

DAFTAR ISI

| | |
|--|-------------|
| HALAMAN JUDUL | ii |
| HALAMAN PENGESAHAN I | iii |
| HALAMAN PENGESAHAN II..... | iv |
| KATA PENGANTAR..... | v |
| ABSTRAK | vi |
| ABSTRACT | vii |
| DAFTAR ISI..... | viii |
| DAFTAR TABEL | xi |
| DAFTAR GAMBAR..... | xiii |
| DAFTAR LAMPIRAN..... | xv |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah..... | 6 |
| 1.3 Batasan Masalah | 6 |
| 1.4 Tujuan Penelitian | 7 |
| BAB II TINJAUAN PUSTAKA..... | 8 |
| 2.1 Tuberkulosis..... | 8 |
| 2.2 Konsep Data Mining dan <i>Machine Learning</i> | 16 |
| 2.3 Proses <i>Pre-processing</i> | 18 |
| 2.3.1. Penanganan <i>Missing Value</i> | 19 |
| 2.3.1. Encoding Data Kategorikal | 20 |
| 2.3.2. Normalisasi Data (<i>Standard Scaling</i>)..... | 21 |
| 2.4 Pembagian Data Latih (<i>Training Set</i>) dan Data Uji (<i>Testing Set</i>) | 22 |
| 2.5 Penyeimbangan Data dengan SMOTE (<i>Synthetic Minority Over-sampling Technique</i>)..... | 23 |
| 2.6 <i>Classification and Regression Tree</i> (CART)..... | 26 |
| 2.6.1 Pembentukan Pohon Klasifikasi | 28 |
| 2.6.2 Pemangkasan Pohon Klasifikasi (<i>Pruning</i>)..... | 35 |
| 2.7 Konsep Dasar <i>Boosting</i> | 38 |

| | | |
|--|--|-----------|
| 2.8 | <i>Gradient Boosting Classifier</i> | 39 |
| 2.8.1 | Pengertian Algoritma <i>Gradient Boosting</i> | 39 |
| 2.8.2 | Rumus dan Mekanisme <i>Gradient Boosting</i> | 40 |
| 2.8.3 | <i>Hyperparameter</i> Utama pada <i>Gradient Boosting</i> | 50 |
| 2.9 | <i>Hyperparameter Tuning</i> | 51 |
| 2.9.1 | <i>Grid Search</i> | 53 |
| 2.9.2 | OPTUNA (<i>Optimization Using Tree-structured Parzen Estimator</i>) | 54 |
| 2.10 | Evaluasi Kinerja Model Klasifikasi | 55 |
| 2.10.1 | <i>Confussion Matrix</i> | 56 |
| 2.10.2 | Metrik Evaluasi Kinerja Model | 57 |
| 2.11 | Interpretabilitas Model | 59 |
| 2.11.1 | Konsep <i>Explainable AI (XAI)</i> | 60 |
| 2.11.2 | Teori Nilai Shapley | 62 |
| 2.11.3 | SHAP (<i>Shapley Additive Explanations</i>) | 63 |
| 2.11.4 | LIME (<i>Local Interpretable Model-Agnostic Explanations</i>) | 68 |
| BAB III METODE PENELITIAN | | 72 |
| 3.1 | Jenis dan Sumber Data | 72 |
| 3.2 | Variabel Penelitian | 72 |
| 3.3 | Tahapan Analisis Data | 73 |
| 3.4 | Diagram Alir Analisis Data | 75 |
| BAB IV HASIL DAN PEMBAHASAN | | 77 |
| 4.1 | Deskripsi Umum Dataset | 77 |
| 4.2 | Pengecekan <i>Missing Value</i> | 79 |
| 4.3 | Encoding Data Kategorikal | 81 |
| 4.4 | Normalisasi Data dengan <i>Standard Scaling</i> | 81 |
| 4.5 | Pembagian Data Latih (<i>Training Set</i>) dan Data Uji (<i>Testing Set</i>) | 82 |
| 4.6 | Penanganan Data Tidak Seimbang Menggunakan SMOTE | 83 |
| 4.7 | Pembentukan Pohon CART sebagai Dasar <i>Weak Learner</i> pada <i>Gradient Boosting</i> | 85 |
| 4.8 | Pembangunan Model <i>Gradient Boosting</i> | 91 |
| 4.9 | <i>Hyperparameter Tuning</i> Menggunakan <i>Grid Search</i> | 102 |
| 4.10 | <i>Hyperparameter Tuning</i> Menggunakan <i>OPTUNA</i> | 103 |

| | | |
|-----------------------|---|------------|
| 4.11 | Evaluasi Perfoma Model..... | 104 |
| 4.11.1 | <i>Confusion Matrix</i> | 104 |
| 4.11.2 | Metrik Evaluasi..... | 105 |
| 4.11.3 | Kurva <i>Receiver Operating Characteristic</i> (ROC)..... | 107 |
| 4.12 | Interpretabilitas Model <i>Gradient Boosting</i> | 108 |
| 4.12.1 | Interpretasi SHAP pada Model <i>Gradient Boosting</i> | 108 |
| 4.12.2 | Interpretasi LIME pada Model <i>Gradient Boosting</i> | 116 |
| BAB V | PENUTUP | 119 |
| 5.1. | Kesimpulan | 119 |
| 5.2. | Saran | 120 |
| DAFTAR PUSTAKA | | 121 |
| LAMPIRAN | | 124 |