Journal of Computer and Information Systems Ampera
Vol. 7, No. 2, May 2026 e-ISSN: 2775-2496
https://journal-computing.org/index.php/journal-cisa/index

Analisis Perbandingan Algoritma Machine Learning Pada Penyakit Liver dengan Menggunakan RapidMiner

Syeila Ayu Rizkita[1], Sazila Azka Adzkia[2], Puteri Marchanda Izzati[3], Wita Aulia Susanti[4]

[1,2,3,4] Faculty of Information Technology , Adhirajasa Reswara Sanjaya University, Bandung, Idnonesia
Email: [1]syeilaayr30@gmail.com , [2]sazilaazka@gmail.com , [3]putrimarchandaaa@gmail.com , [3]witaaulia27@gmail.com

Abstract

Hati merupakan organ yang memegang peranan sangat penting dalam tubuh manusia. Permasalahan yang muncul adalah sulitnya mengidentifikasi penyakit liver secara dini, bahkan setelah penyakit tersebut menyebar pun tetap sulit untuk dideteksi. Penelitian ini bertujuan untuk melakukan analisis perbandingan algoritma machine learning dalam memprediksi penyakit liver menggunakan perangkat lunak *RapidMiner*. Ada 10 Algoritma yang dibandingkan. Hasil penelitian menunjukkan bahwa algoritma *SVM (PSO)* memiliki akurasi tertinggi dalam klasifikasi penyakit liver, dengan ratio 60:40 sebesar 96,86%, diikuti oleh algoritma *k-NN* dengan akurasi 81,71%. Temuan ini menunjukkan bahwa *SVM (PSO)* dan *k-NN* merupakan algoritma terbaik untuk klasifikasi penyakit liver berdasarkan dataset *ILPD* dari *UCI Repository* dan aplikasi RapidMiner versi 10.3. Studi ini merekomendasikan penelitian selanjutnya untuk menggunakan perangkat lunak lain guna mendapatkan hasil perbandingan yang lebih baik. Dengan demikian, hasil penelitian ini memberikan kontribusi penting dalam pengembangan metode prediksi penyakit liver menggunakan algoritma *machine learning* dan perangkat lunak *RapidMiner*.

Keywords: k-NN, Liver, Machine Learning, RapidMiner , SVM

1. PENDAHULUAN

Tubuh kita mempunyai beberapa organ penting yang masing-masing mempunyai fungsi yang sangat bermanfaat, salah satunya adalah hati. Hati merupakan organ terbesar dalam tubuh kita dan mempunyai fungsi yang sangat penting. Fungsi hati adalah menyimpan glikogen, membantu

This work is licensed under a Creative Commons Attribution 4.0 International License.

pembentukan dan sekresi empedu, mensintesis urea, serta berperan dalam metabolisme kolesterol dan lemak.Fungsi utama hati adalah mendetoksifikasi atau menetralisir racun.Hati melindungi tubuh kita dari berbagai racun yang dapat membahayakan kesehatan kita.Penyakit ini dapat mempengaruhi hati dan mencegahnya berfungsi secara normal, sehingga dapat menyebabkan kematian.Penyakit liver sudah ada sejak zaman dahulu dan tersebar luas di masyarakat [1]. Terdapat peningkatan kasus penyakit hati secara global, baik karena faktor gaya hidup, pola makan, atau faktor-faktor lainnya. Jika penyakit liver menyerang dan merusak hati, maka kemampuan tubuh dalam menetralisir racun yang masuk ke dalam tubuh perlahan menurun sehingga menimbulkan kerugian bagi tubuh jika tidak segera ditangani [2].

Machine learning dapat menjadi alat yang efektif untuk membantu dalam diagnosis dan prediksi penyakit hati. Adanya ketertarikan dalam menggunakan algoritma *machine learning* untuk mendukung diagnosa medis. Penerapan algoritma ini diharapkan dapat membantu dokter dalam melakukan identifikasi penyakit hati lebih cepat dan akurat. Diagnosis penyakit hati seringkali rumit dan memerlukan analisis mendalam terhadap berbagai faktor. *Machine learning* dapat menjadi alat yang efektif untuk membantu mendiagnosis dan memprediksi penyakit hati. Penulis tertarik menggunakan algoritma *machine learning* untuk mendukung diagnosis medis. Penerapan algoritma ini diharapkan dapat membantu dokter mengidentifikasi penyakit liver dengan lebih cepat dan akurat [3].

Penelitian ini bertujuan untuk membandingkan kinerja beberapa algoritma *machine learning* yang digunakan untuk memprediksi penyakit liver. Penelitian ini menggunakan *RapidMiner* sebagai platform dan algoritma-algoritma yang dibandingkan meliputi *Decision Tree, Random Forest, Random Tree,Linear Regression, k-NN, Support Vector Machine (PSO), Neural Network, Support Vector Machine, Logistic Regression, dan LDA* [4].

Penelitian sebelumnya mungkin telah menunjukkan bahwa berbagai algoritma Penelitian ini bertujuan untuk membandingkan kinerja beberapa algoritma *machine learning* yang digunakan untuk memprediksi penyakit liver. Penelitian ini menggunakan *RapidMiner* sebagai platform dan algoritma-algoritma yang dibandingkan meliputi *Decision Tree, Random Forest, Random Tree,Linear Regression, k-NN, Support Vector Machine (PSO), Neural Network, Support Vector Machine, Logistic*

Regression, dan LDA. *Machine learning* dapat memberikan hasil yang berbeda-beda untuk masalah klasifikasi penyakit hati. berikut adalah beberapa penelitian terdahulu yang disertakan

Pada penelitian [5] penulis menyajikan pengembangan sistem deteksi penyakit hati otomatis menggunakan teknik pendeteksian support vector machine dan random forest. Studi ini menggunakan data dari uji coba Mayo Clinic pada penyakit sirosis bilier primer (PBC) hati yang dilakukan antara tahun 1974 dan 1984. Hasilnya menunjukkan bahwa model *Support Vector Machine (SVM)* yang diusulkan mencapai akurasi 78,3% dalam mendeteksi penyakit hati. Studi ini juga menganalisis pentingnya fitur-fitur berbeda dalam diagnosis penyakit hati. Artikel tersebut mengusulkan sistem deteksi penyakit hati otomatis menggunakan model random forest, yang mencapai akurasi 78,3% dalam mendiagnosis penyakit hati. Sistem ini membantu dalam proses pengambilan keputusan diagnosis penyakit hati dan dapat membantu para ahli medis dalam memprediksi penyakit hati .

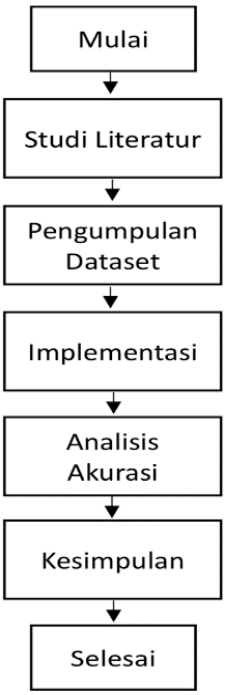
Penelitian [6] penulis menyatakan penggunaan algoritma *Decision Tree* untuk mendiagnosis penyakit liver menggunakan dataset ILPD menghasilkan akurasi sebesar 72,67%. Hanya dua variabel, yaitu SPGT_AA dan Usia (Age), yang diidentifikasi sebagai faktor krusial dalam menentukan penyakit liver. Penelitian ini memberikan kontribusi dalam memahami faktor-faktor yang mempengaruhi penyakit liver dan metode pemodelan klasifikasi .

Dalam jurnal [7] penulis menyimpulkan bahwa penggunaan teknik *machine learning*, terutama algoritma pohon keputusan, dapat memberikan prediksi penyakit hati yang lebih akurat dibandingkan dengan algoritma klasifikasi lainnya.Hasil eksperimen memberikan hasil yang lebih baik dibandingkan dengan algoritma klasifikasi lainnya, dengan tingkat akurasi sebesar 81%. Studi ini menggunakan berbagai algoritma klasifikasi seperti *Naive Bayes, Random Forest, Support Vector Machine, dan Artificial Neural Network* untuk membandingkan kinerja masing-masing algoritma. Dalam prediksi penyakit hati, identifikasi penyakit pada tahap awal dianggap penting untuk meningkatkan perawatan yang dipersonalisasi bagi pasien di India.

Dalam konteks ini, penggunaan algoritma *machine learning* menjadi fokus penelitian yang bertujuan untuk melakukan analisis perbandingan beberapa algoritma untuk prediksi penyakit hati menggunakan aplikasi *RapidMiner*.Dengan membandingkan kinerja algoritma ini, penulis berharap dapat menemukan pendekatan yang efektif dalam memprediksi penyakit hati, yang dapat memberikan [8].

2. METODE

Penelitian yang dilakukan meliputi pengolahan dataset ILPD menggunakan bantuan aplikasi. Dalam hal ini aplikasi bantuan yang dimaksud adalah *Rapidminer* versi 10.3. Penelitian ini akan menggunakan pendekatan kuantitatif untuk membandingkan berbagai algoritma *machine learning* yang digunakan dalam deteksi atau analisis penyakit liver. Tahapan yang digunakan dalam penelitian ini adalah



Gambar 1. Tahapan Metode

a. Studi Literatur

Untuk memulai tahap pertama, peneliti terlibat dalam proses ekstensif dalam melakukan studi literatur, yang melibatkan pencarian, pembacaan, dan pemahaman penelitian sebelumnya yang berkaitan dengan pokok bahasan yang diteliti dengan cermat [9]

b. Pengumpulan Dataset

Setelah melakukan penelitian literatur, penulis melanjutkan untuk mengambil dan mengumpulkan data penyakit hati dan menggunakannya dalam penelitian ini.

c. Implementasi

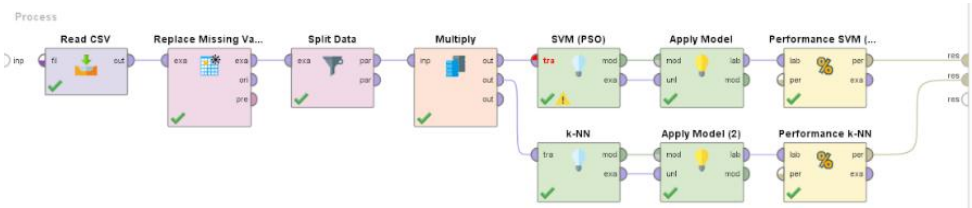
Proses selanjutnya adalah melakukan pengolahan data menggunakan 10 algoritma dengan menggunakan aplikasi *RapidMiner* 10.3.

d. Analisis akurasi

Proses terakhir adalah hasilnya. Hasil penelitian ini akan menghasilkan kesimpulan yang kemudian menjadi pengetahuan

3. HASIL DAN PEMBAHASAN

Hasil penelitian yang dilakukan pada pembahasan ini akan menguraikan proses cara mengklasifikasikan dataset penyakit liver dengan menggunakan beberapa algoritma yaitu *Decision Trees, Random Forests, Random Trees, Linear Regression, K-NN, SVM, Neural Networks, SVM (PSO), Regresi Logistik dan LDA*. Pengelompokan ini dilakukan melalui proses pengujian menggunakan *Machine Learning* yaitu perangkat lunak *RapidMiner* Studio 10.3. Saat proses dijalankan, dibutuhkan waktu kurang lebih 30 detik hingga hasil klasifikasi muncul, waktu yang diperlukan itu tergantung spesifikasi laptop yang Anda miliki.



Gambar 2. Klasifikasi pada algoritma *SVM (PSO)* dan *k-NN*

Klasifikasi penyakit hati diimplementasikan melalui algoritma *machine learning* di *RapidMiner* 10.3 menggunakan operator yang disediakan di *RapidMiner*. Operator dan penjelasannya adalah:

1. *Read CSV*, operator yang digunakan untuk membaca file *CSV* yang di import. Dataset *Indian Liver Patient Dataset (ILPD)* diambil dari *UCI Machine Learning Repository* dan sudah dalam format CSV.
2. *Replace Missing Value,* operator untuk proses pengisian data yang hilang (missing value) pada suatu dataset.. Artinya, nilai yang kosong atau tidak lengkap akan diubah ke nilai lain. Hal ini dilakukan untuk melengkapi data. Sehingga data yang kosong pada atribut "Albumin and Gobulin Ratio" akan diisi dengan nilai 0. 10 feature yaitu *Age, Gender, TB, DB, Alkphos, SGPT, SGOT, TP, ALB, A/G Ratio* [10].
3. *Multiply,* suatu operator yang digunakan untuk menggabungkan beberapa operator agar dapat berjalan secara bersamaan.
4. *Split data*, dilakukan untuk memisahkan data latih dan data uji serta menyesuaikan rasionya sesuai kebutuhan
5. *Apply Model :* Digunakan untuk menerapkan model yang sebelumnya dilatih menggunakan data pelatihan untuk menguji data.
6. *Performance :* Akhir dari proses yang akan menunjukkan hasilnya dari data yang diperoleh dengan nilai akurasinya ditunjukkan pada 145able performance[11]

Hasil akurasi berdasarkan **Tabel 1** menunjukkan akurasi terbaik dari beberapa algoritma yang digunakan. Dengan membandingkan hasil tersebut maka dapat disimpulkan bahwa 2 algoritma yaitu *SVM (PSO)* mempunyai akurasi tertinggi dalam klasifikasi penyakit *liver*, dengan ratio 60:40 yaitu sebesar 96,86%, dan algoritma *k-NN* sebesar 81,71%, dengan ratio adalah 90:10 di antara algoritma yang digunakan [12]

Tabel 1. Hasil akurasi algoritma

Algoritma	Tanpa Split Data	Training 50	Training 60	Training 70	Training 80	Training 90
Decision Tree	77.19%	76.37%	84.00%	79.41%	77.73%	77.71%
Random Forest	79.43%	81.85%	82.29%	80.39%	79.87%	79.43%
Random Tree	75.64%	75.68%	75.71%	75.74%	75.59%	75.62%
Linear Regression	75.64%	75.68%	75.71%	75.74%	75.59%	75.62%
k-NN	81.30%	81.16%	79.14%	80.88%	80.94%	81.71%
SVM	75.64%	75.68%	75.71%	75.74%	75.59%	75.62%
Neural Network	75.64%	75.68%	75.43%	75.74%	75.59%	75.62%
SVM (PSO)	95.03%	93.49%	96.86%	95.83%	95.50%	95.24%
Logistic Regression	75.30%	75.34%	76.00%	75.74%	75.37%	75.62%
LDA	75.81%	75.68%	75.43%	75.74%	75.59%	75.43%

Ketika skema di *RapidMiner* selesai dijalankan, disini penulis akan menunjukkan hasil akurasi dari 2 algoritma tertinggi dan melakukan perhitungan manual klasifikasi Liver. Hasil pengukuran akurasi, presisi dan recall masing-masing algoritma adalah:

1. K-NN

accuracy: 81.71%

	true Female	true Male	class precision
pred. Female	54	22	71.05%
pred. Male	74	375	83.52%
class recall	42.19%	94.46%	

Gambar 3. Hasil menggunakan Algoritma *k-NN*

$$Akurasi = \frac{54 + 375}{54 + 375 + 22 + 74}$$

$$= \frac{429}{525} = 81{,}71\% \quad (1)$$

$$Presisi = \frac{58}{58+25}$$

$$= \frac{58}{83} = 69{,}88\% \quad (2)$$

$$Recall = \frac{58}{58+84}$$

$$= \frac{58}{142} = 40{,}85\% \quad (3)$$

2. SVM (PSO)

accuracy: 96.86%

	true Female	true Male	class precision
pred. Female	74	0	100.00%
pred. Male	11	265	96.01%
class recall	87.06%	100.00%	

Gambar 4. Hasil menggunakan Algoritma *SVM (PSO)*

$$Akurasi = \frac{74+265}{74+265+11+0}$$

$$= \frac{339}{350} = 96{,}86\% \quad (1)$$

$$Presisi = \frac{74}{74+0}$$

$$= \frac{74}{74} = 100{,}00\% \quad (2)$$

$$Recall = \frac{74}{74+11}$$

$$= \frac{74}{85} = 87{,}06\% \quad (3)$$

4. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa analisis perbandingan algoritma *Decision Tree, Random Forest, Random Tree, Linear Regression, k-NN, SVM, Neural Network, SVM (PSO), Logistic Regression* dan *LDA* untuk menyelesaikan permasalahan klasifikasi pasien menderita penyakit liver. Penelitian yang dilakukan meliputi pengolahan dataset *Indian Liver Patient Dataset (ILPD)* yang diambil dari UCI Repository menggunakan bantuan aplikasi *Rapidminer* versi 10.3.

Dari hasil analisis dan perbandingan yang telah dilakukan, algoritma *SVM (PSO)* dengan menggunakan ratio 60:40 menghasilkan akurasi yang lebih baik sebesar 96,86% dan algoritma *K-NN* sebesar 81,30%. Dengan demikian algoritma *SVM (PSO)* dan *K-NN* dapat dikatakan algoritma terbaik untuk klasifikasi penyakit liver dibandingkan dengan algoritma *Decision Tree, Random Forest, Random Tree, Linear Regression, SVM, Neural Network, Logistic Regression* dan *LDA.*

Pada penelitian ini dilakukan uji coba menggunakan satu software, sehingga pada penelitian selanjutkan dapat dilakukan uji coba menggunakan software lain untuk mendapatkan hasil perbandingan yang lebih beragam agar menjadi rinci dalam mengklasifikasi penyakit *liver*.

REFERENSI

[1] A. P. Ayudhitama and U. Pujianto, “JIP (Jurnal Informatika Polinema) ANALISA 4 ALGORITMA DALAM KLASIFIKASI PENYAKIT LIVER MENGGUNAKAN RAPIDMINER”.
[2] D. R. Fadli, “Liver,” https://www.halodoc.com/kesehatan/liver.
[3] E. Pusporani, S. Qomariyah, and Irhamah, “Klasifikasi Pasien Penderita Penyakit Liver dengan Pendekatan Machine Learning,” 2019.
[4] F. Lia, D. Cahyanti, F. Sarasati, W. Astuti, and E. Firasari, “KLASIFIKASI DATA MINING DENGAN ALGORITMA MACHINE LARNING UNTUK PREDIKSI PENYAKIT LIVER,” 2023. [Online]. Available: https://ojs.uniska-bjm.ac.id/index.php/JIT

[5] T. A. Assegie, R. Subhashni, N. K. Kumar, J. P. Manivannan, P. Duraisamy, and M. F. Engidaye, "Random forest and support vector machine-based hybrid liver disease detection," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 3, pp. 1650–1656, Jun. 2022, doi: 10.11591/eei.v11i3.3787.

[6] I. Setiawati, A. P. Wibowo, A. Hermawan, M. Teknologi, I. Universitas, and T. Yogyakarta, "IMPLEMENTASI DECISION TREE UNTUK MENDIAGNOSIS PENYAKIT LIVER," 2019.

[7] SCAD College of Engineering and Technology and Institute of Electrical and Electronics Engineers, *Proceedings of the International Conference on Trends in Electronics and Informatics (ICOEI 2018) : 11-12, May 2018*. 2018.

[8] R. Tuntun, K. Kusrini, and K. Kusnawi, "Analisis Perbandingan Kinerja Algoritma Klasifikasi dengan Menggunakan Metode K-Fold Cross Validation," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, p. 2111, Oct. 2022, doi: 10.30865/mib.v6i4.4681.

[9] P. Algoritma K-means Pada Kasus HIV Di Jawa Barat Untuk Pengelompokan Berdasarkan Tingkat Penyebaran Di Setiap Kabupaten, K. Menggunakan RapidMiner, D. Parameswari Syifa Dewi, and A. Voutama, "Application of K-Means Algorithm on HIV Cases in West Java for Clustering Based on The Level of Spread in Each District / City Using RapidMiner," *JTSI*, vol. 4, no. 2, pp. 289–300, 2023.

[10] "Replace Missing Values," RapidMiner Altair Company. Accessed: Jan. 10, 2024. [Online]. Available: https://academy.rapidminer.com/learn/video/replace-missing-values

[11] Z. Zukhri, "Hasil dan Pembahasan," 2018, Accessed: Jan. 10, 2024. [Online]. Available: https://dspace.uii.ac.id/bitstream/handle/123456789/10286/05.4%20bab%204.pdf?isAllowed=y&sequence=8

[12] A. B. Wibisono and A. Fahrurozi, "PERBANDINGAN ALGORITMA KLASIFIKASI DALAM PENGKLASIFIKASIAN DATA PENYAKIT JANTUNG KORONER," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 24, no. 3, pp. 161–170, 2019, doi: 10.35760/tr.2019.v24i3.2393.