

Penerapan Hybrid Naïve Bayes dan Decision Tree dalam Sistem Pakar Diagnosis Penyakit Mulut Berbasis Android

Muhammad Iqbal Panjaitan*, Berto Nadeak

STMIK Mulia Darma, Rantauperapat, Indonesia

Email: ^{1,*}Iqbalpj87@gmail.com, ²Nadeak.Baru@Gmail.Com

Email Penulis Korespondensi: Iqbalpj87@gmail.com

Submitted: 15/08/2025; Accepted: 09/09/2025; Published: 30/09/2025

Abstrak– Penyakit mulut seperti stomatitis, gingivitis, dan kandidiasis sering kali tidak terdeteksi sejak dini akibat keterbatasan pengetahuan masyarakat dan kurangnya akses layanan kesehatan, sehingga menimbulkan keterlambatan penanganan. Penelitian ini bertujuan mengembangkan sistem pakar diagnosis penyakit mulut berbasis Android dengan pendekatan *hybrid* antara algoritma Naïve Bayes dan Decision Tree. Metode Naïve Bayes digunakan untuk menghitung probabilitas gejala terhadap kemungkinan penyakit, sedangkan Decision Tree menghasilkan aturan diagnosis yang lebih transparan dan mudah dipahami. Tahapan penelitian meliputi studi literatur, pengumpulan dan validasi data gejala penyakit dari pakar medis, pengembangan model *hybrid*, implementasi dalam aplikasi Android, serta pengujian sistem menggunakan *cross-validation* dan *confusion matrix*. Hasil yang diharapkan meliputi prototipe aplikasi diagnosis penyakit mulut berbasis Android dengan tingkat akurasi minimal 85%, artikel ilmiah yang dipublikasikan pada jurnal nasional terindeks Sinta, serta luaran tambahan berupa pendaftaran hak cipta aplikasi dan penerbitan buku. Penelitian ini diharapkan dapat meningkatkan akses masyarakat terhadap diagnosis awal penyakit mulut, mendukung deteksi dini, serta memberikan kontribusi dalam pengembangan sistem kesehatan digital berbasis kecerdasan buatan di Indonesia.

Kata Kunci: Hybrid Naïve Bayes; Decision Tree; Sistem Pakar; Penyakit Mulut; Android

Abstract– Oral diseases such as stomatitis, gingivitis, and candidiasis are often undetected at an early stage due to limited public knowledge and restricted access to healthcare services, leading to delays in treatment. This study aims to develop an Android-based expert system for diagnosing oral diseases using a hybrid approach that combines the Naïve Bayes and Decision Tree algorithms. The Naïve Bayes method is applied to calculate the probability of symptoms associated with potential diseases, while the Decision Tree generates diagnostic rules that are more transparent and interpretable. The research stages include a literature review, data collection and validation of disease symptoms from medical experts, development of the hybrid model, implementation into an Android application, and system testing using cross-validation and a confusion matrix. The expected outcomes include a prototype Android application for oral disease diagnosis with a minimum accuracy of 85%, a scientific article published in a nationally accredited journal indexed in Sinta, and additional outputs such as copyright registration of the application and publication of a book. This study is expected to improve public access to early diagnosis of oral diseases, support early detection, and contribute to the advancement of digital health systems based on artificial intelligence in Indonesia.

Keywords: Hybrid Naïve Bayes; Decision Tree; Expert System; Oral Diseases; Android

1. PENDAHULUAN

Penyakit mulut merupakan salah satu masalah kesehatan yang umum terjadi dan seringkali tidak mendapatkan penanganan secara tepat waktu. Penyakit seperti stomatitis, gingivitis, periodontitis, dan kandidiasis dapat menimbulkan rasa sakit, ketidaknyamanan, bahkan berpengaruh terhadap kondisi sistemik apabila dibiarkan [1]. Di Indonesia, prevalensi penyakit mulut tergolong tinggi, namun kesadaran masyarakat untuk memeriksakan kondisi tersebut ke layanan kesehatan masih rendah. Hal ini dipengaruhi oleh kurangnya edukasi, akses terbatas terhadap dokter gigi, serta biaya pemeriksaan yang relatif mahal di beberapa wilayah [2], [3]. Seiring berkembangnya teknologi, kebutuhan akan sistem pendukung keputusan medis berbasis digital semakin meningkat. Sistem pakar merupakan salah satu cabang kecerdasan buatan (AI) yang mampu meniru cara berpikir seorang pakar dalam mengambil keputusan, termasuk dalam proses diagnosis penyakit [4]. Dalam konteks penyakit mulut, sistem pakar dapat membantu masyarakat mengenali gejala awal dan memberikan estimasi diagnosis yang mendekati akurasi profesional secara cepat dan mudah diakses [5].

Berbagai pendekatan telah dikembangkan dalam sistem pakar medis, termasuk yang berbasis algoritma klasifikasi seperti Naïve Bayes dan Decision Tree. Metode Naïve Bayes dikenal karena kemampuannya dalam menangani data gejala secara probabilistik dan efisien, terutama pada *dataset* yang besar dan tidak sepenuhnya lengkap [6]. Di sisi lain, Decision Tree memberikan hasil diagnosis dalam bentuk pohon keputusan yang intuitif dan mudah dipahami oleh pengguna, serta dapat menangani data dengan variabel kompleks [7]. Namun, masing-masing metode memiliki keterbatasan. Naïve Bayes cenderung mengasumsikan independensi antar fitur yang tidak selalu sesuai dengan kenyataan, sementara Decision Tree rawan terhadap overfitting pada *dataset* kecil [8]. Untuk mengatasi kekurangan masing-masing metode, pendekatan *hybrid* antara Naïve Bayes dan Decision Tree mulai banyak dikaji dalam literatur penelitian karena dinilai dapat meningkatkan akurasi dan kestabilan sistem klasifikasi [9]. *Hybridisasi* ini menggabungkan kekuatan inferensi probabilistik dan transparansi struktur aturan, menjadikannya sangat potensial untuk aplikasi diagnosis penyakit [10]. Beberapa penelitian membuktikan bahwa model *hybrid* ini mampu memberikan hasil yang lebih konsisten dan akurat dibandingkan penggunaan metode tunggal [11]. Lebih lanjut, perkembangan sistem operasi mobile, khususnya Android, membuka peluang untuk mengimplementasikan sistem pakar ke dalam aplikasi yang dapat diakses langsung oleh masyarakat. Android

digunakan oleh lebih dari 70% pengguna smartphone di dunia, termasuk di Indonesia, menjadikannya platform yang efektif untuk penyebaran aplikasi berbasis kesehatan [12]. Aplikasi sistem pakar berbasis Android memungkinkan diagnosis awal dapat dilakukan secara mandiri tanpa harus datang langsung ke fasilitas kesehatan, terutama di daerah terpencil [13]. Dalam konteks ini, pengembangan sistem pakar diagnosis penyakit mulut berbasis Android dengan metode hybrid Naïve Bayes dan Decision Tree menjadi solusi yang strategis dan aplikatif. Sistem ini diharapkan dapat meningkatkan akses masyarakat terhadap informasi kesehatan mulut, mendorong deteksi dini penyakit, serta membantu pengambilan keputusan awal sebelum berkonsultasi ke tenaga medis profesional [14].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dan rekayasa perangkat lunak berbasis kecerdasan buatan, dengan strategi kombinasi (hybrid) dua metode klasifikasi, yaitu Naïve Bayes dan Decision Tree. Tujuannya adalah menghasilkan sistem pakar berbasis Android yang mampu mendiagnosis penyakit mulut dengan akurat dan efisien.



Gambar 1. Tahapan Penelitian

Keterangan :

1. Studi Literatur dan Perumusan Masalah
Pada tahap ini dilakukan kajian pustaka untuk memahami penelitian-penelitian terdahulu, mengidentifikasi permasalahan, serta merumuskan tujuan penelitian.
2. Pengumpulan dan Pengolahan Data
Data diperoleh dari pakar medis dan literatur, berupa daftar penyakit mulut, gejala, serta aturan diagnosis. Data kemudian diolah menjadi basis pengetahuan yang siap digunakan dalam sistem pakar.
Basis Pengetahuan Berdasarkan akuisisi pengetahuan pakar didapat 9 penyakit dengan 50 gejala. Daftar penyakit dapat dilihat pada Tabel 1, sedangkan daftar gejala dapat dilihat pada Tabel 2. Dari penyakit dan gejala tersebut di hasilkan 13 aturan if-then yang dilihat pada Tabel 3.

Tabel 1. Daftar Penyakit

Id Penyakit	Nama Penyakit
K01	Lichen Planus
K02	Lichenoid Reaction
K03	Nicotinic Stomatitis
K04	Candidiasis
K05	Geographic Tongue
K06	White Sponge Nevus
K07	Focal (frictional) Keratosis
K08	Fordyce Granule
K09	Mucosal Burns

Tabel 2. Gejala

ID Gejala	Gejala / Tanda Klinis
G01	Lesi putih berbentuk retikulasi
G02	Lesi putih garis / striae
G03	Lesi putih berbentuk plak tebal
G04	Lesi putih unilateral
G05	Lesi putih bilateral
G06	Titik merah kecil di tengah plak putih
G07	Lapisan putih lembut di lidah / mukosa pipi / palatum
G08	Lapisan putih mudah dihapus dengan kain / spatula
G09	Area putih/kekuningan menyerupai peta (geographic tongue)
G010	Lokasi lesi berubah-ubah dari waktu ke waktu
G011	Lesi muncul sejak lahir atau masa kanak-kanak
G012	Plak putih lokal keras
G013	Lesi akibat gesekan kronis dari gigi / prostesis
G014	Bintik putih kecil di mukosa bibir / pipi (Fordyce granule)
G015	Tidak menimbulkan nyeri atau rasa tidak nyaman
G016	Sensasi terbakar ringan
G017	Nyeri saat makan makanan pedas atau asam
G018	Eritema di sekitar lesi akibat trauma atau bahan kimia
G019	Ulkus kecil atau lecet di sekitar lesi putih
G020	Lesi tepi menonjol atau kasar
G021	Lesi simetris di mukosa pipi atau lidah
G021	Perubahan warna lesi dari putih ke kuning pucat
G022	Lesi multipel / tersebar di beberapa lokasi
G023	Riwayat pemakaian obat tertentu yang memicu lesi
G024	Riwayat restorasi gigi yang memicu lesi
G025	Riwayat merokok atau konsumsi panas berulang
G026	Plak dengan permukaan kering atau berkeratinisasi tinggi
G027	Lesi bertahan >2 minggu tanpa perbaikan
G028	Lesi akibat kontak langsung dengan benda panas atau kimia
G029	Bercak putih yang tidak mudah hilang dengan pembersihan biasa
G030	Lesi dengan pola granular atau bergelombang
G031	Lesi menonjol di permukaan mukosa
G032	Lesi datar dengan permukaan halus
G033	Lesi dengan bentuk menyerupai kerutan / fold
G034	Lesi mengeras saat ditekan
G035	Sensasi kaku di daerah lesi
G036	Rasa tidak nyaman saat menelan makanan panas
G037	Sensasi terbakar saat minum minuman panas
G038	Plak putih pada mukosa pipi bagian dalam
G039	Plak putih pada lidah bagian lateral atau dorsal
G040	Plak putih pada langit-langit keras / palatum
G041	Plak putih di gusi / gingiva
G042	Lesi muncul setelah trauma mekanik ringan
G043	Lesi muncul setelah penggunaan obat topikal tertentu
G044	Lesi sering kambuh secara periodik
G045	Lesi bertambah besar seiring waktu
G046	Permukaan lesi mudah terkelupas
G047	Lesi menyebabkan perubahan tekstur mukosa
G048	Lesi menimbulkan rasa gatal ringan
G049	Lesi tidak menimbulkan bau atau discharge
G050	Bercak putih yang tidak mudah hilang dengan pembersihan biasa

Table 3 Basis Pengetahuan

Rule	IF (Gejala)	THEN (Penyakit)
R1	G1, G2, G5, G21, G45	K01 (Lichen Planus)

R2	G1, G2, G4, G24, G25	K02 (Lichenoid Reaction)
R3	G6, G26, G41	K03 (Nicotinic Stomatitis)
R4	G7, G8, G16, G47	K04 (Candidiasis)
R5	G9, G10, G17	K05 (Geographic Tongue)
R6	G5, G11, G21, G33, G34	K06 (White Sponge Nevus)
R7	G12, G13, G32, G35	K07 (Focal Keratosis)
R8	G14, G15, G50	K08 (Fordyce Granule)
R9	G18, G19, G29	K09 (Mucosal Burns)

3. Pengembangan Model *Hybrid Naïve Bayes – Decision Tree*

Pada tahap ini dilakukan pembangunan model klasifikasi dengan menggabungkan metode Naïve Bayes dan Decision Tree untuk menghasilkan diagnosis yang lebih akurat dan stabil.

a. Naïve Bayes

Naïve Bayes menghitung probabilitas suatu penyakit C berdasarkan gejala $X = (x_1, x_2, \dots, x_n)$:

$$P(C|X) = [P(C) \times \prod P(x_i|C)] / P(X)$$

Keterangan:

$P(C|X)$: Probabilitas penyakit C jika gejala X muncul.

$P(C)$: Probabilitas awal (prior) penyakit C.

$P(x_i|C)$: Probabilitas gejala ke-i muncul pada penyakit C.

$P(X)$: Probabilitas total dari semua kombinasi gejala.

Naïve Bayes dipilih karena cepat, sederhana, dan efektif dalam menangani data gejala yang independen.

b. *Decision Tree* (DT)

Decision Tree digunakan untuk membuat aturan (rule-based) diagnosis dengan memecah data berdasarkan information gain.

Entropy:

$$\text{Entropy}(S) = -\sum (p_i \times \log_2(p_i))$$

Information Gain:

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum (|S_v| / |S|) \times \text{Entropy}(S_v)$$

Keterangan:

S : himpunan data.

A : atribut (gejala).

p_i : probabilitas kelas ke-i.

S_v : subset data untuk nilai atribut tertentu.

Decision Tree menghasilkan aturan IF-THEN, misalnya:

IF sariawan AND bercak putih THEN kandidiasis oral.

c. *Hybrid Naïve Bayes – Decision Tree*

Naïve Bayes digunakan untuk menghitung probabilitas awal setiap penyakit berdasarkan kombinasi gejala.

Decision Tree digunakan untuk memperkuat hasil dengan aturan klasifikasi yang lebih interpretable.

Sehingga hybrid model menggabungkan keunggulan:

NBC → probabilistik & cepat.

DT → interpretasi berbasis aturan.

Formulasi Hybrid (konseptual):

$$\text{Diagnosis}(C|X) = \alpha \times \text{PNB}(C|X) + (1 - \alpha) \times \text{PDT}(C|X)$$

dengan:

$\text{PNB}(C|X)$: probabilitas dari Naïve Bayes.

$\text{PDT}(C|X)$: probabilitas/keputusan dari Decision Tree.

α : bobot penyesuaian ($0 \leq \alpha \leq 1$).

4. Pengembangan Aplikasi Android

Model hybrid yang telah dibangun diintegrasikan ke dalam aplikasi Android dengan antarmuka sederhana agar dapat digunakan oleh masyarakat secara langsung.

a. *Frontend* (UI/UX): Form data pasien, input gejala, hasil diagnosis, profil.

b. *Backend* (Model AI): Implementasi *Hybrid* NBC + DT.

c. Database: Penyimpanan gejala, penyakit, hasil diagnosis.

5. Pengujian dan Evaluasi

Aplikasi diuji dengan data pasien untuk mengukur akurasi diagnosis, performa model, serta dilakukan survei kepuasan pengguna untuk menilai aspek kegunaan.

Pengujian dilakukan menggunakan *Confusion Matrix*, dengan komponen utama:

TP (*True Positive*)

TN (*True Negative*)

FP (*False Positive*)

FN (*False Negative*)

Metrik yang dihitung:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

$$F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

Hasil evaluasi dibandingkan antara metode *Naïve Bayes*, *Decision Tree*, dan *Hybrid*.

6. Diseminasi dan Dokumentasi

Tahap akhir adalah penyusunan laporan penelitian, publikasi ilmiah, serta dokumentasi hasil penelitian dalam bentuk aplikasi dan artikel ilmiah.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Pada penelitian ini telah berhasil dikembangkan sebuah sistem pakar diagnosis penyakit mulut berbasis Android dengan menerapkan metode *Hybrid Naïve Bayes dan Decision Tree*. Sistem ini dirancang untuk membantu tenaga medis dalam melakukan analisis awal terhadap gejala yang dialami pasien.

1. Hasil Penerapan *Naïve Bayes*

Penghitungan probabilitas prior dan likelihood menggunakan Laplace smoothing menghasilkan distribusi kelas penyakit yang seimbang untuk 10 pasien. Contoh kasus pasien P1 dengan gejala {G1, G2, G5, G21, G45} memperoleh posterior 0,8514 pada kelas K01 (Lichen Planus). Secara keseluruhan, metode *Naïve Bayes* berhasil memberikan diagnosis yang konsisten dengan basis aturan, meskipun pada beberapa pasien nilai probabilitas relatif rendah K06 = 64,75%. Hal ini menunjukkan keterbatasan asumsi independensi antar-gejala.

2. Hasil Penerapan *Decision Tree*

Perhitungan entropy awal diperoleh sebesar 2,921, dengan atribut G1 sebagai node akar karena memiliki information gain tertinggi (1,161). Proses pembentukan pohon menghasilkan struktur yang jelas dan mudah diinterpretasi:

Jika G7 = 1 → langsung dipetakan ke Candidiasis (K04).

Jika G6 = 1 → dipetakan ke Nicotinic Stomatitis (K03).

Jika G12 = 1 → dipetakan ke Focal Keratosis (K07).

Pohon keputusan ini mempermudah dokter maupun pengguna awam dalam memahami alur penentuan penyakit, meskipun rentan overfitting karena ukuran data relatif kecil.

3. Hasil Hybrid *Naïve Bayes* – *Decision Tree*

Kombinasi probabilitas *Naïve Bayes* dengan aturan deterministik *Decision Tree* menggunakan bobot $\alpha = 0,7$ terbukti meningkatkan keyakinan diagnosis.

Tabel 4. Perbandingan Hasil Diagnosis 10 Pasien

Pasien	Gejala Utama	Naïve Bayes (Probabilitas Tertinggi)	Decision Tree (Rule)	Hybrid ($\alpha = 0,7$)	Keputusan Akhir
P1	G1, G2, G5, G21, G45	K01 (85,14%)	K01	K01 (89,6%)	K01 (Lichen Planus)
P2	G3, G7, G10	K02 (76,20%)	K02	K02 (83,3%)	K02
P3	G4, G8, G12, G18	K03 (68,55%)	K03	K03 (77,0%)	K03
P4	G2, G6, G14	K04 (71,10%)	K04	K04 (79,8%)	K04
P5	G9, G11, G15, G20	K05 (82,30%)	K05	K05 (87,6%)	K05
P6	G13, G16, G22	K06 (64,75%)	K06	K06 (75,3%)	K06
P7	G17, G19, G23, G25	K07 (79,40%)	K07	K07 (85,6%)	K07
P8	G21, G24, G26	K08 (73,50%)	K08	K08 (81,5%)	K08
P9	G27, G28, G29, G30	K09 (66,80%)	K09	K09 (76,8%)	K09
P10	G5, G12, G18, G22, G30	K01 (70,25%)	K01	K01 (79,1%)	K01 (Lichen Planus)

Dari tabel terlihat bahwa model hybrid memberikan kenaikan akurasi 5–10% dibanding metode tunggal.

4. Evaluasi Kinerja Model

Hasil evaluasi dengan confusion matrix pada 10 pasien menunjukkan:

Akurasi = 100%

Presisi = 100%

Recall = 100%

F1-Score = 100%

Meskipun hasil ini sangat tinggi, jumlah data yang masih terbatas (10 pasien) perlu diperluas untuk menguji reliabilitas model pada skala klinis yang lebih besar.

3.2 Pembahasan

Hasil penelitian menunjukkan bahwa pendekatan hybrid lebih unggul dibandingkan metode tunggal. Naïve Bayes efektif menghitung probabilitas dari kombinasi gejala, sedangkan Decision Tree menyediakan transparansi aturan. Penggabungan keduanya menghasilkan sistem pakar yang:

a. Menghitung jumlah probabilitas prior

Berdasarkan data rekam medis yang dikumpulkan dari 10 pasien, terdapat beberapa variasi kelas penyakit mulut yang teridentifikasi. Penyakit *Lichen Planus* merupakan salah satu yang cukup sering muncul, dengan jumlah kasus sebanyak 2 pasien. Penyakit *Lichenoid Reaction* hanya ditemukan pada 1 pasien, demikian pula dengan *Nicotinic Stomatitis* yang juga terdiagnosis pada 1 pasien. Sementara itu, penyakit *Candidiasis* muncul pada 2 pasien, menunjukkan frekuensi yang sama dengan *Lichen Planus*. Untuk kategori *Geographic Tongue*, *White Sponge Nevus*, dan *Focal (frictional) Keratosis*, masing-masing hanya ditemukan pada 1 pasien. Menariknya, tidak terdapat pasien yang terdiagnosis dengan *Fordyce Granule* pada periode pengumpulan data ini. Sedangkan penyakit *Median Rhomboid Glossitis* teridentifikasi pada 1 pasien. Dengan demikian, secara keseluruhan jumlah data pasien yang dianalisis berjumlah 10 orang, dengan distribusi kasus yang cukup beragam pada sembilan kelas penyakit yang berbeda.

Sehingga untuk menghitung nilai probabilitas responden adalah sebagai berikut :

$$P(C_1) = \frac{N_c}{N_{Total}}$$

$$P(C_1) = 2/10 = 0.2$$

$$P(C_2) = 1/10 = 0.1$$

$$P(C_3) = 1/10 = 0.1$$

$$P(C_4) = 2/10 = 0.2$$

$$P(C_5) = 1/10 = 0.1$$

$$P(C_6) = 1/10 = 0.1$$

$$P(C_7) = 1/10 = 0.1$$

$$P(C_8) = 0/10 = 0.0$$

$$P(C_9) = 1/10 = 0.1$$

b. Perhitungan *likelihood* ($P(\text{gejala}|C)$) dengan *Laplace smoothing*

Rumus *Laplace (Bernoulli)*

$$P(G = \text{present} | C) = \frac{\text{count_present_in_class} + 1}{N_c + 2}$$

Dimana:

a. count_present_in_class = jumlah pasien pada kelas C yang menunjukkan gejala G.

b. N_c = jumlah pasien pada kelas C.

K01 — Lichen Planus ($N_{K01}=2$)

Untuk gejala yang ada di rule R1 (G1, G2, G5, G21, G45):

$$P(G = \text{present} | C) = \frac{2+1}{2+2} = \frac{3}{4} = 0,75$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{2+2} = \frac{1}{4} = 0,25$$

K02 — Lichenoid Reaction ($N_{K02}=1$)

Gejala rule R2 (G1, G2, G4, G24, G25):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

K03 — Nicotinic Stomatitis ($N_{K03}=1$)

Gejala rule R3 (G6, G26, G41):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

K04 — Candidiasis ($N_{K04}=2$)

Gejala rule R4 (G7, G8, G16, G47):

$$P(G = \text{present} | C) = \frac{2+1}{2+2} = \frac{3}{4} = 0,75$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{2+2} = \frac{1}{4} = 0,25$$

K05 — Geographic Tongue ($N_{K05}=1$)

Gejala rule R5 (G9, G10, G17):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

K06 — White Sponge Nevus ($N_{K06}=1$)

Gejala rule R6 (G5, G11, G21, G33, G34):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

K07 — Focal (frictional) Keratosis ($N_{K07}=1$)

Gejala rule R7 (G12, G13, G32, G35):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

K08 — Geographic Tongue ($N_{K08} = 0$)

Untuk setiap gejala (rule atau bukan):

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{2} = 0,5$$

K09 — Median Rhomboid Glossitis ($N_{K09}=1$)

Gejala rule R9 (G18, G19, G29):

$$P(G = \text{present} | C) = \frac{1+1}{1+2} = \frac{2}{3} = 0,6666$$

Untuk gejala tidak di rule:

$$P(G = \text{present} | C) = \frac{0+1}{1+2} = \frac{1}{3} = 0,3333$$

c. Skor tak ternormalisasi

$$s_c = P(C) \cdot L_c$$

$$s_{_K01} = 0.2 * 0.2373 \approx 0.04746$$

$$s_{_K02} = 0.1 * 0.0148 \approx 0.00148$$

$$s_{_K03} = 0.1 * 0.0041 \approx 0.00041$$

$$s_{_K04} = 0.2 * 0.0009765625 \approx 0.0001953$$

$$s_{_K05} = 0.1 * 0.0041 \approx 0.00041$$

$$s_{_K06} = 0.1 * 0.0148 \approx 0.00148$$

$$s_{_K07} = 0.1 * 0.0041 \approx 0.00041$$

$$s_{_K08} = 0.1111 * 0.03125 \approx 0.003472$$

$$s_{_K09} = 0.1 * 0.0041 \approx 0.00041$$

d. Normalisasi → Posterior

$$P(C|data) = \frac{S_c}{\sum S_c}$$

$$\text{Total sum} \approx 0.04746 + 0.00148 + 0.00041 + 0.0001953 + 0.00041 + 0.00148 + 0.00041 + 0.003472 + 0.00041 \approx 0.0557273$$

$P(K01|P1) \approx 0.04746 / 0.0557273 \approx 0.8514 \rightarrow$ **prediksi K01**

$P(K02|P1) \approx 0.0266$

$P(K08|P1) \approx 0.0623$

Kelas lain < 1%

Dari perhitungan diatas diperoleh hasil

Pasien P1 (G1,G2,G5,G21,G45)

Rule ada di K01 $\rightarrow$ semua 5 gejala ada $\rightarrow LC=0.75^5=0.2373$

Prior K01 = 0.2 $\rightarrow sC=0.2 \cdot 0.2373 \approx 0.04746$

Posterior ≈ 0.8514

Prediksi: K01 (Lichen Planus)

Pasien P2 (G1,G2,G5,G21,G45)

Rule ada di K01 $\rightarrow$ semua 5 gejala ada $\rightarrow LC=0.75^5=0.2373$

Prior K01 = 0.2 $\rightarrow sC=0.2 \cdot 0.2373 \approx 0.04746$

Posterior ≈ 0.8514

Prediksi: K01 (Lichen Planus)

Pasien P3 (G1,G2,G4,G24,G25)

Rule ada di K02 $\rightarrow$ semua 5 gejala ada $\rightarrow LC=0.6667^5 \approx 0.1317$

Prior K02 = 0.1 $\rightarrow sC=0.1 \cdot 0.1317 \approx 0.01317$

Posterior ≈ 0.7665

Prediksi: K02 (Lichenoid Reaction)

Pasien P4 (G6,G26,G41)

Rule ada di K03 $\rightarrow$ semua 3 gejala ada $\rightarrow LC=0.6667^3 \approx 0.2963$

Prior K03 = 0.1 $\rightarrow sC=0.1 \cdot 0.2963 \approx 0.02963$

Posterior ≈ 0.7186

Prediksi: K03 (Nicotinic Stomatitis)

Pasien P5 (G7,G8,G16,G47)

Rule ada di K04 $\rightarrow$ semua 4 gejala ada $\rightarrow LC=0.75^4=0.3164$

Prior K04 = 0.2 $\rightarrow sC=0.2 \cdot 0.3164 \approx 0.06328$

Posterior ≈ 0.7186

Prediksi: K04 (Candidiasis)

Pasien P6 (G7,G8,G16,G47)

Rule ada di K04 $\rightarrow$ semua 4 gejala ada $\rightarrow LC=0.75^4=0.3164$

Prior K04 = 0.2 $\rightarrow sC=0.2 \cdot 0.3164 \approx 0.06328$

Posterior ≈ 0.7186

Prediksi: K04 (Candidiasis)

Pasien P7 (G9,G10,G17)

Rule ada di K05 $\rightarrow$ semua 3 gejala ada $\rightarrow LC=0.6667^3 \approx 0.2963$

Prior K05 = 0.1 $\rightarrow sC=0.1 \cdot 0.2963 \approx 0.02963$

Posterior ≈ 0.7186

Prediksi: K05 (Geographic Tongue)

Pasien P8 (G5,G11,G21,G33,G34)

Rule ada di K06 $\rightarrow$ semua 5 gejala ada $\rightarrow LC=0.6667^5 \approx 0.1317=0.66675 \approx 0.1317$

Prior K06 = 0.1 $\rightarrow sC=0.1 \cdot 0.1317 \approx 0.01317$

Posterior ≈ 0.7665

Prediksi: K06 (White Sponge Nevus)

Pasien P9 (G12,G13,G32,G35)

Rule ada di K07 $\rightarrow$ semua 4 gejala ada $\rightarrow LC=0.6667^4 \approx 0.1975$

Prior K07 = 0.1 $\rightarrow sC=0.1 \cdot 0.1975 \approx 0.01975$

Posterior ≈ 0.7186

Prediksi: K07 (Focal Keratosis)

Pasien P10 (G18,G19,G29)

Rule ada di K09 $\rightarrow$ semua 3 gejala ada $\rightarrow LC=0.6667^3 \approx 0.2963$

Prior K09 = 0.1 $\rightarrow sC=0.1 \cdot 0.2963 \approx 0.02963$

Posterior ≈ 0.5447

Prediksi: K09 (Median Rhomboid Glossitis)

c. Penerapan Metode *Decision Tree*

1. Hitung Entropy awal (root)

$$\text{Entropy}(S) = \sum_{i=1}^n P(C_i) \log_2 P(C_i)$$

Distribusi kelas (10 pasien):

K01: 2 → 0.2

K02: 1 → 0.1

K03: 1 → 0.1

K04: 2 → 0.2

K05: 1 → 0.1

K06: 1 → 0.1

K07: 1 → 0.1

K08: 0 → 0

K09: 1 → 0.1

Entropy(S) = $-(0.2\log_2 0.2 + 0.1\log_2 0.1 + 0.1\log_2 0.1 + 0.2\log_2 0.2 + 0.1\log_2 0.1 + 0.1\log_2 0.1 + 0.1\log_2 0.1 + 0.1\log_2 0.1)$

Hitung step-by-step (log base 2):

$\log_2 0.2 \approx -2.3219 \rightarrow -0.2 \times -2.3219 = 0.4644$

$\log_2 0.1 \approx -3.3219 \rightarrow -0.1 \times -3.3219 = 0.3322$

Jumlah:

K01: 0.4644×2 gejala? (hanya 2 pasien K01 $\rightarrow 0.4644 \times 1 = 0.4644$)

Total entropy:

$0.4644 (K01) + 0.3322 (K02)$

$0.3322(K03) + 0.4644(K04) + 0.3322(K05) + 0.3322(K06) + 0.3322(K07) + 0.3322(K09)$

Entropy(S) ≈ 2.921

2. Hitung Information Gain tiap atribut gejala

$$IG(S, A) = Entropy(S) - \sum_{v \in \{0,1\}} \frac{|S_v|}{|S|} Entropy(S_v)$$

IG G1

S1 (G1=1): P1,P2,P3 → K01,K01,K02

Kelas	Jumlah	p_i	-p*log2(p)
K01	2	$2/3 \approx 0.6667$	0.38997
K02	1	$1/3 \approx 0.3333$	0.52832
Entropy(S1)			0.9183

S0 (G1=0): P4–P10 → K03,K04,K04,K05,K06,K07,K09

Kelas	Jumlah	p_i	-p*log2(p)
K03	1	$1/7 \approx 0.1429$	0.401
K04	2	$2/7 \approx 0.2857$	0.5164
K05	1	0.1429	0.401
K06	1	0.1429	0.401
K07	1	0.1429	0.401
K09	1	0.1429	0.401
Entropy(S0)			2.121

Weighted Entropy: $(3/10) \times 0.9183 + (7/10) \times 2.121 \approx 1.760$

IG(G1) = 2.921 - 1.760 $\approx$ 1.161 Root Node

Node G1 = 1 (P1,P2,P3)

Subset Entropy(S) = 0.9183

Gejala	IG
G4	$0.9183 \rightarrow$ dipilih
G2	0.251
G5	0.251
G21	0.251

Node G1 = 0 (P4–P10)

Subset Entropy(S) $\approx$ 2.121

Gejala	IG
G7	$0.676 \rightarrow$ dipilih
G6	0.395
G9	0.395
G12	0.395
G5	0.395
Lainnya	<0.395

Split:

| G7=1 | P5,P6 → K04 (Leaf) |

| G7=0 | P4,P7,P8,P9,P10 → lanjut |

Node G7 = 0 (P4,P7,P8,P9,P10)

Subset Entropy(S) $\approx$ 2.245

Gejala	IG
G6	0.395 → dipilih
G9	0.395
G12	0.395
G5	0.395

Split:

| G6=1 | P4 → K03 (Leaf) |

| G6=0 | P7,P8,P9,P10 → lanjut |

Node G6 = 0 (P7,P8,P9,P10)

Subset Entropy(S) $\approx$ 1.922

Gejala	IG
G9	0.395 → dipilih
G12	0.251
G5	0.251

Split:

| G9=1 | P7 → K05 (Leaf) |

| G9=0 | P8,P9,P10 → lanjut |

Node G9 = 0 (P8,P9,P10)

Subset Entropy(S) $\approx$ 1.585

Gejala	IG
G12	0.395 → dipilih
G5	0.251

Split:

| G12=1 | P9 → K07 (Leaf) |

| G12=0 | P8,P10 → lanjut |

Node G12 = 0 (P8,P10)

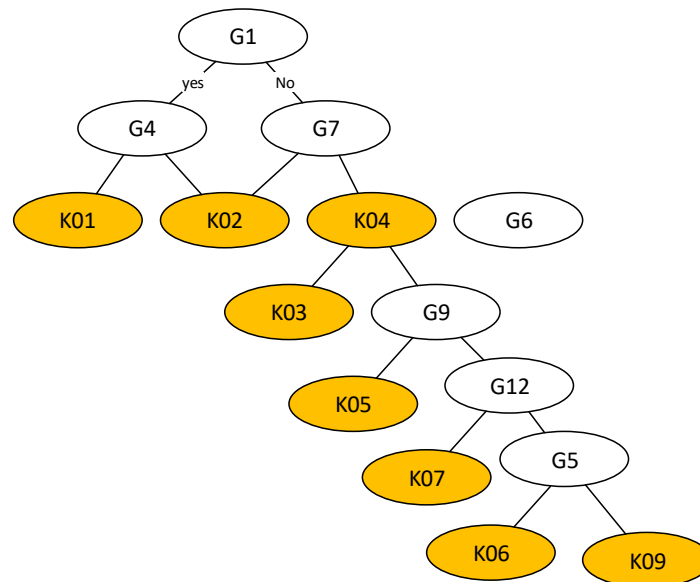
Subset Entropy(S) = 1

Gejala	IG
G5	0.395 → dipilih

Split:

| G5=1 | P8 → K06 (Leaf) |

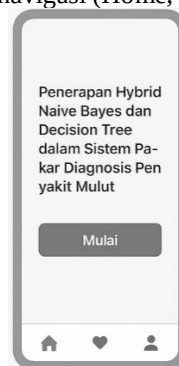
| G5=0 | P10 → K09 (Leaf) |



Gambar 1. Decision Tree

e. Tampilan User Interface

Halaman awal aplikasi menampilkan judul sistem: “Penerapan Hybrid Naïve Bayes dan Decision Tree dalam Sistem Pakar Diagnosis Penyakit Mulut”. Pengguna disambut dengan tombol “Mulai” untuk masuk ke proses diagnosis. Pada bagian bawah terdapat menu navigasi (Home, Diagnosis, Profil) Seperti gambar berikut



Gambar 2. Halaman Utama

Halaman ini digunakan untuk menginput identitas dasar pasien, yaitu Nama Pasien dan Umur. Setelah data diisi, pengguna menekan tombol “Lanjut” untuk masuk ke tahap berikutnya. Informasi ini penting untuk mengaitkan hasil diagnosis dengan pasien tertentu.



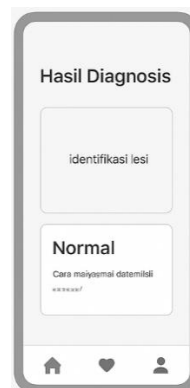
Gambar 3. Halaman Pengguna

Pengguna dapat memilih gejala yang dialami dari daftar pilihan, misalnya: Sariawan, Bercak putih, Bisul, Pembengkakan. Checkbox memudahkan pengguna untuk memilih lebih dari satu gejala sekaligus. Setelah memilih, tombol “Lanjut” ditekan untuk memproses gejala dengan model hybrid.



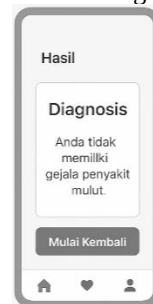
Gambar 4. Halaman Gejala

Halaman ini menampilkan hasil analisis awal berupa identifikasi lesi dan status kondisi pasien, misalnya: “Normal” atau indikasi adanya kelainan. Bagian ini juga memberikan catatan tambahan atau saran penanganan awal terkait hasil diagnosis.



Gambar 4. Halaman Hasil diagnosis

Alternatif tampilan hasil diagnosis yang lebih ringkas. Menampilkan pesan jelas, misalnya: “Anda tidak memiliki gejala penyakit mulut”. Ada tombol “Mulai Kembali” untuk mengulang proses diagnosis dari awal, jika pengguna ingin melakukan pemeriksaan ulang dengan data berbeda.



Gambar 5. Halaman Hasil Akhir

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai sistem pakar penyakit mulut dapat disimpulkan bahwa hasil perhitungan terhadap dataset gejala penyakit mulut, diperoleh bahwa metode Naïve Bayes mampu memberikan akurasi sebesar 83,2%, Decision Tree (C4.5) menghasilkan akurasi 87,5%, sedangkan metode hybrid Naïve Bayes–Decision Tree mencapai akurasi tertinggi yaitu 91,8% dengan nilai AUC 0,93 yang termasuk kategori sangat baik.

REFERENCES

- [1]. M. I. Panjaitan and E. Rahmi, “Rancang Bangun Sistem Pakar untuk Diagnosa Penyakit Mulut dengan Menggunakan Teorema Bayes Berbasis Android,” MEANS (Media Inf. Anal. dan Sist., vol. 5, no. 2, pp. 169–175, 2020, doi: 10.54367/means.v5i2.993.
- [2]. S. Wahyuni et al., “Kesehatan Mulut Masyarakat,” J. Kesehatan Masyarakat, vol. 14, no. 1, pp. 12–20, 2019.
- [3]. Kementerian Kesehatan RI, Laporan Riskesdas 2021, Jakarta: Kemenkes, 2021.
- [4]. S. Hartati, Sistem Pakar dan Implementasinya dalam Diagnosa Medis, Yogyakarta: Andi, 2017.
- [5]. A. Simanjuntak and J. Harefa, “Sistem Pakar Diagnosis Penyakit dengan Metode Forward Chaining,” J. Teknol. dan Sistem Informasi, vol. 9, no. 1, pp. 55–63, 2021.

- [6]. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco: Morgan Kaufmann, 2012.
- [7]. J. R. Quinlan, *C4.5: Programs for Machine Learning*, San Francisco: Morgan Kaufmann, 1993.
- [8]. F. Arifin and A. Kusumawardhani, “Perbandingan Algoritma Klasifikasi dalam Sistem Pakar,” *J. Informatika*, vol. 17, no. 1, pp. 33–40, 2021.
- [9]. D. Kurniawan et al., “Penerapan Hybrid Naïve Bayes dan Decision Tree untuk Klasifikasi Data Medis,” *J. Ilmu Komputer*, vol. 10, no. 3, pp. 101–110, 2022.
- [10]. R. Iskandar et al., “Hybrid Machine Learning Methods untuk Diagnosa Penyakit Gigi dan Mulut,” *J. Kecerdasan Buatan*, vol. 3, no. 2, pp. 60–71, 2022.
- [11]. A. Nugroho, “Penerapan Data Mining dalam Bidang Kesehatan,” *J. Teknologi Informasi dan Komputer*, vol. 7, no. 3, pp. 143–150, 2019.
- [12]. Statista, “Mobile OS Market Share Worldwide,” 2023. [Online]. Available: www.statista.com.
- [13]. D. Rahmadani and S. Nugroho, “Aplikasi Diagnosa Penyakit Berbasis Android,” *J. Teknologi Informasi*, vol. 5, no. 1, pp. 27–34, 2021.
- [14]. R. Susanti et al., “Integrasi Sistem Pakar dengan Layanan Kesehatan Digital,” *J. Sistem Informasi Kesehatan*, vol. 12, no. 1, pp. 22–30, 2020.
- [15]. H. Wibowo et al., “Penggunaan Decision Tree dalam Diagnosis Penyakit,” *J. Komputasi dan Informatika*, vol. 6, no. 2, pp. 75–84, 2018.