

ABSTRAK

Tuberkulosis (TB) paru merupakan salah satu penyakit menular yang masih menjadi permasalahan kesehatan sehingga diperlukan upaya deteksi dan klasifikasi yang efektif. Penelitian ini bertujuan untuk mengklasifikasikan penyakit TB paru menggunakan algoritma *Classification and Regression Tree* (CART) dan *Extreme Gradient Boosting* (XGBOOST) serta membandingkan kinerja kedua algoritma tersebut. Data yang digunakan berupa 3.732 data rekam medis pasien dari Rumah Sakit Khusus Paru Sumatera Utara dengan atribut umur, jenis kelamin, demam, batuk lebih dari 2 minggu, berat badan turun, keringat malam, sesak napas, nafsu makan turun, riwayat kontak TB, dan diabetes mellitus. Tahapan penelitian meliputi *preprocessing*, transformasi data, pembagian data *training* dan *testing* dengan rasio 80:20, pembentukan model CART dan XGBOOST, serta evaluasi menggunakan *Accuracy*, *Precision*, *Recall*, *F1-Score*, *Confusion Matrix*, dan *10-Fold Cross Validation*. Hasil penelitian menunjukkan bahwa CART memperoleh *Accuracy* sebesar 69,86%, *Precision* sebesar 72,36%, *Recall* sebesar 58,94%, dan *F1-Score* sebesar 64,96%. Sementara itu, XGBOOST memperoleh *Accuracy* sebesar 70,64%, *Precision* sebesar 74,47%, *Recall* sebesar 57,95%, dan *F1-Score* sebesar 65,18%. Berdasarkan hasil evaluasi, XGBOOST dipilih sebagai model terbaik karena memperoleh *Accuracy*, *Precision*, dan *F1-Score* yang lebih tinggi dibandingkan CART, meskipun nilai *Recall* CART sedikit lebih tinggi. Model XGBOOST selanjutnya diimplementasikan ke dalam sistem berbasis web untuk membantu proses klasifikasi awal TB paru berdasarkan gejala pasien. Penelitian ini menunjukkan bahwa CART dan XGBOOST dapat diterapkan sebagai pendekatan klasifikasi TB paru berbasis data rekam medis.

Kata kunci: Tuberkulosis Paru, Klasifikasi, CART, XGBOOST, *Machine Learning*.