

**IMPLEMENTASI ALGORITMA SMOTE DAN KLASIFIKASI
DECISION TREE UNTUK MENDETEKSI KECURANGAN
TRANSAKSI ONLINE**

SKRIPSI



Oleh :

NORA MINA

20081010042

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2024**

**IMPLEMENTASI ALGORITMA SMOTE DAN KLASIFIKASI
DECISION TREE UNTUK MENDETEKSI KECURANGAN
TRANSAKSI ONLINE**

SKRIPSI

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar
Sarjana Komputer Program Studi Informatika



Oleh :

NORA MINA

20081010042

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2024**

LEMBAR PENGESAHAN SKRIPSI

**Judul : IMPLEMENTASI ALGORITMA SMOTE DAN KLASIFIKASI
DECISION TREE UNTUK MENDETEKSI KECURANGAN
TRANSAKSI ONLINE**

Oleh : NORA MINA

NPM 20081010042

**Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Kamis, Tanggal 04 Juli 2024**

Mengetahui

Dosen Pembimbing

Dosen Penguji

1.



Eka Prakarsa Mandyartha, S.T, M.Kom

NIP. 19880525 2018031 001

1.



Budi Nugroho S.Kom, M.Kom

NIP. 19800907 2021211 005

2.



Agung Mustika Rizki, S.Kom, M.Kom

NIP. 19930725 202203 1008

2.



Andreas Nugroho Sihananto, S.Kom, M.Kom

NPT. 211199 00 412271

Menyetujui

Dekan

Fakultas Ilmu Komputer

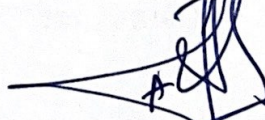


Prof. Dr. Ir. Novirina Hendrasarie, MT.

NIP. 19681126 199403 2 001

Koordinator Program Studi

Informatika



Fetty Tri Anggraeny, S.Kom.M.Kom

NIP. 19820211 2021212 005

SURAT PERNYATAAN BEBAS DARI PLAGIASI

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Nora Mina

NPM : 20081010042

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

“ IMPLEMENTASI ALGORITMA SMOTE DAN KLASIFIKASI *DECISION TREE* UNTUK MENDETEKSI KECURANGAN”

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dari juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.

Surabaya, 02 Juli 2024

Hormat saya,



Nora Mina

NPM. 20081010042



IMPLEMENTASI ALGORITMA SMOTE DAN KLASIFIKASI *DECISION TREE* UNTUK MENDETEKSI KECURANGAN TRANSAKSI ONLINE

Nama Mahasiswa : Nora Mina

NPM : 20081010042

Program Studi : Informatika

Dosen Pembimbing : Eka Prakarsa Mandyartha, ST, M.Kom

Agung Mustika Rizki, S.Kom, M.Kom

Abstrak

Industri keuangan mengalami transformasi signifikan dengan meningkatnya transaksi online melalui platform *e-commerce*, perbankan digital, dan layanan pembayaran online. Namun, bersamaan dengan kenyamanan yang ditawarkan, muncul tantangan serius dalam bentuk kecurangan transaksi online, yang dapat merugikan konsumen dan merusak integritas sistem pembayaran. Deteksi kecurangan transaksi online menjadi krusial untuk melindungi konsumen dan menjaga integritas ekosistem bisnis online. Permasalahan ketidakseimbangan dataset dapat mempengaruhi kinerja prediksi klasifikasi. Untuk menyeimbangkan data, pendekatan *oversampling* yang digunakan dalam penelitian menggunakan metode SMOTE. Tujuan dari penelitian ini ialah menganalisis kinerja algoritma SMOTE dan klasifikasi *decision tree* dalam menangani masalah ketidakseimbangan data transaksi curang. Penelitian dilakukan dengan 3 skenario yang dibedakan dari tanpa dan dengan SMOTE, penambahan kelas minoritas setengah dan sama dengan kelas mayoritas. Hasil penelitian menunjukkan tanpa menggunakan metode SMOTE pada *split* 60% data latih dan 40% data uji mendapatkan nilai *recall* 32% menunjukkan model sangat kurang baik dalam mengklasifikasikan kedua kelas. Sedangkan penggunaan metode SMOTE menunjukkan seluruh nilai metrik sempurna dan stabil pada nilai 99% dan mendapatkan nilai *recall* hingga 100%. Namun dengan nilai metrik yang sempurna penambahan data sintesis dari SMOTE yang digunakan kurang dalam mempertahankan data asli sehingga meningkatkan jumlah *false negatives*. Hasil evaluasi *k-fold cross validation* pada seluruh model skenario pengujian mendapatkan nilai akurasi yang juga tinggi dengan nilai 99%. Maka dapat disimpulkan, pemilihan model ini sangat cocok dalam dataset transaksi online.

Kata kunci: transaksi online, SMOTE, klasifikasi, *Decision Tree*.

KATA PENGANTAR

Segala puja dan puji syukur penulis panjatkan kepada Allah SWT yang Maha Esa atas segala nikmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir skripsi yang berjudul “Implementasi Algoritma SMOTE dan Klasifikasi *Decision Tree* untuk Mendeteksi Kecurangan Transaksi Online”. Penyusunan skripsi ini bertujuan untuk memperoleh gelar Sarjana Informatika pada Program Ilmu Komputer di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Selama penelitian dan penulisan skripsi ini, penulis menghadapi banyak hambatan. Namun, berkat bantuan, dorongan, dan bimbingan dari berbagai pihak, skripsi ini dapat terselesaikan dengan baik.

Penulis menganggap skripsi ini sebagai karya terbaik yang dapat dipersembahkan. Meskipun demikian, penulis menyadari bahwa masih mungkin terdapat kekurangan di dalamnya. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca pada umumnya.

Surabaya, 02 Juli 2024

Penulis

UCAPAN TERIMA KASIH

Segala rasa puja dan puji syukur penulis panjatkan atas kehadiran Allah SWT, yang telah memberikan kesempatan untuk penulis menyelesaikan tugas akhir ini, dengan berjudul “Implementasi Algoritma SMOTE dan Klasifikasi *Decision Tree* untuk Mendeteksi Kecurangan Transaksi Online”. Dan tidak juga lupa shalawat serta salam saya haturkan kepada Nabi besar Nabi Muhammad SAW.

Ada suka maupun duka yang penulis rasakan dalam menyelesaikan tugas akhir ini guna mencapai gelar sarjana (S.Kom) pada Program Studi Informatika. Penulis menggunakan semua kemampuan dan kemauannya yang kuat, upaya yang luar biasa, serta kesabaran yang tiada habisnya dalam menyelesaikan tugas akhir ini. Dan dengan tulus, penulis ingin mengucapkan terima kasih dan penghargaan setinggi-tingginya kepada semua pihak yang telah memberikan bantuan dan bimbingan, sehingga mempermudah penulis dalam menyusun tugas akhir ini, diantaranya:

1. Kepada Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Kepada Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Kepada Ibu Fetty Tri Anggraeny, S.Kom. M.Kom selaku koordinator program studi Informatika dari Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Kepada Bapak Eka Prakarsa Mandyartha, ST, M.Kom, selaku dosen pembimbing pertama yang senantiasa sabar membimbing peneliti dan memberikan dukungan, ilmu, serta waktunya sehingga peneliti bisa menyelesaikan tugas akhir tepat pada waktunya.
5. Kepada Bapak Agung Mustika Rizki, S.Kom, M.Kom, selaku dosen pembimbing kedua yang membantu memberi arahan dan dukungan untuk peneliti lebih memahami dan mendalami dalam pengerjaan tugas akhir.

6. Kepada Ibu Hulah pintu surgaku, beliau sangat berperan penting dalam menyelesaikan tugas akhir ini. Beliau yang selalu memberikan semangat dan doa yang tiada hentinya untuk penulis dapat menyelesaikan penelitian ini. Terima kasih atas cinta, doa, dan kasih sayangnya yang tak terhingga.
7. Kepada Alm Abu Bakar, terima kasih atas dedikasi, motivasi dan dukungan. Semoga karya penelitian ini dapat membanggakan beliau .
8. Kepada seseorang yang tak kalah penting kehadirannya, Fauzia. Terima kasih telah menjadi bagian dari perjalanan hidup saya dan berkontribusi banyak dalam penulisan karya ini, baik tenaga, waktu, maupun materi kepada saya. Terima kasih atas dukungan dan menghibur peneliti, mendengarkan keluh kesah dan memberikan semangat untuk tidak mudah menyerah. Semoga Allah SWT membalas seluruh kebaikan yang sudah diberikan.
9. Kepada seluruh teman-teman seperjuangan dalam menyelesaikan skripsi yang telah menemani proses pengerjaan skripsi, memberi semangat, dan doa kepada penulis.
10. Kepada sahabat SMA peneliti yang selalu memberikan dukungan dan semangat dengan cara meluangkan waktu disela kesibukan yang dimiliki.

DAFTAR ISI

KATA PENGANTAR.....	v
UCAPAN TERIMA KASIH	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan ..	4
1.4. Manfaat	4
1.5. Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA	6
2.1. Penelitian Terdahulu	6
2.2. Landasan Teori.....	8
2.2.1. Transaksi Online	8
2.2.2. Kecurangan	9
2.2.3. Data Mining	9
2.2.4. <i>Exploratory Data Analysis</i>	10
2.2.5. Klasifikasi	11
2.2.6. Phyton	11
2.2.7. Ketidakseimbangan Data	12
2.2.7. SMOTE	13
2.2.8. <i>Decision Tree</i>	14
2.2.9. <i>Confusion Matrix</i>	15
2.2.10. <i>K-fold Cross Validation</i>	17
BAB III METODOLOGI	19
3.1. Alur Penelitian	19
3.1.1. Studi Literatur.....	19
3.1.2. Pengambilan Data.....	19
3.1.3. Olah Data	19
3.1.4. Pengujian.....	20

3.2. Metode yang Diusulkan	20
3.2.1. Dataset	20
3.2.2. <i>Exploratory Data Analysis</i> (EDA)	22
3.2.3. Preprocessing Data	23
3.2.4. SMOTE (<i>Synthetic Minority Oversampling Technique</i>)	23
3.2.5. <i>Splitting Data</i>	26
3.2.6. Klasifikasi <i>Decision Tree</i>	26
3.2.7. Cofusion Matrix	30
3.2.8. Evaluasi.....	30
3.3. Skenario Pengujian	31
BAB IV HASIL DAN PEMBAHASAN	32
4.1. Hasil.....	32
4.1.1. Data.....	32
4.1.2. Exploratory Data Analysis (EDA).....	32
4.1.3. Preprocessing Data.....	43
4.1.4. SMOTE (<i>Synthetic Minority Oversampling Technique</i>)	46
4.1.5. <i>Splitting Data</i>	58
4.1.6. Klasifikasi <i>Decision Tree</i>	58
4.1.7. Evaluasi.....	79
4.2. Pembahasan.....	87
4.2.1. Analisis Parameter <i>Decision Tree</i>	88
4.2.2. Analisis Berdasarkan Pohon Keputusan	88
4.2.3. Analisis Berdasarkan Nilai Metrik Performa.....	89
4.2.4. Analisis Berdasarkan <i>false positives</i> dan <i>false negatives</i>	90
4.2.5. Analisis Penambahan <i>missing value</i>	91
4.2.6. Analisis Model Pemandangan	92
BAB V PENUTUP.....	94
5.1. Kesimpulan	94
5.2. Saran	96
DAFTAR PUSTAKA.....	97

DAFTAR TABEL

Tabel 3.1 Kolom Dataset	21
Tabel 3.2 Jumlah Data	22
Tabel 3.3 Contoh Dataset	28
Tabel 4.1 Jumlah Sampel SMOTE	56
Tabel 4.2 Hasil Nilai Metrik Evaluasi Skenario 1	74
Tabel 4.3 Hasil Nilai Metrik Evaluasi Skenario 1	76
Tabel 4.4 Hasil Nilai Metrik Evaluasi Skenario 3	78
Tabel 4.5 Evaluasi Skenario 1	82
Tabel 4.6 Evaluasi Skenario 2	84
Tabel 4.7 Evaluasi Skenario 3	86
Tabel 4.8 Analisis Model Perbandingan.....	86

DAFTAR GAMBAR

Gambar 2.1 Visualisasi Undersampling & Oversampling	13
Gambar 2.2 Iterasi k-Fold Cross Validation	18
Gambar 3.1 Alur Penelitian	19
Gambar 3. 2 Metode yang Diusulkan	20
Gambar 3.3 Diagram Plot Dataset Tidak Seimbang	22
Gambar 3.4 Missing Value Data.....	23
Gambar 4.1 Dataset.....	32
Gambar 4.2 Hasil Ubah Nama Kolom Dataset	33
Gambar 4.3 Hasil Ubah Nilai Kolom Penerima	34
Gambar 4.4 Diagram Plot Kategori Penerima	34
Gambar 4.5 Diagram Plot Jenis Pembayaran	35
Gambar 4.6 Diagram Plot Jenis Pembayaran Transaksi Curang	36
Gambar 4.7 Diagram Plot Jumlah Nominal Transaksi Curang	37
Gambar 4.8 Data Transaksi Tidak Sesuai	40
Gambar 4.9 Hasil Update Data yang Sesuai	41
Gambar 4.10 Korelasi Antar Variabel	42
Gambar 4.11 Hasil Data dengan EDA	43
Gambar 4.12 Hasil Preprocessing	46
Gambar 4.13 Flowchart inti SMOTE.....	50
Gambar 4.14 Diagram Batang Sampel Rasio 0.5	56
Gambar 4.15 Diagram Batang Sampel Rasio 1.0	57
Gambar 4.16 Data Resampled Tidak Curang	57
Gambar 4.17 Tree Skenario 1	74
Gambar 4.18 Pemeriksaan Overfitting Skenario 1	75
Gambar 4.19 Tree Skenario 2	76
Gambar 4.20 Pemeriksaan Overfitting Skenario 2	77
Gambar 4.21 Tree Skenario 3	78
Gambar 4.22 Pemeriksaan Overfitting Skenario 3	79
Gambar 4.23 Confusion Matrix k = 3 (1)	82
Gambar 4.24 Confusion Matrix k = 5 (1)	83
Gambar 4.25 Confusion Matrix k = 7 (1)	83
Gambar 4.26 Confusion Matrix k = 3 (2)	84
Gambar 4.27 Confusion Matrix k = 5 (2)	85
Gambar 4.28 Confusion Matrix k = 7 (2)	85

Gambar 4.29 Confusion Matrix $k = 3$ (3)	86
Gambar 4.30 Confusion Matrix $k = 5$ (3)	87
Gambar 4.31 Confusion Matrix $k = 7$ (3)	87
Gambar 4.32 Hasil Klasifikasi	88
Gambar 4.33 Report Klasifikasi Skenario 1 Split ke 2	89
Gambar 4.34 Report Klasifikasi Skenario 2 Split ke 2	90
Gambar 4.35 Diagram perbandingan FP FNSkenario	91
Gambar 4.36 Hasil Klasifikasi tanpa SMOTE	92
Gambar 4.37 Hasil Klasifikasi dengan SMOTE	92