



Klasifikasi Penyakit *Hiperkolesterol* Menggunakan *Algoritma Decision Tree C4.5*

Aprisarita Liya^{a,1,*}, Yonal Supit^{b,2}, Andi Muhammad Islah^{b,3}

^{ab} STMIK Catur Sakti Kendari, D.I Panjaitan, Kendari and 9316, Indonesia
¹ aprisaritalya@gmail.com ; ² yonalsupit@gmail.com* ; ³ andiislah@gmail.com

* corresponding author

ABSTRACT

ARTICLE INFO

Penyakit hiperkolesterol menimbulkan tantangan serius dalam bidang kesehatan, mengingat kompleksitas faktor risiko yang terlibat. Dalam upaya menangani masalah ini, penerapan metode machine learning (ML) menonjol sebagai pendekatan yang inovatif. Tujuan utama penelitian ini adalah untuk menginvestigasi kemampuan ML dalam meningkatkan klasifikasi penyakit hiperkolesterol. Metode penelitian melibatkan analisis data pasien dengan kecepatan dan ketepatan tinggi, memungkinkan identifikasi pola-pola kompleks yang mendasari hiperkolesterol. Hasilnya, ML memberikan kontribusi signifikan dalam deteksi dini, memungkinkan pengembangan rencana perawatan yang lebih spesifik dan efisien. Selain itu, metode ini membuka peluang untuk pendekatan personal dalam perawatan dengan memperhitungkan faktor risiko individu, seperti riwayat keluarga dan gaya hidup. Prediksi risiko kardiovaskular yang diperoleh dari model ML juga menjadi landasan untuk penilaian risiko individual, yang dapat membimbing tindakan pencegahan yang lebih tepat. Manfaat lainnya mencakup pengelolaan data pasien yang terpusat, meningkatkan koordinasi perawatan, dan efisiensi pertukaran informasi di antara penyedia layanan kesehatan. Implikasi hasil penelitian ini tidak hanya terbatas pada peningkatan efektivitas diagnosis, tetapi juga melibatkan aspek kualitas hidup pasien yang ditingkatkan. Diskusi mendalam tentang hasil menunjukkan bahwa penerapan ML mendorong pendekatan perawatan yang lebih responsif dan personal, melibatkan pemantauan pasien yang lebih efisien dan tanggung jawab dalam manajemen jangka panjang. Dengan demikian, penelitian ini memberikan kontribusi pada evolusi sistem kesehatan menuju pendekatan holistik yang lebih terperinci dan adaptif terhadap kompleksitas penyakit hiperkolesterol.

Article history

Received: 7 September 2023

Revised: 21 November 2023

Accepted: 27 November 2023

Keywords

Hiperkolesterol

Stroke

Atribut

Algoritma C4.5

Decision Tree

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



I Pendahuluan

Dalam beberapa tahun terakhir, penelitian di bidang kesehatan telah mengalami perkembangan pesat, terutama dalam pemakaian teknologi *artificial intelligence* (AI) untuk diagnosis penyakit. Salah satu pendekatan yang menonjol adalah penggunaan Algoritma C4.5, yang merupakan Algoritma pembelajaran mesin untuk pengambilan keputusan dan klasifikasi. Meskipun telah banyak penelitian yang menggali potensi Algoritma C4.5 dalam berbagai konteks, implementasinya pada klasifikasi



penyakit masih terbatas. Penting untuk mencatat bahwa Algoritma C4.5 memiliki keunggulan dalam kemampuannya untuk mengatasi dataset kompleks dengan variabel yang bervariasi. Dalam konteks penyakit tertentu, seperti hiperkolesterol, di mana faktor risiko dapat sangat bervariasi, implementasi C4.5 dapat menjadi solusi yang menjanjikan untuk meningkatkan akurasi klasifikasi dan memberikan dasar yang lebih kuat untuk pengambilan keputusan klinis. Meskipun beberapa penelitian telah mencoba menerapkan Algoritma C4.5 pada bidang kesehatan secara umum, masih ada kekurangan dalam konteks klasifikasi penyakit tertentu seperti hiperkolesterol. Oleh karena itu, penelitian ini diarahkan untuk mengisi kesenjangan tersebut dengan mengeksplorasi potensi dan batasan implementasi Algoritma C4.5 dalam konteks spesifik klasifikasi penyakit hiperkolesterol. Dengan memfokuskan perhatian pada Algoritma C4.5, diharapkan penelitian ini tidak hanya dapat memberikan kontribusi pada pemahaman umum tentang penerapan kecerdasan buatan dalam kesehatan, tetapi juga menyoroti potensi khusus Algoritma ini dalam meningkatkan klasifikasi penyakit hiperkolesterol. Melalui pemahaman yang lebih mendalam tentang kinerja Algoritma C4.5 dalam konteks ini, diharapkan dapat membuka jalan untuk pengembangan model klasifikasi yang lebih efektif dan relevan untuk meningkatkan pengelolaan penyakit hiperkolesterol.

Berdasarkan Data (Riskesdas, 2018) di Indonesia berdasarkan tempat tinggal menunjukkan bahwa kadar kolesterol yang tinggi di atas normal pada penduduk perkotaan lebih besar dari pada pedesaan, penduduk pada kota sebesar 38,5% sedangkan penduduk pedesaan sebesar 32,1%. Menurut Hasil Survei Kesehatan Rumah Tangga (SKRT) tahun 2004, prevalensi hiperkolesterol usia 25 - 34 Tahun ialah 9,3%, sedangkan pada usia 55 - 64 sebanyak 15,5% (TIDARRIA, 2015). Prevalensi hiperkolesterol di Jawa Timur 30,38% (Riskesdas, 2018). Kadar kolesterol pada lansia baik laki - laki juga perempuan biasanya cenderung semakin tinggi. Berdasarkan KaderKes (2021) di Desa Keboan Sikep, Kecamatan Gedangan Kabupaten Sidoarjo sekitar 4 orang lansia yang menderita kolesterol dari jumlah anggota posyandu 15 orang lansia yang usianya rata - rata 60 - 70 Tahun.

Hiperkolesterol berdasarkan penyebabnya dapat diklarifikasi yaitu disebabkan oleh faktor genetik, usia, jenis kelamin, kegiatan fisik, pola makan serta obesitas (Matfin, 2003 dalam Lerebulan et al., 2021). Hiperkolesterol tidak memberikan tanda - tanda yang khas, tak jarang seseorang baru mengetahui terkena hiperkolesterol saat melakukan pemeriksaan kesehatan. Tanda - tanda yang sering ditemui yaitu sering pusing pada kepala bagian belakang, tengkuk serta pundak terasa pegal, kesemutan pada tangan dan kaki bahkan terdapat keluhan dada sebelah kiri terasa nyeri mirip tertusuk. Apabila kolesterol tinggi lalu menyumbat saluran pembuluh darah, maka oksigen serta nutrisi pada darah tak akan mampu mencapai jantung sebagai akibatnya menyebabkan serangan jantung dan bila peredaran darah ke otak terputus (sebab kolesterol tinggi), maka akan terjadi serangan stroke (Miachitra, 2021). Menurut data (World Health Organization dalam Amalia & Isnaeni, 2017) menunjukkan lebih dari 50% serangan jantung dan 20% serangan stroke disebabkan oleh kadar kolesterol tinggi. Hiperkolesterol banyak diderita lansia sebab faktor usia yang semakin usang badan akan semakin malas untuk bergerak, sebagai akibatnya kolesterol di dalam tubuh akan menumpuk pada hati, oleh sebab itu diperlukan gerak yang seimbang antara menjaga pola makan serta olahraga supaya terhindar dari kolesterol berlebih [1].

2 Implementasi Metode Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk menghasilkan pohon keputusan. Alur kerja pada algoritma C4.5 adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang mempunyai prioritas tertinggi atau yang mempunyai nilai gain tertinggi berdasarkan nilai atribut entropy sebagai sumbu klasifikasi atribut. Pada tahap ini algoritma C4.5 mempunyai 2 aturan kerja yaitu: Pembuatan pohon keputusan, dan pembuatan aturan (rule model). Aturan yang dibentuk dari pohon keputusan akan membentuk suatu kondisi berupa if maka [2]. Algoritma C4.5 mengolah data dari pohon keputusan menjadi aturan sederhana yang mudah dipahami. Selain itu dalam penerapannya algoritma C4.5 mudah untuk mengetahui atribut paling berpengaruh [3]. Pembelajaran Mesin telah digunakan dalam berbagai aplikasi dalam beberapa tahun terakhir. Algoritma pembelajaran mesin digunakan dalam pemrosesan bahasa, kontrol robot, optimasi kombinatorial, pengenalan suara, pengenalan tulisan tangan, pengenalan wajah, analisis data medis, dll [4].

4.1. Algoritma Decision Tree

Decision tree atau pohon keputusan merupakan proses mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule dan menyederhanakan. Pohon keputusan termasuk dalam kategori data mining klasifikasi [5].

Algoritma C4.5 menggunakan struktur pengumpulan data dengan jumlah yang cukup besar yang dijadikan sebagai bentuk sebuah himpunan rekaman yang lebih efisien dan menggunakan rangkain aturan keputusan yang sesuai [6]. Algoritma C4.5 yang diterapkan dalam upaya pembangunan pohon harus memperhatikan beberapa hal berikut, diantaranya:

- a. Memilih atribut yang menjadi akar
- b. Membuat cabang pada setiap nilai
- c. Membagi kasus pada cabang tersebut
- d. Mengulangi tahapan pada seluruh cabang hingga setiap kasus mempunyai kelas yang selaras.

Sedangkan Menurut [7], dalam pembangunan decision tree dengan menerapkan algoritma C4.5, beberapa tahapan yang harus dipenuhi yaitu:

- a. Menyiapkan data training.
- b. Menentukan akar dari pohon.
- c. Konsep Entropy diterapkan pada proses megekstrak sebuah kelas. Entropy digunakan pada ketidakaslisan s, dengan rumus yang digunakan, diantaranya :

$$\text{Entropy}(S) = \sum_{i=1}^n -p_i \cdot \log_2 p_i \quad 1$$

Keterangan :

S = Himpunan

n = Jumlah partisi S

Pi = Proporsi Si terhadap S

Menghitung *gain*, untuk menetapkan atribut sebagai akar dengan menghitung nilai gain tertinggi dari semua atribut yang ada:

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=0}^n \frac{|S_i|}{|S|} * \text{Entropy} \quad 2$$

Keterangan :

S = Himpunan Kasus

A = Atribut

N = Jumlah patrisi atribut A

|Si| = Jumlah kasus pada partisi ke-i

|S| = Jumlah kasus dalam S

Algoritma data mining C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokkan dan bersifat prediktif. Dasar algoritma C4.5 adalah pembentukan pohon keputusan (*decision tree*). Adapun tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [8]. Proses penentuan nilai entropy dan nilai gain dari kemungkinan setiap kriteria yang menjadi acuan keputusan yang dilanjutkan dengan proses perangkaian dari hasil keputusan [9].

4.2. Penggunaan Tools Weka Untuk Algoritma C4.5

Penambangan data adalah proses menganalisis data dari berbagai sudut pandang dan merangkum hasilnya menjadi informasi yang berguna. Perangkat lunak penambangan data adalah alat analisis yang memungkinkan pengguna menganalisis data dari berbagai dimensi, mengkategorikan data, dan merangkum hubungan antar data. Secara teknis, data mining adalah proses menemukan korelasi di antara banyak bidang dalam kumpulan data yang sangat besar [10].

4.3. Penyakit Hiperkolesterol

Fenomena sering kali terjadi pada masyarakat yang menderita *Hiperkolesterol* adalah tidak patuh dan tidak tahu pada makanan yang mereka makan, meskipun sudah mengetahui sedang menderita kolesterol namun tetap mengkonsumsi makanan yang mengandung kadar lemak yang tinggi seperti daging serta makanan yang mengandung santan [11]. Kadar kolesterol berlebihan dalam tubuh akan mengakibatkan penumpukan lemak di hepar. Salah satu alternatif obat tradisional yang

dapat menurunkan kadar kolesterol adalah tanaman seledri. Penelitian sebelumnya menyatakan bahwa fraksi air herba seledri dapat menurunkan kadar kolesterol total pada keadaan *hiperkolesterol* [12].

Tabel 1: Klasifikasi LDL dan HDL Kolesterol, Total Kolesterol dan Triglisierida

LDL (Kolesterol Jahat)	
<100 mg/dL	Normal
100 – 159 mg/dL	Sedang
Lebih dari 160 mg/dL	Tinggi
HDL (Kolesterol Baik)	
0 – 34 mg/dL	Rendah
35- 55 mg/dL	Tinggi
TC (Total Kolesterol)	
0- 200 mg/dL	Normal
Lebih dari 200 mg/dL	Tinggi
TGA (Triglisierida)	
100 – 150 mg/dL	Normal
151 – 190 mg/dL	Sedang
191 – 200 mg/dL	Tinggi

Sumber : Data Sekunder Puskesmas Abeli Kota Kendari

- **Kolesterol Jahat** yang tepat dalam tubuh sekitar 60% sampai 70%. LDL membawa kolesterol ke seluruh tubuh yang membutuhkan melalui jaringan arteri. Dia akan mengirimkan kapan saja ketika sel tersebut membutuhkan. Tetapi LDL terlalu banyak, akan menimbun kolesterol pada arteri sehingga menyebabkan plak-plak. Kolesterol jahat berpengaruh dengan kadar lemak jenuh dalam tubuh dan kandungan kolesterol yang kita makan akan. Sehingga ketika kadar kolesterol tinggi, anda harus melakukan diet.
- **Kolesterol Baik** mengandung lemak seperti LDL tetapi mengandung banyak protein. LDL berfungsi sebagai pengantar kolesterol sedangkan HDL berfungsi sebagai pembersih dalam saluran pembuluh arteri. Jadi HDL akan membersihkan LDL yang terlalu tinggi dalam pembuluh darah arteri untuk kembali ke hati dan dicoba untuk didaur ulang kembali. Jika kadar tinggi risiko penyakit jantung sangat kecil tetapi jika HDL rendah akan mengakibatkan penyakit jantung.
- **Kolesterol Total** adalah perhitungan total dari semua jenis kolesterol dalam darah. Kolesterol merupakan senyawa lemak yang diproduksi di hati yang biasanya ditemukan dalam darah. Peningkatan kadar kolesterol terlihat pada gangguan familial dan hipotiroidisme dan dianggap sebagai faktor risiko penyakit jantung.
- **Rigliserida** merupakan salah satu jenis lemak dalam tubuh yang mengandung banyak sekali energy. Ketika anda mengkonsumsi makanan dan energy dalam tubuh sudah tercukupi, sebagian energy akan disimpan sebagai triglisierida. Kadar triglisierida yang normal sangat dibutuhkan oleh tubuh. Tetapi jika terlalu tinggi akan mengakibatkan penyakit jantung, darah tinggi dan diabetes.

a. Berat Badan

Tabel 2 : Kategori Berat Badan Berdasarkan Indeks Masa Tubuh

Keterangan	IMT
Kekurangan berat badan tingkat berat	<17,0
Kekurangan berat badan tingkat ringan	17,0 – 18,5
Berat badan ideal	18,5 – 25,0
Kelebihan berat badan tingkat ringan	25,0 – 27,0
Kelebihan berat badan tingkat berat	>27,0

Sumber : Data Sekunder Puskesmas Abeli Kota Kendari

b. Usia

Tabel 3 : Kategori Usia Berdasarkan Jenis Kelamin

Usia	Laki-Laki			Perempuan		
	LDL	HDL	Normal	LDL	HDL	Normal
< 19 tahun	< 110	> 45	< 170	< 110	> 45	<170
> 20 tahun	< 100	> 40	125-200	< 100	> 45	125-200

Sumber : Data Sekunder Puskesmas Abeli Kota Kendari

c. Perokok

Tabel 43 : Kategori Gaya Hidup Berdasarkan Perokok dan Bukan Perokok

Tidak merokok	1
Tidak merokok tetapi hidup dan bekerja dengan perokok	2
Berhenti merokok selama 3 tahun terakhir	3
Merokok	4
Merokok dan hidup atau bekerja dengan perokok	5

Sumber : Data Sekunder Puskesmas Abeli Kota Kendari

3. Machine Learning dan Data Mining

Algoritma *Machine Learning* (ML) menghadirkan peluang untuk mengembangkan model prediksi risiko yang lebih baik dan lebih dapat digeneralisasikan. Dengan memanfaatkan data berskala besar—dari *electronic health records* (EHRs), misalnya Algoritma tersebut dapat menentukan kombinasi variabel yang dapat memprediksi suatu hasil dengan andal [13]. Machine Learning tersebut dibangun dengan *framework Data Preparation*, serta *framework Data Modelling* yang dapat diluncurkan, dan berjalan secara otomatis. Dengan model pembelajaran mesin, Pemangku Kepentingan dapat membuat studi prediksi dari kumpulan datasets Profil Kesehatan Indonesia yang membantu mereka membuat keputusan [14]. Pendekatan ML bertujuan untuk mengembangkan

model dengan kemampuan generalisasi yang lebih tinggi dan keandalan prediksi yang lebih besar untuk waktu tunggu yang lebih lama [15].

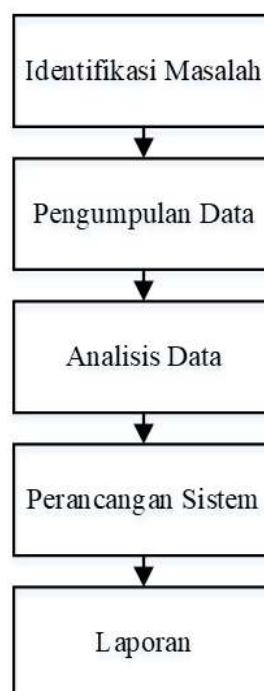
4. Metodologi Penelitian

4.1. Alat Penelitian

Dalam penelitian ini alat yang digunakan adalah tools WEKA dan Bahasa Pemrogram Python yang terdapat pada Google Colab yang bertujuan untuk menguji data pada sistem yang dibangun terkait klasifikasi penyakit *hiperkolesterol* sehingga dapat membantu untuk memodelkan sistem yang dibangun. Weka adalah perangkat lunak open-source yang digunakan dalam bidang data mining dan machine learning. Weka menawarkan berbagai macam alat dan fungsi yang sangat berguna untuk mengelola, menganalisis, dan memodelkan data, sedangkan Python adalah bahasa pemrograman yang sangat populer dalam analisis data dan pengembangan perangkat lunak. Google Colab (*Colaboratory*) adalah lingkungan pengembangan yang berbasis cloud yang memungkinkan Anda untuk menjalankan kode *Python di browser web* tanpa perlu menginstal *Python* atau paket-paket terkait di komputer.

4.2. Tahapan Penelitian

Sebelum melakukan penelitian ini yang pertama dilakukan adalah identifikasi masalah mengumpulkan permasalahan yang ditemukan dan disatukan dalam suatu pertanyaan penelitian. Selanjutnya pertanyaan penelitian ini digunakan sebagai pedoman, penentuan arah atau fokus dari penelitian, kemudian dilanjutkan dengan pengumpulan data melalui studi literatur dan penelitian-penelitian terdahulu yang mendukung proses penelitian ini. Berikut alur penelitian seperti pada gambar 1 di bawah ini:



Gambar 1. Flowchart Tahapan Penelitian

4.3. Identifikasi Masalah

Menurut penelitian awal, hiperkolesterolemia disebabkan oleh kombinasi variabel antara lain Jenis Kelamin, Usia, BMI (*Boddy Mass Index*) dan Status Merokok, Diagnostik penilaian berdasarkan kriteria dimaksud.

4.4. Pengumpulan Data

Penentuan sejumlah objek tertentu sesuai dengan tujuan penelitian untuk fokus pada item yang akan dibahas. Data yang digunakan berasal dari kuesioner yang dikumpulkan sebanyak 30 responden yang terkena penyakit hiperkolesterol.

4.5. Analisis Data

Berikut merupakan data-data pasien penderita hiperkolesterol yang telah diklasifikasikan berdasarkan beberapa atribut yang sudah ditetapkan. Atribut-atribut tersebut diantaranya adalah jenis kelamin, usia, BMI (Body Mass Index) dan Status merokok.

4.6. Implementasi

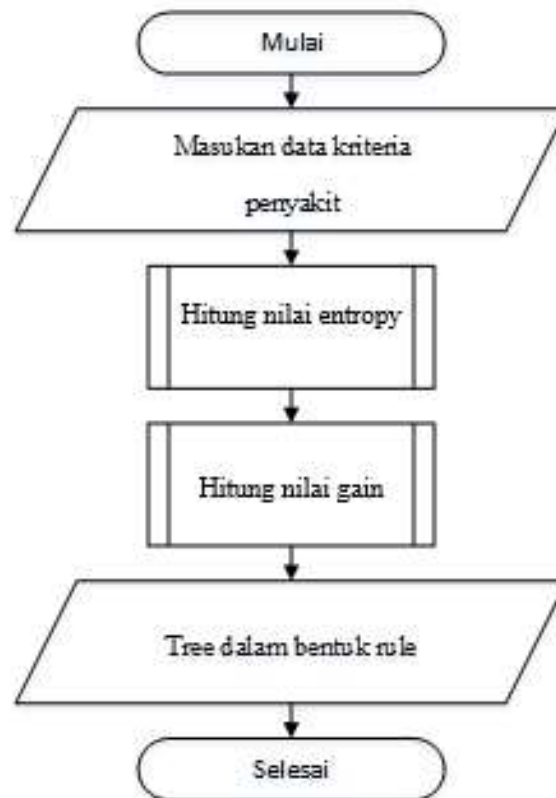
Pada bab implementasi sistem membahas mengenai implementasi dari perancangan sebelumnya dibuat. Implementasi dari sistem ini berupa spesifikasi dari perangkat keras dan perangkat lunak, batasan dan implementasi, implementasi algoritma C4.5 dan implementasi antarmuka sistem. Spesifikasi sistem terdiri dari spesifikasi perangkat keras dan spesifikasi perangkat lunak yang digunakan dalam membangun sistem. Sedangkan implementasi antarmuka terdiri dari halaman-halaman dari sistem yang telah dibuat.

4.7. Pengujian

Pada tahap pengujian hasil akan dilakukan pengujian data dengan menggunakan *algoritma decision tree* C4.5 dan menggunakan software WEKA sengaja digunakan karena mendukung proses pengklasifikasikan objek dan mudah diimplementasikan oleh pengguna secara langsung. Pengguna dapat memuat dataset, memilih algoritma untuk klasifikasi, dan kemudian mendapatkan beberapa plot data yang menunjukkan akurasi dan tingkat kesalahan hasil dari proses klasifikasi.

4.8. Perancangan sistem

Arsitektur perancangan sistem yang akan dibangun seperti gambar 2 dibawah ini yang mana sistem akan bekerja dengan memasuka data kriteria pasien kemudian data diproses dengan metode *decision tree* C4.5. Langkah pertama yaitu dengan menghitung nilai *entropy*, kemudian akan dihitung nilai *gain* dan didapatkan *tree* dalam bentuk *rule*. Keluaran yang dihasilkan sistem berupa jenis penyakit yang diderita pasien. Berikut adalah tahapan perancangan Sistem yang dibangun dengan langkah-langkah sebagai berikut:



Gambar 2. Flowchartt Perancangan Sistem

Perancangan sistem ini adalah tahap kunci dalam pengembangan berbagai jenis aplikasi dan sistem. Melalui Flowchartt perancangan sistem, kita dapat dengan jelas memvisualisasikan bagaimana komponen-komponen sistem saling berinteraksi dan berfungsi bersama untuk mencapai tujuan yang diinginkan. Dalam artikel ini, kami akan menjelaskan langkah-langkah perancangan sistem serta manfaatnya dalam pengembangan aplikasi dan sistem yang efisien dan efektif.

4.9. Alat dan Bahan Penelitian

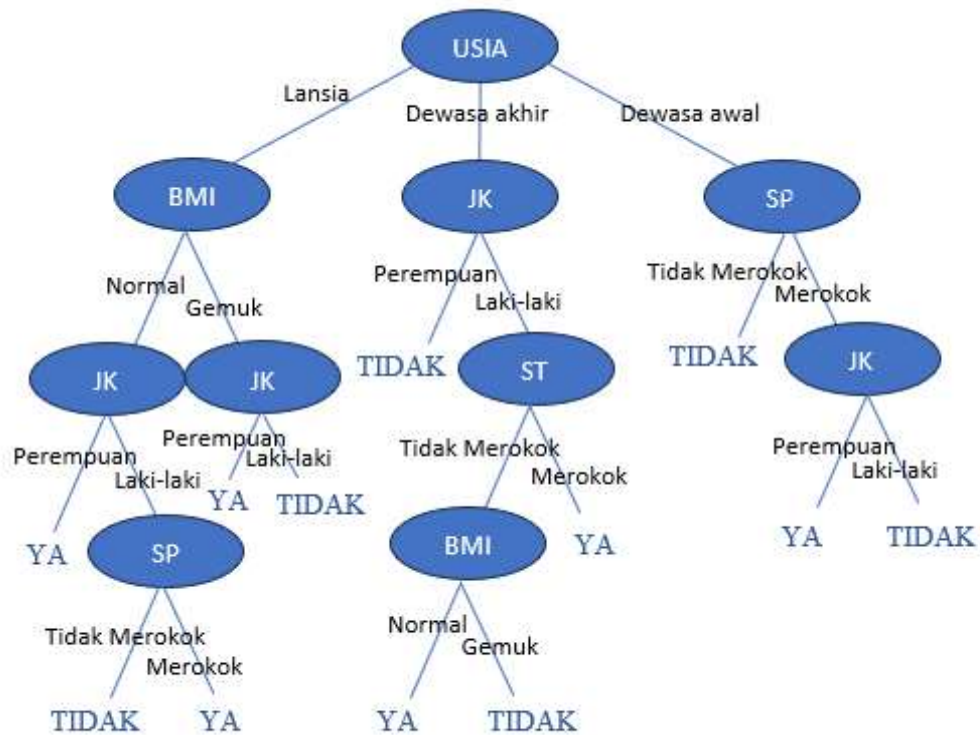
Komputer/Laptop (disarankan spesifikasi terbaru), Web Browser, Tools WEKA, Ms. Excel

4.10. Hasil and Pembahasan

Penelitian ini bersifat kuantitatif dimana data yang dihasilkan akan berbentuk angka. Dari data yang didapat dilakukan analisis dengan menggunakan software excel. Data dikumpulkan dengan kuesioner sebanyak 30 responden yang terkena penyakit hiperkolesterol. Penyebaran kuesioner ini dilakukan secara tertutup dengan menggunakan *skala likert* 1-5. Kuesioner yang dibuat dengan variabel yang diteliti memiliki rata 4-3 item pertanyaan. Responden yang dikategorikan berdasarkan beberapa kelompok berdasarkan umur, jenis kelamin, gaya hidup, kesehatan, dan pengetahuan dan kesadaran.

4.11. Pemodelan Algoritma Decision Tree C4.5

Pemodelan algoritma *Decision Tree C4.5* terdiri dari 4 atribut yang merupakan atribut penyakit hiperkolesterol dan class yang merupakan hasil akhir prediksi. Model dari algoritma *Decision Tree C4.5* yaitu berupa pohon keputusan, agar lebih mudah dalam membuat pohon keputusan, langkah pertama adalah menghitung jumlah class yang berpotensi terkena penyakit *hiperkolesterol* dari masing-masing class berdasarkan atribut yang telah ditentukan dengan menggunakan data latih.



Gambar 3. Klasifikasi Penyakit Hiperkolesterol Menggunakan Algoritma Decision Tree C4.5

Penyakit *hiperkolesterol* merupakan masalah kesehatan yang semakin mendapat perhatian karena potensi risiko yang ditimbulkannya. Dalam upaya untuk mengatasi masalah ini, penelitian ilmiah terus berkembang, salah satunya dengan mengimplementasikan teknologi seperti algoritma *Decision Tree C4.5*. Artikel ini akan membahas langkah-langkah dan hasil dari penelitian klasifikasi penyakit hiperkolesterol menggunakan algoritma *Decision Tree C4.5*, dengan harapan dapat memberikan kontribusi yang berarti dalam diagnosis dan pengelolaan penyakit ini.

Berdasarkan pohon keputusan pada gambar 3 dapat dilihat bahwa pengaruh yang dominan terhadap penyakit hiperkolesterol adalah usia, karena usia dijadikan sebagai akar atau root dari pohon keputusan tersebut yang dapat menjadi sebuah akar pada pohon keputusan adalah atribut yang memiliki nilai gain ratio tertinggi dibandingkan dengan atribut yang lainnya. Sedangkan pertimbangan terakhir untuk perhitungan rumus algoritma C4.5 yang berpengaruh pada penyakit hiperkolesterol adalah jenis kelamin karena pada node terakhir. Gambar 3. menunjukkan deskripsi lengkap tentang pohon keputusan (*decision tree*) yang terbentuk dari 20 data penyakit *hiperkolesterol* yang berasal dari berbagai Usia yaitu Usia lansia, Usia dewasa akhir dan Usia dewasa awal. Berdasarkan hasil pengujian yang telah dilakukan pada data penyakit hiperkolesterol telah diperoleh klasifikasi penyakit hiperkolesterol. Dan hasil pengujian tersebut dapat digunakan sebagai informasi strategis yang dapat digunakan sebagai informasi strategis yang dapat dijadikan sebagai sebuah pengetahuan.

4.12. Rule yang terbentuk

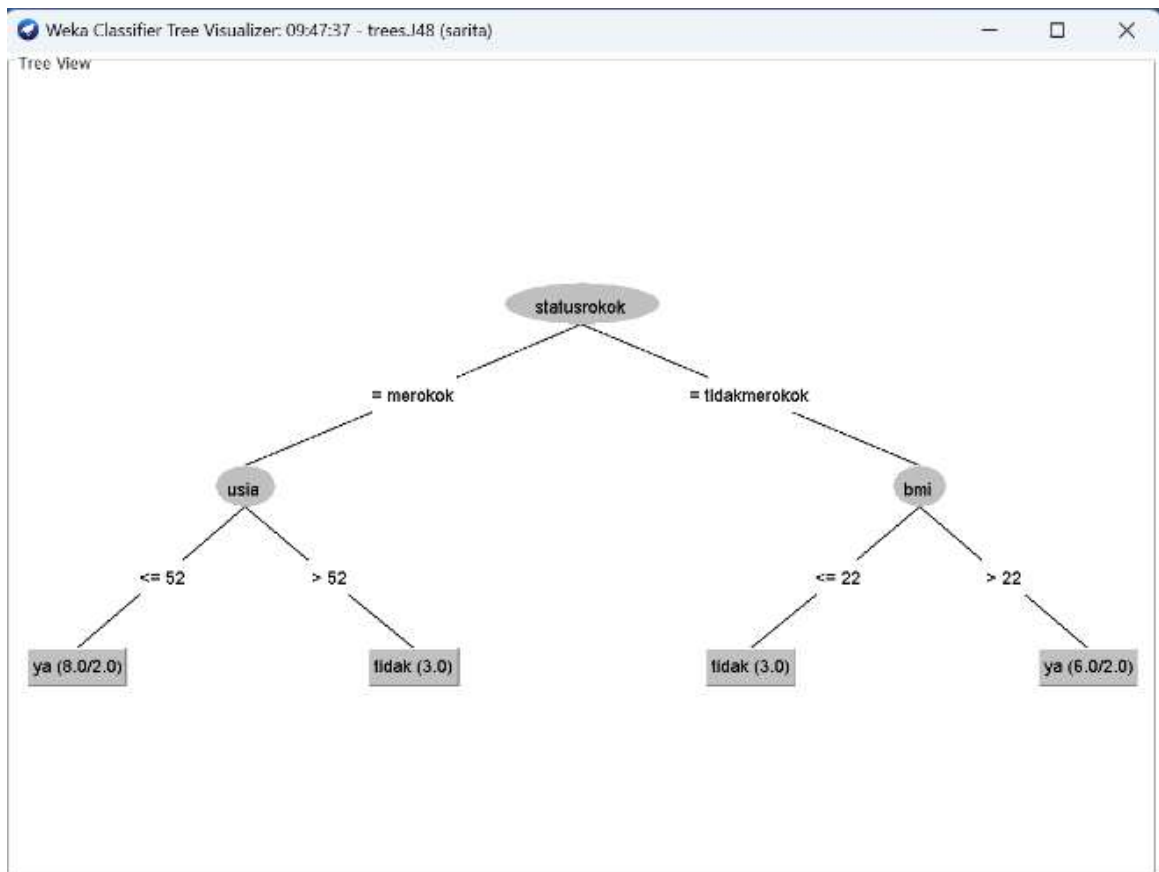
Berikut ini rules yang dihasilkan dari pohon keputusan tersebut:

- IF atribut Usia=lansia AND atribut BMI=normal AND atribut Jenis kelamin= perempuan THEN kategori penyakit hiperkolesterol.
- IF atribut Usia=lansia AND atribut BMI=normal AND atribut Jenis kelamin=laki-laki AND atribut Status perokok=tidak merokok THEN tidak kategori penyakit hiperkolesterol.
- IF atribut Usia=lansia AND atribut BMI=normal AND atribut Jenis kelamin=laki-laki AND atribut Status perokok= merokok THEN kategori penyakit hiperkolesterol.

- d. IF atribut Usia=lansia AND atribut BMI= gemuk AND atribut Jenis kelamin=perempuan THEN kategori penyakit hiperkolesterol.
- e. IF atribut Usia=lansia AND atribut BMI= gemuk AND atribut Jenis kelamin=laki-laki THEN tidak kategori penyakit hiperkolesterol.
- f. IF atribut Usia=dewasa akhir AND atribut Jenis kelamin=perempuan THEN tidak kategori penyakit hiperkolesterol.
- g. IF atribut Usia=dewasa akhir AND atribut Jenis kelamin= laki-laki AND atribut Status perokok= tidak merokok AND atribut BMI=normal THEN kategori penyakit hiperkolesterol.
- h. IF atribut Usia=dewasa akhir AND atribut Jenis kelamin= laki-laki AND atribut Status perokok= tidak merokok AND atribut BMI=gemuk THEN tidak kategori penyakit hiperkolesterol.
- i. IF atribut Usia=dewasa akhir AND atribut Jenis kelamin= laki-laki AND atribut Status perokok= merokok THEN kategori penyakit hiperkolesterol.
- j. IF atribut Usia=dewasa awal AND atribut Statut perokok=tidak merokok THEN tidak kategori penyakit hiperkolesterol.
- k. IF atribut Usia=dewasa awal AND atribut Statut perokok=merokok AND atribut Jenis kelamin=perempuan THEN kategori penyakit hiperkolesterol.
- l. IF atribut Usia=dewasa awal AND atribut Statut perokok=merokok AND atribut Jenis kelamin=laki-lakin THEN tidak kategori penyakit hiperkolesterol.

4.13.Penerapan Algoritma C4.5 pada WEKA

Hasil pohon keputusan yang dihasilkan oleh perangkat lunak WEKA merupakan inti dari analisis data dan pengambilan keputusan berbasis data. Pohon keputusan adalah alat yang sangat efektif dalam mengungkap pola dan hubungan dalam data yang kompleks. Dalam tulisan ini, kami akan menjelajahi hasil dari penggunaan perangkat lunak WEKA untuk membangun pohon keputusan, serta bagaimana hasil tersebut dapat membantu pemahaman dan pengambilan keputusan yang lebih baik berikut adalah hasil dari penerapan *algoritma decision tree* menggunakan aplikasi WEKA.



Gambar 4. Hasil Pohon Keputusan Pada Software WEKA

Gambar di atas merupakan hasil pengujian menggunakan Software Weka 3.8. Dapat dilihat bahwa pengujian software Weka 3.8 pada kriteria “Status Perokok” yang menjadi akar pertama dari pohon keputusan.

5 Kesimpulan

Dalam rangkaian penelitian ini, implementasi Algoritma C4.5 membawa kontribusi positif yang signifikan terhadap klasifikasi penyakit hiperkolesterol. Hasil penelitian ini memberikan pemahaman mendalam tentang bagaimana Algoritma C4.5 dapat memainkan peran kunci dalam mengatasi tantangan kompleksitas hiperkolesterol, meningkatkan akurasi diagnosis, dan memberikan dasar yang lebih kuat untuk pengambilan keputusan klinis. Secara khusus, penelitian ini menunjukkan bahwa Algoritma C4.5 mampu mengidentifikasi pola-pola kompleks dalam dataset hiperkolesterol. Proses pembentukan pohon keputusan oleh Algoritma ini memungkinkan penangkapan fitur-fitur kritis yang mendefinisikan kondisi hiperkolesterol. Keberhasilan ini memberikan bukti signifikan bahwa Algoritma C4.5 dapat menjadi instrumen yang efektif dalam menganalisis data kompleks yang terkait dengan penyakit ini. Selain itu, hasil penelitian menunjukkan bahwa implementasi C4.5 mampu memberikan hasil yang lebih personal dalam konteks perawatan hiperkolesterol. Dengan mempertimbangkan faktor risiko individual seperti riwayat keluarga dan gaya hidup, Algoritma ini menghasilkan klasifikasi yang lebih relevan dengan karakteristik unik pasien. Hal ini memiliki dampak positif pada pengelolaan penyakit, karena memungkinkan penyedia layanan kesehatan untuk menyusun rencana perawatan yang lebih adaptif dan responsif terhadap kebutuhan spesifik pasien.

Prediksi risiko kardiovaskular yang diperoleh dari model C4.5 menambah dimensi proaktif pada manajemen penyakit. Hasilnya memberikan pemahaman yang lebih dalam tentang risiko pasien secara individual, memungkinkan penyedia layanan kesehatan untuk mengambil langkah-langkah pencegahan yang lebih tepat. Dengan adanya informasi ini, dapat dihasilkan rencana intervensi yang lebih terukur dan spesifik, menciptakan peluang untuk mencegah perkembangan penyakit ke tingkat yang lebih serius.

Kemudahan interpretasi model C4.5 menjadi nilai tambah yang signifikan. Dalam konteks klinis, di mana kejelasan dan pemahaman model sangat penting, Algoritma ini menawarkan solusi yang bersahabat dengan praktisi kesehatan. Pohon keputusan yang dihasilkan secara intuitif dapat digunakan oleh berbagai latar belakang, memudahkan adopsi dan implementasi dalam lingkungan perawatan kesehatan. Dalam keseluruhan, penelitian ini memberikan kontribusi penting pada pemahaman kita tentang potensi dan batasan implementasi Algoritma C4.5 dalam klasifikasi penyakit hiperkolesterol. Hasilnya memberikan dasar yang kuat untuk mengakui bahwa C4.5 bukan hanya sebuah alat analisis yang andal tetapi juga alat yang dapat memenuhi kebutuhan spesifik dan variabilitas individual yang terkait dengan hiperkolesterol. Namun, seperti halnya dengan setiap metode analisis, ada pertimbangan yang perlu diakui. Meskipun C4.5 menunjukkan keberhasilan dalam menangani kompleksitas hiperkolesterol, potensi overfitting dan sensitivitas terhadap noise tetap menjadi tantangan. Selain itu, penting untuk diperhatikan bahwa penelitian ini membutuhkan dataset yang cukup besar untuk memberikan hasil yang lebih generalisasi. Oleh karena itu, pengembangan lebih lanjut pada skala yang lebih luas dan dengan data yang lebih bervariasi dapat memberikan wawasan tambahan tentang keefektifan C4.5 dalam situasi klinis yang lebih beragam.

Implementasi Algoritma C4.5 dalam klasifikasi penyakit hiperkolesterol dapat dianggap sebagai langkah maju yang berarti dalam penelitian ini. Kemampuannya untuk mengatasi kompleksitas dan memberikan hasil yang personal memberikan landasan yang kuat untuk penerapan lebih lanjut dalam lingkungan kesehatan. Sebagai alat yang dapat diakses dan dapat diinterpretasikan, C4.5 menawarkan potensi untuk memberikan manfaat yang nyata bagi praktisi kesehatan dan, lebih penting lagi, untuk meningkatkan perawatan dan manajemen pasien dengan hiperkolesterol.

References

- [1] A. P. Anggrayni, "Asuhan Keperawatan Gerontik Dengan Nyeri Akut Pada Klien Gout Arthritis di Dusun Rowoglagah Desa Sidomulyo Kecamatan Deket Kabupaten Lamongan," 2020.
- [2] R. Muttaqien, "Implementation of Data Mining Using C4.5 Algorithm for Predicting Customer Loyalty of PT. Pegadaian (Persero) Pati Area Office," *International Journal*, vol. 02, no. 03, 2021.
- [3] D. E. Yanti, L. Framesti, and A. Desiani, "PERBANDINGAN ALGORITMA C4.5 DAN SVM DALAM KLASIFIKASI PENYAKIT ANEMIA," *JIP*, vol. 9, no. 4, pp. 427–434, Aug. 2023, doi: 10.33795/jip.v9i4.1381.
- [4] D. Chatterjee and S. Chandran, "Prediction and Classification of Heart Disease using AML and Power BI.," in *Proceedings of the 11th International Joint Conference on Computational Intelligence*, Vienna, Austria: SCITEPRESS - Science and Technology Publications, 2019, pp. 508–515. doi: 10.5220/0008381505080515.
- [5] D. Devina, A. A. Supianto, and W. Purnomo, "Aplikasi Data Mining Menggunakan Algoritma C4.5 untuk Memprediksi Ketepatan Lulus Mahasiswa Berdasarkan Faktor Demografi".
- [6] S. Saefudin and S. Lestari, "Sistem Pendukung Keputusan Untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Metode Algoritma C4.5," *protekinfo*, vol. 2, pp. 40–43, Jan. 2017, doi: 10.30656/protekinfo.v2i0.48.
- [7] Gina Sonia and A. Indriyani, "Analisis Penerapan Data Mining Dengan Metode Algoritma C4.5 Untuk Pendataan Karyawan Tetap Di Koni Sumatera Utara," *CoSIE*, pp. 48–55, Jan. 2023, doi: 10.55537/cosie.v2i1.523.
- [8] A. N. Z. Hidayah and A. F. Rozi, "Penerapan Data Mining Dalam Menentukan Kinerja Karyawan Terbaik Dengan Menggunakan Metode Algoritma C4.5 (Studi Kasus : Universitas Mercu Buana Yogyakarta)," *ARTIFICIAL INTELLIGENCE*, no. 2, 2021.
- [9] A. A. Damanik, Z. Azhar, and A. Sapta, "Penerapan Data Mining Mengklasifikasi Pola Nasabah Menggunakan Algoritma C4.5 Pada Pegadaian Tanjungbalai," *Jurnal Teknologi dan Sistem Informasi*, vol. 2, no. 1, 2022.
- [10] K. Sumathi, S. Kannan, and K. Nagarajan, "Data Mining: Analysis of student database using Classification Techniques," *IJCA*, vol. 141, no. 8, pp. 22–27, May 2016, doi: 10.5120/ijca2016909703.
- [11] D. S. Tarlik, K. W. R. Putra, and R. Riesmiyatiningdyah, "Studi Kasus Penerapan Asuhan Keperawatan Gerontik Pada Penderita Hiperkolesterol Dengan Pendekatan Keluarga Binaan Di Desa Keboan Sikep Gedangan Kabupaten Sidoarjo," *IJoHVE*, vol. 1, no. 2, pp. 74–83, Nov. 2022, doi: 10.36720/ijohve.v1i2.451.
- [12] A. Dwinanda and N. Afriani, "Pengaruh Jus Seledri (*Apium graveolens* L.) terhadap Gambaran Mikroskopis Hepar Tikus (*Rattus norvegicus*) yang Diinduksi Diet Hiperkolesterol".
- [13] A. Ward *et al.*, "Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population," *npj Digit. Med.*, vol. 3, no. 1, p. 125, Sep. 2020, doi: 10.1038/s41746-020-00331-1.
- [14] R. G. Wardhana, G. Wang, and F. Sibuea, "Penerapan Machine Learning Dalam Prediksi Tingkat Kasus Penyakit Di Indonesia," *JOISM*, vol. 5, no. 1, pp. 40–45, Jul. 2023, doi: 10.24076/joism.2023v5i1.1136.
- [15] S. Ardabili *et al.*, "COVID-19 Outbreak Prediction with Machine Learning," *Algorithms*, vol. 13, no. 10, p. 249, Oct. 2020, doi: 10.3390/a13100249.