

DAFTAR ISI

HALAMAN PENGESAHAN PEMBIMBING	ii
HALAMAN PENGESAHAN PENGUJI.....	iii
SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR	iv
PERNYATAAN BEBAS PLAGIASI	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Tahapan Penelitian	4
1.7 Sistematika Penulisan	5
BAB II TINJAUAN LITERATUR.....	6
2.1 Penyakit Kardiovaskular dan Polifarmasi.....	6
2.2 Enzim Sitokrom P450 (CYP450).....	6
2.3 Kriteria Aktivitas Biologis dan Pelabelan Data	8
2.4 Kemoinformatika dan Representasi Molekul	9
2.5 Algoritma Machine Learning.....	11
2.5.1 <i>Random Forest</i>	11
2.5.2 <i>Extreme Gradient Boosting (XGBoost)</i>	12
2.5.3 <i>Support Vector Machine (SVM)</i>	13
2.6 Teknik Penyeimbangan dan Optimasi Model	14
2.7 Metrik Evaluasi Performa	17
2.8 Interpretasi Model dengan SHAP	18
2.8.1 Konsep Dasar Shapley Value.....	18
2.8.2 Teknik Visualisasi dan Analisis SHAP	19
2.9 Penelitian Terdahulu.....	20

BAB III METODOLOGI PENELITIAN	23
3.1 <i>Business Understanding</i>	24
3.2 <i>Data Understanding</i>	25
3.2.1 Sumber Data	25
3.2.2 Karakteristik Atribut <i>Dataset</i>	25
3.2.3 Audit Kualitas dan Penyusutan Data	26
3.2.4 Distribusi Kelas dan Analisis Ketidakseimbangan.....	27
3.3 <i>Data Preparation</i>	27
3.3.1 <i>Data Cleaning & Labeling</i>	28
3.3.2 Integrasi dan Deduplikasi Data	28
3.3.3 Ekstraksi Fitur	29
3.3.4 <i>Data Splitting</i>	30
3.3.5 Ilustrasi Penelusuran Data	30
3.4 <i>Modeling</i>	31
3.4.1 Representasi dan Dimensi Fitur.....	31
3.4.2 Algoritma <i>Random Forest</i>	31
3.4.3 Algoritma <i>Extreme Gradient Boosting</i> (XGBoost)	31
3.4.4 Algoritma <i>Support Vector Machine</i> (SVM).....	32
3.4.5 Strategi Penanganan Ketidakseimbangan Data	32
3.4.6 <i>Stratified K-Fold Cross Validation</i>	33
3.4.7 Implementasi <i>Pipeline</i> SMOTE, dan <i>Cross Validation</i>	33
3.4.8 Optimasi <i>Hyperparameter</i> dengan Optuna.....	34
3.5 <i>Evaluation</i>	35
3.5.1 Skenario Pengujian	35
3.5.2 Metrik Kinerja Model.....	36
3.5.3 <i>Decision Threshold Optimization</i>	36
3.6 <i>Interpretation</i> (SHAP Analysis)	37
3.6.1 Pendekatan <i>TreeExplainer</i> untuk Interpretasi Global.....	37
3.6.2 Mekanisme Pemetaan Fitur dan Visualisasi Substruktural.....	37
3.6.3 Validasi Relevansi Biologis	37
3.7 <i>Deployment</i>	38

BAB IV HASIL PENGUJIAN DAN PEMBAHASAN	39
4.1 Implementasi.....	39
4.1.1 Deskripsi Alat dan Lingkungan Eksekusi.....	39
4.1.2 Tahapan Implementasi	40
4.2 Hasil	46
4.2.1 <i>Data Preparation</i>	47
4.2.2 <i>Modeling</i>	48
4.2.3 <i>Evaluation</i>	50
4.2.4 <i>Interpretation</i>	54
4.2.5 <i>Deployment</i>	60
4.3 Pembahasan.....	61
4.3.1 Analisis Komparatif Algoritma dalam Ruang Fitur Dimensi Tinggi	61
4.3.2 Efektivitas <i>Decision Threshold Optimization</i>	62
4.3.3 Analisis Transparansi Logika Model Melalui Global SHAP	63
4.3.4 Analisis Kesalahan dan Batasan Domain Aplikasi.....	64
BAB V PENUTUP	65
5.1 Kesimpulan	65
5.2 Saran	65
DAFTAR PUSTAKA	67
LAMPIRAN	69