

Penerapan Algoritma *Decision Tree* Dengan Optimasi Parameter Dalam Memprediksi Reaksi Autoimun Akibat Obat

Saifudin¹, Nuzul Imam Fadlilah², Ahmad Nouvel³, Sunanto⁴

^{1,2,3,4}Universitas Bina Sarana Informatika

Jl. HR Bunyamin No 106 Pabuwara, Purwokerto Utara, Banyumas, Indonesia

Email: saifudin.sfn@bsi.ac.id¹, nuzul.nfh@bsi.ac.id², ahmad.avl@bsi.ac.id³, sunanto.sun@bsi.ac.id⁴

Artikel Info : Diterima : 17 Juni 2025 | Direvisi : 24 Juli 2025 | Disetujui : 25 Juli 2025

Abstrak, Dalam perkembangan teknologi medis terkini, machine learning, khususnya penerapan algoritma *Decision Tree*, menjadi pendekatan yang semakin populer untuk analisis data kesehatan berskala besar. *Decision Tree* dikenal karena kemampuannya mengidentifikasi pola tersembunyi dalam data klinis dengan interpretasi yang mudah dipahami tenaga medis. Melalui proses optimasi parameter, akurasi model dapat ditingkatkan secara signifikan, memungkinkan prediksi yang lebih tepat terhadap kemungkinan reaksi autoimun akibat penggunaan obat tertentu. Penggunaan model prediktif berbasis *Decision Tree* dengan parameter yang dioptimalkan bukan hanya memperkuat pengambilan keputusan klinis, tetapi juga membuka jalan bagi praktik pengobatan yang lebih personal dan presisi. Optimasi parameter digunakan untuk eksekusi terhadap seluruh variasi parameter yang ditetapkan melalui subprocessesnya. Hasil akhirnya mencatat performa prediktif yang optimal sebesar 77,50% dengan 98,28% lebih presisi untuk kelas “true=0” dibandingkan dengan kelas “true=1”.

Kata kunci: autoimun, algoritma *Decision Tree*, optimasi parameter

Abstract, In the latest developments in medical technology, machine learning, especially the application of the *Decision Tree* algorithm, is becoming an increasingly popular approach for large-scale health data analysis. *Decision Tree* is known for its ability to identify hidden patterns in clinical data with interpretation that is easy for medical professionals to understand. Through the process of parameter optimization, the accuracy of the model can be significantly improved, allowing for more precise predictions of possible autoimmune reactions due to the use of certain drugs. The use of *Decision Tree*-based predictive models with optimized parameters not only strengthens clinical decision-making, but also paves the way for more personalized and precise treatment practices. Parameter optimization is used for the execution of all parameter variations that are set through its subprocesses. The final result recorded an optimal predictive performance of 77.50% with 98.28% more precision for the "true=0" class compared to the "true=1" class.

Keyword: autoimmunne, algoritma *Decision Tree*, optimize parameter

PENDAHULUAN

Reaksi autoimun yang dipicu oleh konsumsi obat-obatan merupakan tantangan serius dalam dunia medis. Fenomena ini terjadi ketika sistem kekebalan tubuh secara keliru menyerang jaringan sehat akibat respons terhadap zat aktif dalam obat. Menurut data dari World Allergy Organization (2023), sekitar 7–10% reaksi obat serius berujung pada gangguan autoimun, yang dalam kasus tertentu dapat mengancam jiwa pasien. Penyakit autoimun mempengaruhi sekitar 10% populasi dunia. Sebagian besar adalah penyakit kompleks yang disebabkan oleh interaksi antara faktor genetik, epigenetik dan lingkungan. Selain itu, penyakit autoimun telah menjadi penyakit kronis terbesar ketiga, setelah penyakit kardiovaskular dan kanker (Fortes et al., 2024). *Adverse drug reactions* (ADRs) mencakup 3% hingga 6% dari semua pasien rawat inap dan terjadi pada 10% hingga 15% pasien rawat inap dan hingga 25% pasien rawat jalan. Alergi obat relatif jarang terjadi, mencakup kurang dari 10% dari semua ADRs. Alergi obat terjadi pada 1% hingga 2% dari semua pasien rawat inap dan 3% hingga 5% pasien rawat inap, tetapi insiden alergi obat yang sebenarnya di masyarakat, dan di antara anak-anak dan orang dewasa, tidak diketahui (Thong et al., 2021). Oleh karena itu, kemampuan untuk memprediksi reaksi semacam ini secara dini sangat penting dalam menunjang keselamatan pasien dan efisiensi pengobatan.



Dalam perkembangan teknologi medis terkini, machine learning, khususnya penerapan algoritma *Decision Tree*, menjadi pendekatan yang semakin populer untuk analisis data kesehatan berskala besar. *Decision Tree* dikenal karena kemampuannya mengidentifikasi pola tersembunyi dalam data klinis dengan interpretasi yang mudah dipahami tenaga medis. Melalui proses optimasi parameter, akurasi model dapat ditingkatkan secara signifikan, memungkinkan prediksi yang lebih tepat terhadap kemungkinan reaksi autoimun akibat penggunaan obat tertentu.

Penggunaan model prediktif berbasis *Decision Tree* dengan parameter yang dioptimalkan bukan hanya memperkuat pengambilan keputusan klinis, tetapi juga membuka jalan bagi praktik pengobatan yang lebih personal dan presisi. Penelitian ini akan mengulas lebih lanjut bagaimana teknik ini dapat menjadi solusi efektif dalam menghadapi tantangan reaksi autoimun yang dipicu oleh terapi farmakologis.

METODE PENELITIAN

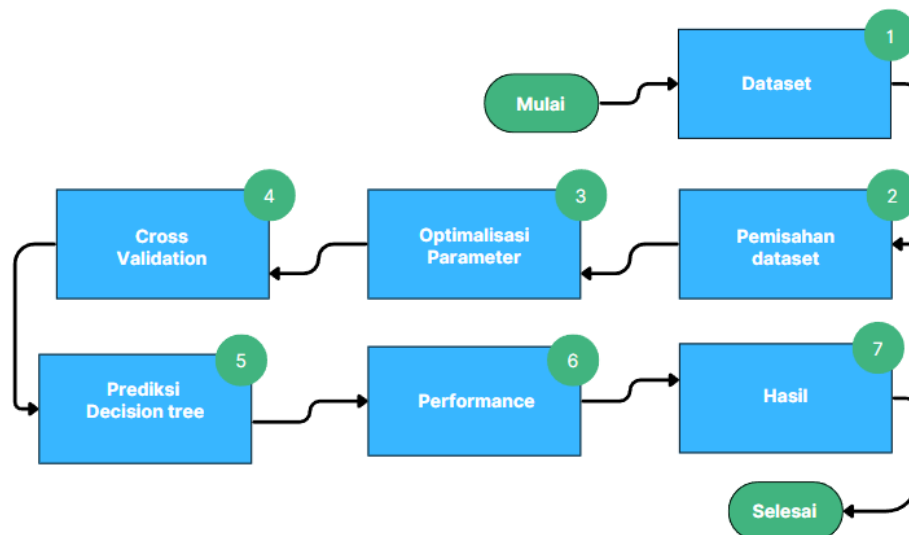
Autoimun menggambarkan bagaimana perubahan yang diinduksi patogen dalam sistem kekebalan tubuh menyebabkan kecenderungan jangka panjang (Ming, 2025).

Decision Tree merupakan metode machine learning yang dapat menghasilkan model prediktif berdasarkan serangkaian keputusan logis. Dalam bidang kesehatan, penggunaan algoritma *Decision Tree* dapat membantu dalam mengklasifikasikan risiko seseorang untuk terkena penyakit jantung berdasarkan data yang relevan (Nafi'ah & Fatah, 2024). Algoritma *Decision Tree* mempunyai kemampuan untuk menangani data yang kompleks dan memberikan hasil yang akurat dengan interpretasi yang mudah dipahami (Fauzi et al., 2025). Pohon keputusan telah diterapkan sejak nol untuk mendukung prediksi waktu nyata menggunakan kumpulan data dinamis. Metode ekstraksi unsur gaya kasar digunakan untuk mengidentifikasi atribut paling berpengaruh yang mempengaruhi atribut dalam dataset (P, 2025).

Optimize Parameters sebagai operator bertingkat melakukan eksekusi terhadap seluruh variasi parameter yang ditetapkan melalui subprosesnya. Setelah proses evaluasi selesai, operator ini akan mengoutputkan nilai parameter optimal via port parameter set yang disediakan.

Autoimun akibat obat merupakan reaksi obat merugikan terkait kekebalan non-IgE yang menimbulkan tantangan substansial dalam toksikologi prediktif karena sifatnya yang idiosinkratik, patogenesis kompleks, dan manifestasi klinis yang beragam (Huang et al., 2025).

Data mengenai autoimun diambil dari *uci machine learning* diprediksi menggunakan algoritma *Decision Tree*, dataset ini terdiri dari 477 baris data yang memiliki sebanyak 195 atribut yang dibutuhkan dalam penelitian ini dengan target label nilai 1 berarti positif obat dan nilai 0 berarti negatif obat (Xiaojie Huang, 2025).



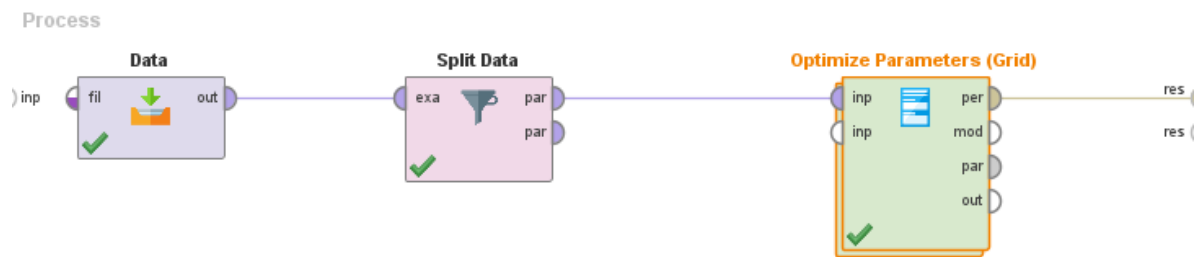
Gambar 1. Langkah-langkah penelitian

Langkah-langkah dalam memprediksi dalam penelitian ini sebagai berikut: diawali dengan pembacaan dataset yang dikumpulkan dari *uci machine learning* kemudian dataset dibagi menjadi data training dan *data testing* menggunakan *split data*, setelah terpisah data tersebut dilakukan optimasi parameter (grid) dengan pemilihan parameter: *Decision Tree minimal gain* dengan range minimal 0,01, maksimum 1 dan langkah 50, dengan parameter tersebut akan tercipta sebanyak 204 kombinasi. Pada prosesnya akan dijalankan untuk semua kombinasi nilai-nilainya berulang-ulang sampai 204 kali. Didalam optimasi parameter akan dilaksanakan *cross validation* dan diambil log. Didalam *cross validation* dilaksanakan prediksi menggunakan algoritma *Decision Tree*, *Decision Tree* digunakan untuk membangun model, kemudian *Apply Model* mengeksekusinya pada *data testing* atau input

baru. Bagian terakhir adalah Performance yang menganalisis kualitas model lewat perhitungan metrik performa seperti tingkat akurasi.

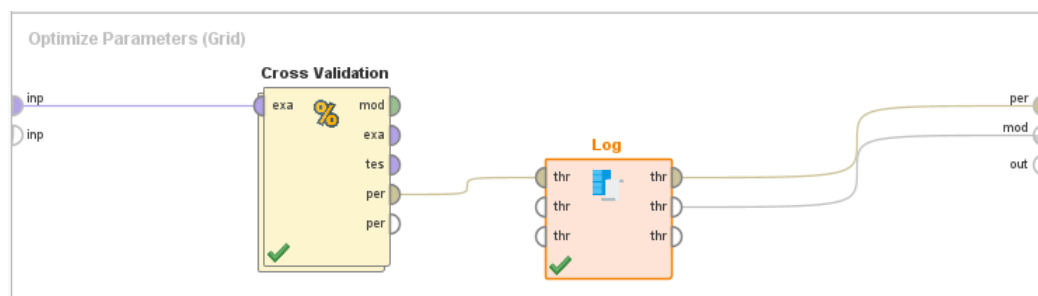
HASIL DAN PEMBAHASAN

Penelitian ini menggunakan rapidminer yang dimulai dengan pembacaan data csv yang dikumpulkan dari *uci machine learning*, kemudian dilakukan pemecahan data (*split data*) menjadi dua bagian, yaitu data *training* dan *data testing* yang nantinya akan digunakan sebagai pembanding pada saat dilakukan prediksi dengan *Decision Tree*. Selanjutnya akan dilakukan optimasi parameter untuk mendukung hasil akurasi prediksi.



Gambar 2. Langkah optimasi parameter (grid)

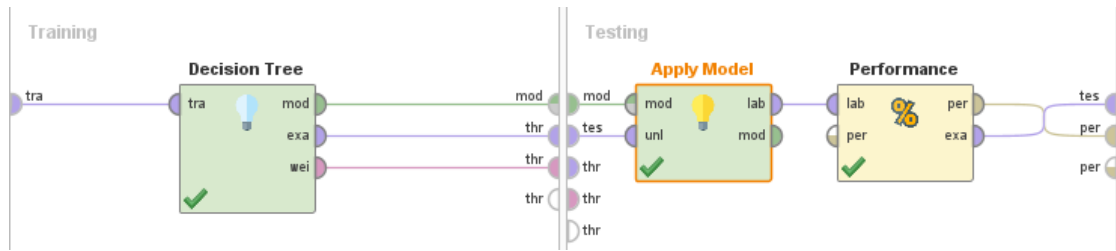
Dalam proses optimasi parameter dilakukan *cross validation*, *cross validation* digunakan untuk menguji kinerja algoritma atau model *Decision Tree* yang digunakan dalam penelitian ini. *Cross validation* dapat digunakan untuk menyempurnakan parameter model statistik dan *machine learning*, untuk mencegah *overfitting*, untuk membandingkan algoritme pembelajaran, dan untuk memperkirakan kesalahan generalisasi model prediktif (Berrar, 2025). *Cross validation* digunakan dalam eksperimen untuk mendapatkan performa model yang bervariasi, sementara variasi dalam klasifikasi biasanya kecil dan umumnya dikaitkan dengan fluktuasi acak (Chen & Martel, 2023).



Gambar 3. Proses didalam optimasi parameter

Dalam penggunaan *Decision Tree* untuk klasifikasi dengan cara memisahkan nilai yang termasuk dalam kelas yang berbeda, untuk memisahkannya dan mengurangi kesalahan dengan cara yang optimal untuk kriteria parameter yang dipilih. Pembuatan node baru diulang sampai kriteria berhenti terpenuhi. Prediksi untuk label kelas Atribut ditentukan tergantung pada sebagian besar, contoh yang mencapai daun ini selama generasi, sedangkan estimasi untuk nilai numerik diperoleh dengan rata-rata nilai dalam daun. *Decision Tree* menerapkan konsep dasar yaitu membagi dataset menjadi dua bagian. Bagian yang lebih besar digunakan untuk *data training* dan yang lebih kecil digunakan untuk mengujinya. Hasil dari penggunaan operator *Decision Tree* adalah tingkat akurasi yang dapat dilihat dalam gambar 6.

Operator *Apply Model* pertama kali dilatih dalam *data training*, yang seringkali merupakan algoritma pembelajaran. selanjutnya, *Apply Model* dapat diterapkan pada *data testing*. Mempunyai tujuan untuk mendapatkan prediksi data yang tidak terlihat atau untuk mengubah data dengan menerapkan model prapemrosesan. Dataset tempat model diterapkan, harus kompatibel dengan Atribut model. berarti dataset memiliki angka, urutan, jenis, dan peran Atribut yang sama dengan Dataset yang digunakan untuk menghasilkan model.



Gambar 4. Penggunaan algoritma *Decision Tree*

Dalam operator performance digunakan untuk pemrosesan *data training* dan *data testing*, dan menghasilkan tingkat akurasi performance vektor yang dapat dilihat pada gambar 5.

PerformanceVector

```
PerformanceVector:
accuracy: 77.50% +/- 2.64% (micro average: 77.49%)
ConfusionMatrix:
True:   0    1
0:     285   81
1:       5   11
```

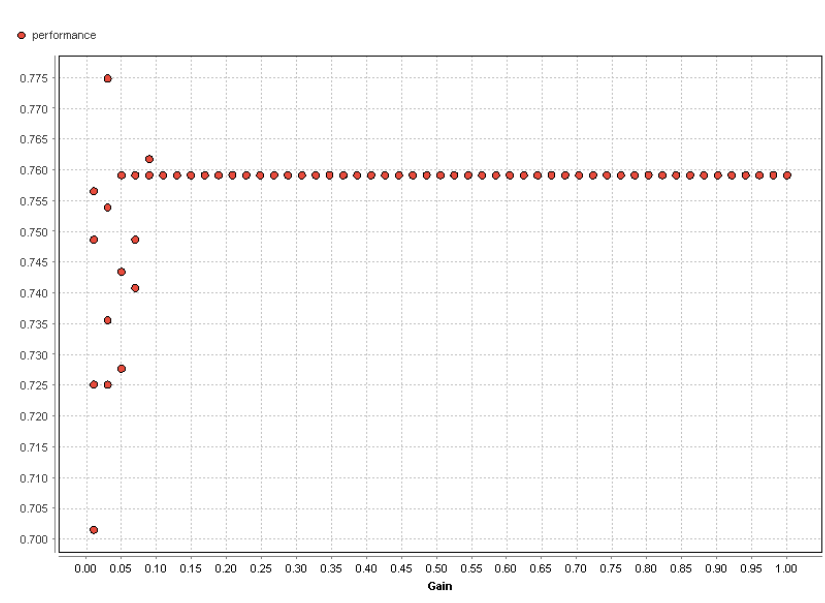
Gambar 5. Performance Vektor

accuracy: 77.50% +/- 2.64% (micro average: 77.49%)

	true 0	true 1	class precision
pred. 0	285	81	77.87%
pred. 1	5	11	68.75%
class recall	98.28%	11.96%	

Gambar 6. Hasil Akurasi prediksi algoritma *Decision Tree*

Dalam penelitian ini digunakan operator log yang bertujuan untuk menyimpan semua informasi selama proses penerapan algoritma *Decision Tree* termasuk berapa banyak pengulangan yang dibutuhkan, jumlah operator yang digunakan termasuk nilai parameter yang digunakan selama proses berlangsung.



Gambar 7. Hasil Log dari performa dan gain

KESIMPULAN

Temuan penelitian mengungkap keunggulan algoritma *Decision Tree* dalam prediksi reaksi autoimun akibat obat. Dengan akurasi tinggi, sistem ini mampu menganalisis risiko berdasarkan variabel-variabel yang ada dalam dataset. Alur pengolahan datanya terstruktur mulai dari pembacaan data, dilanjutkan pemisahan dataset hingga dioptimasi parameternya. Optimasi parameter digunakan untuk eksekusi terhadap seluruh variasi parameter yang ditetapkan melalui subprosesnya. Hasil akhirnya mencatat performa prediktif yang optimal sebesar 77,50% dengan 98,28% lebih presisi untuk kelas “true=0” dibandingkan dengan kelas “true=1”.

DAFTAR PUSTAKA

- Berrar, D. (2025). Cross-Validation. In S. Ranganathan, M. Cannataro, & A. M. Khan (Eds.), *Encyclopedia of Bioinformatics and Computational Biology (Second Edition)* (Second Edition, pp. 638–644). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-323-95502-7.00032-4>
- Chen, J., & Martel, A. (2023). *Cross-Validation Is All You Need: A Statistical Approach To Label Noise Estimation*. <https://doi.org/10.48550/arXiv.2306.13990>
- Fauzi, I., Nurmakarim, H., Pasha, R., & Rozikin, C. (2025). Comparative Study Of Spam Email Classification Decision Tree Between Using Cart And J48. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9, 4032–4036. <https://doi.org/10.36040/jati.v9i3.13533>
- Fortes, A., Ferreira, L., Monteiro, J., Cunha, A., & Dias, C. (2024). Microbiota Transplantation And Its Role In Autoimmune Diseases: Literature Review. *Revista Contemporânea*, 4, e3272. <https://doi.org/10.56083/RCV4N2-042>
- Huang, L., Liu, P., & Huang, X. (2025). InterDIA: Interpretable prediction of drug-induced autoimmunity through ensemble machine learning approaches. *Toxicology*, 511, 154064. <https://doi.org/https://doi.org/10.1016/j.tox.2025.154064>
- Ming, Z. (2025). Pathogen-Induced Autoimmune Imprinting: Immune-Repertoire Alteration Syndrome. *Journal of Mosaic of Autoimmunity*, 1(1), 6. <https://doi.org/10.53941/jmai.2025.100006>
- Nafi'ah, L., & Fatah, Z. (2024). Implementasi Algoritma Decision Tree Untuk Pendeteksian Penyakit Jantung. *JUSIFOR : Jurnal Sistem Informasi Dan Informatika*, 3, 160–165. <https://doi.org/10.70609/jusifor.v3i2.5729>
- P, M. (2025). “A Dual Approach: Naïve Bayes and Decision Tree for Malnutrition Prediction .” *International Journal Of Scientific Research In Engineering And Management*, 09, 1–9. <https://doi.org/10.55041/IJSREM48471>
- Thong, B., Vervloet, D., & Torres Jaen, M. J. (2021). *Drug Allergies*. <https://www.worldallergy.org/component/content/article/drug-allergies-thong-b-vervloet-d-torres-jaen-mj-updated-2021?catid=16&Itemid=101>
- Xiaojie Huang. (2025, May 1). *Drug induced Autoimmunity Prediction*. https://archive.ics.uci.edu/dataset/1104/drug_induced_autoimmunity_prediction