

## 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>xi</b>
<b>DAFTAR GAMBAR</b>	<b>xii</b>
<b>DAFTAR LAMBANG</b>	<b>xiv</b>
<b>INTISARI</b>	<b>xv</b>
<b>ABSTRACT</b>	<b>xvi</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. Pembatasan Masalah	3
1.4. Tinjauan Pustaka	4
1.5. Metodologi Penelitian	5
1.6. Sistematika Penulisan	6
<b>II DASAR TEORI</b>	<b>7</b>
2.1. Kecurangan ( <i>Fraud</i> )	7
2.2. Indikator <i>Fraud</i>	9
2.2.1. Teori Segitiga Kecurangan ( <i>Fraud Triangle</i> )	9
2.2.2. Indikator Kecurangan ( <i>Red Flags</i> )	10
2.2.3. Deteksi Kecurangan ( <i>Fraud Detection</i> )	11
2.3. Mencegah dan Menginvestigasi Kecurangan ( <i>Fraud</i> )	13
2.3.1. Mencegah Kecurangan ( <i>Fraud</i> )	13
2.3.2. Menginvestigasi Kecurangan ( <i>Fraud</i> )	14
2.4. Perasuransian	14
2.5. <i>Machine Learning</i>	15
2.5.1. <i>Supervised Learning</i>	16
2.5.2. <i>Unsupervised Learning</i>	16
2.5.3. <i>Reinforcement Learning</i>	17

2.6. <i>Data Mining</i> . . . . .	17
2.6.1. Teknik <i>Data Mining</i> . . . . .	17
2.6.2. Teknik <i>Sampling</i> . . . . .	18
2.7. Evaluasi Model . . . . .	19
2.8. Teori <i>Framework Data Mining</i> dan Algoritma . . . . .	23
2.8.1. <i>Knowledge Discovery in Databases (KDD)</i> . . . . .	23
2.8.2. <i>Naive Bayes</i> . . . . .	24
2.8.3. <i>Decision Tree</i> . . . . .	25
2.8.4. <i>K-Nearest Neighbor</i> . . . . .	25
2.8.5. <i>Logistic Regression</i> . . . . .	26
<b>III METODE ANALISIS</b> . . . . .	<b>29</b>
3.1. Tahapan Analisis . . . . .	30
3.2. Teknik Pengumpulan Data . . . . .	34
<b>IV HASIL DAN PEMBAHASAN</b> . . . . .	<b>36</b>
4.1. Deskripsi Data . . . . .	36
4.2. Analisis Eksploratif Berfokus pada Variabel Target <i>FraudFound_P</i> . . . . .	46
4.3. <i>Preprocessing</i> Data . . . . .	59
4.4. Pemodelan dan Evaluasi . . . . .	64
4.4.1. Algoritma <i>Naive Bayes</i> . . . . .	64
4.4.2. Algoritma <i>K-Nearest Neighbors (KNN)</i> . . . . .	68
4.4.3. Algoritma <i>Decision Tree</i> . . . . .	72
4.4.4. Algoritma <i>Logistic Regression</i> . . . . .	76
4.5. Perbandingan Performa Algoritma . . . . .	80
4.5.1. Perbandingan <i>Confusion Matrix</i> . . . . .	80
4.5.2. Perbandingan Metrik Evaluasi . . . . .	83
4.5.3. Perbandingan Kurva ROC . . . . .	88
4.6. Pengaruh <i>Oversampling</i> SMOTE terhadap Performa Model . . . . .	92
4.7. Estimasi Regresi Logistik Variabel . . . . .	94
4.8. Visualisasi Proses Kerja Algoritma Terbaik . . . . .	99
<b>V PENUTUP</b> . . . . .	<b>104</b>
5.1. Kesimpulan . . . . .	104
5.2. Saran . . . . .	105
<b>DAFTAR PUSTAKA</b> . . . . .	<b>107</b>
<b>A PROGRAM <i>FLOWCHART</i> ALUR PENELITIAN</b> . . . . .	<b>113</b>
<b>B PROGRAM <i>INPUT</i> DATASET ASLI</b> . . . . .	<b>116</b>
<b>C PROGRAM <i>OVERSAMPLING</i> DATASET</b> . . . . .	<b>117</b>

<b>D</b>	<b>PROGRAM DISTRIBUSI KELAS SEBELUM <i>OVERSAMPLING</i></b>	<b>120</b>
<b>E</b>	<b>PROGRAM DISTRIBUSI KELAS SETELAH <i>OVERSAMPLING</i></b>	<b>121</b>
<b>F</b>	<b>PROGRAM HISTOGRAM DAN BARPLOT VARIABEL</b>	<b>123</b>
<b>G</b>	<b>PROGRAM HUBUNGAN DENGAN VARIABEL FRAUDFOUND_P</b>	<b>128</b>
<b>H</b>	<b>PROGRAM <i>NAIVE BAYES</i> SEBELUM <i>OVERSAMPLING</i></b>	<b>130</b>
<b>I</b>	<b>PROGRAM <i>NAIVE BAYES</i> SETELAH <i>OVERSAMPLING</i></b>	<b>135</b>
<b>J</b>	<b>PROGRAM KNN SEBELUM <i>OVERSAMPLING</i></b>	<b>140</b>
<b>K</b>	<b>PROGRAM KNN SETELAH <i>OVERSAMPLING</i></b>	<b>145</b>
<b>L</b>	<b>PROGRAM <i>DECISION TREE</i> SEBELUM <i>OVERSAMPLING</i></b>	<b>150</b>
<b>M</b>	<b>PROGRAM <i>DECISION TREE</i> SETELAH <i>OVERSAMPLING</i></b>	<b>156</b>
<b>N</b>	<b>PROGRAM <i>LOGISTIC REGRESSION</i> SEBELUM <i>OVERSAMPLING</i></b>	<b>161</b>
<b>O</b>	<b>PROGRAM <i>LOGISTIC REGRESSION</i> SETELAH <i>OVERSAMPLING</i></b>	<b>166</b>
<b>P</b>	<b>PROGRAM PERBANDINGAN AVERAGE PERFORMANCE METRICS MODEL ALGORITMA MACHINE LEARNING SEBELUM <i>OVERSAMPLING</i></b>	<b>171</b>
<b>Q</b>	<b>PROGRAM PERBANDINGAN AVERAGE PERFORMANCE METRICS MODEL ALGORITMA MACHINE LEARNING SETELAH <i>OVERSAMPLING</i></b>	<b>174</b>
<b>R</b>	<b>PROGRAM TABEL PERBANDINGAN AVERAGE PERFORMANCE METRICS</b>	<b>177</b>
<b>S</b>	<b>PROGRAM PERBANDINGAN KURVA ROC MODEL ALGORITMA MACHINE LEARNING SEBELUM <i>OVERSAMPLING</i></b>	<b>179</b>
<b>T</b>	<b>PROGRAM PERBANDINGAN KURVA ROC MODEL ALGORITMA MACHINE LEARNING SETELAH <i>OVERSAMPLING</i></b>	<b>183</b>
<b>U</b>	<b>PROGRAM AVERAGE PERFORMANCE METRICS SETIAP MODEL</b>	<b>187</b>
<b>V</b>	<b>PROGRAM ESTIMASI REGRESI LOGISTIK SEBELUM <i>OVERSAMPLING</i></b>	<b>189</b>
<b>W</b>	<b>PROGRAM ESTIMASI REGRESI LOGISTIK SETELAH <i>OVERSAMPLING</i></b>	<b>191</b>
<b>X</b>	<b>PROGRAM VISUALISASI ALGORITMA <i>DECISION TREE</i> SEBELUM <i>OVERSAMPLING</i></b>	<b>193</b>
<b>Y</b>	<b>PROGRAM VISUALISASI ALGORITMA <i>DECISION TREE</i> SETELAH <i>OVERSAMPLING</i></b>	<b>195</b>