

**ANALISIS DAMPAK SMOTE DAN KOMPARASI
METODE SELEKSI FITUR PADA KLASIFIKASI
N-STAGE ADENOKARSINOMA PARU BERBASIS
RADIOMIK**



LAPORAN RESEARCH & TECHNOLOGY PROGRAM

**PANDU AJI WICAKSONO
00000078666**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026**

**ANALISIS DAMPAK SMOTE DAN KOMPARASI
METODE SELEKSI FITUR PADA KLASIFIKASI
N-STAGE ADENOKARSINOMA PARU BERBASIS
RADIOMIK**



LAPORAN RESEARCH & TECHNOLOGY PROGRAM

**PANDU AJI WICAKSONO
00000078666**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Pandu Aji Wicaksono

Nomor Induk Mahasiswa : 00000078666

Program Studi : Informatika

menyatakan dengan sesungguhnya bahwa Laporan saya yang berjudul:

ANALISIS DAMPAK SMOTE DAN KOMPARASI METODE SELEKSI
FITUR PADA KLASIFIKASI N-STAGE ADENOKARSINOMA PARU
BERBASIS RADIOMIK

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain. Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian daftar pustaka.

Jika di kemudian hari terbukti ditemukan penyimpangan dan penyalahgunaan dalam pelaksanaan maupun dalam penulisan laporan, saya bersedia menerima konsekuensi untuk dinyatakan **TIDAK LULUS**. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 31 Desember 2025



(Pandua Aji Wicaksono)

HALAMAN PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Pandu Aji Wicaksono
NIM : 00000078666
Program Studi : Informatika
Judul Laporan : Analisis Dampak SMOTE dan Komparasi Metode
Seleksi Fitur pada Klasifikasi N-Stage Adenokarsinoma Paru Berbasis
Radiomik

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas/Laporan/Project/Tugas Akhir*(pilih salah satu) sebagai berikut (beri tanda centang yang sesuai):

- ☐ Menggunakan AI sebagaimana diizinkan untuk membantu dalam menghasilkan ide-ide utama saja
- ☐ Menggunakan AI sebagaimana diizinkan untuk membantu menghasilkan teks pertama saja
- ☒ Menggunakan AI untuk menyempurnakan sintaksis dan tata bahasa untuk pengumpulan tugas
- ☐ Karena tidak diizinkan: Tidak menggunakan bantuan AI dengan cara apa pun dalam pembuatan tugas

Saya juga menyatakan bahwa:

- (1) Menyerahkan secara lengkap dan jujur penggunaan perangkat AI yang diperlukan dalam tugas melalui Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)
- (2) Mengakui telah menggunakan bantuan AI dalam tugas saya baik dalam bentuk kata, paraphrase, penyertaan ide atau fakta penting yang disarankan oleh AI dan saya telah menyantumkan dalam sitasi serta referensi
- (3) Terlepas dari pernyataan di atas, tugas ini sepenuhnya merupakan karya saya sendiri

Tangerang, 31 Desember 2025



(Pandua Aji Wicaksono)

HALAMAN PERNYATAAN KEABSAHAN PERUSAHAAN

Nama : Pandu Aji Wicaksono
NIM : 00000078666
Program Studi : Informatika
Fakultas : Teknik dan Informatika

menyatakan bahwa saya melaksanakan kegiatan di:

Nama Perusahaan/Organisasi : LPPM Universitas Multimedia Nusantara
Alamat : Jl. Scientia Boulevard, Gading Serpong, Kel. Curug
Sangereng, Kec. Kelapa Dua, Kab. Tangerang, Prop. Banten 15810, Indonesia
Email Perusahaan/Organisasi : fti@umn.ac.id

1. Perusahaan/Organisasi tempat saya melakukan kegiatan dapat di validasi keberadaannya.
2. Jika dikemudian hari, terbukti ditemukan kecurangan/penyimpangan data yang tidak valid di perusahaan/organisasi tempat saya melakukan kegiatan, maka:
 - a. Saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.
 - b. Saya bersedia menerima semua sanksi yang berlaku sebagaimana ditetapkan dalam peraturan yang berlaku di Universitas Multimedia Nusantara.

Pernyataan ini saya buat dengan sebenar-benarnya dan digunakan sebagaimana mestinya.

Tangerang, 19 Desember 2025

Mengetahui


(Marinda Vasty Overbeek)

Menyatakan


(Pandua Aji Wicaksono)

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan dibawah ini:

Nama : Pandu Aji Wicaksono
NIM : 00000078666
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Dampak SMOTE dan Komparasi
Metode Seleksi Fitur pada Klasifikasi N-Stage Adenokarsinoma Paru
Berdasarkan Radiomik

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk memublikasikan hasil laporan saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh civitas academica UMN/ publik. Saya menyatakan bahwa laporan yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia memublikasikan hasil laporan ini ke dalam repositori Knowledge Center karena ada data yang bersifat konfidensial.
- ☐ Lainnya, pilih salah satu:
- ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara.
 - ☐ Embargo publikasi laporan dalam kurun waktu 3 tahun.

Tangerang, 31 Desember 2025


(Pandua Aji Wicaksono)

Halaman Persembahan / Motto

"If you want something you never had, do something you never did."

(Pandji Pragiwaksono)



KATA PENGANTAR

Puji syukur dipanjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan rahmat-Nya, sehingga laporan *Research & Technology* dengan judul "Analisis Dampak SMOTE dan Komparasi Metode Seleksi Fitur pada Klasifikasi N-Stage Adenokarsinoma Paru Berbasis Radiomik" ini dapat diselesaikan dengan baik. Penulisan laporan ini merupakan salah satu syarat guna memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Saya menyadari bahwa penyelesaian laporan ini tidak lepas dari bantuan serta bimbingan berbagai pihak, mulai dari masa perkuliahan hingga tahap akhir penyusunan. Sebagai bentuk apresiasi, saya menyampaikan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Dennis Gunawan, S.Kom., M.Sc., selaku Sekretaris Program Studi Informatika Universitas Multimedia Nusantara.
5. Bapak Yustinus Widya Wiratama, S.Kom., M.Sc., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya laporan ini.
6. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Pembimbing Lapangan yang telah meluangkan waktu dan perhatian dalam memberikan bimbingan, arahan, pelajaran, serta motivasi selama proses penelitian ini.
7. Rekan-rekan peneliti yang telah memberikan dukungan melalui diskusi produktif, pertukaran gagasan, serta kerja sama yang konstruktif selama berlangsungnya proses penelitian ini.
8. Grup "KANDAR BERJAYE" yang telah memberikan dukungan moral, informasi bermanfaat, serta keceriaan yang menjadi penyemangat selama penyusunan laporan ini.

9. Orang tua, keluarga, dan teman-teman yang senantiasa memberikan dukungan moral dan motivasi, sehingga laporan magang ini dapat diselesaikan.

Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang informatika medis, serta menjadi referensi bagi penelitian selanjutnya. Penulis menyadari adanya keterbatasan dalam kajian ini, sehingga saran dan kritik yang bersifat membangun sangat diharapkan guna penyempurnaan di masa yang akan datang.

Tangerang, 31 Desember 2025



Pandu Aji Wicaksono



ANALISIS DAMPAK SMOTE DAN KOMPARASI METODE SELEKSI FITUR PADA KLASIFIKASI N-STAGE ADENOKARSINOMA PARU BERBASIS RADIOMIK

Pandu Aji Wicaksono

ABSTRAK

Adenokarsinoma paru adalah subtype kanker paru yang paling umum, di mana tahapan penyebaran pada kelenjar getah bening (*N-Stage*) sangat penting dalam pengobatan dan prognosis pasien. Pendekatan radiomik pada citra *Computed Tomography* (CT) memberikan solusi diagnosis non-invasif, namun terkendala oleh dimensi fitur tinggi dan ketidakseimbangan kelas yang dapat menyebabkan bias dalam model. Penelitian ini menganalisis penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) dan evaluasi tiga metode seleksi fitur (mRMR, Boruta, dan L1-LASSO) pada model klasifikasi *Machine Learning* dan *Deep Learning* (XGBoost, SVM, TabNet). Menggunakan dataset TCGA-LUAD yang diproses dengan PyRadiomics, hasil eksperimen menunjukkan bahwa SMOTE meningkatkan sensitivitas model, terutama pada SVM. XGBoost dengan seleksi fitur mRMR dan SMOTE menghasilkan performa optimal (*Accuracy* 0,95, *F1-Macro* 0,96). Penelitian ini menyimpulkan bahwa integrasi teknik penyeimbangan data dan seleksi fitur yang tepat sangat penting untuk sistem pendukung keputusan klinis dalam penentuan *N-Stage* kanker paru.

Kata kunci: Adenokarsinoma Paru, *N-Stage*, *Radiomics*, Seleksi Fitur, SMOTE



**ANALYSIS OF THE IMPACT OF SMOTE AND COMPARISON OF
FEATURE SELECTION METHODS ON N-STAGE LUNG
ADENOCARCINOMA CLASSIFICATION BASED ON RADIOMICS**

Pandu Aji Wicaksono

ABSTRACT

Lung adenocarcinoma is the most prevalent subtype of lung cancer, where the staging of lymph node involvement (N-Stage) plays a crucial role in treatment strategies and patient prognosis. Radiomics approaches using Computed Tomography (CT) images offer a non-invasive diagnostic solution but are challenged by high feature dimensionality and class imbalance, which can lead to model bias. This study analyzes the implementation of the Synthetic Minority Over-sampling Technique (SMOTE) and evaluates three feature selection methods (mRMR, Boruta, and LI-LASSO) across Machine Learning and Deep Learning classification models (XGBoost, SVM, and TabNet). Utilizing the TCGA-LUAD dataset processed with PyRadiomics, experimental results indicate that SMOTE enhances model sensitivity, particularly in SVM. XGBoost combined with mRMR feature selection and SMOTE yielded optimal performance (Accuracy 0.95, F1-Macro 0.96). This study concludes that the integration of data balancing techniques and appropriate feature selection is essential for developing reliable clinical decision support systems for lung cancer N-Stage determination.

Keywords: Feature Selection, Lung Adenocarcinoma, N-Stage, Radiomics, SMOTE

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN ORISINALITAS	ii
PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)	iii
HALAMAN PENGABSAHAN	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
HALAMAN PERSEMBAHAN/MOTO	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR KODE	xv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Urgensi Penelitian	3
1.5 Luaran Penelitian	4
1.6 Manfaat Penelitian	4
BAB 2 TINJAUAN PUSTAKA	5
2.1 Adenokarsinoma Paru	5
2.1.1 Karakteristik Patologis dan Molekuler	5
2.1.2 Perbandingan Paru-Paru Normal dan Adenokarsinoma Paru	6
2.1.3 N-Stage Adenokarsinoma Paru	8
2.2 Segmentasi TotalSegmentator	9
2.3 Rekayasa Fitur	10
2.3.1 Ekstraksi Fitur	10
2.3.2 SMOTE	12
2.3.3 Seleksi Fitur	13
2.4 XGBoost	14
2.5 Support Vector Machine (SVM)	17
2.6 TabNet (<i>Attentive Interpretable Tabular Learning</i>)	19
2.7 Hyperparameter Tuning	20
2.8 Metriks Evaluasi	22
BAB 3 METODE PENELITIAN	25
3.1 Pengumpulan Data	26
3.1.1 Data Klinis	26
3.1.2 Data Gambar	26
3.2 Pre-proses data dan Segmentasi Citra	27
3.2.1 Pre-proses Data	27
3.2.2 Segmentasi Citra	28
3.3 Resampling Citra, Ekstraksi Fitur, dan Normalisasi	28
3.4 Seleksi Fitur dan SMOTE	29
3.4.1 Seleksi Fitur	29

3.4.2	SMOTE	30
3.5	Pengembangan Model dan Evaluasi	31
BAB 4	HASIL DAN PEMBAHASAN	33
4.1	Pengumpulan Data	33
4.1.1	Dataset Citra	33
4.1.2	Dataset Klinis	33
4.2	Pre-proses dan Segmentasi Citra	34
4.2.1	Pre-proses	34
4.2.2	Segmentasi Citra	34
4.3	Resampling Citra, Ekstraksi Fitur, dan Normalisasi	35
4.3.1	Resampling Citra	35
4.3.2	Ekstraksi Fitur	37
4.3.3	Normalisasi	38
4.4	Seleksi Fitur dan SMOTE	39
4.4.1	SMOTE	39
4.4.2	Perbandingan Seleksi Fitur Tanpa SMOTE dan dengan SMOTE	40
4.5	Pengembangan Model dan Evaluasi Model	42
4.5.1	Pengembangan Model	42
4.5.2	Evaluasi Model Menggunakan Seluruh Fitur	42
4.5.3	Evaluasi Model Menggunakan Seleksi Fitur Tanpa SMOTE	43
4.5.4	Evaluasi Model Menggunakan SMOTE Tanpa Seleksi Fitur	47
4.5.5	Evaluasi Model Menggunakan SMOTE dan Seleksi Fitur	49
4.5.6	Diskusi Hasil Evaluasi	53
BAB 5	SIMPULAN SARAN	56
5.1	Simpulan	56
5.2	Saran	57
DAFTAR PUSTAKA		58



DAFTAR TABEL

Tabel 2.1	Deskriptor N Kanker Paru: Definisi, Prognosis (Edisi 8 & 9), dan Terapi	9
Tabel 4.1	Rekapitulasi Perbandingan Performa Model pada Seluruh Skenario	53



DAFTAR GAMBAR

Gambar 2.1	Perbedaan Paru-Paru Normal dan Adenokarsinoma Paru . . .	6
Gambar 2.2	Confusion Matrix Multiclass	24
Gambar 3.1	Alur Penelitian Dampak SMOTE dan Komparasi Metode Seleksi Fitur pada Klasifikasi N-Stage Adenokarsinoma Paru Berbasis Radiomik	25
Gambar 3.2	Proses Pengambilan Data Klinis TCGA-LUAD	26
Gambar 3.3	Proses Pengambilan Data Citra TCGA-LUAD	27
Gambar 3.4	Proses Preproses Data TCGA-LUAD	28
Gambar 3.5	Proses Segmentasi Data Citra TCGA-LUAD	28
Gambar 3.6	Proses Resampling Citra, Ekstraksi Fitur, dan Normalisasi	29
Gambar 3.7	Proses Seleksi Fitur dengan SMOTE dan Tanpa SMOTE	30
Gambar 3.8	Proses SMOTE	31
Gambar 3.9	Proses Pengembangan Model dan Evaluasi	32
Gambar 4.1	Struktur Folder untuk Pengumpulan Data Citra	33
Gambar 4.2	<i>File Klinis Phenotype</i> TCGA-LUAD	34
Gambar 4.3	Detail Jumlah Data Gambar dan Pasien	34
Gambar 4.4	Contoh <i>Output</i> Segmentasi	35
Gambar 4.5	Visualisasi Sebelum Melakukan <i>Resampling</i>	36
Gambar 4.6	Visualisasi Setelah Melakukan <i>Resampling</i>	37
Gambar 4.7	Visualisasi Seleksi Fitur Tanpa SMOTE	41
Gambar 4.8	Visualisasi Seleksi Fitur SMOTE	41
Gambar 4.9	Visualisasi Evaluasi Seluruh Fitur	43
Gambar 4.10	Confussion Matrix Seluruh Fitur	43
Gambar 4.11	Visualisasi Evaluasi mRMR Tanpa SMOTE	44
Gambar 4.12	Confussion Matrix mRMR Tanpa SMOTE	44
Gambar 4.13	Visualisasi Evaluasi Boruta Tanpa SMOTE	45
Gambar 4.14	Confussion Matrix Boruta Tanpa SMOTE	46
Gambar 4.15	Visualisasi Evaluasi L1-LASSO Tanpa SMOTE	47
Gambar 4.16	Confussion Matrix L1-LASSO Tanpa SMOTE	47
Gambar 4.17	Visualisasi Evaluasi SMOTE Tanpa Seleksi Fitur	48
Gambar 4.18	Confussion Matrix SMOTE Tanpa Seleksi Fitur	48
Gambar 4.19	Visualisasi Evaluasi mRMR SMOTE dan Seleksi	50
Gambar 4.20	Confussion Matrix mRMR SMOTE dan Seleksi	50
Gambar 4.21	Visualisasi Evaluasi Boruta Tanpa SMOTE	51
Gambar 4.22	Confussion Matrix Boruta Tanpa SMOTE	51
Gambar 4.23	Visualisasi Evaluasi L1-LASSO Tanpa SMOTE	52
Gambar 4.24	Confussion Matrix L1-LASSO Tanpa SMOTE	53

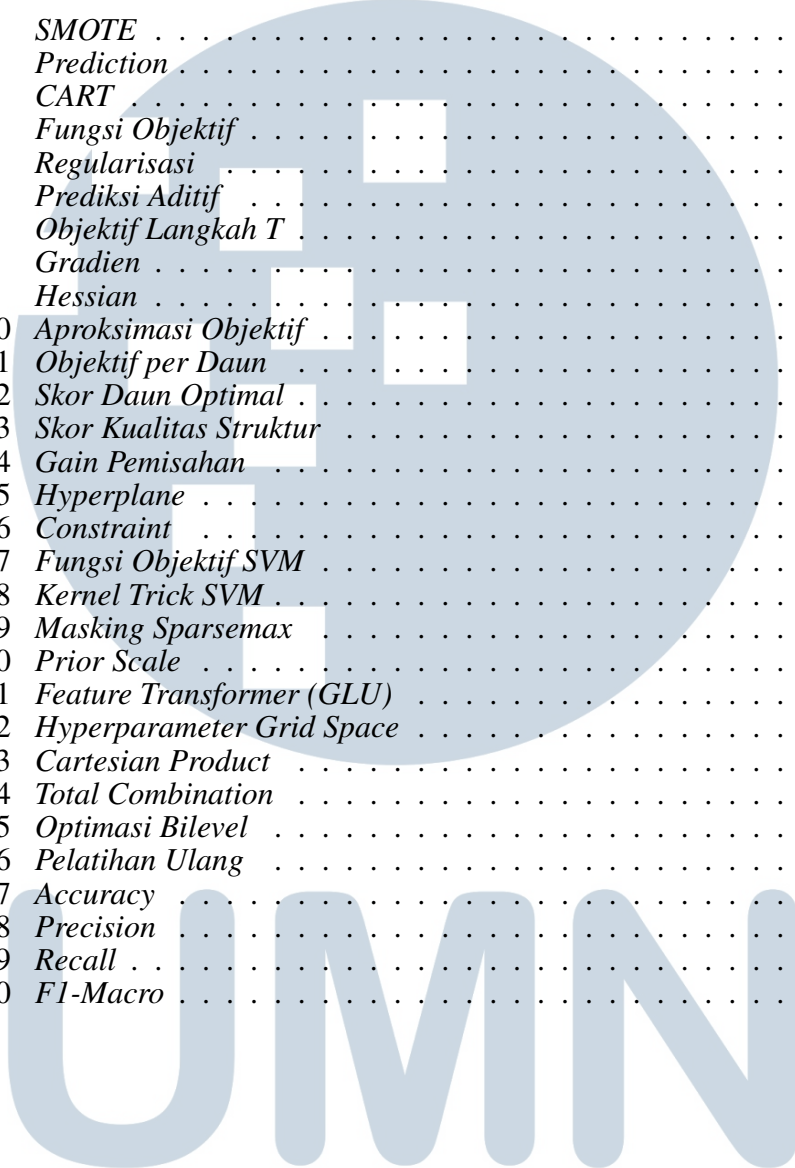
DAFTAR KODE

Kode 4.1	Perintah Eksekusi Segmentasi Nodul Paru Menggunakan TotalSegmentator	35
Kode 4.2	<i>Settings</i> untuk Isotropik Resampling	35
Kode 4.3	Potongan Kode Normalisasi	38
Kode 4.4	Potongan Kode <i>Train Test Split</i>	39
Kode 4.5	Potongan Kode SMOTE	40



DAFTAR RUMUS

Rumus 2.1	<i>SMOTE</i>	12
Rumus 2.2	<i>Prediction</i>	14
Rumus 2.3	<i>CART</i>	15
Rumus 2.4	<i>Fungsi Objektif</i>	15
Rumus 2.5	<i>Regularisasi</i>	15
Rumus 2.6	<i>Prediksi Aditif</i>	16
Rumus 2.7	<i>Objektif Langkah T</i>	16
Rumus 2.8	<i>Gradien</i>	16
Rumus 2.9	<i>Hessian</i>	16
Rumus 2.10	<i>Aproksimasi Objektif</i>	16
Rumus 2.11	<i>Objektif per Daun</i>	16
Rumus 2.12	<i>Skor Daun Optimal</i>	17
Rumus 2.13	<i>Skor Kualitas Struktur</i>	17
Rumus 2.14	<i>Gain Pemisahan</i>	17
Rumus 2.15	<i>Hyperplane</i>	18
Rumus 2.16	<i>Constraint</i>	18
Rumus 2.17	<i>Fungsi Objektif SVM</i>	18
Rumus 2.18	<i>Kernel Trick SVM</i>	19
Rumus 2.19	<i>Masking Sparsemax</i>	19
Rumus 2.20	<i>Prior Scale</i>	20
Rumus 2.21	<i>Feature Transformer (GLU)</i>	20
Rumus 2.22	<i>Hyperparameter Grid Space</i>	21
Rumus 2.23	<i>Cartesian Product</i>	21
Rumus 2.24	<i>Total Combination</i>	21
Rumus 2.25	<i>Optimasi Bilevel</i>	21
Rumus 2.26	<i>Pelatihan Ulang</i>	22
Rumus 2.27	<i>Accuracy</i>	22
Rumus 2.28	<i>Precision</i>	23
Rumus 2.29	<i>Recall</i>	23
Rumus 2.30	<i>F1-Macro</i>	23


 UNIVERSITAS
 MULTIMEDIA
 NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	PRO-STEP 01 Cover Letter	61
Lampiran 2	PRO-STEP 02 Research Card	62
Lampiran 3	PRO-STEP 03 Daily Task	63
Lampiran 4	PRO-STEP 04 Verification Form	76
Lampiran 5	PRO-STEP 05 Letter of Acceptance	77
Lampiran 6	Form Bimbingan	78
Lampiran 7	Turnitin	79
Lampiran 8	Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI) . .	86
Lampiran 9	Research Paper	88
Lampiran 10	Surat Kerja Sama Penelitian Institusi	95

