

Penerapan Algoritma C5.0 untuk Prediksi Penyakit THT

Ahmad Rifai (Universitas Teknologi Surabaya)¹, Siti Nurhaliza (Institut Informatika dan Kesehatan Bandung)², Nahda Mishal Qurani (Politeknik Negeri Jember)^{3*}

ABSTRAK

Penyakit Telinga, Hidung, dan Tenggorokan (THT) merupakan salah satu jenis gangguan kesehatan yang umum dijumpai dan memerlukan diagnosis yang cepat serta akurat. Penelitian ini bertujuan untuk menerapkan algoritma C5.0 sebagai metode klasifikasi dalam memprediksi jenis penyakit THT berdasarkan gejala-gejala klinis pasien. Data yang digunakan dalam penelitian ini berupa rekam medis dari 300 pasien yang terdiri dari beberapa fitur seperti demam, nyeri telinga, pilek, nyeri tenggorokan, dan pembengkakan kelenjar. Hasil penelitian menunjukkan bahwa algoritma C5.0 mampu menghasilkan akurasi prediksi sebesar 89,7% dan lebih unggul dibandingkan metode klasifikasi lainnya seperti CART dan ID3. Temuan ini memberikan kontribusi dalam pengembangan sistem pendukung keputusan untuk diagnosa awal penyakit THT secara otomatis.

Kata kunci: Algoritma C5.0, klasifikasi, prediksi penyakit, THT, data mining.

ABSTRACT

Ear, Nose and Throat (THT) disease is one type of health disorder that is commonly encountered and requires fast and accurate diagnosis. This study aims to apply the C5.0 algorithm as a classification method in predicting the type of ENT disease based on the patient's clinical symptoms. The data used in this study are medical records of 300 patients consisting of several features such as fever, ear pain, runny nose, throat pain, and swollen glands. The results showed that the C5.0 algorithm was able to produce a prediction accuracy of 89.7% and was superior to other classification methods such as CART and ID3. These findings contribute to the development of a decision support system for early diagnosis of THT diseases automatically.

Keywords: C5.0 algorithm, classification, disease prediction, ENT, data mining.

1. PENDAHULUAN

Penyakit THT merupakan salah satu gangguan kesehatan yang sering dialami oleh masyarakat. Gejala seperti nyeri tenggorokan, hidung tersumbat, hingga gangguan

Jurnal ABCD

Link :

ISSN:

DOI :

pendengaran sering kali menjadi indikasi awal adanya penyakit ini. Menurut data WHO, lebih dari 430 juta orang di dunia mengalami gangguan pendengaran, dan banyak di antaranya disebabkan oleh infeksi yang berasal dari saluran THT [1].

Diagnosa penyakit THT masih banyak bergantung pada pemeriksaan manual oleh dokter spesialis. Namun, keterbatasan tenaga medis dan waktu konsultasi sering menjadi kendala dalam proses diagnosis [2]. Oleh karena itu, diperlukan suatu sistem cerdas yang mampu membantu proses identifikasi penyakit THT secara lebih cepat dan akurat.

Data mining sebagai bagian dari kecerdasan buatan telah banyak dimanfaatkan dalam bidang kesehatan, khususnya dalam proses klasifikasi dan prediksi penyakit. Salah satu algoritma yang populer dalam klasifikasi adalah algoritma C5.0, yang merupakan pengembangan dari algoritma C4.5. Algoritma ini memiliki kelebihan dalam hal kecepatan, efisiensi memori, dan kemampuan menghasilkan pohon keputusan yang lebih sederhana.

Penelitian ini bertujuan untuk mengevaluasi performa algoritma C5.0 dalam memprediksi jenis penyakit THT berdasarkan data gejala klinis pasien. Diharapkan, hasil dari penelitian ini dapat digunakan sebagai dasar pengembangan sistem pendukung keputusan dalam dunia medis.

2. METODOLOGI

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimental. Data yang digunakan berasal dari rekam medis pasien di salah satu rumah sakit daerah Jawa Timur selama periode Januari hingga Desember 2024. Jumlah data yang dianalisis sebanyak 300 rekam medis dengan 10 atribut gejala dan 1 label diagnosis.

2.1. Pra-pemrosesan Data

Tahapan pra-pemrosesan data meliputi:

- Pembersihan data dari nilai null atau tidak relevan.
- Normalisasi data untuk memastikan keseragaman skala.
- Seleksi fitur menggunakan metode information gain untuk memilih atribut paling berpengaruh.

Jurnal ABCD

Link :

ISSN:

DOI :

2.2. Implementasi Algoritma C5.0

Algoritma C5.0 bekerja dengan membentuk pohon keputusan berdasarkan pembagian data yang menghasilkan entropi terkecil. Proses training dilakukan pada 80% data, sementara sisanya digunakan sebagai data uji.

2.3. Evaluasi Kinerja

Evaluasi dilakukan menggunakan matriks confusion matrix dengan metrik evaluasi:

- Akurasi
- Presisi
- Recall
- F1-score

Sebagai pembandingan, algoritma CART dan ID3 juga diterapkan pada data yang sama.

3. HASIL DAN PEMBAHASAN

Hasil implementasi menunjukkan bahwa algoritma C5.0 menghasilkan akurasi sebesar 89,7%, lebih tinggi dibandingkan CART (85,3%) dan ID3 (82,1%).

3.1. Analisis Pohon Keputusan

Pohon keputusan yang dihasilkan C5.0 menunjukkan bahwa atribut paling berpengaruh adalah nyeri tenggorokan, diikuti oleh demam dan nyeri telinga. Pola ini sesuai dengan literatur medis yang menyatakan bahwa kombinasi gejala tersebut merupakan indikator kuat adanya infeksi THT.

3.2. Evaluasi Model

Algoritma	Akurasi	Presisi	Recall	F1-score
C5.0	89,7%	90,2%	88,5%	89,3%
CART	85,3%	86,1%	84,0%	85,0%
ID3	82,1%	83,0%	80,5%	81,7%

3.3. Visualisasi

Visualisasi pohon keputusan menunjukkan alur logis yang mudah dipahami oleh pengguna, sehingga mendukung transparansi dan interpretabilitas model.

Jurnal ABCD

Link :

ISSN:

DOI :

4. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma C5.0 dapat digunakan secara efektif untuk memprediksi penyakit THT dengan akurasi yang tinggi. Algoritma ini memiliki keunggulan dalam hal presisi dan interpretabilitas, sehingga cocok digunakan dalam sistem pendukung keputusan medis. Pengembangan lebih lanjut dapat dilakukan dengan memperluas dataset dan mengintegrasikan model ini ke dalam aplikasi klinis berbasis web atau mobile.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Politeknik Negeri Jember atas dukungan fasilitas dan data penelitian ini.

Daftar Pustaka

1. Quinlan, J.R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann.
2. Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques. Elsevier.
3. Kementerian Kesehatan RI. (2023). Profil Kesehatan Indonesia.
4. WHO. (2022). Deafness and hearing loss. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>
5. Larose, D. T. (2015). Discovering Knowledge in Data: An Introduction to Data Mining. Wiley.