

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMBANG	xiii
INTISARI	xiv
ABSTRACT	xv
I PENDAHULUAN	1
1.1. Tujuan dan Manfaat Penelitian	3
1.2. Tinjauan Pustaka	3
1.3. Metodologi	4
1.4. Sistematika Penulisan	5
II DASAR TEORI	7
2.1. Asuransi	7
2.2. Asuransi Jiwa	8
2.3. Polis	9
2.4. Variabel Random	11
2.4.1. Variabel Random Diskrit	11
2.4.2. Variabel Random Kontinu	12
2.5. Probabilitas	12
2.6. Distribusi Bernoulli	13
2.7. Maximum Likelihood Estimation (MLE)	13
2.8. Regresi Logistik	15
2.9. <i>Machine Learning</i>	18
2.10. <i>Decision Tree</i>	20
2.11. Pertidaksamaan Hoeffding	22
2.12. <i>One-Hot Encoding</i>	23
2.13. <i>Synthetic Minority Oversampling Technique</i>	24

2.14. Normalisasi Data	25
III METODE ENSEMBLE LEARNING	26
3.1. Random Forest	27
3.1.1. <i>Bootstrap Sampling</i>	28
3.1.2. <i>Bagging (Bootstrap Aggregating)</i>	29
3.2. XGBoost	31
3.2.1. <i>Boosting</i>	31
3.2.2. <i>Gradient Boosting</i>	32
3.2.3. <i>Extreme Gradient Boosting (XGBoost)</i>	34
3.3. <i>Stacking</i>	41
3.4. Optimisasi <i>Hyperparameter</i>	42
3.4.1. Hyperparameter pada Random Forests	42
3.4.2. Hyperparameter pada XGBoost	44
3.5. Uji Performa	46
3.5.1. <i>Confusion Matrix</i>	46
3.5.2. <i>Kurva Receiver Operating Characteristic—Area Under the Curve (ROC-AUC)</i>	48
IV STUDI KASUS	50
4.1. Deskripsi Data	50
4.1.1. Statistik Deskriptif <i>Entry Age</i>	51
4.1.2. Statistik Deskriptif <i>Benefit</i>	52
4.1.3. Statistik Deskriptif <i>Premium</i>	53
4.1.4. Statistik Deskriptif <i>Substandard Risk</i>	54
4.1.5. Statistik Deskriptif <i>Duration Months</i>	55
4.2. Pra-Pemrosesan Data	55
4.2.1. <i>Missing Data</i>	56
4.2.2. Deskripsi Data	56
4.2.3. Normalisasi Data	57
4.2.4. Korelasi Antar Variabel	57
4.2.5. <i>One-Hot Encoding</i>	58
4.2.6. Uji Signifikansi Variabel Independen	58
4.2.7. Partisi Data	60
4.2.8. <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	60
4.3. Pemodelan dengan Regresi Logistik	60
4.3.1. Uji Asumsi Regresi Logistik	61
4.3.2. Regresi Logistik	62

4.4.	Pemodelan dengan <i>Ensemble Learning</i>	65
4.4.1.	Random Forest	65
4.4.2.	XGBoost	68
4.4.3.	Stacking	71
4.5.	Perbandingan Performa Model	74
4.6.	<i>Feature Importance</i>	75
V	PENUTUP	78
5.1.	Kesimpulan	78
5.2.	Saran	79
	DAFTAR PUSTAKA	80
A	<i>CODING PYTHON</i>	83