

**DETEKSI ANOMALI PADA GENERATIF AUGMENTASI DATA  
TRANSAKSI KEUANGAN DENGAN MENGGUNAKAN  
GENERATIVE ADVERSARIAL NETWORKS**

**TESIS**

Oleh :

**GEGE SATYA PUTRA  
NIM. 211231020**



**UNIVERSITAS  
MIKROSKIL**

**PROGRAM STUDI S-2 TEKNOLOGI INFORMASI  
FAKULTAS INFORMATIKA  
UNIVERSITAS MIKROSKIL  
MEDAN  
2024**

**ANOMALY DETECTION ON GENERATIVE AUGMENTATION OF  
FINANCIAL TRANSACTION DATA USING GENERATIVE  
ADVERSARIAL NETWORKS**

**THESIS**

**By :**

**GEGE SATYA PUTRA  
NIM. 211231020**



**UNIVERSITAS  
MIKROSKIL**

**MAJOR OF S-2 INFORMATION TECHNOLOGY  
FACULTY OF INFORMATICS  
UNIVERSITAS MIKROSKIL  
MEDAN  
2024**

LEMBARAN PENGESAHAN

DETEKSI ANOMALI PADA GENERATIF AUGMENTASI DATA  
TRANSAKSI KEUANGAN DENGAN MENGGUNAKAN  
GENERATIVE ADVERSARIAL NETWORKS

TESIS

Diajukan untuk Melengkapi Persyaratan Guna  
Mendapatkan Gelar Magister  
Program Studi S-2 Magister Teknologi Informasi

Oleh:

GEGE SATYA PUTRA  
NIM. 211231020

Disetujui Oleh:

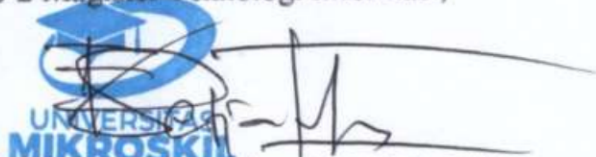
Dosen Pembimbing,



Dr. Ronsen Purba, M.Sc.

Medan, 13 Agustus 2024  
Diketahui dan Disahkan Oleh:

Ketua Program Studi  
S-2 Magister Teknologi Informasi,



Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D.

## HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-2 Magister Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Gege Satya Putra  
NIM : 211231020

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir / Tesis \*) dengan judul dan tempat penelitian sebagai berikut:

Judul Tugas Akhir / Tesis \*) : DETEKSI ANOMALI PADA GENERATIF AUGMENTASI DATA TRANSAKSI KEUANGAN DENGAN MENGGUNAKAN GENERATIVE ADVERSARIAL NETWORKS

Tempat Penelitian : -  
Alamat Tempat Penelitian : -  
No. Telp. Tempat Penelitian : -

Sehubungan dengan Tesis tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 13 Agustus 2024  
Saya yang membuat pernyataan,



Gege Satya Putra

# DETEKSI ANOMALI PADA GENERATIF AUGMENTASI DATA TRANSAKSI KEUANGAN DENGAN MENGGUNAKAN GENERATIVE ADVERSARIAL NETWORKS

## Abstrak

*Seiring kemajuan teknologi, perkembangan kejahatan pada transaksi kartu kredit juga semakin mutakhir. Deteksi fraud terhadap kejahatan transaksi kartu kredit sudah dilakukan dengan menggunakan machine learning, namun memiliki beberapa tantangan dalam hal ketidakseimbangan data. Penelitian ini mengusulkan pendekatan dengan menggunakan generative adversarial networks untuk melakukan augmentasi data. Dengan pendekatan ini, diharapkan dapat menghasilkan kualitas data sintetis yang mendekati karakteristik data aslinya dan dapat mengatasi masalah ketidakseimbangan data. Untuk melakukan deteksi anomali, digunakan model autoencoder pada dataset yang sudah dikombinasi dengan data sintetis. Model autoencoder yang menggunakan data hasil augmentasi mengalami peningkatan pada classification report dan performa metrik. Pada perbandingan data normal terhadap fraud sebesar 90% memiliki hasil paling baik dengan nilai precision 96%, recall 90%, f1-score 93% dan accuracy 96%. Untuk performa model menghasilkan nilai ROC sebesar 94%, MCC sebesar 91% dan G-mean sebesar 94%.*

**Kata kunci:** fraud, anomali, augmentasi, generatif, GAN, Autoencoder

## Abstract

*As technology advances, the development of crime in credit card transactions is also increasingly sophisticated. Fraud detection of credit card transaction crimes has been done using machine learning, but it has some challenges in terms of data imbalance. This research proposes an approach using generative adversarial networks to perform data augmentation. With this approach, it is expected to produce synthetic data quality that is close to the characteristics of the original data and can overcome the problem of data imbalance. To perform anomaly detection, an autoencoder model is used on the dataset that has been combined with synthetic data. The autoencoder model that uses augmented data has improved classification report and metric performance. In the comparison of normal data to fraud at 90%, the best results are precision 96%, recall 90%, f1-score 93% and accuracy 96%. For model performance, the ROC value is 94%, MCC is 91% and G-mean is 94%.*

**Keywords:** fraud, anomaly, augmentation, generative, GAN, Autoencoder

## KATA PENGANTAR

Ucapan syukur kepada Allah Tuhan Yang Maha Esa karena berkat rahmatnya penulis bisa menyelesaikan penelitian tesis yang berjudul “DETEKSI ANOMALI PADA GENERATIF AUGMENTASI DATA TRANSAKSI KEUANGAN DENGAN MENGGUNAKAN GENERATIVE ADVERSARIAL NETWORKS”.

Tesis ini merupakan salah satu syarat untuk mendapatkan gelar Magister Teknologi Informasi (M.T.I) pada program studi S-2 Magister Teknologi Informasi Universitas Mikroskil Medan.

Dalam penulisan tesis ini penulis banyak menerima bantuan, bimbingan, nasehat, dukungan dan dorongan semangat dari berbagai pihak. Oleh karena itu ucapan terima kasih yang tidak terhingga penulis sampaikan kepada”

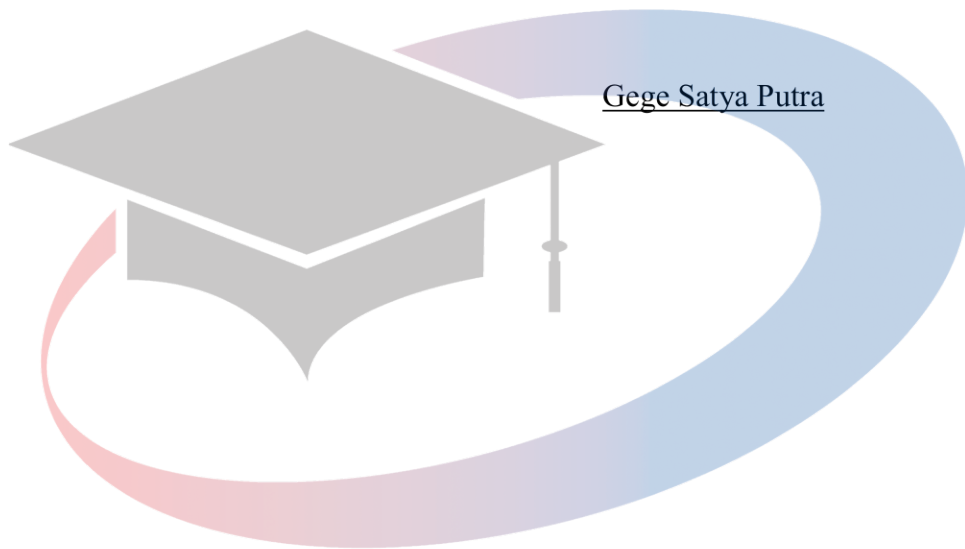
1. Bapak Dr. Ronsen Purba, M.Sc. selaku Dosen Pembimbing yang telah meluangkan waktu, pemikiran, saran penulisan dan memberikan kesempatan yang besar dalam penyelesaian tesis.
2. Bapak Muhammad Fermi Pasha, B.Sc., M.Sc.,PhD, selaku Pendamping Pembimbing yang telah meluangkan waktu dan memberikan kesempatan yang besar dalam penyelesaian tesis ini.
3. Bapak Dr. Farisya Setiadi, S.T., M.T.I., selaku Dosen Penguji yang telah meluangkan waktu dan memberikan uji dan kritik ilmiah pada tesis ini.
4. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
5. Bapak Ng Poi Wong, S.Kom., M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
6. Bapak Erwin Setiawan Panjaitan, M.M.S.I., Ph.D., selaku Ketua Program Studi S-2 Magister Teknologi Informasi Fakultas Informatika Universitas Mikroskil Medan.
7. Orang tua tercinta Bapak Alm. H. Abdul Djalil Basnury, Ibu Hj. Kartini Simatupang yang selalu membisikkan doa terbaik untuk anak-anaknya. Keluarga dan Saudara kandung Miranda, Meutia, Moslem, Alm Roeslan, Lulu dan Tonna. Juga kepada istri tercinta Juliarni dan anakku Muhammad Aga Putra
8. Bapak/Ibu Dosen, Staff dan teman-teman Angkatan Tahun 2021 Universitas Mikroskil yang senantiasa memberikan dorongan dan semangat
9. Ibu Direktur selaku Pimpinan dan teman-teman di Politeknik Wilmar Bisnis Indonesia yang telah memberikan semangat dan motivasi

10. Kepada pihak-pihak yang tidak bisa penulis sebutkan satu per satu yang telah memberikan bantuan, baik secara langsung maupun tidak langsung

Penulis menyadari penelitian tesis ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang bersifat membangun sangat penulis harapkan dari para pembaca. Akhir kata, penulis berharap semoga penelitian tesis ini bermanfaat bagi para pembacanya.

Medan, 13 Agustus 2024

Penulis,



Gege Satya Putra

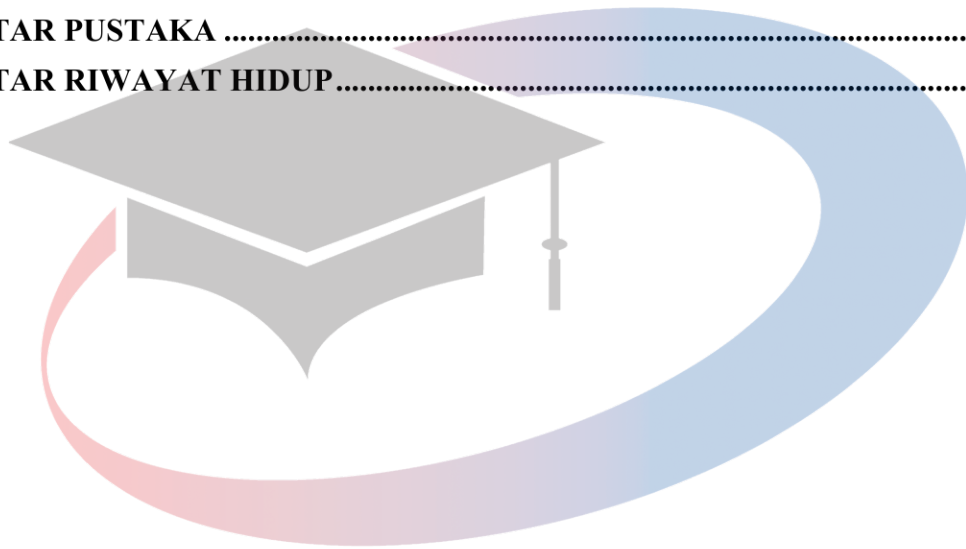
UNIVERSITAS  
MIKROSKIL

# DAFTAR ISI

<b>ABSTRAK</b> .....	<b>i</b>
<b>KATA PENGANTAR</b> .....	<b>ii</b>
<b>DAFTAR ISI</b> .....	<b>iv</b>
<b>DAFTAR GAMBAR</b> .....	<b>vi</b>
<b>DAFTAR TABEL</b> .....	<b>vii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan .....	3
1.4 Manfaat .....	3
1.5 Ruang Lingkup .....	4
<b>BAB II KAJIAN LITERATUR</b> .....	<b>5</b>
2.1 Tinjauan Pustaka .....	5
2.1.1 <i>Fraud</i> pada Transaksi Kartu Kredit .....	5
2.1.2 Deteksi <i>Fraud</i> Menggunakan <i>Machine Learning</i> (ML).....	6
2.1.3 <i>Imbalanced Data</i> (Ketidakseimbangan Data) .....	7
2.1.4 GAN dan CTGAN .....	8
2.1.5 <i>Autoencoders</i> (AE).....	10
2.2 Penelitian Terdahulu .....	11
2.3 Kerangka Konseptual .....	14
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>15</b>
3.1 Analisis Masalah .....	15
3.2 Rancangan Penelitian .....	17
3.3 Data Yang Digunakan .....	20
3.4 Alat-alat Penelitian.....	22
3.5 Teknik Analisis Data.....	23
3.5.1 Kombinasi dan Perubahan Nilai <i>hyperparameter</i> .....	23
3.5.2 Evaluasi Model.....	24
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>27</b>
4.1 Hasil .....	27
4.1.1 Membaca <i>Dataset</i> .....	27
4.1.2 Normalisasi <i>Dataset</i> .....	30
4.1.2 Augmentasi Data .....	34
4.1.3 Pemodelan <i>Autoencoder</i> .....	35
4.1.4 Deteksi <i>Fraud</i> Pada <i>Dataset</i> Asli ( <i>Highly Imbalanced</i> ) .....	39



4.1.5 Deteksi <i>Fraud</i> Pada <i>Dataset</i> Dengan Generatif Data Menggunakan CTGAN.	46
4.2 Pembahasan.....	48
4.2.1 Evaluasi Hasil <i>Classification Report</i> ( <i>Precision</i> , <i>Recall</i> , <i>F1-Score</i> dan <i>Accuracy</i> ) .....	48
4.2.2 Evaluasi Performa Model.....	50
4.2.2 Perbandingan Model Autoencoder Dengan DNN dan TCN.....	51
4.2.3 Evaluasi <i>Confussion Matrix</i> .....	52
<b>BAB V PENUTUP .....</b>	<b>54</b>
5.1 Kesimpulan .....	54
5.2 Saran.....	54
<b>DAFTAR PUSTAKA .....</b>	<b>55</b>
<b>DAFTAR RIWAYAT HIDUP .....</b>	<b>59</b>



# UNIVERSITAS MIKROSKIL

## DAFTAR GAMBAR

Gambar 2. 1 Data Fraud Secara Global.....	5
Gambar 2. 2 Arsitektur <i>Conditional Tabular</i> GAN [6].....	9
Gambar 2. 3 Arsitektur Jaringan Saraf Tiruan <i>Autoencoders</i> [27].....	10
Gambar 2. 4 Kerangka Konseptual.....	14
Gambar 3. 1 Kerugian Akibat Tindakan Fraud pada Transaksi Kartu Kredit [35].....	15
Gambar 3. 2 Tahapan Penelitian.....	18
Gambar 3. 3 Bagan Alir Klasifikasi <i>Fraud</i> dengan <i>Autoencoders</i> .....	20
Gambar 4. 1 Komposisi Tipe Data .....	28
Gambar 4. 2 Menampilkan Data dari <i>Dataset</i> .....	28
Gambar 4. 3 Proporsi Data Normal dan Data <i>Fraud</i> .....	29
Gambar 4. 4 Proses Identifikasi <i>Missing Values</i> .....	30
Gambar 4. 5 <i>Code Python</i> Ekstraksi <i>hour</i> pada kolom <i>trans_date_trans_time</i> .....	31
Gambar 4. 6 Aktifitas Transaksi Normal/ <i>Fraud</i> Berdasarkan Jam.....	31
Gambar 4. 7 Distribusi Umur dan Aktifitas Normal/ <i>Fraud</i> .....	32
Gambar 4. 8 Tampilan <i>Dataset</i> Setelah Normalisasi .....	34
Gambar 4. 9 Hasil Akhir Normalisasi <i>Dataset</i> .....	34
Gambar 4. 10 <i>Listing Code</i> Arsitektur Model 1 <i>Encoder-Decoder</i> .....	36
Gambar 4. 11 <i>Listing Code</i> Arsitektur Model 2 <i>Encoder-Decoder</i> .....	37
Gambar 4. 12 <i>Listing Code</i> <i>Autoencoder</i> Model 1 Dengan <i>Hyperparameter Regularization</i> dan <i>Dropout</i> .....	38
Gambar 4. 13 <i>Listing Code</i> <i>Autoencoder</i> Model 1 Dengan <i>Hyperparameter Regularization</i> dan <i>Dropout</i> .....	38
Gambar 4. 14 <i>Confussion Matrix</i> Partisi 50:50 .....	52
Gambar 4. 15 <i>Confussion Matrix</i> Partisi 80:20 .....	53
Gambar 4. 16 <i>Confussion Matrix</i> Partisi 90:10 .....	53

UNIVERSITAS  
MIKROSKIL

## DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu mengenai deteksi anomali dengan <i>Machine Learning/Deep Learning</i> .....	11
Tabel 3. 1 Sampel <i>Dataset</i> Transaksi Kartu Kredit.....	21
Tabel 3. 2 Daftar Atribut Kolom pada <i>Dataset</i> Transaksi Kartu Kredit .....	21
Tabel 3. 3 <i>Confusion Matrix</i> pada Deteksi Fraud Kartu Kredit .....	24
Tabel 4. 1 Struktur Data .....	27
Tabel 4. 2 Rangkuman Statistik <i>Dataset</i> .....	29
Tabel 4. 3 Perbandingan Rangkuman Statistik Kolom <i>Amount</i> .....	30
Tabel 4. 4 <i>Dataset</i> Setelah Normalisasi.....	33
Tabel 4. 5 Model <i>Summary CTGANSynthesizer</i> .....	35
Tabel 4. 6 Partisi Sampel Normal dan Fraud.....	36
Tabel 4. 7 <i>Classification Report Dataset</i> Asli Partisi 50:50 .....	39
Tabel 4. 8 <i>Classification Report Dataset</i> Asli Partisi 80:20 .....	40
Tabel 4. 9 <i>Classification Report Dataset</i> Asli Partisi 90:10 .....	40
Tabel 4. 10 ROC, MCC dan <i>G-Mean</i> Pada <i>Dataset</i> Asli Partisi 50:50.....	41
Tabel 4. 11 ROC, MCC dan <i>G-Mean</i> Pada <i>Dataset</i> Asli Partisi 80:20.....	42
Tabel 4. 12 ROC, MCC dan <i>G-Mean</i> Pada <i>Dataset</i> Asli Partisi 90:10.....	43
Tabel 4. 13 <i>Classification Report Dataset</i> Asli Partisi 50:50 Menggunakan <i>Hyperparameter</i> .....	43
Tabel 4. 14 <i>Classification Report Dataset</i> Asli Partisi 80:20 Menggunakan <i>Hyperparameter</i> .....	44
Tabel 4. 15 <i>Classification Report Dataset</i> Asli Partisi 90:10 Menggunakan <i>Hyperparameter</i> .....	44
Tabel 4. 16 ROC, MCC dan <i>G-Mean</i> Partisi 50:50 Menggunakan <i>Hyperparameter</i> .....	45
Tabel 4. 17 ROC, MCC dan <i>G-Mean</i> Partisi 80:20 Menggunakan <i>Hyperparameter</i> .....	45
Tabel 4. 18 ROC, MCC dan <i>G-Mean</i> Partisi 90:10 Menggunakan <i>Hyperparameter</i> .....	46
Tabel 4. 19 Hasil <i>Classification Report</i> dan Performa Model Pada Data Generatif Dengan Perbandingan 50:50 .....	46
Tabel 4. 20 Hasil <i>Classification Report</i> dan Performa Model Pada Data Generatif Dengan Perbandingan 80:20 .....	47
Tabel 4. 21 Hasil <i>Classification Report</i> dan Performa Model Pada Data Generatif Dengan Perbandingan 90:10 .....	47
Tabel 4. 22 Perbandingan <i>Precision, Recall, F1-score</i> dan <i>Accuracy</i> Pada <i>Dataset</i> Dengan Perbandingan 50:50 .....	49
Tabel 4. 23 Perbandingan <i>Precision, Recall, F1-score</i> dan <i>Accuracy</i> Pada <i>Dataset</i> Dengan Perbandngan 80:20 .....	49
Tabel 4. 24 Perbandingan <i>Precision, Recall, F1-score</i> dan <i>Accuracy</i> Pada <i>Dataset</i> Dengan Perbandngan 90:10 .....	50
Tabel 4. 25 Perbandingan Performa Model ROC, MCC dan <i>G-Mean</i> .....	50
Tabel 4. 26 Perbandingan <i>Classification Report</i> AE, DNN dan TCN Pada <i>Dataset</i> Asli ..	51
Tabel 4. 27 Perbandingan Performa Model AE, DNN dan TCN Pada <i>Dataset</i> Asli .....	51

Tabel 4. 28 Perbandingan *Classification Report* AE, DNN dan TCN Pada Generatif Dataset..... 51

Tabel 4. 29 Perbandingan Performa Model AE, DNN dan TCN Pada Generatif Dataset.. 52

