

ABSTRAK

Tuberkulosis (TBC) merupakan salah satu penyakit infeksi paru yang dapat menyerang semua kelompok usia, mulai dari anak-anak hingga lanjut usia, dan prevalensinya masih tinggi di Indonesia. Penyakit ini disebabkan oleh *Mycobacterium tuberculosis* dan ditandai dengan gejala khas seperti batuk berdahak berkepanjangan lebih dari dua minggu, kadang disertai darah, penurunan berat badan, hilang nafsu makan, serta sesak napas. Penelitian ini bertujuan membangun dan mengevaluasi model klasifikasi TBC menggunakan algoritma *Gradient Boosting* berbasis data klinis pasien di Puskesmas Gayamsari. Data yang digunakan terdiri dari 255 observasi pasien dengan dua kelas klasifikasi, yaitu TBC Positif dan TBC Negatif. Faktor-faktor yang digunakan dalam proses klasifikasi meliputi jenis kelamin, umur, demam, penurunan berat badan, hilang nafsu makan, sesak napas, sakit dada, kondisi lemas, lama batuk, dan jenis batuk. Dataset dibagi menjadi data latih dan data uji dengan rasio 80:20. Proses pemodelan mencakup optimasi hyperparameter menggunakan *Grid Search* dan *Optimization Using Tree-structured Parzen Estimator* (OPTUNA). Evaluasi model dilakukan menggunakan *confusion matrix* dan ROC-AUC. Algoritma *Gradient Boosting* dipilih karena kemampuannya dalam menangani hubungan nonlinier antar variabel pada data klinis. Interpretasi model dilakukan menggunakan metode *SHapley Additive exPlanations* (SHAP) untuk mengidentifikasi kontribusi global variabel dan *Local Interpretable Model-Agnostic Explanations* (LIME) untuk penjelasan lokal pada satu observasi pasien. Hasil penelitian menunjukkan bahwa model *Gradient Boosting* menghasilkan *F1-score* awal sebesar 96%, kemudian setelah dilakukan optimasi *hyperparameter* menggunakan *Grid Search* dan OPTUNA, nilai *F1-score* menjadi 95% pada kedua metode dikarenakan parameter awal model *baseline* telah memberikan kinerja yang optimal. Analisis SHAP menunjukkan bahwa variabel lama batuk, umur, hilang nafsu makan, dan penurunan berat badan merupakan faktor yang memberikan kontribusi terbesar dalam proses klasifikasi TBC. Temuan ini menunjukkan bahwa penggunaan algoritma *Gradient Boosting* yang dikombinasikan dengan metode interpretabilitas SHAP dan LIME mampu menghasilkan model klasifikasi yang akurat dan mudah dipahami, sehingga berpotensi mendukung pengambilan keputusan klinis dalam proses diagnosis TBC.

Kata Kunci: Tuberkolosis, *Gradient Boosting*, *Hyperparameter Optimization*, SHAP, LIME, Klasifikasi Klinis