

Literature Review: Klasifikasi Penyakit Paru-paru Menggunakan Metode *Decision Tree*

Angga Rakhmansyah¹, Perani Rosyani^{1*}

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspipetek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: 1anggafsn@gmail.com, 1*dosen00837@unpam.ac.id

(* : coressponding author)

Abstrak– Peningkatan kasus penyakit paru-paru global menghadirkan tantangan serius dalam bidang kesehatan yang memerlukan sistem deteksi awal untuk penanganan optimal. Penelitian ini mengkaji implementasi algoritma Decision Tree dalam mengklasifikasikan berbagai jenis penyakit paru-paru berdasarkan analisis komprehensif dari studi-studi terkini. Metodologi yang diterapkan menggunakan pendekatan Systematic Literature Review (SLR) dengan menganalisis secara mendalam lima publikasi ilmiah terpilih yang terbit dalam rentang waktu 2023-2024. Hasil evaluasi menunjukkan bahwa algoritma Decision Tree mendemonstrasikan performa yang menjanjikan dalam klasifikasi kondisi paru-paru dengan rentang akurasi 56,7% hingga 99,67%. Temuan penelitian mengindikasikan bahwa optimalisasi algoritma Decision Tree dapat dicapai melalui integrasi teknik preprocessing data yang tepat dan seleksi fitur yang cermat. Berdasarkan analisis yang dilakukan, dapat disimpulkan bahwa Decision Tree merupakan metode yang dapat diandalkan untuk klasifikasi penyakit paru-paru, khususnya ketika diimplementasikan dengan konfigurasi parameter yang dioptimasi dan dataset yang proporsional.

Kata Kunci: Sistem Klasifikasi Penyakit Paru-Paru, Algoritma Decision Tree, Kajian Sistematis Literatur, Machine Learning

Abstract– The global increase in lung disease cases presents a serious healthcare challenge requiring early detection systems for optimal treatment. This study examines the implementation of the Decision Tree algorithm in classifying various types of lung diseases based on a comprehensive analysis of recent studies. The methodology employs a Systematic Literature Review (SLR) approach by thoroughly analyzing five selected scientific publications published between 2023-2024. Evaluation results demonstrate that the Decision Tree algorithm shows promising performance in lung condition classification with accuracy ranges from 56.7% to 99.67%. Research findings indicate that Decision Tree algorithm optimization can be achieved through the integration of appropriate data preprocessing techniques and careful feature selection. Based on the analysis conducted, it can be concluded that Decision Tree is a reliable method for lung disease classification, particularly when implemented with optimized parameter configurations and proportional datasets.

Keywords: Lung Disease Classification System, Decision Tree Algorithm, Systematic Literature Review, Machine Learning

1. PENDAHULUAN

Penyakit paru-paru adalah salah satu masalah kesehatan global yang cukup mengkhawatirkan, dengan kanker paru-paru sebagai penyebab utama kematian di antara berbagai jenis kanker. *World Health Organization (WHO)* menyatakan kanker paru-paru menyebabkan jutaan kematian setiap tahunnya, dengan angka yang terus meningkat terutama di negara-negara berkembang. Deteksi dini menjadi faktor krusial dalam meningkatkan tingkat kelangsungan hidup pasien, namun proses diagnosis konvensional seringkali memakan waktu lebih lama dan sangat membutuhkan pada keahlian para ahli medis.

Seiring dengan perkembangan teknologi, metode *machine learning* telah menawarkan solusi yang menjanjikan dalam membantu proses diagnosis penyakit paru-paru. Di antara berbagai algoritma yang ada, Decision Tree muncul menjadi salah satu metode yang sering kali digunakan karena beberapa keunggulannya. Metode ini menawarkan kemudahan interpretasi hasil klasifikasi, kemampuan untuk menangani berbagai jenis data (baik numerik maupun kategorikal), proses pembelajaran yang efisien, dan visualisasi keputusan yang mudah dipahami oleh para praktisi medis.

Implementasi Decision Tree dalam klasifikasi penyakit paru-paru telah menunjukkan hasil yang beragam dalam berbagai penelitian. Beberapa studi menunjukkan tingkat akurasi yang tinggi, sementara yang lain menghadapi tantangan dalam hal optimasi model dan pemilihan fitur. Hal ini menunjukkan pentingnya melakukan kajian komprehensif terhadap implementasi metode ini untuk memahami faktor-faktor yang mempengaruhi keberhasilannya.

Penelitian ini bertujuan untuk menganalisis efektivitas metode Decision Tree dalam klasifikasi penyakit paru-paru, mengidentifikasi faktor-faktor yang mempengaruhi akurasi klasifikasi, dan memberikan rekomendasi untuk pengembangan sistem klasifikasi yang lebih baik. Melalui studi literature review, semoga dapat memberikan penjelasan yang lebih mendalam mengenai potensi dan tantangan dalam penggunaan Decision Tree untuk klasifikasi penyakit paru-paru, serta menjadi acuan untuk penelitian selanjutnya dalam bidang ini.

2. METODOLOGI PENELITIAN

2.1 Kerangka Kerja *Systematic Literature Review*

Systematic Literature Review (SLR) merupakan metodologi penelitian terstruktur yang diimplementasikan untuk mengumpulkan, mengevaluasi, dan menganalisis secara sistematis hasil-hasil penelitian yang relevan dengan topik tertentu. Proses SLR dimulai dengan menentukan pertanyaan penelitian yang jelas, diikuti dengan menetapkan kriteria seleksi literatur yang akan dipertimbangkan. Selanjutnya, dilakukan pencarian literatur dari berbagai sumber terpercaya, diikuti dengan penilaian kualitas studi yang relevan. Data dari studi yang terpilih kemudian diekstraksi dan dianalisis untuk memperoleh temuan yang komprehensif. Akhirnya, kesimpulan ditarik dan rekomendasi diberikan berdasarkan bukti yang ditemukan. SLR bertujuan untuk menyajikan gambaran yang objektif, mengurangi bias, dan memberikan bukti yang lebih kuat dari penelitian yang ada.

2.2 Kriteria Seleksi Literatur

Inclusion criteria adalah syarat-syarat yang harus dipenuhi oleh studi agar dapat dimasukkan dalam analisis, seperti topik yang relevan, metodologi yang sesuai, atau tahun publikasi tertentu. Oleh karena itu, untuk memastikan kualitas dan relevansi studi yang dianalisis, kriteria seleksi yang ketat diterapkan

2.2.1 Kriteria Inklusi

- Publikasi dalam rentang 2023-2024
- Fokus utama pada implementasi Decision Tree
- Aplikasi spesifik dalam diagnosis penyakit paru-paru
- Tersedia dalam format full-text
- Terpublikasi dalam bahasa Indonesia atau Inggris

2.2.2 Kriteria Eksklusi

- Studi tanpa validasi empiris
- Penelitian dengan dataset tidak terverifikasi
- Review non-sistematis
- Publikasi dalam bentuk abstrak saja

2.3 Strategi Pengumpulan Data

Proses pengumpulan data dilaksanakan melalui tahapan sistematis:

2.3.1 Identifikasi Sumber

- Database primer: Google Scholar
- Database sekunder: Directory of Open Access Journals
- Repository institusional terakreditasi

2.3.2 Protokol Pencarian

- Penggunaan kata kunci terstruktur
- Implementasi filter temporal
- Aplikasi kriteria seleksi bertahap

2.3.3 Dokumentasi Pencarian

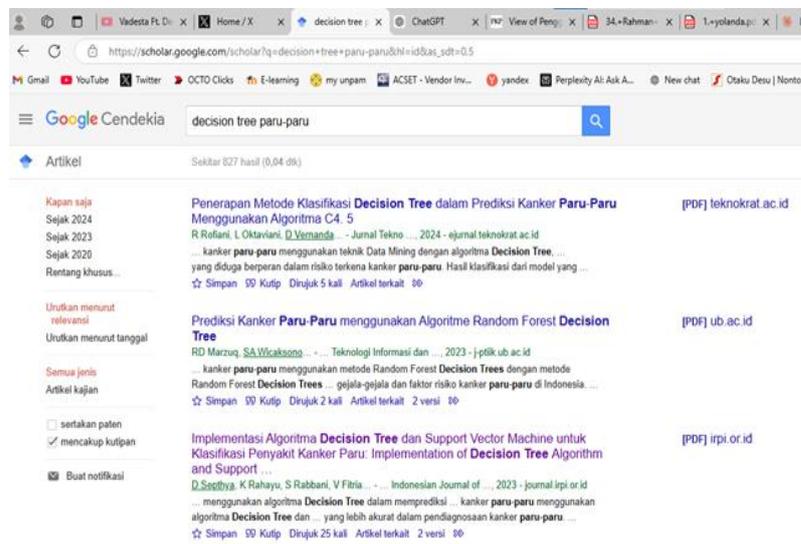
- Pencatatan hasil query
- Tracking proses seleksi
- Validasi duplikasi

2.4 Langkah Pengambilan Data

Metode yang digunakan pada pengumpulan data dalam program aplikasi ini adalah sebagai berikut:

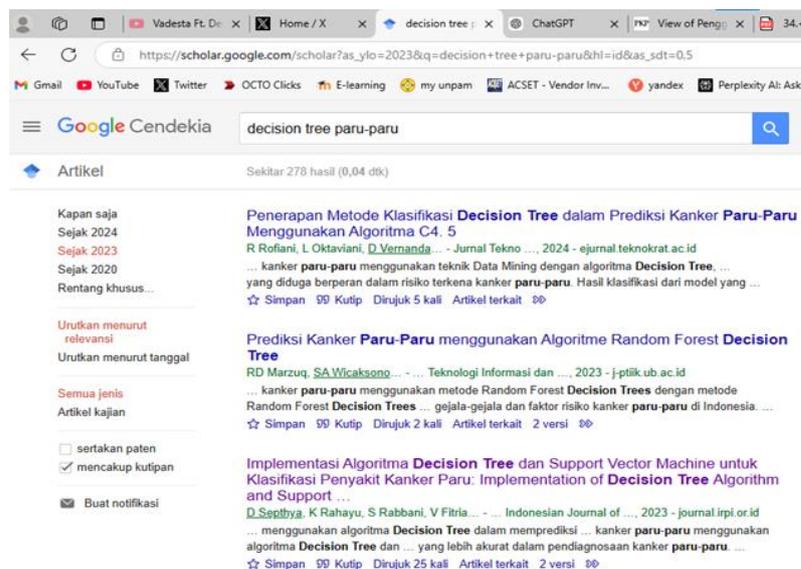
Pengambilan data menggunakan <https://scholar.google.com/> sebagai sumber website.

1. Buka browser kemudian masukan URL <https://scholar.google.com/>.
2. Pada kolom penelusuran gunakan *keyword* “decision tree paru-paru”



Gambar 1. Hasil Pencarian Tanpa Filter

3. Pada bagian kiri tampilan website, pilih “Sejak 2023” untuk menampilkan data yang dipublikasi dari 2023 sampai dengan saat ini (2024).



Gambar 2. Hasil Pencarian dengan Tahun Publikasi 2023-2024

3. ANALISA DAN PEMBAHASAN

Tabel 1. *Literature Review*

No.	Author dan Tahun	Metode Penelitian	Tujuan Penelitian	Kesimpulan
1	Dhini Septhya, Kharsma Rahayu, Salsabila Rabbani, Vindi Fitria, Rahmaddeni, Yuda Irowan, Regiolina, Hayami (2023)	Support Vector Machine dan Decision Tree	Membandingkan dua algoritma untuk Klasifikasi Penyakit Kanker Paru-paru, dengan algoritma Decision Tree dan Support Vector Machine	Algoritma Support Vector Machine (SVM) akurasinya lebih baik, dengan konfigurasi menggunakan splitting data 80:20 dengan akurasi 62,3% sedangkan algoritma Decision Tree memiliki tingkat akurasi 56,7% dengan splitting data 80:20
2	Riska Rofani, Lulu Oktaviani, Dwi Vernnda, Tri Hendriwan (2023)	Decision Tree C4.5	Mengembangkan model prediksi kanker paru-paru untuk mengidentifikasi individu beresiko tinggi secara dini	Algoritma C4.5 akurasinya cukup tinggi dalam memprediksi kanker paru-paru, dengan kebiasaan merokok sebagai faktor paling signifikan
3	Jatnika Fahmi Idris, Rafid Ramadhani, Muhammad Malik Mutoffar (2024)	Random Forest, Decission Tree, Naive Bayes, Support Vector Machines (SVM), Logistic Regresion, K-Nearest Neighbor(kNn)	Menentukan algoritma machine learning dengan tingkat akurasi tertinggi dalam mengklasifikasikan pasien yang terkena atau tidak terkena kanker paru-paru.	Algoritma Random Forest memiliki akurasi paling tinggi dengan nilai 97.48%, Decission Tree 95,16%, SVM sebesar 55.27% kNN sebesar 77.437 %, Naive Bayes sebesar 87.053% dan Logistic Regression sebesar 94.993%.
4	S. Udit Krisna, A.N. Barat Laksman, T. Arcana, K. Raja, dan M. Ayadurai (2024)	Decision Tree, Model VGG16 Convolutional Neural Network (CNN)	Memprediksi tingkat penyakit kanker paru-paru pada pasien berdasarkan gejala dan riwayat medis, serta identifikasi tipe kanker paru-paru menggunakan CT scan untuk pasien resiko sedang dan tinggi	Akurasi dengan model Decision Tree sangat baik dengan nilai 99,67% untuk klasifikasi risiko kanker paru-paru, sedangkan model VGG16 CNN akurasinya 92,53% dalam mengklasifikasikan tipe kanker. Hasil ini lebih baik daripada menggunakan metode yang sebelumnya.
5	Aqila Darin Makkyah, Muhammad Faisal (2023)	Decision Tree, Exploratory Data Analysis (EDA)	Untuk mendeteksi kanker paru-paru secara dini, klasifikasikan kanker paru-paru dengan menggunakan metode Decision Trees yang sangat akurat.	Hasil penelitian menunjukkan bahwa model Decision Tree sangat akurat dalam mengklasifikasikan kanker paru-paru. Model ini menghasilkan nilai recall tinggi untuk kelas positif "Yes" dan nilai rendah untuk kelas negatif "No".

4. KESIMPULAN

Berdasarkan analisis sistematis terhadap lima jurnal penelitian terkini, dapat disimpulkan bahwa metode Decision Tree menunjukkan potensi yang signifikan dalam klasifikasi penyakit paru-paru. Tingkat akurasi yang bervariasi dari 56,7% hingga 99,67% mengindikasikan bahwa efektivitas metode ini sangat bergantung pada berbagai faktor implementasi. Faktor-faktor seperti kualitas dataset, teknik preprocessing data, dan pemilihan parameter model memegang peranan penting dalam menentukan keberhasilan klasifikasi.

Dalam perbandingannya dengan metode machine learning lainnya seperti Support Vector Machine (SVM) dan Random Forest, Decision Tree menunjukkan keunggulan terutama dalam hal interpretabilitas hasil dan kemampuan menangani berbagai jenis data. Hal ini menjadikan Decision Tree sebagai pilihan yang tepat untuk aplikasi medis, di mana transparansi dalam pengambilan keputusan sangat diperlukan. Namun, perlu diperhatikan bahwa optimalisasi model tetap diperlukan untuk mencapai hasil yang optimal.

Penelitian ini juga mengungkapkan bahwa pendekatan hybrid, yang mengkombinasikan Decision Tree dengan teknik lain, dapat meningkatkan akurasi klasifikasi. Beberapa studi menunjukkan bahwa penggunaan teknik preprocessing data yang tepat dan pemilihan fitur yang relevan dapat secara signifikan meningkatkan performa model. Selain itu, pertimbangan terhadap faktor-faktor klinis dalam pemilihan fitur juga terbukti memberikan kontribusi positif terhadap akurasi klasifikasi.

Untuk pengembangan penelitian selanjutnya, disarankan untuk memfokuskan pada penggunaan dataset yang lebih besar dan seimbang, mengeksplorasi kombinasi Decision Tree dengan teknik ensemble, serta mempertimbangkan integrasi faktor-faktor klinis yang lebih komprehensif dalam proses klasifikasi. Hal ini diharapkan dapat meningkatkan kehandalan sistem klasifikasi penyakit paru-paru dan memberikan kontribusi yang lebih besar dalam bidang diagnosis medis.

REFERENCES

- Anggraeni, F., Kristiawan, N., Lutfiati, R., Dirgantara, Y., & Rosyani, P. (2023). Prediksi Cuaca Yang Akan Datang Menggunakan Metode Data Mining. *NEWTON: Jurnal Matematika, Fisika, Algoritma dan Sains*, 73-83.
- Anggraini, Y., Indra, M., Khoirusofi, M., Azis, I. N., & Rosyani, P. (2023). Systematic Literature Review: Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Forward Chaining. *BINER: Jurnal Ilmu Komputer, Teknik Dan Multimedia*, 1-7.
- Aqila, A., & Faisal, M. (2023). Lung Cancer EDA Classification Using the Decision Trees Method in Python. *Informatics and Software Engineering*, 8-13.
- Artaningsih, E. Y., & Muhajir, A. (2023). Komparasi Algoritma Klasifikasi C4. 5 Dan C4. 5 Berbasis Particle Swarm Optimization Untuk Evaluasi Penentuan Kelayakan Pemberian Kredit Koperasi Di PT. Indah Kiat Pulp & Paper TBK. *Scientia Sacra: Jurnal Sains, Teknologi dan Masyarakat*, 29-48.
- Doni, A., Fadli, A., Maulana, R. H., Putri, V. Y., & Rosyani, P. (2023). Analisis Metode Backward Chaining pada Sistem Pakar: Systematic Literature Review. *JURIHUM: Jurnal Inovasi dan Humaniora*, 10-16.
- Idris, J. F., Ramadhani, R., & Mutoffar, M. M. (2024). KLASIFIKASI PENYAKIT KANKER PARU MENGGUNAKAN PERBANDINGAN ALGORITMA MACHINE LEARNING. *Jurnal Media Akademik (JMA)*.
- Krishna, S. U., Lakshman, A. N., Archana, T., Raja, K., & Ayyadurai, M. (2024). Lung Cancer Prediction and Classification Using Decision Tree and VGG16 Convolutional Neural Networks. *The Open Biomedical Engineering Journal*.
- Lestari, A., Wijaya, H., Riyadi, N. S., & Rosyani, P. (2023). Systematic Literature Review: Sistem Pakar Diagnosis Penyakit Pada Manusia Menggunakan Metode Backward Channing. *Buletin Ilmiah Ilmu Komputer dan Multimedia (BIIKMA)*, 71-77.



- Perdana, W., Honi, R. A., Wibowo, B., & Rosyani, P. (2023). Studi Literature Review: Perbandingan Metode Klasifikasi Kecerdasan Buatan Pada Computer Vision. *JURIHUM: Jurnal Inovasi dan Humaniora*, 192-195.
- Rofiani, R., Oktaviani, L., Vernanda, D., & Hendriawan, T. (2024). Penerapan Metode Klasifikasi Decision Tree dalam Prediksi Kanker Paru-Paru Menggunakan Algoritma C4. 5. *Jurnal Tekno Kompak*, 126-139.
- Septhya, D., Rahayu, K., Rabbani, S., Fitria, V., Rahmadden, R., Irawan, Y., & Hayami, R. (2023). Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Penyakit Kanker Paru: Implementation of Decision Tree Algorithm and Support Vector Machine for Lung Cancer Classification. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 15-19.