



Literatur Review: Klasifikasi Penyakit Paru-paru Menggunakan Metode Decision Tree

Nuraina¹, Saiyah Awaliyah², Santi Ayu Purnamawati³, Vera Oktaviani⁴, Perani Rosyani^{5*}

^{1,2,3,4,5}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email: ¹nurainirusdy4@gmail.com, ²saiyah.awlyh266@gmail.com, ³santiayupurnamawati81@gmail.com, ⁴veraoktaviani223@gmail.com, ^{5*}dosen00837@unpam.ac.id

(* : coresponding author)

Abstrak - Kanker paru-paru merupakan penyebab utama kematian di dunia, dengan banyaknya kasus yang baru terdeteksi pada stadium lanjut, sehingga pengobatan menjadi kurang efektif. Dalam upaya untuk meningkatkan deteksi dini dan prediksi risiko kanker paru-paru, pendekatan berbasis machine learning telah menarik perhatian dalam penelitian medis. Artikel ini mengulas penerapan metode *Decision Tree* dalam klasifikasi penyakit paru-paru, yang memungkinkan identifikasi pola-pola penting dalam gejala-gejala awal penyakit tersebut. Metode *Decision Tree* digunakan untuk membangun model klasifikasi berdasarkan berbagai faktor risiko yang memengaruhi timbulnya kanker paru-paru, seperti kebiasaan merokok, usia, dan faktor lingkungan. Review ini juga membahas berbagai teknik yang digunakan untuk meningkatkan akurasi model, termasuk seleksi fitur dan penggunaan Cross Validation. Berbagai penelitian yang mengaplikasikan *Decision Tree* untuk prediksi kanker paru-paru menunjukkan hasil yang beragam, dengan sebagian besar penelitian mencapai akurasi yang cukup baik, yang menandakan potensi metode ini dalam aplikasi klinis. Penelitian lebih lanjut berfokus pada pengembangan algoritma yang lebih kuat, seperti Random Forest dan Support Vector Machine (SVM), serta integrasi dengan data klinis dan genetika untuk meningkatkan prediksi. Artikel ini bertujuan untuk memberikan gambaran menyeluruh tentang penerapan *Decision Tree* dalam klasifikasi kanker paru-paru, serta tantangan dan peluang yang ada dalam pemanfaatan machine learning untuk deteksi dini kanker paru-paru.

Kata Kunci: Klasifikasi, Penyakit Paru – Paru, Literatur Review, Metode *Decision Tree*

Abstract - Lung cancer is a leading cause of death in the world, with many cases only detected at an advanced stage, making treatment less effective. In an effort to improve early detection and risk prediction of lung cancer, machine learning-based approaches have attracted attention in medical research. This article reviews the application of the *Decision Tree* method in lung disease classification, which enables the identification of important patterns in the early symptoms of the disease. The *Decision Tree* method is used to build a classification model based on various risk factors that influence the onset of lung cancer, such as smoking habits, age, and environmental factors. This review also discusses various techniques used to improve the accuracy of the model, including feature selection and the use of cross validation. Studies applying the *Decision Tree* for lung cancer prediction have shown mixed results, with most studies achieving reasonably good accuracy, signaling the potential of this method in clinical applications. Further research focuses on the development of more robust algorithms, such as Random Forest and Support Vector Machine (SVM), as well as integration with clinical data and genetics to improve prediction. This article aims to provide a comprehensive overview of the application of *Decision Tree* in lung cancer classification, as well as the challenges and opportunities that exist in utilizing machine learning for lung cancer early detection.

Keywords: Classification, Lung Disease, Literature Review, *Decision Tree* Method

1. PENDAHULUAN

Kanker paru-paru adalah salah satu penyakit yang paling mematikan di dunia, dengan angka kematian yang terus meningkat setiap tahunnya. Menurut data dari Organisasi Kesehatan Dunia (WHO), lebih dari 1,8 juta orang meninggal akibat penyakit ini setiap tahun. Salah satu tantangan terbesar dalam menangani kanker paru-paru adalah deteksi dini, di mana banyak kasus baru ditemukan pada stadium lanjut, sehingga mengurangi peluang kesembuhan. Dengan meningkatnya angka kejadian dan dampak yang ditimbulkan, penting untuk mencari solusi yang efektif dalam deteksi dan klasifikasi penyakit ini.

Dalam upaya untuk mengatasi masalah ini, teknologi machine learning telah muncul sebagai alat yang sangat menjanjikan. Metode *Decision Tree*, khususnya, telah banyak digunakan dalam klasifikasi penyakit karena kemampuannya untuk memberikan hasil yang jelas dan mudah



dipahami. Decision Tree bekerja dengan membangun model berdasarkan data yang ada, mengidentifikasi faktor-faktor risiko yang berkontribusi terhadap perkembangan kanker paru-paru, seperti kebiasaan merokok, usia, dan riwayat kesehatan. Keunggulan utama dari metode ini adalah kemampuannya untuk menyajikan informasi dalam bentuk visual yang intuitif, memudahkan para profesional medis dalam mengambil keputusan.

Dalam penelitian ini, kami mengangkat judul "Klasifikasi Penyakit Paru-paru Menggunakan Metode Decision Tree" untuk mengeksplorasi lebih dalam tentang penerapan metode ini dalam mendeteksi kanker paru-paru. Melalui Studi Literatur Review (SLR), kami bertujuan untuk mengumpulkan dan menganalisis berbagai penelitian yang telah dilakukan terkait penggunaan Decision Tree dalam klasifikasi penyakit paru-paru. Dengan pendekatan ini, kami ingin mengidentifikasi tren, tantangan, serta efektivitas metode ini dalam meningkatkan akurasi deteksi dini kanker paru-paru.

Hasil yang diharapkan dari penelitian ini adalah pemahaman yang lebih baik mengenai bagaimana metode Decision Tree dapat diterapkan dalam praktik klinis, serta identifikasi faktor-faktor kunci yang dapat membantu dalam pengambilan keputusan medis. Selain itu, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam pengembangan teknik-teknik baru untuk mendeteksi dan mengklasifikasikan kanker paru-paru, serta membuka jalan bagi penelitian lanjutan di bidang kesehatan dan teknologi informasi.

Dengan menyoroti pentingnya deteksi dini dan penggunaan teknologi modern dalam pengelolaan kanker paru-paru, penelitian ini diharapkan dapat memberikan manfaat tidak hanya bagi dunia medis tetapi juga bagi pasien dan masyarakat luas.

2. METODOLOGI PENELITIAN

2.1 Tinjauan Literatur Sistematis

Tinjauan Literatur Sistematis (Systematic Literature Review) merupakan pendekatan yang terstruktur dan sistematis dalam mengumpulkan serta mengevaluasi penelitian yang relevan terkait topik tertentu. Dalam konteks penelitian ini, fokusnya adalah pada penerapan metode Decision Tree dalam klasifikasi penyakit paru-paru. Dengan melakukan tinjauan ini, diharapkan dapat diperoleh pemahaman yang lebih mendalam mengenai kontribusi metode ini dalam deteksi dini kanker paru-paru (Triandini et al., 2019).

2.2 Pertanyaan Penelitian

Pertanyaan penelitian adalah langkah awal yang penting untuk memandu arah studi ini. Pertanyaan-pertanyaan yang diajukan bertujuan untuk menggali lebih dalam mengenai efektivitas dan tantangan dalam penggunaan metode Decision Tree. Berikut adalah pertanyaan yang dirumuskan:

Tabel 1. Pertanyaan Peneliti

No	Pertanyaan Peneliti
1.	Faktor risiko apa saja yang dapat diidentifikasi melalui metode Decision Tree untuk memprediksi kanker paru-paru?
2.	Seberapa akurat model Decision Tree dalam mengklasifikasikan pasien kanker paru-paru dibandingkan dengan metode klasifikasi lainnya?
3.	Apa saja kendala yang dihadapi dalam penerapan metode Decision Tree untuk deteksi dini kanker paru-paru?



2.3 Proses Pencarian

Proses pencarian referensi dan bahan yang relevan untuk menjawab pertanyaan penelitian dilakukan secara sistematis. Pencarian ini dilakukan menggunakan mesin pencari, seperti Google Chrome, dengan mengakses situs <https://scholar.google.co.id/> atau <https://journal.unnes.ac.id/nju/index.php/sji> serta database akademik lainnya. Dengan pendekatan ini, diharapkan dapat ditemukan studi-studi yang berkaitan dengan topik yang diteliti.

2.4 Kriteria Inklusi dan Eksklusi

Pada tahap ini, dilakukan evaluasi terhadap kelayakan data untuk Tinjauan Literatur Sistematis. Studi-studi yang dipilih harus memenuhi kriteria berikut (Triandini et al., 2019):

Data yang digunakan harus berasal dari periode 2014 hingga 2024.

Data harus dapat diakses melalui situs <https://scholar.google.co.id/>.

Fokus data harus terkait dengan penerapan metode Decision Tree dalam klasifikasi penyakit paru-paru.

2.5 Pengumpulan Data

Pengumpulan informasi yang relevan merupakan langkah penting dalam proses penelitian ini. Berikut adalah langkah-langkah yang dilakukan untuk mengumpulkan data:

Mengakses halaman website situs <https://scholar.google.co.id/> Memasukkan kata kunci “Klasifikasi Penyakit Paru-paru Menggunakan Metode Decision Tree” pada kolom pencarian.

Menggunakan filter untuk memilih rentang tahun 2014 hingga 2024, sehingga hanya studi-studi yang relevan yang akan ditampilkan.

2.6 Analisis Data

Analisis data akan difokuskan pada studi-studi yang dipublikasikan antara tahun 2014 hingga 2024. Pada tahap ini, hasil dan temuan dari penelitian yang relevan akan dianalisis untuk mengidentifikasi tren, metodologi yang umum digunakan, serta efektivitas metode Decision Tree dalam klasifikasi penyakit paru-paru (Triandini et al., 2019). Dengan cara ini, diharapkan dapat diperoleh wawasan yang lebih jelas tentang potensi dan tantangan yang ada dalam penerapan metode ini.

Melalui langkah-langkah yang terstruktur ini, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan terhadap pemahaman tentang penerapan metode Decision Tree dalam klasifikasi penyakit paru-paru, serta membuka peluang untuk penelitian lebih lanjut di bidang ini.

3. ANALISA DAN PEMBAHASAN

Pada bagian ini berisi temuan, hasil serta pembahasan dari topik penelitian

Tabel 2. Hasil Temuan Artikel Relevan

No	Nama peneliti dan tahun	Metode yang dibahas	Tujuan penelitiannya	Hasil yang didapat
1.	Riska Rofiani ¹ , Luluk Oktaviani, Dwi Vernanda, Tri Hendriawan, 2024.	Klasifikasi Decision Tree menggunakan algoritma C4.5.	Penelitian ini bertujuan untuk mengembangkan model prediksi penyakit paru-paru menggunakan algoritma C4.5 dalam konteks data mining. Dengan menggabungkan teknik	Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu membangun model prediksi penyakit paru-paru dengan tingkat akurasi yang memadai. Proses analisis data, seleksi atribut, dan transformasi data memastikan kualitas dataset

			data mining dan algoritma C4.5.	yang digunakan dalam pengembangan model. Atribut yang terpilih, seperti usia, jenis kelamin, kebiasaan merokok, konsumsi alkohol, dan faktor-faktor lainnya, secara signifikan berkontribusi terhadap kemampuan prediktif model
2.	Afly Dwi Marzuq, Satrio Agung Wicaksono, dan Nanang Yudi Setiawan.2023	Dalam penelitian ini, penulis mengeksplorasi penggunaan algoritme Random Forest Decision Trees untuk memprediksi kanker paru-paru. Metode ini berfungsi untuk menganalisis gejala awal yang dialami oleh pasien kanker paru-paru dan membantu dalam penentuan label kelas yang tepat. Selain itu, penelitian ini juga menerapkan teknik evaluasi model melalui Confusion Matrix dan K-Fold Cross Validation untuk memastikan keakuratan dan validitas hasil.	Tujuan utama dari penelitian ini adalah untuk memprediksi kemungkinan terjadinya kanker paru-paru dengan memanfaatkan gejala awal dan faktor risiko. Penelitian ini juga bertujuan untuk memberikan dukungan kepada tenaga medis dalam pengambilan keputusan yang lebih baik, serta meningkatkan kesadaran masyarakat akan pentingnya deteksi dini kanker paru-paru.	Hasil dari penelitian ini menunjukkan bahwa model algoritme Random Forest Decision Trees mampu mencapai akurasi tertinggi sebesar 0,904 pada pengujian menggunakan Confusion Matrix. Rata-rata akurasi yang diperoleh adalah 0,813. Selain itu, saat menggunakan K-Fold Cross Validation, model ini menunjukkan rata-rata akurasi tertinggi sebesar 0,889 dengan metode 5-fold. Penelitian ini menekankan pentingnya pemilihan fitur yang relevan berdasarkan gejala dan faktor risiko kanker paru-paru di Indonesia untuk meningkatkan akurasi prediksi.
3.	Abdul Muis Alfatah, Riza Arifudin, Much Aziz Musli. 2018	Menggunakan Decision Tree forward chaining dan perhitungan diagnosis	Tujuan penelitian ini adalah mengimplementasikan metode Decision Tree dan Dempster Shafer pada sistem pakar	Hasil akhirnya berupa suatu fakta atau pun kombinasi fakta dengan nilai bobot terbesar. Setelah mengimplementasikan 65 data rekam medis yang



		dengan metode Dempster-Shafer.	untuk mendiagnosis penderita penyakit paru-paru dan mengetahui tingkat akurasi sistemnya.	diambil dari Puskesmas Tegowanu menggunakan metode Dempster-Shafer dan Decision Tree, dapat diketahui akurasi hasil perhitungan Dempster-Shafer dengan diagnosis dari dokter adalah sebanyak 83,08%, dengan 54 data menghasilkan diagnosis yang sama dengan dokter.
4.	Dhini Septhya, Kharisma Rahayu, Salsabila Rabbani, Vindi Fitria, Rahmaddeni, Yuda Irawan, Regiolina Hayami. 2023	Jurnal ini membahas dua metode klasifikasi utama yang digunakan untuk mendeteksi kanker paru-paru yaitu: Decision Tree dan support vector machine (SVM)	<p>a. Membandingkan kinerja algoritma Decision Tree dan Support Vector Machine dalam klasifikasi kanker paru.</p> <p>b. Menilai pengaruh penggunaan metode seleksi fitur (forward selection) terhadap akurasi model klasifikasi.</p> <p>c. Memberikan wawasan tentang algoritma mana yang lebih efektif dalam mendeteksi kanker paru, yang merupakan salah satu penyebab utama kematian di dunia.</p>	<p>a. Algoritma SVM dengan penggunaan forward selection memberikan akurasi tertinggi, yaitu 62,3% dengan rasio pembagian data 80:20.</p> <p>b. Tabel yang disajikan dalam jurnal menunjukkan perbandingan akurasi dari berbagai kombinasi algoritma dan metode seleksi fitur. Misalnya, pada rasio 80:20, Decision Tree tanpa forward selection mencapai akurasi 56,7%, sedangkan SVM tanpa forward selection mencapai 54,2%. Namun, ketika forward selection diterapkan pada SVM, akurasi meningkat menjadi 62,3%.</p> <p>Hasil ini menunjukkan bahwa penggunaan teknik seleksi fitur dapat memberikan dampak positif yang signifikan terhadap kinerja model klasifikasi, terutama dalam konteks deteksi kanker paru.</p>
5.	Faurika, Ahsanun Naseh Khudori2, M Syauqi Haris. 2024.	Penelitian ini menggunakan algoritma Decision Tree (ID3) dalam penerapan machine learning untuk	Penelitian ini bertujuan untuk memprediksi kanker paru-paru dengan memanfaatkan algoritma Decision Tree, sehingga membantu dalam diagnosa dini yang	Penelitian ini menghasilkan model dengan akurasi, presisi, recall, dan F1-Score sebesar 100%, menunjukkan bahwa model ini sangat efektif dalam memprediksi risiko kanker



		prediksi kanker paru-paru.	dapat menurunkan angka kematian akibat kanker paru-paru.	paru-paru berdasarkan data pasien(717-1462-1-SM).
--	--	----------------------------	--	---

4. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa metode Decision Tree memiliki peran penting dan potensi yang menjanjikan dalam mendukung klasifikasi dan deteksi dini kanker paru-paru. Kanker paru-paru menjadi penyebab utama kematian dengan banyak kasus yang terlambat terdeteksi sehingga mengurangi efektivitas pengobatan. Oleh karena itu, solusi berbasis machine learning seperti Decision Tree diharapkan dapat membantu identifikasi faktor risiko sejak dini dan meningkatkan prediksi diagnosis, yang pada akhirnya mendukung pengambilan keputusan klinis yang lebih baik.

Melalui tinjauan literatur yang dilakukan, ditemukan bahwa metode Decision Tree mampu mengidentifikasi pola-pola penting dari gejala dan faktor risiko kanker paru-paru, seperti kebiasaan merokok, usia, konsumsi alkohol, dan faktor lingkungan. Dalam penelitian yang direview, Decision Tree menunjukkan akurasi yang baik dalam membangun model prediksi berdasarkan faktor-faktor tersebut. Peningkatan akurasi model juga dicapai melalui teknik-teknik pendukung seperti seleksi fitur yang memprioritaskan variabel yang relevan dan penggunaan Cross Validation untuk memastikan ketepatan hasil prediksi. Penelitian-penelitian menunjukkan bahwa algoritma Decision Tree cukup efektif, dengan model yang dibangun menghasilkan akurasi yang baik dan mudah diinterpretasikan dalam bentuk visual yang intuitif.

Namun, tantangan tetap ada dalam penerapan metode ini, terutama jika dibandingkan dengan algoritma lain seperti Support Vector Machine (SVM) dan Random Forest, yang dalam beberapa studi menunjukkan performa akurasi lebih tinggi untuk kasus tertentu. Terlepas dari keterbatasan tersebut, metode Decision Tree tetap memiliki keunggulan utama pada kemampuannya menghasilkan model yang transparan dan mudah dipahami oleh tenaga medis, yang penting untuk diterapkan dalam skenario klinis. Oleh karena itu, pengembangan lebih lanjut diusulkan untuk menggabungkan metode ini dengan algoritma lain, memperkaya dataset menggunakan data klinis dan genetika, serta memanfaatkan teknologi terbaru untuk analisis data lebih komprehensif.

REFERENCES

- Faurika, F. K. (2024). Implementasi Decision Tree Untuk Prediksi Kanker Paru-Paru. *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 9(1), 94-106.
- Gunadi, I. A. (2022). MENDIAGNOSA PENYEBAB PENYAKIT STROKE PADA PASIEN MENGGUNAKAN METODE BACKWARD CHAINING. *BISIK: Jurnal Ilmu Komputer, Hukum, Kesehatan dan Sosial Humaniora*, 1(3: Oktober), 254-263.
- Marzuq, R. D. (2023). Prediksi Kanker Paru-Paru menggunakan Algoritme Random Forest Decision Tree. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(7), 3448-3456.
- Ramadhan, M. R. (2023). IMPLEMENTASI DATA MINING PADA DATASET PRAKIRAAN CUACA MENGGUNAKAN ALGORITMA C4. 5. *ALKHAWARIZMI: Jurnal Matematika, Algoritma dan Sains*, 1(1), 131-140.
- Rofiani, R. O. (2024). Penerapan Metode Klasifikasi Decision Tree dalam Prediksi Kanker Paru-Paru Menggunakan Algoritma C4. 5. *Jurnal Tekno Kompak*, 18(1), 126-139.
- Santosa, I. R. (2018). Implementasi Algoritma Decision Tree C. 45 Untuk Diagnosa Penyakit Tuberculosis (Tb). *Jurnal Ilmiah NERO Vol*, 3(3).
- Sathyanarayanan, P. P. (2019). Identification of breast cancer using the Decision Tree algorithm. In 2019 IEEE International conference on system, computation, automation and networking (ICSCAN) (pp. 1-6). IEEE.
- Septhya, D. R. (2023). Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Penyakit Kanker Paru. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 15-19.
- Sholih, M. G. (2019). Risk factors of lung cancer in Indonesia: A qualitative study. *Journal of Advanced Pharmacy Education and Research*, 9(2-2019), 41-40.



JRIIN : Jurnal Riset Informatika dan Inovasi

Volume 2, No. 8 Januari Tahun 2025

ISSN 3025-0919 (media online)

Hal 1444-1450

Suryana, A. S. (2023). Sistem Pakar Penyakit Ginjal pada Manusia Menggunakan Metode Forward Chaining. RIIN: Jurnal Riset Informatika dan Inovasi, 1(1), 277-281.