

DAFTAR ISI

HALAMAN PENGESAHAN.....	III
HALAMAN PERNYATAAN	IV
DAFTAR GAMBAR	X
DAFTAR TABEL.....	XIII
INTISARI.....	XIV
ABSTRACT.....	XV
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Persamaan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
BAB III LANDASAN TEORI.....	17
3.1 <i>Preprocessing</i>	17
3.2 Feature Selection.....	18
3.3 <i>Ensemble learning</i>	22
3.3.1 <i>Bagging (Bootstrap Aggregating)</i>	22
3.3.2 <i>Boosting</i>	23
3.3.3 <i>Stacking</i>	24
3.1.4 Voting.....	25
3.4 <i>Support Vector Machine (SVM)</i>	27
3.5 <i>Random Forest (RF)</i>	29
3.6 <i>Gradient Boosting Machine (GBM)</i>	30
3.7 XGBoost.....	32
3.8 <i>Confusion matrix</i>	33
3.8.1 Struktur <i>Confusion matrix</i>	33
3.8.2 Metrik Kinerja.....	34
3.9 <i>Explainable Artificial Intelligence (XAI)</i>	35
3.9.1 LIME.....	35
3.9.2 SHAP (<i>SHAPley Additive Explanations</i>)	37
3.9.3 ANCHORS	39
3.10 Validasi Klinis XAI dengan CLIX-M	41
BAB IV METODOLOGI PENELITIAN	43
4.2 Deskripsi Penelitian	43
4.3 Dataset.....	43

4.3.1	Sumber Data	43
4.3.2	Proses Pengumpulan Data	43
4.4	Rancangan Penelitian	46
4.4.1	Eksploratory Data Analisis (EDA).....	47
4.4.2	<i>Split</i> Dataset	48
4.4.3	<i>Data preprocessing</i>	48
4.4.4	<i>Feature Selection</i>	48
4.4.5	Rancangan Pengujian dan Evaluasi Model.....	48
4.4.6	Implementasi XAI	52
4.4.7	Validasi Interpretasi oleh Dokter Ahli	56
4.5	Alat dan Bahan	58
4.5.1	Alat.....	58
BAB V IMPLEMENTASI		59
5.1	<i>Split</i> Dataset	59
5.2	<i>Exploratory Data Analysis</i> (EDA)	59
5.2.1	Data <i>Outlier</i>	59
5.2.2	<i>Missing value</i>	60
5.2.3	Distribusi Label.....	60
5.2.4	Distribusi Data	60
5.2.5	Statistik Numerik	61
5.3	Preprocessing	61
5.3.1	<i>Label enCoder</i>	61
5.3.2	Fungsi Oulier	61
5.3.3	<i>Embedded Feature Selector (Random Forest)</i>	62
5.3.4	<i>Pipeline Feature Selection</i>	62
5.3.5	Fungsi Metrix Evaluasi.....	63
5.3.6	Evaluasi hasil	63
5.4	Penerapan Fitur Terbaik	64
5.5	Tuning Model.....	64
5.6	<i>Ensemble</i> klasifikasi berbasis <i>Soft voting</i>	65
5.7	<i>Ensemble</i> klasifikasi <i>Stacking</i>	66
5.8	Augmentasi Data.....	67
5.8.1	MACR.....	67
5.8.1.1	Simulasi MCAR.....	67
5.8.1.2	Imputation median	67
5.8.1.3	Eksperimen tiga tingkat missing dan ringkasan	68
5.8.1.4	Evaluasi ketahanan model	68
5.8.2	MAR	69
5.8.2.1	Penjelasan Konsep dan Mapping MAR.....	69
5.8.2.2	Simulasi Kehilangan Data dengan Fungsi	70
5.8.2.3	Eksperimen MAR per target.....	71

5.8.2.4	Eksperimen MAR kombinasi (fungsi run_mar_threshold_combo_experiment)	72
5.8.2.5	Ringkasan Statistik dan Evaluasi Stabilitas Model.....	74
5.9	Penerapan XAI	75
5.9.1	LIME.....	75
5.9.2	SHAP	76
5.9.3	ANCHOR.....	77
BAB VI	HASIL DAN PEMBAHASAN.....	79
6.1	<i>Split</i> Dataset	79
6.2	EDA	80
6.2.1	Distribusi data.....	80
6.2.2	Label	81
6.2.3	<i>Missing Value</i>	81
6.3	<i>Outlier</i>	82
6.4	<i>Standarisasi/ Scaling</i>	83
6.5	<i>Feature Selection</i>	84
6.5.1	Filter SelectKBest (ANOVA F-test)	84
6.5.2	<i>Wrapper: Sequential Feature Selector (SFS forward)</i>	85
6.5.3	<i>Metode Embedded</i>	86
6.6	Pemodelan dan Evaluasi	88
6.6.1	SVM.....	88
6.6.2	Model <i>Random Forest</i>	88
6.6.3	<i>Gradient Boosting (GBM)</i>	89
6.6.4	XGBoost	89
6.6.5	<i>Ensemble (Soft voting)</i>	90
6.6.6	<i>Ensemble Stacking</i>	90
6.7	Pemilihan Model yang Diterapkan pada XAI.....	91
6.8	Pengujian <i>Robustness</i> Model	93
6.8.1	MCAR (<i>Missing Completely At Random</i>)	93
6.8.2	MAR (<i>Missing At Random</i>)	95
6.9	Penerapan XAI	96
6.9.1	LIME.....	97
6.9.2	SHAP	98
6.9.3	ANCHOR.....	100
6.10	Modul XAI.....	102
6.10.1	Penerapan XAI.....	102
6.10.2	Hasil Penerapan XAI	103
6.10.3	Hasil Validasi Dokter Ahli	106
6.10.4	Perbedaan Proses Klinis dan Proses Pembuatan Dataset WDBC.....	110
6.10.5	Analisis Overfitting Data	111
BAB VII	KESIMPULAN DAN SARAN	113
DAFTAR PUSTAKA		115