

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	<b>i</b>
<b>HALAMAN PENGESAHAN</b>	<b>ii</b>
<b>HALAMAN PERNYATAAN</b>	<b>iii</b>
<b>HALAMAN PERSEMBAHAN</b>	<b>iv</b>
<b>HALAMAN MOTTO</b>	<b>v</b>
<b>PRAKATA</b>	<b>vi</b>
<b>DAFTAR ISI</b>	<b>viii</b>
<b>DAFTAR TABEL</b>	<b>xii</b>
<b>DAFTAR GAMBAR</b>	<b>xiv</b>
<b>DAFTAR LAMBANG</b>	<b>xvi</b>
<b>INTISARI</b>	<b>xvii</b>
<b>ABSTRACT</b>	<b>xviii</b>
<b>I PENDAHULUAN</b>	<b>1</b>
1.1. Latar Belakang dan Perumusan Masalah	1
1.2. Tujuan dan Manfaat Penelitian	2
1.3. Tinjauan Pustaka	2
1.4. Metodologi Penelitian	4
1.5. Sistematika Penulisan	4
<b>II LANDASAN TEORI</b>	<b>6</b>
2.1. Asuransi Kendaraan Bermotor	6
2.2. <i>Fraud</i> Pada Asuransi	6
2.3. Variabel Acak	7
2.3.1. Variabel Acak Diskrit	7
2.3.2. Variabel Acak Kontinu	8
2.3.3. Properti Nilai Ekspektasi Lainnya	9
2.4. Peluang	9
2.5. <i>Z-Score</i>	10
2.6. Winsorisasi	10
2.7. Seleksi Fitur	11
2.7.1. Uji <i>Chi-Square</i>	11
2.7.2. <i>Mutual Information</i>	12
2.8. <i>One-Hot Encoding</i>	13
2.9. Model Logistik	14

2.9.1. Metode Estimasi Koefisien . . . . .	14
2.9.2. Bentuk Umum Logistik Multivariabel . . . . .	15
2.10. <i>Machine Learning</i> . . . . .	16
2.10.1. <i>Supervised Learning</i> . . . . .	17
2.10.2. <i>Ensemble Learning</i> . . . . .	17
2.11. Pohon Keputusan . . . . .	19
2.11.1. Komponen Dalam Pohon Keputusan . . . . .	19
2.11.2. Pohon Klasifikasi . . . . .	20
2.12. <i>Bagging</i> . . . . .	22
2.13. <i>Boosting</i> . . . . .	23
2.14. Pengaturan <i>Hyperparameter</i> . . . . .	25
2.14.1. <i>Random Search</i> . . . . .	26
2.15. Metrik Evaluasi . . . . .	27
2.15.1. <i>Confusion Matrix</i> . . . . .	27
2.15.2. Akurasi . . . . .	28
2.15.3. Akurasi Seimbang . . . . .	28
2.15.4. <i>Precision</i> . . . . .	28
2.15.5. <i>Recall/Sensitivity</i> . . . . .	28
2.15.6. <i>Specificity</i> . . . . .	29
2.15.7. <i>F1 Score</i> . . . . .	29
2.15.8. AUC dan ROC . . . . .	29
<b>III METODOLOGI PENELITIAN . . . . .</b>	<b>31</b>
3.1. <i>Extreme Gradient Boosting</i> . . . . .	31
3.1.1. Fungsi Objektif . . . . .	32
3.1.2. <i>Gradient Tree Boosting</i> . . . . .	34
3.1.3. <i>Shrinkage and Column Subsampling</i> . . . . .	38
3.1.4. Algoritma Pencarian Pemisahan . . . . .	39
3.1.5. Algoritma XGBoost . . . . .	40
3.1.6. <i>Hyperparameter XGBoost</i> . . . . .	41
3.2. <i>Random Forest</i> . . . . .	43
3.2.1. Algoritma <i>Random Forest</i> . . . . .	43
3.2.2. <i>Hyperparameter Random Forest</i> . . . . .	43
3.3. <i>Explainable Artificial Intelligence</i> . . . . .	44
3.3.1. <i>Shapley Value</i> . . . . .	45
3.3.2. SHAP ( <i>SHapley Additive exPlanations</i> ) . . . . .	49
3.3.3. <i>Tree SHAP</i> . . . . .	51

3.3.4. Interpretasi SHAP . . . . .	53
<b>IV STUDI KASUS . . . . .</b>	<b>55</b>
4.1. Deskripsi Data . . . . .	55
4.1.1. Statistik Deskriptif Variabel Numerik . . . . .	57
4.1.2. Distribusi Jumlah Polis Berdasarkan Variabel Kategori . . . . .	59
4.2. Pra-Pemrosesan Data . . . . .	66
4.2.1. Penanganan Data Hilang dan Penghapusan Variabel Identitas . . . . .	66
4.2.2. Penanganan Pencilan . . . . .	68
4.3. Seleksi Fitur . . . . .	69
4.3.1. Uji <i>Chi-Square</i> . . . . .	69
4.3.2. <i>Mutual Information</i> . . . . .	70
4.4. Pemberian Label Variabel Target . . . . .	71
4.5. <i>One-Hot Encoding</i> Variabel Kategorik . . . . .	72
4.6. Pembagian Data Latih dan Data Uji . . . . .	72
4.7. Permodelan XGBoost . . . . .	73
4.7.1. Model Awal XGBoost . . . . .	73
4.7.2. Pengaturan <i>Hyperparameter</i> XGBoost . . . . .	74
4.7.3. Kinerja Model XGBoost . . . . .	77
4.7.4. Interpretasi SHAP Pada Model XGBoost . . . . .	80
4.8. Permodelan <i>Random Forest</i> . . . . .	83
4.8.1. Model Awal <i>Random Forest</i> . . . . .	83
4.8.2. Pengaturan <i>Hyperparameter Random Forest</i> . . . . .	84
4.8.3. Kinerja Model <i>Random Forest</i> . . . . .	87
4.8.4. Interpretasi SHAP Pada Model <i>Random Forest</i> . . . . .	90
4.9. Perbandingan Model XGBoost dan <i>Random Forest</i> . . . . .	92
4.9.1. Kinerja Model . . . . .	92
4.9.2. Karakteristik Polis Asuransi yang <i>Fraud</i> . . . . .	94
4.10. Analisis Komparatif Model Lain . . . . .	95
4.11. Sistem Peringatan Untuk Deteksi Polis <i>Fraud</i> . . . . .	97
<b>V PENUTUP . . . . .</b>	<b>99</b>
5.1. Kesimpulan . . . . .	99
5.2. Saran . . . . .	99
<b>DAFTAR PUSTAKA . . . . .</b>	<b>100</b>
<b>A PROGRAM PYTHON: PERSIAPAN <i>LIBRARY</i> . . . . .</b>	<b>104</b>
<b>B PROGRAM PYTHON: PRA-PEMROSESAN DATA . . . . .</b>	<b>106</b>
<b>C PROGRAM PYTHON: PERMODELAN XGBOOST . . . . .</b>	<b>116</b>

<b>D</b>	<b>PROGRAM PYTHON: PERMODELAN <i>RANDOM FOREST</i></b>	<b>123</b>
<b>E</b>	<b>PROGRAM PYTHON: SISTEM PERINGATAN DETEKSI <i>FRAUD</i></b>	<b>130</b>
<b>F</b>	<b>PROGRAM PYTHON: PERMODELAN K-NN</b>	<b>138</b>
<b>G</b>	<b>PLOT <i>TREE</i> XGBOOST</b>	<b>142</b>
<b>H</b>	<b>PLOT <i>TREE</i> <i>RANDOM FOREST</i></b>	<b>145</b>