

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model klasifikasi berbasis pembelajaran mesin menggunakan algoritma ID3 Decision Tree yang diimplementasikan dari scratch untuk memprediksi risiko katarak berdasarkan sepuluh variabel klinis, yaitu usia, riwayat keluarga, kekeruhan lensa, penurunan ketajaman penglihatan, sensitivitas cahaya, perubahan warna mata, penglihatan ganda, tekanan intraokular, hasil pemeriksaan slitlamp, dan visus. Model ini mencapai kinerja yang sangat baik dengan akurasi 95,5%, presisi 0.8686, recall 0.8700, dan F1-score 0.8684, menggunakan Information Gain sebagai splitting criteria dengan struktur pohon yang sederhana (kedalaman maksimum 3 level). Analisis feature importance menunjukkan bahwa Ketajaman_Penglihatan memiliki kontribusi terbesar dengan 56.51%, diikuti oleh Usia_Pasien dengan 35.11%, mengkonfirmasi pentingnya kedua faktor ini dalam diagnosis katarak secara medis. Sistem ini dikembangkan menggunakan antarmuka interaktif yang memungkinkan prediksi secara real-time dilengkapi dengan visualisasi intuitif dan rekomendasi medis. Dengan implementasi algoritma ID3 yang transparan dan interpretable, sistem ini diharapkan dapat menjadi alat bantu yang efisien bagi tenaga medis dalam melakukan diagnosis awal katarak dengan tingkat akurasi tinggi yang mendukung pengambilan keputusan medis berbasis data.

Kata Kunci: Katarak, Klasifikasi, *Machine Learning*, *Random Forest*

ABSTRACT

This research aims to develop a machine learning-based classification model using ID3 Decision Tree algorithm implemented from scratch to predict cataract risk based on ten clinical variables: age, family history, lens opacity, visual acuity decline, light sensitivity, eye color changes, double vision, intraocular pressure, slit-lamp examination results, and visual acuity. The model achieved excellent performance with 95.5% accuracy, 0.8686 precision, 0.8700 recall, and 0.8684 F1-score, utilizing Information Gain as splitting criteria with a simple tree structure (maximum depth of 3 levels). Feature importance analysis revealed that Visual Acuity had the largest contribution at 56.51%, followed by Patient Age at 35.11%, confirming the medical significance of these two factors in cataract diagnosis. The system was developed with an interactive interface that enables real-time prediction accompanied by intuitive visualization and medical recommendations. With the transparent and interpretable implementation of the ID3 algorithm, this system is expected to serve as an efficient tool for medical practitioners in conducting initial cataract diagnosis with high accuracy that supports data-driven medical decision-making.

Keywords: Cataract, Classification, Machine Learning, Random Forest