



Penerapan Decision Tree untuk Klasifikasi Penyakit Berdasarkan Data Rekam Medis

Hendri Purnomo¹, Randi Estian Pambudi², Rendi Irawan³

¹, # Fakultas Ilmu Komputer, ²IIB Darmajaya Lampung

², # Fakultas Ilmu Komputer, ²STMIK ALSIBA.Ogan Komering Ilir Sumatra Selatan

¹ hendrialie@damajaya.ac.id, ² randiestian@damajaya.ac.id, ³ ren.irawan33@gmail.com

ABSTRACT

This research discusses the application of the Decision Tree algorithm for disease classification based on medical record data. The patients' medical records are analyzed using this algorithm to identify significant patterns in disease diagnosis. The resulting model is evaluated using metrics such as accuracy, precision, and recall, showing adequate results with an accuracy of 91.55%. Visualization in the form of a decision tree highlights the key attributes influencing classification, such as chest pain type, cholesterol level, and patient age. This study emphasizes the effectiveness of the Decision Tree algorithm in supporting data-driven medical decision support systems, although careful data processing is required to address the challenges of medical record data quality.

Keywords: Decision Tree, Disease Classification, Medical Records, Precision, Recall, Decision Support Systems, Data Mining

ABSTRAK

Penelitian ini membahas penerapan algoritma **Decision Tree** untuk klasifikasi penyakit berdasarkan data rekam medis. Data rekam medis pasien dianalisis menggunakan algoritma ini untuk mengidentifikasi pola yang signifikan dalam diagnosis penyakit. Model yang dihasilkan dievaluasi menggunakan metrik seperti akurasi, presisi, dan recall, menunjukkan hasil yang memadai dengan akurasi mencapai 91.55%. Visualisasi dalam bentuk pohon keputusan menunjukkan atribut utama yang memengaruhi klasifikasi, seperti tipe nyeri dada, kadar kolesterol, dan usia pasien. Penelitian ini menyoroti efektivitas algoritma *Decision Tree* dalam mendukung sistem pendukung keputusan medis berbasis data, meskipun memerlukan pengolahan data yang cermat untuk mengatasi tantangan kualitas data rekam medis.

Kata Kunci: Decision Tree, Klasifikasi Penyakit, Data Rekam Medis, Presisi, Recall, Sistem Pendukung Keputusan, Data Mining

I. PENDAHULUAN

Perkembangan teknologi informasi telah membawa dampak signifikan terhadap berbagai bidang, termasuk kesehatan. Salah satu inovasi yang sangat bermanfaat dalam mendukung layanan kesehatan adalah penerapan teknologi berbasis data untuk mendukung diagnosis dan pengambilan keputusan medis. Rekam medis, yang berisi data kesehatan pasien seperti riwayat penyakit, hasil pemeriksaan laboratorium, dan data vital lainnya, merupakan salah satu sumber informasi penting yang dapat diolah untuk mendapatkan wawasan yang lebih mendalam tentang kondisi pasien. Dalam konteks ini, teknik data mining menjadi alat yang sangat potensial untuk menganalisis data rekam medis guna mendukung proses klasifikasi dan prediksi penyakit[1].

Decision Tree merupakan salah satu algoritma data mining yang populer digunakan dalam masalah klasifikasi karena kemampuannya untuk menghasilkan model yang mudah dipahami dan diinterpretasikan. Algoritma ini bekerja dengan membangun struktur pohon berdasarkan pembagian data menjadi subset yang semakin homogen sesuai dengan atribut yang paling signifikan. Dalam kasus diagnosis penyakit, Decision Tree dapat membantu mengidentifikasi pola-pola tertentu dari data pasien untuk memprediksi kemungkinan keberadaan penyakit tertentu, seperti penyakit jantung, diabetes, atau kanker[2].

Penerapan Decision Tree dalam klasifikasi penyakit memiliki beberapa keunggulan, antara lain kemampuan menangani data dengan berbagai jenis atribut (numerik dan kategorikal), serta hasil model yang dapat divisualisasikan dalam bentuk pohon keputusan, sehingga memudahkan interpretasi oleh praktisi kesehatan. Selain itu, algoritma ini relatif efisien dalam hal waktu komputasi, menjadikannya pilihan yang layak untuk diterapkan dalam sistem pendukung keputusan medis[3].

Namun, penggunaan Decision Tree dalam klasifikasi penyakit tidak terlepas dari tantangan. Salah satu tantangan utama adalah kualitas data rekam medis itu sendiri, yang

sering kali mengandung data yang hilang atau tidak konsisten. Oleh karena itu, preprocessing data menjadi langkah penting dalam memastikan bahwa model yang dihasilkan memiliki tingkat akurasi yang tinggi. Selain itu, pemilihan atribut yang relevan juga menjadi faktor kunci dalam meningkatkan kinerja model[4].

Penelitian ini bertujuan untuk menerapkan algoritma Decision Tree dalam klasifikasi penyakit berdasarkan data rekam medis. Studi ini akan mengeksplorasi bagaimana algoritma Decision Tree dapat digunakan untuk mengidentifikasi pola-pola yang signifikan dalam data pasien dan menghasilkan model yang akurat untuk mendukung diagnosis awal penyakit. Dengan pendekatan ini, diharapkan hasil penelitian dapat memberikan kontribusi dalam pengembangan sistem pendukung keputusan berbasis data di bidang kesehatan[5]

II. TINJAUAN PUSTAKA

2.1. Penyakit Jantung

Penyakit jantung merupakan penyakit yang sangat berbahaya karena salah satu penyakit yang mematikan, terutama penyakit jantung koroner yang jumlah penderitanya sangat banyak dibandingkan dengan jenis penyakit jantung lainnya. Penyakit jantung koroner terjadi akibat adanya penyumbatan pembuluh arteri oleh plak yang menghambat suplai oksigen dan nutrisi ke jantung. [6]Kemunculan plak yang berupa timbunan lemak atau kalsium melalui proses secara bertahap. Biasanya, diawali dengan kekakuan pembuluh darah atau biasa di sebut dengan aterosklerosis, kemudian penyempitan pembuluh darah,

2.1.2. Gejala Penyakit Jantung

Penyakit jantung koroner atau arteri koroner dianggap sangat berbahaya karena dapat menimbulkan serangan jantung mendadak yang berujung kematian. Serangan jantung terjadi akibat terhambatnya aliran darah menuju jantung sehingga suplai oksigen dan nutrisi di otot jantung dan jaringan disekitar jantung berkurang.

Tidak seperti otot tubuh lainnya, otot jantung tidak memiliki kemampuan beregenerasi. Apabila terdapat saja kerusakan maka akan berakibat fatal bagi tubuh. Semakin lama serangan jantung terjadi semakin banyak pula kerusakan pada jantung. Karena itu penting bagi kita untuk mengenali gejala-gejala dari penyakit jantung sehingga dapat memberikan pertolongan dengan segera. Gejala pnuakit jantung koroner secaa umum tidak dikenali oleh orang awam. Mereka terkadang menyepelekan dan menganggapnya wajar. Penderita baru menyadari bahwa dirinya terkena penyakit jantung koroner ketika kondisinya sudah parah. Bahkan, tak jarang dari mereka pada akhirnya harus meregang nyawa karena keterlambatan penanganan [7].

2.1.3. Tipe Penyakit Jantung

Beberapa tipe dari penyakit jantung antara lain[8]:

1. Gagal jantung

Gagal jantung disebut juga gagal jantung kongestif, yang berarti jantung tidak memompa darah sebagaimana mestinya. Gagal jantung tidak berarti jantung berhenti bekerja, tetapi kebutuhan tubuh akan darah dan oksigen tidak terpenuhi.

2. Arrhythmia

Aritmia adalah detak jantung yang tidak normal, jantung berdetak terlalu lambat, terlalu cepat, atau tidak teratur. Ketika denyut jantung kurang dari 60 denyut permenit hal ini dinamakan bradiakardia. Takikardia adalah ketika denyut jantung lebih dari 100 denyut permenit. Aritmia dapat mempengaruhi kinerja jantung.

3. Masalah Katup Jantung

Beberapa jenis masalah katup jantung diantaranya adalah stenosis, regurgitasi, dan prolaps katup mitral. Stenosis adalah keadaan dimana katup jantung tidak cukup terbuka untuk memungkinkan darah mengalir sebagaimana mestinya. Regurgitasi terjadi ketika katup jantung tidak menutup dengan benar dan memungkinkan darah bocor. Prolaps katup mitral terjadi ketika katup leaflet menonjol atau prolaps kembali ke ruang atas.

2.1.4. Data Mining

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan-perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada prediksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu[9]Kaka's data mining meramalkan tren dan sifat-sifat perilaku bisnis yang sangat berguna untuk mendukung pengambilan keputusan penting. Analisis yang diotomatisasi yang dilakukan oleh data mining melebihi yang dilakukan oleh sistem pendukung keputusan tradisional yang sudah banyak digunakan. Secara khusus, koleksi metode yang dikenal sebagai 'data mining' menawarkan metodologi dan solusi teknis untuk mengatasi analisis data medis dan konstruksi prediksi model[3].Berdasarkan tugas dan tujuan analisis, proses data mining dapat dibagi menjadi dua kategori utama, Tergantung pada adanya target variabel dan metode belajar (learning) yaitu antara proses belajar yang diawasi (supervised) dan tanpa pengawasan (unsupervised) [10]

2.1.5. Klasifikasi.

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri dapat berupa aturan jika-maka (if-then), berupa pohon keputusan (decision tree), jaringan saraf tiruan (neural network).

Klasifikasi adalah urutan yang sangat penting dalam data komunitas pertambangan. Klasifikasi adalah salah satu prediksi teknik data mining yang membuat prediksi tentang data nilai menggunakan hasil yang diketahui yang ditemukan dari kumpulan data yang berbeda. Masalah akurasi dari banyak algoritma klasifikasi adalah diketahui mengalami penurunan informasi saat dihadapi dengan data yang tidak seimbang, misalnya ketika distribusi sampel lintas kelas sangat miring[11]

Dalam klasifikasi, ada variabel kategoris target, seperti braket pendapatan, yang, misalnya, dapat dipartisi menjadi tiga kelas atau kategori: berpenghasilan tinggi, menengah pendapatan, dan pendapatan rendah. Model

data mining memeriksa satu set besar catatan, masing-masing catatan yang berisi informasi tentang variabel target serta satu set input atau prediktor variable. Contoh tugas klasifikasi dalam bisnis dan penelitian meliputi:

- Menentukan apakah transaksi kartu kredit tertentu adalah penipuan
- Menempatkan siswa baru pada jalur tertentu yang berkaitan dengan kebutuhan khusus
- Menilai apakah aplikasi hipotek adalah risiko kredit yang baik atau buruk
- Mendiagnosis apakah ada penyakit tertentu
- Menentukan apakah surat wasiat ditulis oleh almarhum yang sebenarnya, atau curang oleh orang lain
- Mengidentifikasi apakah perilaku keuangan atau pribadi tertentu menunjukkan kemungkinan ancaman teroris

Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Jaringan Saraf Tiruan [1]

2.1.6. Algoritma C4.5

Algoritma C4.5 adalah sebuah algoritma yang berfungsi untuk membangun decision tree (pohon keputusan). Algoritma C4.5 dan pohon keputusan merupakan dua model yang tidak terpisahkan. Algoritma C4.5 adalah salah satu dari algoritma klasifikasi yang kuat dan banyak digunakan atau di implementasikan untuk pengklasifikasian dalam berbagai hal. Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan (1996) sebagai versi perbaikan dari algoritma Iterative Dichotomiser 3 (ID3). Serangkaian perbaikan dilakukan pada algoritma ID3 mencapai puncaknya dengan menghasilkan sebuah sistem praktis dan simple yang berpengaruh untuk pembentukan pohon keputusan. Perbaikan tersebut meliputi metode untuk menangani data kontinew, mengatasi missing data, dan melakukan pemangkasan pohon [9]

2.1.7. Confusion Matrix

Pengujian dengan Confusion Matrix Pada tahap ini pengujian model penelitian dilakukan dengan metode Confusion Matrix yang

mempresentasikan hasil evaluasi model dengan menggunakan tabel matrik, Jika dataset terdiri dari 2 kelas, kelas pertama dianggap positif dan kelas kedua dianggap negatif. Evaluasi menggunakan confusion matrix menghasilkan nilai Akurasi, Precision, Recall, serta F-Measure[12]. Akurasi dalam klasifikasi merupakan presentasi ketepatan record data diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Precision merupakan proposisi yang diprediksi positif yang juga positif benar pada data sebenarnya. Recall merupakan proporsi kasus positif yang sebenarnya diprediksi positif secara benar. True Positive (TP) merupakan jumlah record positif dalam dataset yang diklasifikasikan positif. True Negative (TN) merupakan jumlah record negatif dalam dataset yang diklasifikasikan positif[13].

False Positive (FP) merupakan jumlah record negatif dalam dataset yang diklasifikasikan positif. False Negative (FN) merupakan jumlah record positif dalam dataset yang diklasifikasikan negatif. Berikut adalah persamaan model Confusion Matrix. [14]

Accuracy adalah jumlah perbandingan data yang benar dengan jumlah keseluruhan data.

$$\text{Akurasi} = \frac{(TN + TP)}{(TN + FN + FP + TP)} \dots (2.4)$$

Precision digunakan untuk mengukur seberapa besar proporsi dari kelas data positif yang berhasil diprediksi dengan benar dari keseluruhan hasil prediksi kelas positif.

$$\text{precision} = \frac{TP}{(TP + FP)} \dots (2.5)$$

Recall digunakan untuk menunjukkan presentase kelas data positif yang berhasil diprediksi benar dari keseluruhan data kelas positif.

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

III. METODOLOGI

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan teknik data mining untuk menganalisis data rekam medis dan membangun model klasifikasi penyakit

menggunakan algoritma Decision Tree. Metodologi penelitian dirancang untuk memastikan hasil yang valid dan dapat diandalkan, dengan tahapan sebagai berikut:

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data rekam medis pasien, yang berisi atribut-atribut relevan seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, kadar gula darah, dan diagnosis penyakit. Dataset diperoleh dari sumber data terbuka seperti <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> atau dataset kesehatan lainnya yang relevan. Jika diperlukan, data akan disimulasikan dengan mengikuti pola umum yang terdapat pada data kesehatan.

2. Preprocessing Data

Preprocessing data dilakukan untuk memastikan kualitas data yang optimal sebelum dianalisis. Tahapan preprocessing meliputi[15]:

1. Pembersihan Data (Data Cleaning):
 - Mengisi nilai yang hilang (missing values) menggunakan metode statistik, seperti mean atau median.
 - Menghapus data duplikat atau outlier yang tidak relevan.
2. Transformasi Data:
 - Normalisasi data numerik agar atribut memiliki skala yang seragam.
 - Encoding untuk atribut kategorikal, seperti jenis kelamin (0 = wanita, 1 = pria).
3. Pembagian Data:
 - Dataset dibagi menjadi data pelatihan (80%) dan data pengujian (20%) menggunakan teknik stratified sampling untuk memastikan distribusi kelas tetap seimbang.

3. Penerapan Algoritma Decision Tree

1. Pemilihan Algoritma:
 - Algoritma Decision Tree, seperti CART (Classification and Regression Tree), dipilih karena kemampuannya dalam

menghasilkan model yang interpretatif dan efektif untuk klasifikasi.

2. Pelatihan Model:

- Model Decision Tree dilatih menggunakan data pelatihan. Proses pelatihan melibatkan pembuatan pohon keputusan berdasarkan atribut yang memiliki tingkat kepentingan (information gain atau Gini index) tertinggi.

3. Pengujian Model:

- Model diuji menggunakan data pengujian untuk mengevaluasi performa berdasarkan metrik akurasi, precision, recall, dan F1-score.

4. Evaluasi Model

Evaluasi model dilakukan untuk menilai kemampuan algoritma dalam mengklasifikasikan penyakit secara akurat. Metrik evaluasi yang digunakan meliputi:

- Akurasi: Proporsi prediksi yang benar terhadap total data.
- Precision: Proporsi prediksi positif yang benar terhadap seluruh prediksi positif.
- Recall: Proporsi data positif yang berhasil diklasifikasikan dengan benar.
- Confusion Matrix: Matriks untuk menunjukkan distribusi prediksi model.

5. Visualisasi Model

Hasil model Decision Tree akan divisualisasikan dalam bentuk pohon Keputusan. Visualisasi ini bertujuan untuk memudahkan interpretasi pola-pola yang diidentifikasi dalam data.

6. Analisis dan Interpretasi Hasil

Hasil penelitian dianalisis untuk mengidentifikasi pola-pola signifikan yang ditemukan oleh model. Selain itu, penelitian ini juga membahas kelebihan dan keterbatasan algoritma Decision Tree dalam konteks klasifikasi penyakit berdasarkan data rekam medis.

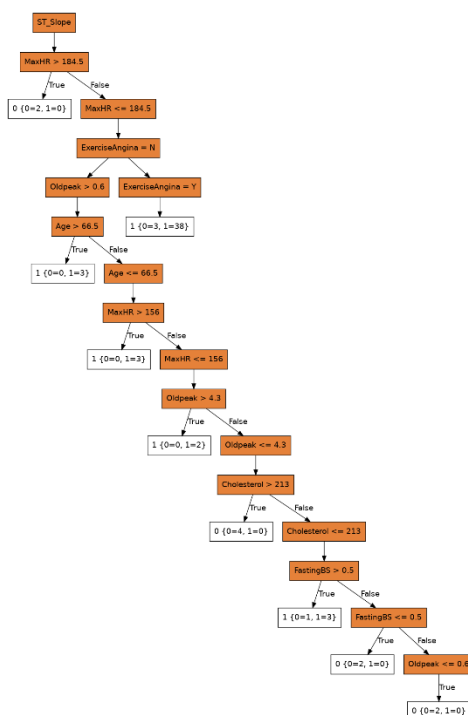
7. Kesimpulan dan Rekomendasi

Berdasarkan hasil analisis, penelitian ini akan memberikan kesimpulan terkait efektivitas algoritma Decision Tree dalam klasifikasi penyakit. Selain itu, rekomendasi akan diberikan untuk pengembangan sistem pendukung keputusan medis berbasis data di masa depan.

IV. HASIL DAN PEMBAHASAN

Model Decision Tree yang dihasilkan dari proses pelatihan dan pengujian menggunakan dataset rekam medis menunjukkan hasil berikut:

Visualisasi Pohon Keputusan: Model menghasilkan pohon keputusan dengan atribut utama yang memengaruhi klasifikasi penyakit, seperti jenis nyeri dada (Chest Pain Type), kadar kolesterol (Cholesterol), dan usia pasien (Age). Struktur pohon menunjukkan bahwa atribut Chest Pain Type memiliki pengaruh terbesar, diikuti oleh atribut lainnya seperti Resting Blood Pressure dan Fasting Blood Sugar.



Gambar 1 Pohon Keputusan

Pohon keputusan yang dibuat memiliki tiga cabang utama berdasarkan atribut *ST_Slope*: *Down*, *Flat*, dan *Up*. Setiap cabang membagi data lebih lanjut menggunakan berbagai atribut seperti *MaxHR*, *ExerciseAngina*, *Oldpeak*, *Age*, dan *Cholesterol*, untuk menentukan output akhir berupa label biner (0 atau 1). Cabang *ST_Slope* = *Down* menunjukkan pemisahan awal berdasarkan *MaxHR*, diikuti dengan atribut seperti *Oldpeak* dan *Age*, untuk membedakan hasil. Cabang *ST_Slope* = *Flat* mengintegrasikan atribut *Sex* dan *ChestPainType* untuk pengambilan keputusan, sementara cabang *ST_Slope* = *Up* menekankan atribut seperti *Oldpeak* dan *ExerciseAngina*. Setiap simpul daun menyatakan hasil akhir dengan distribusi data (misalnya, {0=2, 1=0}) yang memberikan wawasan tentang pengaruh atribut terhadap prediksi. Pohon ini menunjukkan struktur yang jelas dalam mengidentifikasi pola berdasarkan kriteria kesehatan yang relevan.

Metrik Evaluasi Model: Berdasarkan hasil evaluasi terhadap data pengujian, performa model Decision Tree dapat dirangkum sebagai berikut:

- Akurasi: 91,55%
- Precision: 93,43%
- Recall: 91,13%

Confusion Matrix

Tabel 1 Confusion Matrix

Precision (Presisi)		Ketepatan prediksi positif. Seberapa sering prediksi positif benar-benar positif?
* pred 0	89.35%	Dari semua yang diprediksi sebagai kelas 0, 89.35% benar-benar kelas 0.
* pred 1	93.43%	Dari semua yang diprediksi sebagai kelas 1, 93.43% benar-benar kelas 1.
Recall		Kelengkapan prediksi positif. Seberapa banyak dari kelas positif yang berhasil diidentifikasi?
* true 0	92.07%	Dari semua data yang sebenarnya kelas 0, 92.07% berhasil diprediksi sebagai kelas 0.
* true 1	91.13%	Dari semua data yang sebenarnya kelas 1, 91.13% berhasil diprediksi sebagai kelas 1.

Tabel menunjukkan evaluasi performa model berdasarkan metrik *Precision* dan *Recall*, yang menggambarkan kemampuan model dalam memprediksi kelas positif dan mengidentifikasi data sesuai kelasnya. Presisi untuk kelas 0 adalah **89.35%**, yang berarti dari semua prediksi kelas 0, sebanyak 89.35% adalah benar-benar kelas 0. Presisi untuk kelas 1 mencapai **93.43%**, menunjukkan tingkat akurasi yang tinggi dalam memprediksi kelas 1. Recall untuk kelas 0 adalah **92.07%**, yang berarti model berhasil mengidentifikasi 92.07% dari semua data yang sebenarnya kelas 0. Sementara itu, *recall* untuk kelas 1 adalah **91.13%**, menunjukkan model mampu mengenali sebagian besar data yang benar-benar kelas 1. Kombinasi presisi dan recall ini mencerminkan keseimbangan performa model dalam membuat prediksi yang tepat dan lengkap untuk kedua kelas.

PEMBAHASAN

Kombinasi metrik presisi dan recall penting untuk memahami kekuatan dan kelemahan model. Dalam kasus ini, presisi yang tinggi untuk kelas 1 (**93.43%**) menunjukkan bahwa model jarang memberikan prediksi positif palsu untuk kelas tersebut, sehingga cocok digunakan dalam konteks di mana kesalahan positif palsu harus diminimalkan, seperti dalam diagnosis penyakit atau deteksi penipuan.

Namun, recall yang sedikit lebih rendah untuk kelas 1 (**91.13%**) menunjukkan bahwa ada sebagian kecil data positif (kelas 1) yang gagal teridentifikasi. Ini dapat menjadi perhatian jika konteks aplikasi lebih mengutamakan mengidentifikasi semua kasus positif, seperti dalam sistem peringatan bencana.

Sementara itu, untuk kelas 0, presisi dan recall yang masing-masing sebesar **89.35%** dan **92.07%** menunjukkan bahwa model memiliki kemampuan yang baik dalam memprediksi dan mengidentifikasi data negatif. Namun, presisi yang sedikit lebih rendah dibanding kelas 1 mengindikasikan adanya beberapa prediksi negatif palsu, yang mungkin perlu diperhatikan tergantung pada konteks penggunaannya

V. PENUTUP

Secara keseluruhan, model menunjukkan performa yang baik dengan presisi dan recall yang tinggi untuk kedua kelas, mencerminkan kemampuan prediksi yang andal. Dengan presisi lebih tinggi pada kelas 1 dan recall yang cukup seimbang di kedua kelas, model ini dapat diterapkan secara efektif di berbagai konteks, terutama jika prioritas adalah meminimalkan kesalahan prediksi positif palsu. Namun, untuk aplikasi yang sangat kritis, seperti mendeteksi semua kasus positif tanpa terlewat, diperlukan optimisasi tambahan untuk meningkatkan recall, khususnya pada kelas 1. Evaluasi tambahan dengan metrik seperti F1-score atau AUC-ROC dapat memberikan wawasan lebih lanjut untuk memastikan keseimbangan optimal antara presisi dan recall sesuai kebutuhan operasi

DAFTAR PUSTAKA

- [1] T. Retnasari and E. Rahmawati, "Diagnosa Prediksi Penyakit Jantung Dengan Model Algoritma Naïve Bayes Dan Algoritma C4.5," *Konf. Nas. Ilmu Sos. Teknol.*, pp. 7-12 Retnasari, T., Rahmawati, E. (2017). *Diagnos*, 2017.
- [2] D. K. Destiani *et al.*, "Klasifikasi Sinyal Ecgagal Jantung Menggunakan Wavelet Dan Jst Propagasi Balik Dengan Modifikasi Gradien Konjugat Polak-Ribiere Heart Failure Ecg Signal Classification Using Wavelet and Ann Backpropagation With Polak-Ribiere Conjugate Gradient," vol. 5, no. 1, pp. 1811–1818, 2018.
- [3] J. Informatika *et al.*, "Penerapan Algoritma Naïve Bayes Untuk Mengetahui Pasien Penyakit Gagal Jantung Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," vol. 2, no. September, 2022.
- [4] Y. Yuliani, "Algoritma Random Forest Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung Menggunakan Seleksi Fitur Bestfirst," *Infotek J. Inform. dan Teknol.*, vol. 5, no. 2, pp. 298–306, 2022, doi: 10.29408/jit.v5i2.5896.
- [5] W. S. Dharmawan, "I N F O R M a T I K a Dalam Prediksi Penyakit Jantung," *J. Inform. Manaj. dan Komput.*, vol. 13, no. 2, pp. 31–41, 2021.

- [6] E. Nurlia and U. Enri, “Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5,” *J. Tek. Inform. Musirawas)* Elin Nurlia, vol. 6, no. 1, p. 42, 2021.
- [7] H. M. Nawawi, J. J. Purnama, and A. B. Hikmah, “Komparasi Algoritma Neural Network Dan Naïve Bayes Untuk Memprediksi Penyakit Jantung,” *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 189–194, 2019, doi: 10.33480/pilar.v15i2.669.
- [8] D. A. M. Reza, A. M. Siregar, and Rahmat, “Penerapan Algoritma K-Nearest Neighbord Untuk Prediksi Kematian Akibat Penyakit Gagal Jantung,” *Sci. Student J. Information, Technol. Sci.*, vol. III, no. 1, pp. 105–112, 2022.
- [9] R. Wajhillah, “Particle Swarm Optimization Untuk Prediksi Penyakit Jantung,” vol. I, no. 1, pp. 26–36, 2014.
- [10] D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [11] Oskar, Fransiska Ria, Rupina, and N. P, “Pengelolaan Data Penyakit Jantung Dengan Menggunakan Metode Naive Bayes,” *J. Sains Dan Komput.*, vol. 8, no. 02, pp. 49–54, 2024, doi: 10.61179/jurnalinfact.v8i02.531.
- [12] M. Metode, K. N. Dan, and L. Regression, “Implementasi data mining untuk memprediksi penyakit jantung menggunakan metode k-nearest neighbor dan logistic regression,” vol. 5, pp. 493–501, 2022, doi: 10.37600/tekinkom.v5i2.698.
- [13] R. Fadnavis, K. Dhore, D. Gupta, J. Waghmare, and D. Kosankar, “Heart disease prediction using data mining,” *J. Phys. Conf. Ser.*, vol. 1913, no. 1, 2021, doi: 10.1088/1742-6596/1913/1/012099.
- [14] M. C. Arta, N. Anwar, Y. A. Putri, and M. Asroll, “Implementasi Prediksi Penyakit Jantung Menggunakan Data Mining Untuk Dunia Kesehatan Pre-processing,” vol. 10, no. 01, pp. 42–48, 2024.
- [15] I. S. B. Azhar and W. K. Sari, “Penerapan Data Mining Dan Teknologi Machine Learning Pada Klasifikasi Penyakit Jantung,” *JSI J. Sist. Inf.*, vol. 14, no. 1, pp. 2560–2568, 2022, doi: 10.18495/jsi.v14i1.16140.