Karies Lanjut
L12	1	0	0.50	1.32	Karies Lanjut
L13	1	1	1.20	0.47	Karies Lanjut
L14	0	1	0.50	1.03	Karies Lanjut
L15	1	0	1.20	1.60	Karies Lanjut

Tabel 4. 4 Hasil Normalisasi Z-Score untuk Data Uji

ID Pasien	Gigi Ngilu (G1)	Bercak Putih (G2)	Tingkat Nyeri (Z-A1)	Lama Keluhan (Z-A2)
U01	1	1	-0.88	-0.94
U02	1	0	0.50	0.19
U03	0	1	1.20	0.75

4.1.4. Proses Klasifikasi Menggunakan *K-Nearest Neighbor* (Studi Kasus)

Setelah seluruh data latih dan data uji berada dalam skala yang seragam, proses inti dari algoritma KNN dapat dijalankan. Bagian ini akan menyajikan studi kasus terperinci untuk mengklasifikasikan pasien pertama dari data uji, yaitu

pasien U01. Sebagai parameter utama algoritma, nilai $K=3$ akan digunakan. Pemilihan nilai K ganjil seperti 3 merupakan praktik umum untuk menghindari kemungkinan hasil voting yang imbang dalam masalah klasifikasi biner atau multi-kelas. Nilai ini dianggap dapat memberikan keseimbangan yang baik antara sensitivitas terhadap pola lokal dan ketahanan terhadap data penciran (*noise*).

a) Perhitungan Jarak *Euclidean*

Jarak Euclidean dihitung untuk mengukur "kedekatan" atau "kemiripan" antara data uji (U01) dengan setiap data dalam set pelatihan (L01 hingga L15).

Rumus yang digunakan adalah sebagai berikut :

$$D(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

Sebagai contoh, berikut adalah perhitungan jarak antara data uji U01 dengan data latih L01:

Data U01: {1, 1, -0.88, -0.94}

Data L01: {1, 1, -1.57, -1.22}

Perhitungan jarak Euclidean antara **U01** dan **L01**:

$$D(U01, L01) = \sqrt{(1 - 1)^2 + (-0.88 - (-1.57))^2 + (-0.94 - (-1.22))^2}$$

$$D(U01, L01) = \sqrt{0 + 0 + 0.4761 + 0.0784}$$

$$D(U01, L01) = \sqrt{0.5545}$$

$$D(x, y) = 0.745$$

b) Pengurutan Jarak dan Penentuan Tetangga Terdekat

Proses perhitungan jarak di atas diulangi untuk menghitung jarak dari U01 ke semua 15 data latih. Hasilnya kemudian diurutkan dari jarak terkecil hingga terbesar untuk mengidentifikasi tetangga terdekat.

Tabel 4. 5 Hasil Perhitungan Jarak Euclidean dan Pengurutan untuk Data Uji U01

ID Pasien Latih	Penyakit (Aktual)	Jarak Euclidean
L04	Karies Awal	0.841
L02	Karies Awal	0.854
L05	Karies Awal	1.054
L01	Karies Awal	1.159
L03	Karies Awal	1.258
L08	Karies Sedang	1.434
L06	Karies Sedang	1.488
L09	Karies Sedang	1.625
L10	Karies Sedang	2.145
L07	Karies Sedang	2.212
L13	Karies Lanjut	2.493
L14	Karies Lanjut	2.701
L12	Karies Lanjut	2.821
L15	Karies Lanjut	3.164
L11	Karies Lanjut	3.208

c) Klasifikasi berdasarkan mayoritas (voting)

Dengan nilai $K=3$, tiga data latih dengan jarak Euclidean terkecil dari U01 dipilih sebagai tetangga terdekat.

Tabel 4. 6 Penentuan Kelas Berdasarkan 3 Tetangga Terdekat untuk U01

ID Pasien Latih	Penyakit (Aktual)	Jarak Euclidean
L04	Karies Awal	0.841
L02	Karies Awal	0.854
L05	Karies Awal	1.054

4.1.5. Evaluasi Kinerja Model pada Seluruh Data Uji

Proses klasifikasi yang telah dirinci pada studi kasus U01 diulangi untuk seluruh data uji lainnya (U02 dan U03). Tabel berikut merangkum hasil prediksi model untuk keseluruhan set data uji, yang akan menjadi dasar untuk evaluasi kinerja secara kuantitatif.

Tabel 4. 7 Ringkasan Hasil Prediksi Model KNN untuk Seluruh Data Uji

ID Pasien Uji	Penyakit Aktual	Prediksi Model	Hasil
U01	Karies Awal	Karies Awal	Benar
U02	Karies Sedang	Karies Sedang	Benar
U03	Karies Lanjut	Karies Lanjut	Benar

Untuk memvisualisasikan performa model klasifikasi secara lebih mendalam, digunakan *Confusion Matrix*. Matriks ini menyajikan perbandingan antara kelas aktual (yang sebenarnya) dengan kelas yang diprediksi oleh model, sehingga dapat diketahui jenis kesalahan yang dibuat oleh model. Matriks ini terdiri dari empat komponen utama untuk setiap kelas:

Tabel 4. 8 Confusion Matrix Hasil Klasifikasi

	Prediksi: Karies Awal	Prediksi: Karies Sedang	Prediksi: Karies Lanjut
Aktual: Karies Awal	1 (TP)	0	0
Aktual: Karies Sedang	0	1 (TP)	0
Aktual: Karies Lanjut	0	0	1 (TP)

Dari matriks di atas, terlihat bahwa semua prediksi yang dibuat oleh model berada pada diagonal utama, yang menandakan bahwa setiap prediksi sesuai dengan kelas aktualnya. Tidak ada kesalahan klasifikasi yang terjadi pada data uji ini.

4.2. Penerapan dalam Sistem Berbasis Web

1. Tampilan Halaman *Login*

Halaman login ini memungkinkan pengguna untuk mengakses Sistem Pendukung Keputusan Karies Gigi dengan memasukkan username dan password.

Setelah login, pengguna dapat menganalisis gejala dan mendapatkan rekomendasi tingkat risiko untuk mendukung diagnosis karies gigi.

Selamat Datang
Login untuk mengakses Sistem Pendukung Keputusan Karies Gigi.

Username

Password

[Lupa Password?](#)

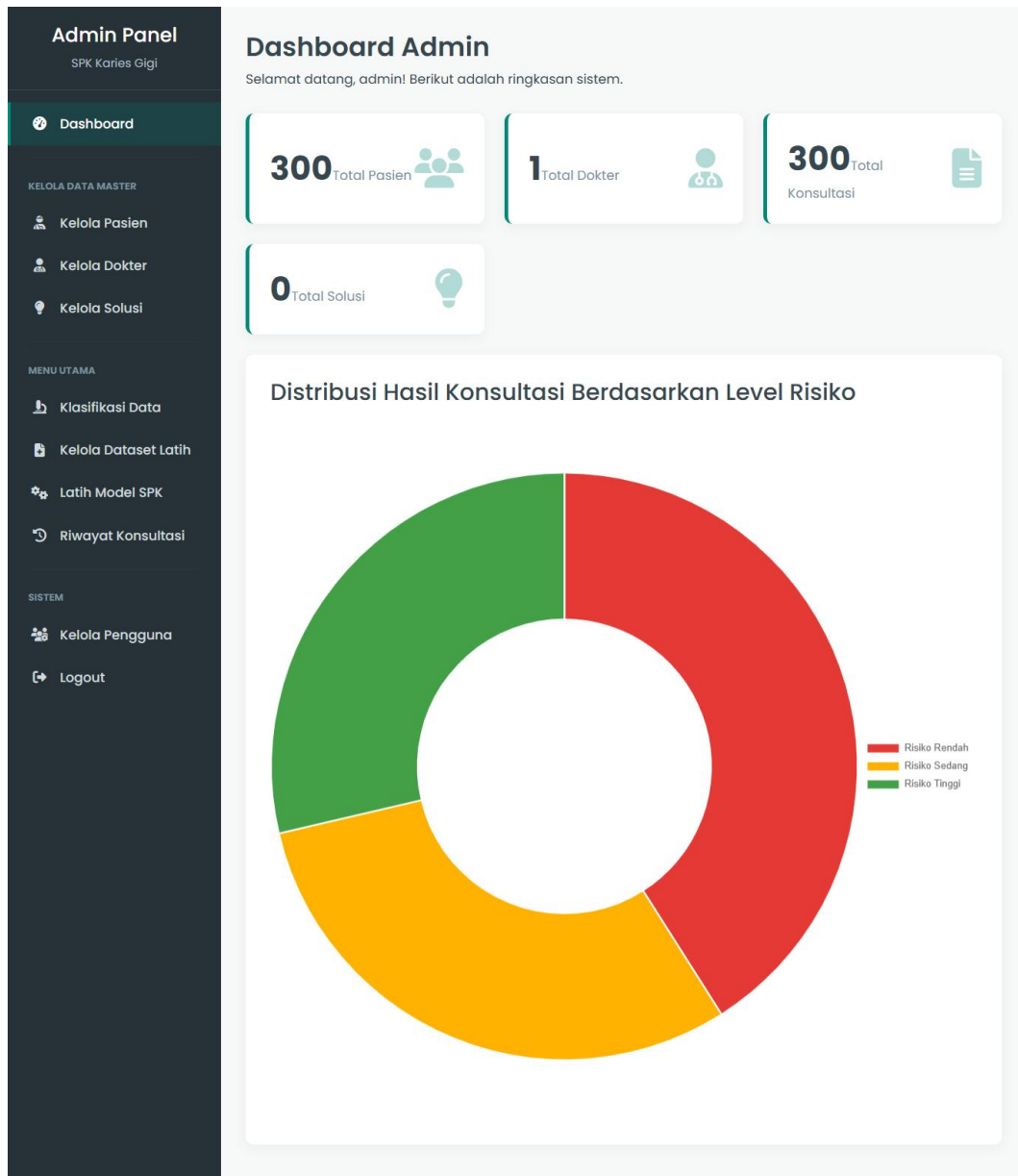
Login

Keputusan Klinis Berbasis Data
Analisis gejala dan dapatkan rekomendasi tingkat risiko karies gigi untuk mendukung diagnosis Anda.

Gambar 4. 1 Tampilan Halaman Login

2. Tampilan Halaman *Dashboard*

Halaman dashboard ini menampilkan informasi ringkas bagi admin, seperti jumlah pasien, dokter, dan konsultasi yang telah dilakukan. Selain itu, terdapat diagram distribusi hasil konsultasi berdasarkan level risiko, yang menunjukkan proporsi risiko rendah, sedang, dan tinggi.



Gambar 4. 2 Tampilan Halaman Dashboard

3. Tampilan Halaman Kelola Data Pasien

Halaman Kelola Data Pasien di dashboard admin memungkinkan admin untuk melihat, mengelola, dan menambahkan data pasien. Tabel ini menampilkan informasi pasien seperti ID, nama, umur, jenis kelamin, pendidikan, dan pekerjaan, dengan opsi untuk mengedit atau menghapus data pasien. Admin dapat menavigasi melalui data pasien yang tersimpan dan melakukan perubahan jika diperlukan.

Admin Panel

SPK Karies Gigi

Dashboard

KELOLA DATA MASTER

Kelola Pasien

Kelola Dokter

Kelola Solusi

MENU UTAMA

Klasifikasi Data

Kelola Dataset Latih

Latih Model SPK

Riwayat Konsultasi

SISTEM

Kelola Pengguna

Logout

Kelola Data Pasien

Daftar Pasien

+ Tambah Pasien

Show

10

entries

Search:

ID	Nama Pasien	Umur	Jenis Kelamin	Pendidikan	Pekerjaan	Aksi
3	Tgk. Unjani Uwais, M.Tl.	22	Perempuan	SMA	Lainnya	
4	Asmuni Zulaika	20	Laki-Laki	Perguruan Tinggi	Pelajar/Mahasiswa	
5	Prabawa Widodo	21	Laki-Laki	SMA	Pelajar/Mahasiswa	
6	Arsipatra Sirait	22	Perempuan	Perguruan Tinggi	Pelajar/Mahasiswa	
7	Tgk. Artanto Tarihoran, S.IP	22	Laki-Laki	Perguruan Tinggi	Lainnya	
8	Jamalia Utami	24	Laki-Laki	Perguruan Tinggi	Pegawai Swasta	
9	Johan Narpati	22	Laki-Laki	Perguruan Tinggi	Pelajar/Mahasiswa	
10	Puti Fitria Manullang, S.Kom	23	Perempuan	SMA	Pelajar/Mahasiswa	
11	Nilam Wahyudin	22	Laki-Laki	Perguruan Tinggi	Lainnya (IRT)	
12	Mulya Anggriawan	21	Laki-Laki	SMA	Pelajar/Mahasiswa	

Showing 1 to 10 of 300 entries

Previous

1

2

3

4

5

...

30

Next

Gambar 4. 3 Tampilan Kelola Data Pasien

4. Tampilan Halaman Tambah Data Pasien

Halaman Tambah Data Pasien memungkinkan admin untuk menambahkan pasien baru ke sistem. Terdapat dua opsi: pertama, unggah data pasien melalui file Excel menggunakan template yang disediakan, dan kedua, tambah pasien secara manual dengan mengisi informasi seperti nama, umur, jenis kelamin, pendidikan terakhir, dan pekerjaan. Admin dapat menyimpan data pasien setelah mengisi form yang diperlukan.

Admin Panel

SPK Karies Gigi

Dashboard

KELOLA DATA MASTER

Kelola Pasien

Kelola Dokter

Kelola Solusi

MENU UTAMA

Klasifikasi Data

Kelola Dataset Latih

Latih Model SPK

Riwayat Konsultasi

SISTEM

Kelola Pengguna

Logout

Tambah Data Pasien

Opsi 1: Tambah Pasien via Unggah Excel

Gunakan fitur ini untuk menambah banyak data pasien sekaligus. Sistem akan melewati pasien dengan nama yang sudah ada.

Unduh Template

Pilih File Excel

Pilih FileTidak ada file yang dipilih

Unggah

Opsi 2: Tambah Pasien secara Manual

Nama Pasien

Umur

Jenis Kelamin

-- Pilih --

Pendidikan Terakhir

Pekerjaan

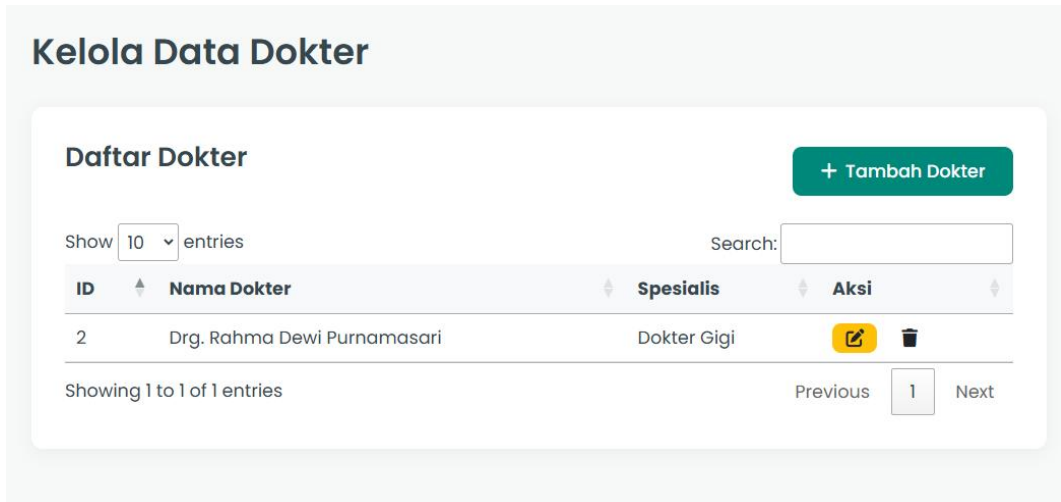
Simpan Pasien

Batal

Gambar 4. 4 Tampilan Tambah Data Pasien

5. Tampilan Kelola Data Dokter

Halaman Kelola Data Dokter menampilkan daftar dokter yang terdaftar dalam sistem, termasuk informasi ID, nama dokter, dan spesialisasi. Admin dapat mengedit atau menghapus data dokter sesuai kebutuhan dengan tombol aksi yang tersedia.



Gambar 4. 5 Tampilan Kelola Data Dokter

6. Tampilan Tambah Data Dokter

Halaman Tambah Data Dokter memungkinkan admin untuk menambahkan data dokter baru ke sistem. Admin dapat memilih untuk unggah data dokter menggunakan file Excel atau menambahkan dokter secara manual dengan mengisi informasi seperti nama dokter dan spesialisasi.

Tambah Data Dokter

Opsi 1: Tambah Dokter via Unggah Excel

Gunakan fitur ini untuk menambah banyak data dokter sekaligus. Sistem akan melewati dokter dengan nama yang sudah ada.

[Unduh Template](#)

Pilih File Excel

Tidak ada file yang dipilih

[Unggah](#)

Opsi 2: Tambah Dokter secara Manual

Nama Dokter

Spesialis

[Simpan Dokter](#) [Batal](#)

Gambar 4. 6 Tampilan Tambah Data Dokter

7. Tampilan Kelola Dataset Latih

Halaman Kelola Dataset Latih ini memungkinkan admin untuk memasukkan data kebiasaan dan gejala pasien yang akan digunakan untuk melatih model prediksi K-Nearest Neighbor (KNN). Admin dapat memilih pasien dan dokter, lalu mengisi kolom dengan informasi terkait kebiasaan pasien, seperti sikat gigi, konsumsi makanan, gejala yang dialami, serta riwayat penyakit sistemik. Setelah data diisi, admin dapat menentukan hasil risiko (target/label) untuk data tersebut dan menyimpannya dalam sistem untuk digunakan dalam pelatihan model.

The screenshot shows the 'Kelola Dataset Latih' form within the 'Admin Panel' of 'SPK Karies Gigi'. The form is divided into several sections:

- Pilih Pasien:** A dropdown menu labeled '-- Pilih Pasien --'.
- Pilih Dokter Pemeriksa:** A dropdown menu labeled '-- Pilih Dokter --'.
- Kuesioner Kebiasaan & Gejala:** A grid of input fields for various symptoms and habits:
 - Infeksi gigi: Ya
 - Sakit Saat Konsumsi: Ya
 - Gigi Berlubang: Ya
 - Perubahan Warna Gigi: Ya
 - Pembengkakan Gusi: Ya
 - Kesulitan Mengunyah: Ya
 - Sakit Kepala: Ya
 - Frekuensi Sakit Kepala: Sering
 - Konsumsi Obat: Ya
 - Gejala Lain: (text input)
 - Riwayat Penyakit Sistemik: (text input)
- Hasil Risiko (Target/Label):** A dropdown menu labeled '-- Tentukan Hasil Risiko untuk Data Latih Ini --'.
- Buttons:** 'Simpan Data Latih' (Save Training Data) and 'Batal' (Cancel).

Gambar 4. 7 Tampilan Kelola Dataset Latih

8. Tampilan Riwayat Konsultasi

Halaman Riwayat Konsultasi SPK menampilkan daftar riwayat konsultasi pasien yang telah dilakukan. Setiap entri mencakup ID pasien, nama pasien, tanggal konsultasi, dan hasil risiko yang diberikan (misalnya, Risiko Rendah, Sedang, atau Tinggi). Admin dapat unggah dataset baru, mengedit data konsultasi, atau menghapus entri yang ada.

Admin Panel

SPK Karies Gigi

Dashboard

KELOLA DATA MASTER

Kelola Pasien

Kelola Dokter

Kelola Solusi

MENU UTAMA

Klasifikasi Data

Kelola Dataset Latih

Latih Model SPK

Riwayat Konsultasi

SISTEM

Kelola Pengguna

Logout

Riwayat Konsultasi SPK

Daftar Riwayat

Unggah Dataset

Hapus Semua Riwayat

Show 10 entries

Search:

ID	Pasien	Tanggal	Hasil Risiko	Aksi
55	Tgk. Unjani Uwais, M.Ti.	23 May 2025 02:17	Risiko Rendah	
56	Asmuni Zulaika	29 May 2025 20:51	Risiko Rendah	
57	Prabawa Widodo	14 Jun 2025 22:57	Risiko Sedang	
58	Arsipatra Sirait	01 Jul 2025 20:49	Risiko Rendah	
59	Tgk. Artanto Tarihoran, S.IP	05 Aug 2025 23:44	Risiko Rendah	
60	Jamalia Utami	29 May 2025 01:18	Risiko Rendah	
61	Johan Narpati	05 Jun 2025 00:43	Risiko Sedang	
62	Puti Fitria Manullang, S.Kom	10 Jun 2025 02:13	Risiko Sedang	
63	Nilam Wahyudin	23 Apr 2025 03:37	Risiko Tinggi	
64	Mulya Anggriawan	12 May 2025 02:07	Risiko Sedang	

Showing 1 to 10 of 298 entries

Previous

1

2

3

4

5

...

30

Next

Gambar 4. 8 Tampilan Riwayat Konsultasi

9. Tampilan Latih Model SPK

Halaman ini memungkinkan admin untuk melatih model Sistem Pendukung Keputusan (SPK) menggunakan data yang telah dimasukkan dalam Kelola Dataset Latih. Admin dapat mengklik tombol "Latih Model SPK" setelah dataset latih diunggah dan data pasien serta gejala telah dimasukkan.

Latih & Laporkan Model KNN Manual SPK

Alur kerja profesional untuk melatih, mengevaluasi, dan menyimpan model prediksi.

Mulai Pelatihan Baru

Tentukan Nilai K Manual (Opsional)

Kosongkan untuk analisis otomatis oleh sistem (rekomendasi)

Gunakan angka ganjil (misal: 3, 5, 7) untuk hasil terbaik.

Latih Model Sekarang

Gambar 4. 9 Tampilan Awal Kelola Model KNN

Ketika tombol ini ditekan, *frontend* akan mengirimkan permintaan ke *backend* untuk memulai pelatihan model menggunakan algoritma yang dipilih, dalam hal ini, K-Nearest Neighbor (KNN). Setelah tombol ditekan, tampilkan status pelatihan kepada admin. Status ini bisa berupa "Proses Latih Dimulai", "Pelatihan Sedang Berlangsung", dan "Pelatihan Selesai" dengan hasil evaluasi (misalnya akurasi, precision, recall, F1-score).

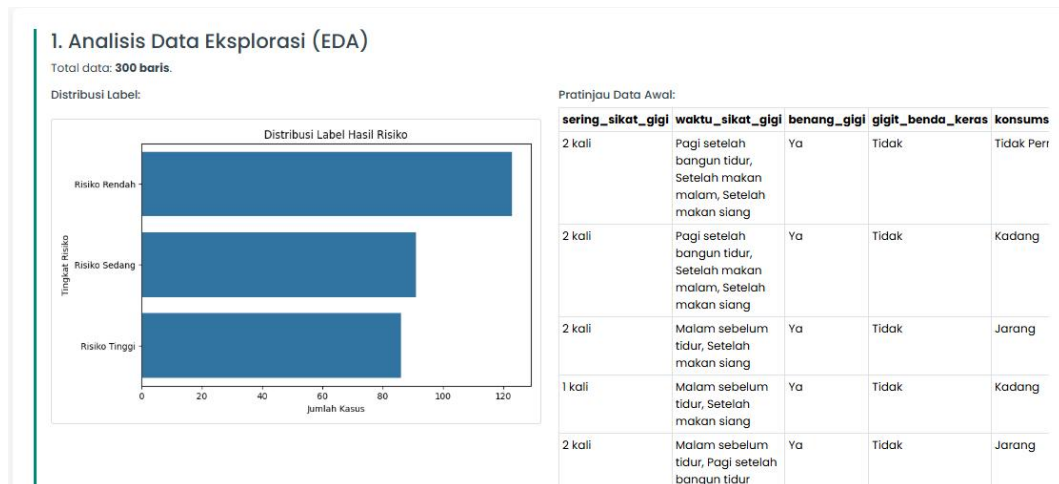
Backend akan melatih model menggunakan algoritma *K-Nearest Neighbor* (KNN) yang sudah diterapkan di backend. Ini melibatkan:

- a. Menentukan nilai k (jumlah tetangga terdekat).
- b. Menghitung jarak antar data menggunakan Euclidean Distance.
- c. Pelatihan Model: Setelah memilih parameter yang sesuai, backend akan melatih model dengan menggunakan dataset latih yang sudah dipersiapkan.

Berikut merupakan hasil dari pelatihan analisis, dapat dilihat pada penjabaran dibawah ini :

- a. Analisis Data Eksplorasi (EDA)

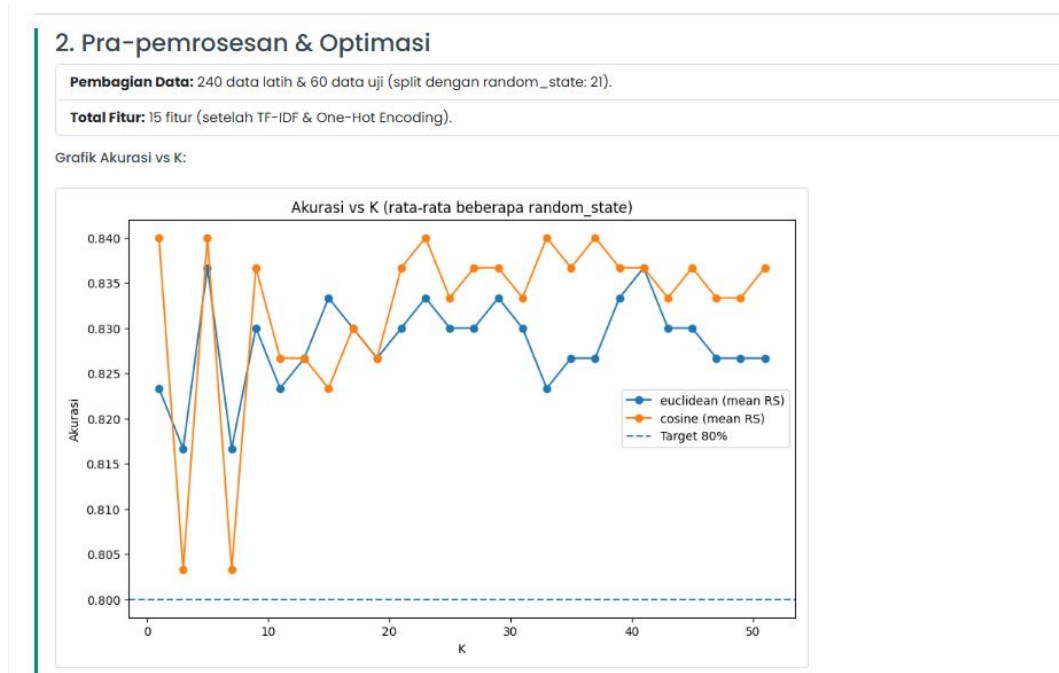
Pada bagian ini, dilakukan analisis distribusi label yang menunjukkan jumlah data untuk setiap kategori risiko (Rendah, Sedang, Tinggi). Visualisasi membantu untuk melihat keseimbangan distribusi antara label yang berbeda, yang memberikan gambaran awal apakah model terlatih akan menghadapi data yang tidak seimbang.



Gambar 4. 10 Bagian EDA dalam Sistem

b. Pra-pemrosesan dan Optimasi

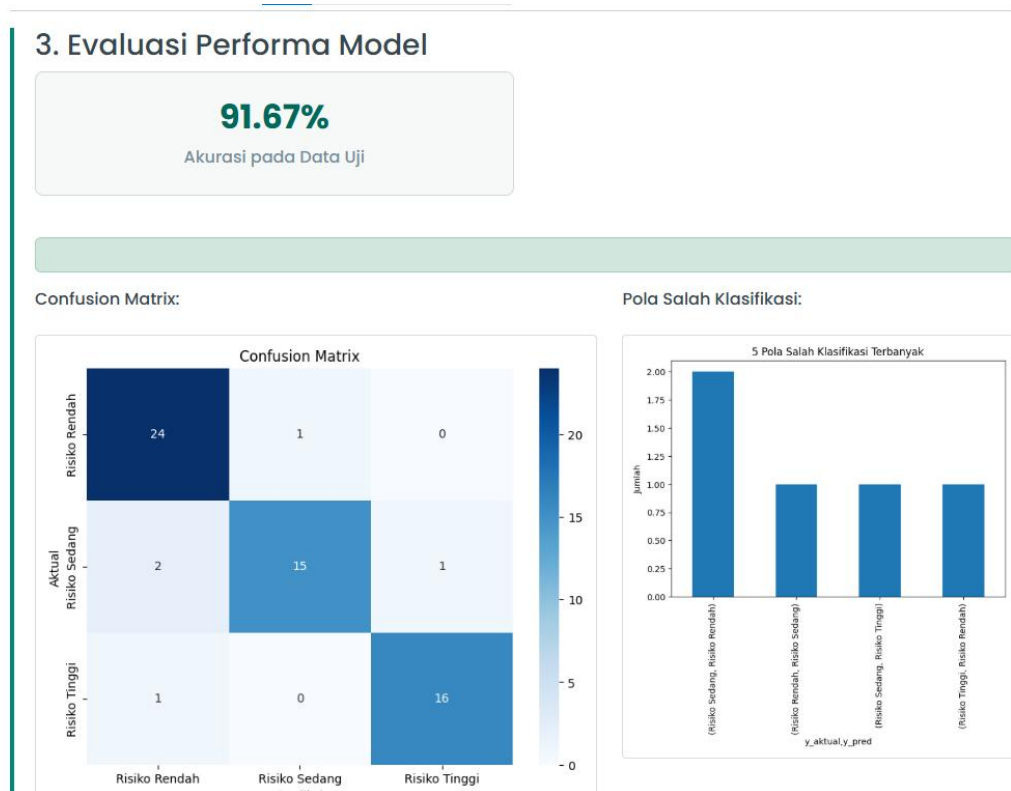
Data dibagi menjadi data latih dan data uji menggunakan teknik split random. Selain itu, dilakukan optimasi hyperparameter, seperti pemilihan nilai k (jumlah tetangga terdekat) yang optimal. Normalisasi data dengan menggunakan Z-Score juga diterapkan untuk memastikan bahwa setiap fitur memiliki skala yang sama.



Gambar 4. 11 Bagian Pra-Pemrosesan & Optimasi

c. Evaluasi Performa Model

Evaluasi menunjukkan bahwa model KNN yang dilatih memiliki akurasi 91.67% pada data uji. Confusion matrix dan classification report menunjukkan kinerja model dalam mengklasifikasikan data ke dalam kategori risiko yang sesuai, dengan hasil precision, recall, dan F1-score yang menunjukkan keseimbangan yang baik antara berbagai metrik.



Gambar 4. 12 Gambar Evaluasi dan Performa Model

Precision untuk setiap kategori menunjukkan tingkat ketepatan model dalam memprediksi kelas yang benar. Untuk Risiko Rendah, *precision* mencapai 88.89%, yang berarti dari semua data yang diprediksi sebagai risiko rendah, 88.89% benar-benar memiliki risiko rendah. Risiko Sedang dan Risiko Tinggi menunjukkan hasil yang lebih tinggi, yaitu 93.75% dan 94.12% masing-masing,

menunjukkan kemampuan model untuk memprediksi dengan akurat untuk kedua kategori ini.

Recall menunjukkan seberapa baik model dapat menemukan seluruh kasus yang sebenarnya ada dalam dataset. Hasil recall yang tinggi, seperti pada Risiko Rendah (96.00%) dan Risiko Tinggi (94.12%), menunjukkan bahwa model mampu mendeteksi sebagian besar kasus positif dalam kategori tersebut. Namun, recall untuk Risiko Sedang sedikit lebih rendah di 83.33%, yang menunjukkan bahwa model mungkin mengalami kesulitan dalam mendeteksi sebagian data yang benar-benar memiliki risiko sedang, meskipun tetap dalam kisaran yang dapat diterima.

F1-Score, yang menggabungkan precision dan recall, memberikan gambaran yang lebih lengkap tentang keseimbangan antara keduanya. F1-score yang tinggi pada semua kategori risiko (92.31% untuk Risiko Rendah, 88.24% untuk Risiko Sedang, dan 94.12% untuk Risiko Tinggi) menunjukkan bahwa model tidak hanya memiliki akurasi yang baik, tetapi juga seimbang dalam mengidentifikasi semua kelas. Rata-rata macro dan weighted averages juga menunjukkan kinerja yang solid, dengan model secara keseluruhan memberikan hasil yang konsisten dan dapat diandalkan dalam klasifikasi data karies gigi berdasarkan risiko.

Tabel 4. 9 Tabel Confusion Matrix

Kelas	Precision	Recall	F1-Score	Support
Risiko Rendah	88.89%	96.00%	92.31%	25
Risiko Sedang	93.75%	83.33%	88.24%	18
Risiko Tinggi	94.12%	94.12%	94.12%	17
macro avg	92.25%	91.15%	91.55%	60
weighted avg	91.83%	91.67%	91.60%	60

d. Detail Prediksi dan Data Uji

Tabel ini menampilkan prediksi model untuk data uji, membandingkan label risiko yang sebenarnya (y_{actual}) dengan label yang diprediksi (y_{pred}). Hasil ini memberikan gambaran tentang seberapa baik model dapat mengklasifikasikan data, termasuk kasus False Positive atau False Negative, yang menunjukkan area perbaikan potensial.

Tabel 4. 10 Tabel Hasil Klasifikasi

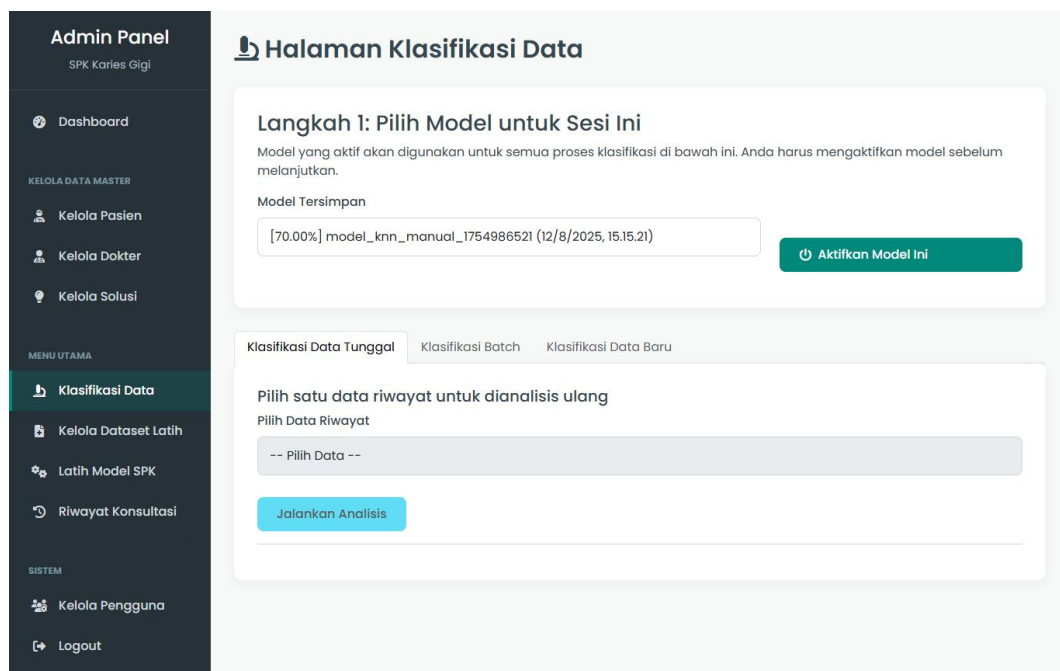
y_ aktual	y_pred	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Sedang	Risiko Sedang	cocok
Risiko Sedang	Risiko Sedang	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Sedang	Risiko Tinggi	tidak cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Sedang	Risiko Sedang	cocok
Risiko Tinggi	Risiko Tinggi	cocok
Risiko Sedang	Risiko Sedang	cocok
Risiko Rendah	Risiko Rendah	cocok
Risiko Rendah	Risiko Sedang	tidak cocok
Risiko Rendah	Risiko Rendah	cocok

10. Tampilan Klasifikasi Data

Halaman Klasifikasi Data ini memungkinkan admin untuk memilih dan mengaktifkan model yang akan digunakan untuk klasifikasi data. Langkah

pertama adalah memilih model yang sudah tersimpan, seperti model_knn_manual yang ditampilkan pada bagian Model Tersimpan. Setelah model dipilih, admin dapat mengklik "Aktifkan Model Ini" untuk mengaktifkan model tersebut, yang akan digunakan dalam proses klasifikasi berikutnya.

Di bawahnya, terdapat opsi untuk memilih jenis klasifikasi data yang ingin dilakukan, seperti Klasifikasi Data Tunggal, Klasifikasi Batch, atau Klasifikasi Data Baru. Admin dapat memilih satu data riwayat untuk dianalisis ulang, lalu menekan "Jalankan Analisis" untuk memulai proses klasifikasi. Halaman ini mempermudah admin dalam menjalankan proses klasifikasi dengan memilih data yang sesuai dan memanfaatkan model yang sudah dilatih sebelumnya.



Gambar 4. 13 Tampilan Klasifikasi Data

4.3. Kelebihan dan Kelemahan Program

4.3.1. Kelebihan Program

Berdasarkan program yang telah dibangun, berikut merupakan kelebihan dari program ini :

1. Program ini dirancang dengan antarmuka yang sederhana dan mudah dipahami, memungkinkan admin dan pengguna lain untuk dengan mudah mengelola data, melatih model, dan melakukan klasifikasi tanpa kesulitan teknis. Hal ini mempermudah penerapan sistem di berbagai lapisan pengguna, seperti dokter dan administrator.
2. Dengan menggunakan algoritma K-Nearest Neighbor (KNN), program ini mampu memberikan klasifikasi risiko dengan akurasi mencapai 91.67% pada data uji. Ini menunjukkan bahwa model yang digunakan dapat memberikan hasil yang cukup akurat untuk mendeteksi risiko karies gigi berdasarkan data pasien.
3. Program ini memungkinkan admin untuk melakukan klasifikasi data tunggal, klasifikasi batch, dan klasifikasi data baru. Fitur ini memberikan fleksibilitas dalam mengelola dan memproses data dalam berbagai format, memungkinkan penggunaan yang lebih efisien dalam berbagai skenario.

4.3.2. Kelemahan Program

Adapun kelemahan dari program yang sudah dibangun adalah sebagai berikut :

1. Meskipun model ini memiliki akurasi yang baik, kinerja model sangat tergantung pada kualitas data latih yang dimasukkan. Jika data latih tidak cukup representatif atau mengandung kesalahan, hasil klasifikasi bisa menjadi kurang akurat, yang dapat mempengaruhi prediksi risiko karies gigi.
2. Walaupun antarmuka mudah digunakan, pengguna dengan latar belakang non-teknis mungkin masih kesulitan memahami metrik evaluasi model seperti

precision dan *recall*. Hal ini dapat membatasi pemahaman mereka tentang seberapa baik model bekerja atau perlu perbaikan lebih lanjut.

3. Meskipun model telah menunjukkan performa yang baik, akurasi dan *precision* yang tinggi pada beberapa kategori menunjukkan bahwa masih ada ruang untuk perbaikan, terutama dalam menangani kelas yang lebih sedikit representasinya, seperti Risiko Sedang. Pembaruan model dengan data lebih banyak dan lebih bervariasi dapat membantu meningkatkan hasil prediksi lebih lanjut.

4.4. Pengujian Program

Pengujian Program dilakukan menggunakan metode *Black Box Testing* untuk memastikan bahwa semua fitur pada sistem berjalan sesuai dengan fungsinya tanpa melihat implementasi internal. Pengujian ini mencakup berbagai skenario, mulai dari login, pengelolaan data pasien dan dokter, hingga klasifikasi data dan pelatihan model. Setiap fitur diuji berdasarkan input yang diberikan dan output yang diharapkan, untuk memastikan bahwa sistem dapat berfungsi dengan baik dan memberikan hasil yang sesuai dengan ekspektasi pengguna.

Tabel 4. 11 Tabel Pengujian Program

No.	Skenario Pengujian	Input	Hasil yang Diharapkan	Status
1	Pengujian Login	Masukkan username dan password yang valid	Pengguna diarahkan ke halaman dashboard	Lulus
2	Pengujian Login (Invalid Credentials)	Masukkan username atau password yang salah	Sistem menampilkan pesan "Username atau Password salah"	Lulus
3	Pengujian Pengelolaan	Masukkan data pasien baru	Data pasien baru berhasil disimpan dan ditampilkan	Lulus

	Data Pasien	(nama, umur, jenis kelamin)		
4	Pengujian Pengelolaan Data Pasien (Edit)	Ubah data pasien yang sudah ada	Data pasien yang diubah berhasil disimpan dan diperbarui	Lulus
5	Pengujian Pengelolaan Data Pasien (Hapus)	Hapus data pasien	Data pasien yang dihapus tidak muncul lagi pada daftar	Lulus
6	Pengujian Pengelolaan Data Dokter	Masukkan data dokter baru (nama, spesialisasi)	Data dokter baru berhasil ditambahkan ke sistem	Lulus
7	Pengujian Pengelolaan Data Dokter (Edit)	Ubah data dokter yang sudah ada	Data dokter yang diubah berhasil disimpan dan diperbarui	Lulus
8	Pengujian Pengelolaan Data Dokter (Hapus)	Hapus data dokter	Data dokter yang dihapus tidak muncul lagi pada daftar	Lulus
9	Pengujian Latih Model	Pilih dataset latih dan klik "Latih Model SPK"	Model berhasil dilatih dan hasil evaluasi ditampilkan	Lulus
10	Pengujian Klasifikasi Data	Pilih data riwayat dan klik "Jalankan Analisis"	Data berhasil diklasifikasikan sesuai model yang aktif	Lulus
11	Pengujian Hasil Klasifikasi	Pilih hasil klasifikasi dan tampilkan laporan	Laporan hasil klasifikasi dengan metrik evaluasi (precision, recall, F1-score) ditampilkan	Lulus

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis, perancangan, dan implementasi sistem yang telah dilakukan, maka dapat ditarik beberapa kesimpulan yang menjawab rumusan masalah sebagai berikut:

1. Pengembangan sistem pakar untuk deteksi dini karies gigi telah berhasil direalisasikan menggunakan algoritma K-Nearest Neighbor (KNN) dalam platform berbasis web. Proses pengembangan dilakukan secara sistematis, dimulai dari pengumpulan data gejala dan atribut dari Klinik Gigi Rahma Dewi, perancangan arsitektur sistem menggunakan pemodelan UML, hingga implementasi antarmuka pengguna yang fungsional. Sistem ini mampu mengolah input data pasien berupa gejala dan atribut terkait, kemudian memprosesnya untuk menghasilkan klasifikasi tingkat risiko karies (Karies Awal, Sedang, atau Lanjut) sebagai output.
2. *K-Nearest Neighbor* (KNN) dalam mendiagnosis karies gigi pada sistem yang dikembangkan menunjukkan performa yang sangat baik. Berdasarkan hasil pengujian pada data uji, model klasifikasi mampu mencapai tingkat akurasi 91.67%. Hal ini didukung oleh penerapan pra-pemrosesan data yang cermat, termasuk normalisasi *Z-Score* untuk fitur numerik dan pemilihan nilai K yang optimal. Hasil evaluasi menggunakan *Confusion Matrix* menunjukkan bahwa model memiliki kemampuan yang solid dalam membedakan kelas risiko rendah, sedang, dan tinggi, yang divalidasi melalui metrik *precision*, *recall*, dan *F1-score*.

5.2 Saran

Meskipun sistem yang dibangun telah berhasil mencapai tujuan penelitian, terdapat beberapa aspek yang dapat dikembangkan lebih lanjut untuk meningkatkan fungsionalitas dan keandalannya di masa depan. Adapun saran yang relevan adalah sebagai berikut:

1. Untuk meningkatkan akurasi dan cakupan diagnostik, disarankan untuk menambah jumlah dan variasi data latih. Data yang digunakan saat ini terbatas pada rekam medis dari Klinik Gigi Rahma Dewi. Penambahan data dari berbagai sumber lain dapat membantu model mengenali pola yang lebih beragam dan meningkatkan kemampuannya dalam melakukan generalisasi pada kasus-kasus baru.
2. Sistem saat ini hanya berfokus pada deteksi karies gigi. Pengembangan selanjutnya dapat diarahkan untuk memperluas cakupan diagnosis dengan menyertakan penyakit mulut dan gigi lainnya. Hal ini akan menjadikan sistem sebagai alat skrining kesehatan gigi yang lebih komprehensif dan bermanfaat.
3. Mengingat sistem yang dikembangkan masih berupa purwarupa (*prototype*), disarankan untuk melakukan validasi klinis secara langsung dengan melibatkan lebih banyak dokter gigi sebagai pakar. Umpan balik dari para ahli dapat digunakan untuk menyempurnakan basis pengetahuan, mengevaluasi output sistem, dan memastikan bahwa rekomendasi yang diberikan relevan dengan praktik klinis di lapangan sebelum sistem diimplementasikan secara penuh.

DAFTAR PUSTAKA

- Anisa, C. (2020). PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK PREDIKSI PENJUALAN OBAT PADA APOTEK KIMIA FARMA ATMO PALEMBANG. *Bina Darma Conference on Computer Science*.
- Binus Student Corner. (2022). *Teknik pre-processing dan classification dalam data science – Master of Industrial Engineering*. Binus University. <https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science/>
- Bisa AI Academy. (2023). *Bisa AI Academy - Course: Machine Learning Dengan Scikit Learn Python*. <https://bisa.ai/course/detail/MzU3/1>
- Cahya Mestika, J., Oktavio Selan, M., & Iqbal Qadafi, M. (2022). Menjelajahi Teknik-Teknik Supervised Learning untuk Pemodelan Prediktif Menggunakan Python. *Buletin Ilmiah*. <https://jurnalmahasiswa.com/index.php/biikma>
- Diastuti, D., Masyhudi, & Duma Krispinus. (2024). *Hubungan Perilaku Kesehatan terhadap Karies Gigi Kriteria ICDAS di Desa*.
- Emalia Saqila, S., Putri Ferina, I., & Iskandar, A. (2023). Analisis Perbandingan Kinerja Clustering Data Mining Untuk Normalisasi Dataset. *Jurnal Sistem Komputer Dan Informatika (JSON)* Hal: 356–, 365(2). <https://doi.org/10.30865/json.v5i2.6919>
- Firahmi, R., Panggabean, G., & Pranata, A. (2020). Sistem Pakar Mendeteksi Kerusakan Quick Truck Menggunakan Metode Dempster Shafer. *Jurnal CyberTech Journal*, 3(1), 106–117. <https://ojs.trigunadharma.ac.id/>
- Firdaus, T., & Yanti, F. (2022a). IMPLEMENTASI METODE NAÏVE BAYES PADA SISTEM PAKAR DIAGNOSA PENYAKIT GIGI BERBASIS WEB. *Dinamika Informatika*, 14(2).
- Firdaus, T., & Yanti, F. (2022b). IMPLEMENTASI METODE NAÏVE BAYES PADA SISTEM PAKAR DIAGNOSA PENYAKIT GIGI BERBASIS WEB. *Dinamika Informatika*, 14(2).
- Hidayat, R., Satriansyah, A., & Nurhayati, M. S. (2022). Penggunaan Metode Waterfall untuk Rancangan Bangun Aplikasi Penyewaan Lapangan Olahraga. *BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 3(1), 9–16. <https://doi.org/10.37148/bios.v3i1.35>
- H.M., A. (2019). *Penerapan Metode C45 Dalam Memprediksi Pola Pembelian Bahan Campuran Olahan Karet (Studi Kasus : PT.Anugrah Sibolga Lestari)*. <http://repository.potensi-utama.ac.id/jspui/jspui/handle/123456789/3504>
- Mahdalena Simanjorang, R., Simangunsong, A., Arifin, M., Yamin, M., Informatika, T., & Pelita Nusantara, S. (2024). Penerapan Sistem Pakar Dalam Diagnosis Dini Penyakit Jantung Dengan Metode Sistem Inferensi Fuzzy. *Jurnal Ilmu Komputer*

- Dan Sistem Informasi (JIKOMSI V, 7(1), 131–142.*
<https://ejournal.sisfokomtek.org/index.php/jikom>
- Mulyati, S., Maulana Husein, S., & Kunci, K. (2020). *RANCANG BANGUN APLIKASI DATA MINING PREDIKSI KELULUSAN UJIAN NASIONAL MENGGUNAKAN ALGORITMA (KNN) K-NEAREST NEIGHBOR DENGAN METODE EUCLIDEAN DISTANCE PADA SMPN 2 PAGEDANGAN sistem dapat memprediksi dan mengklasifikasikan dengan baik dan cepat.* 65–73.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).
- Nur Rahma, D., & Mulyanti, S. (2021). *GAMBARAN ANGKA KEJADIAN KARIES PADA ANAK SEKOLAH DASAR DI KOTA MANADO BALI PALEMBANG.*
<http://dx.doi.org/.....>
- Pratama, A. (2020). *HTML Uncover - Panduan Belajar HTML Untuk Pemula.*
www.duniailkom.com
- Rozi, F., Bagoes, M., & Junianto, S. (2023). Penerapan Machine Learning Untuk Prediksi Harga Saham PT.Telekomunikasi Indonesia Tbk Menggunakan Algoritma K-Nearest Neighbors. *Jurnal Informatika MULTI*, 1(1).
- Rumbaugh, James., Jacobson, Ivar., & Booch, Grady. (2021). *The unified modeling language reference manual.* Addison-Wesley.
- Setiawan, R. (2021). *Memahami Class Diagram Lebih Baik - Dicoding Blog.*
<https://www.dicoding.com/blog/memahami-class-diagram-lebih-baik/>
- Setiawan, S. (2020). *Membicarakan Precision, Recall, dan F1-Score.* Medium.Com.
<https://stevkarta.medium.com/membicarakan-precision-recall-dan-f1-score-e96d81910354>
- Sulistyo, D. B., Saifulloh, S., & Nita, S. (2024). Implementasi Metode Certainty Factor dalam Sistem Pakar Diagnosa Kecanduan Media Sosial. *Digital Transformation Technology*, 4(1), 480–489. <https://doi.org/10.47709/digitech.v4i1.4356>
- Supriyanto, J., Alita, D., & Isnain, A. R. (2023). Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 4(1), 74–80.
<https://doi.org/10.33365/jatika.v4i1.2468>
- Whendasmoro, R. G., & Joseph, J. (2022). Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 872. <https://doi.org/10.30865/jurikom.v9i4.4526>
- Wicaksono, M. S., Arsanti, M., & Hartanto, H. A. (2024). IMPLEMENTATION OF TELEDENTISTRY IN PREVENTING DENTAL AND ORAL DISEASES IN INDONESIA. *Jurnal Teras Kesehatan*, 7(2), 1–10.
<https://doi.org/10.38215/jtkes.v7i2.129>

