

**DETEKSI KANKER PAYUDARA STADIUM AWAL DAN
LANJUT MELALUI DATA EKSPRESI GEN:
PENDEKATAN SELEKSI FITUR DFP**



LAPORAN RESEARCH TECHNOLOGY PROGRAM

**OWEN CHRISTIAN CAHYADI
00000067055**

**PROGRAM STUDI INFORMATIKA
FAKULTASTEKNIKDANINFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026**

**DETEKSI KANKER PAYUDARA STADIUM AWAL DAN
LANJUT MELALUI DATA EKSPRESI GEN:
PENDEKATAN SELEKSI FITUR DFP**



LAPORAN RESEARCH TECHNOLOGY PROGRAM

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**OWEN CHRISTIAN CAHYADI
00000067055**

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Owen Christian Cahyadi
Nomor Induk Mahasiswa : 00000067055
Program Studi : Informatika

Skripsi dengan judul:

**Deteksi Kanker Payudara Stadium Awal dan Lanjut Melalui Data Ekspresi Gen:
Pendekatan Seleksi Fitur Discriminant Fuzzy Pattern**

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 02 Januari 2026



Owen Christian Cahyadi

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan dibawah ini,

Nama : Owen Christian Cahyadi
NIM : 00000067055
Program Studi : Informatika
Judul Laporan : Deteksi Kanker Payudara Stadium
Awal dan Lanjut Melalui Data
Ekspresi Gen: Pendekatan Seleksi
Fitur Discriminant Fuzzy Pattern

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan
Artifisial (AI) dalam pengerjaan Laporan ini sebagai berikut:

- ☒ 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, 02 Januari 2026

Yang menyatakan



Owen Christian Cahyadi



HALAMAN PERNYATAAN KEABSAHAN PERUSAHAAN

Yang bertanda tangan dibawah ini:

Nama : Owen Christian Cahyadi
NIM : 00000067055
Program Studi : Informatika
Fakultas : TEKNIK DAN INFORMATIKA

menyatakan bahwa saya melaksanakan kegiatan di:

Nama Perusahaan/Organisasi : Universitas Brawijaya
Alamat : Jalan Veteran, Ketawanggede, Kecamatan
Lowokwaru, Kota Malang, Jawa Timur
Email Perusahaan/Organisasi : akademik.center@ub.ac.id

1. Perusahaan/Organisasi tempat saya melakukan kegiatan dapat divalidasi keberadaannya.
2. Jika di kemudian hari, terbukti ditemukan kecurangan atau 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.

Mengetahui,



dr. Edvin Perwira Negara

Tangerang, 02 Januari 2026

Menyatakan



Owen Christian Cahyadi

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Owen Christian Cahyadi
NIM : 00000067055
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Deteksi Kanker Payudara Stadium Awal
dan Lanjut Melalui Data Ekspresi Gen:
Pendekatan Seleksi Fitur Discriminant Fuzzy
Pattern

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

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 02 Januari 2026

Yang menyatakan



Owen Christian Cahyadi

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji syukur panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Penelitian ini bertujuan untuk mengembangkan metode deteksi kanker payudara stadium awal dan lanjut menggunakan machine learning melalui data ekspresi gen dengan pendekatan seleksi fitur *Discriminant Fuzzy Pattern* (DFP). Dalam proses penyusunan penelitian ini, sebagai bentuk apresiasi atas segala dukungan yang diberikan, berikut disampaikan ucapan terima kasih kepada pihak-pihak terkait.

Mengucapkan terima kasih

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 Anak Agung Ngurah Ananda Kusuma, B.Eng., M.Eng., Ph.D., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Bapak Moeljono Widjaja, B.Sc., M.Sc., Ph.D., sebagai Pembimbing Lapangan yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya laporan MBKM Penelitian.
6. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga penelitian ini dapat memberikan kontribusi bagi perkembangan ilmu pengetahuan, khususnya dalam bidang informatika dan bioinformatika, serta bermanfaat bagi upaya peningkatan akurasi diagnosis medis pada kanker payudara.

Tangerang, 02 Januari 2026



Owen Christian Cahyadi

**DETEKSI KANKER PAYUDARA STADIUM AWAL DAN
LANJUT MELALUI DATA EKSPRESI GEN:
PENDEKATAN SELEKSI FITUR DFP**

Owen Christian Cahyadi

ABSTRAK

Kanker payudara merupakan penyebab utama kematian akibat kanker pada wanita secara global yang menuntut strategi deteksi dini efektif untuk meningkatkan prognosis pasien. Diagnosis molekuler berbasis data ekspresi gen menawarkan presisi tinggi, namun penerapannya sering terhambat oleh karakteristik data berdimensi tinggi dengan jumlah sampel terbatas (*High Dimensional Low Sample Size*). Metode seleksi fitur konvensional sering kali kurang optimal dalam menangani ketidakpastian biologis data tersebut untuk membedakan stadium kanker secara akurat. Penelitian ini menyajikan integrasi metode seleksi fitur *Discriminant Fuzzy Pattern* (DFP) dan klasifikasi *Support Vector Machine* (SVM) untuk mendeteksi stadium awal dan lanjut menggunakan dataset TCGA-BRCA. Melalui penerapan *Random Under Sampling* untuk menangani ketidakseimbangan kelas, hasil eksperimen menunjukkan bahwa DFP berhasil mereduksi 60.661 gen menjadi 22 gen kandidat paling diskriminatif, di mana 8 gen telah tervalidasi memiliki relevansi biologis dengan mekanisme kanker. Evaluasi model pada skenario pembagian data 90:10 menghasilkan performa optimal dengan akurasi 82%, *F1-Score* makro 0,70, dan nilai AUC sebesar 0,71. Penelitian ini mengimplikasikan bahwa pendekatan logika *fuzzy* efektif dalam menangani kompleksitas data genomik serta berpotensi mengungkap biomarker baru untuk mendukung pengembangan sistem pendukung keputusan klinis yang lebih objektif dan akurat.

Kata kunci: Deteksi Kanker Payudara, *Discriminant Fuzzy Pattern*, Ekspresi Gen, *Machine Learning*, Seleksi Fitur

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DETECTION OF EARLY AND LATE-STAGE BREAST CANCER THROUGH GENE EXPRESSION DATA: DFP FEATURE SELECTION APPROACH

Owen Christian Cahyadi

ABSTRACT

Breast cancer is a leading cause of cancer death among women globally, demanding effective early detection strategies to improve patient prognosis. Molecular diagnosis based on gene expression data offers high precision, but its application is often hindered by High Dimensional Low Sample Size characteristics. Conventional feature selection methods are often suboptimal in handling the biological uncertainty of such data to accurately distinguish cancer stages. This study presents the integration of the Discriminant Fuzzy Pattern (DFP) feature selection method and Support Vector Machine (SVM) classification to detect early and late stages using the TCGA-BRCA dataset. Through the application of Random Under Sampling to handle class imbalance, experimental results show that DFP successfully reduced 60,661 genes to 22 most discriminant candidate genes, of which 8 have been validated to have biological relevance to cancer mechanisms. Model evaluation on a 90:10 data split scenario yielded optimal performance with 82% accuracy, macro F1-Score of 0.70, and an AUC value of 0.71. This research implies that the fuzzy logic approach is effective in handling genomic data complexity and has the potential to reveal new biomarkers to support the development of more objective and accurate clinical decision support systems.

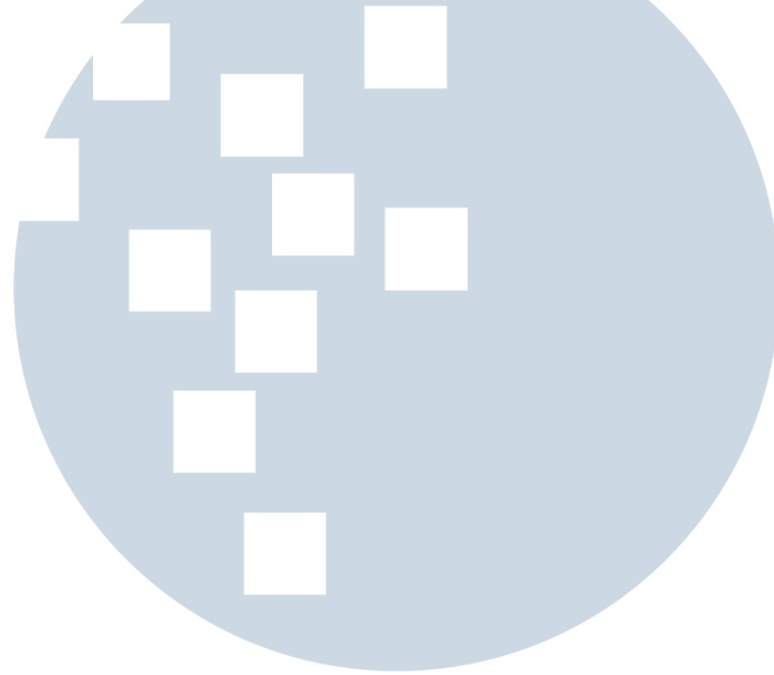
Keywords: *Breast Cancer Detection, Discriminant Fuzzy Pattern, Feature Selection, Gene Expression, Machine Learning*



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
PERNYATAAN PENGGUNAAN BANTUAN AI	iii
HALAMAN PERNYATAAN KEABSAHAN PERUSAHAAN	v
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	vi
HALAMAN PERSEMBAHAN/MOTO	vii
KATA PENGANTAR	viii
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	2
1.3 Batasan Permasalahan	2
1.4 Tujuan Penelitian	3
1.5 Urgensi Penelitian	4
1.6 Luaran Penelitian	4
1.7 Manfaat Penelitian	5
BAB 2 LANDASAN TEORI	6
2.1 Kanker Payudara	6
2.2 Ekspresi Gen (TPM)	7
2.3 Fuzzy Logic	8
2.3.1 Konsep Himpunan Tegas vs. Himpunan Fuzzy	8
2.3.2 Fungsi Keanggotaan (Membership Function)	8
2.3.3 Operator Logika Fuzzy	11
2.3.4 Arsitektur Sistem Inferensi Fuzzy	12
2.4 Feature Selection	13
2.5 Discriminant Fuzzy Pattern	13
2.6 Support Vector Machine (SVM)	14
2.7 Confusion Matrix	17
2.8 Research Gap	19
BAB 3 METODOLOGI PENELITIAN	21
3.1 Studi Literatur	21
3.2 Pengumpulan Data	21
3.3 Pra-Proses Data	22
3.4 Seleksi Fitur	23
3.5 Pembangunan Model	24
3.6 Evaluasi Model	25
3.7 Interpretasi Biomarker	25
BAB 4 HASIL DAN DISKUSI	26
4.1 Implementasi Alur Kerja Penelitian	26
4.1.1 Struktur dan Format Dataset	26
4.1.2 Kategorisasi dan Labeling Kelas	27
4.1.3 Penggabungan Label ke Data Ekspresi Gen	28
4.1.4 Ringkasan Dataset Setelah Praproses	28
4.2 Hasil Tahapan Seleksi Fitur	29
4.2.1 Konfigurasi Parameter Awal	30
4.2.2 Menghitung Parameter Membership (c dan w)	30
4.2.3 Diskritisasi Nilai Ekspresi Gen	32
4.2.4 Pencarian Pola Fuzzy dan Seleksi Fitur Diskriminan	35
4.2.5 Eliminasi Redundansi Fitur Berbasis Korelasi	37
4.2.6 Analisis Redundansi dan Korelasi Antar Fitur	38
4.2.7 Visualisasi Fungsi Keanggotaan Gen Terpilih	39
4.3 Pembangunan Model dan Evaluasi SVM	40
4.3.1 Evaluasi Kinerja Model pada Skenario Pembagian Data 90:10	41

4.3.2	Evaluasi Kinerja Model pada Skenario Pembagian Data 80:20	44
4.4	Analisis dan Interpretasi Kandidat Biomarker	47
BAB 5	SIMPULAN DAN SARAN	50
5.1	Simpulan	50
5.2	Saran	50
DAFTAR PUSTAKA		52



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 2.1	Pengelompokan stadium kanker payudara menurut AJCC edisi ke-8 . . .	7
Tabel 2.2	Confusion Matrix	17
Tabel 2.3	Interpretasi AUC	19
Tabel 2.4	Research gap perbandingan metode seleksi fitur deteksi kanker	19
Tabel 3.1	Deskripsi dataset yang digunakan	22
Tabel 3.2	Hyperparameter tuning DFP	23
Tabel 3.3	Parameter model SVM	24
Tabel 4.1	Dimensi ekspresi gen sebelum dan sesudah transposisi	26
Tabel 4.2	Dimensi data fenotipe sebelum dan sesudah penghapusan fitur null > 50%	27
Tabel 4.3	Distribusi data pada kolom <code>ajcc_pathologic_stage</code>	27
Tabel 4.4	Jumlah sampel sebelum dan sesudah integrasi data	28
Tabel 4.5	Perbandingan statistik 5 sampel sebelum dan sesudah normalisasi	29
Tabel 4.6	Rangkuman hasil praproses dataset	29
Tabel 4.7	Konfigurasi parameter DFP dan seleksi fitur	30
Tabel 4.8	Parameter fungsi keanggotaan (c dan w) untuk 5 gen pertama	32
Tabel 4.9	Hasil diskritisasi nilai ekspresi gen menjadi label linguistik	35
Tabel 4.10	Hasil seleksi fitur diskriminan (DFP) dengan pola berbeda antar kelas	37
Tabel 4.11	Konfigurasi parameter terbaik hasil optimasi	41
Tabel 4.12	Konfigurasi parameter terbaik pada skenario 80:20	44
Tabel 4.13	Daftar 22 gen kandidat biomarker dan validasi literatur	47



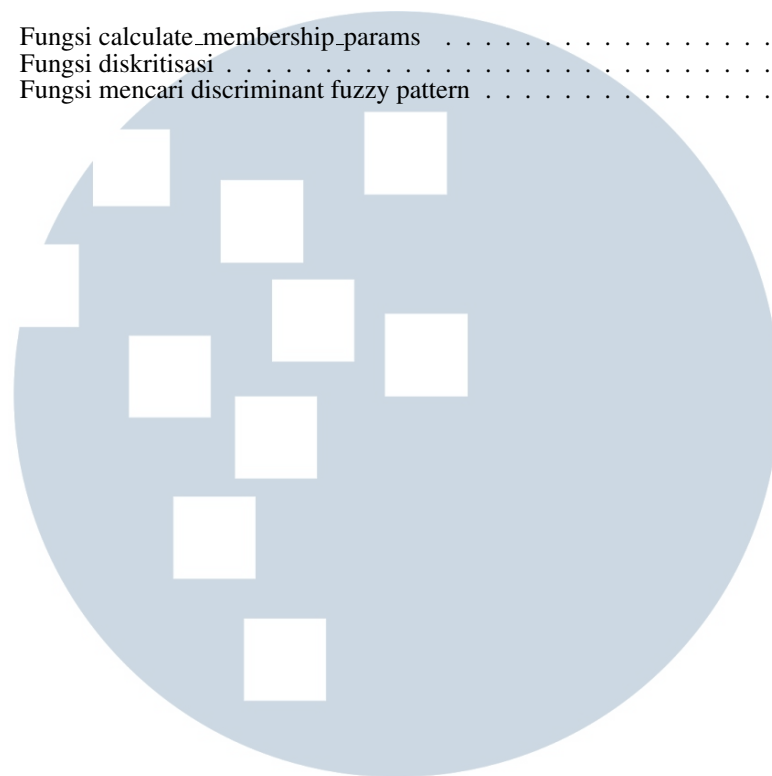
DAFTAR GAMBAR

Gambar 2.1	Representasi kurva keanggotaan segitiga	9
Gambar 2.2	Representasi kurva keanggotaan trapesium	10
Gambar 2.3	Representasi kurva keanggotaan gaussian	11
Gambar 2.4	Diagram blok sistem inferensi fuzzy	12
Gambar 2.5	Hyperplane optimal dengan margin maksimum	15
Gambar 3.1	Research pipeline untuk data gen	21
Gambar 4.1	Visualisasi heatmap matriks korelasi antar fitur terpilih dengan threshold 0.7.	39
Gambar 4.2	Visualisasi fungsi keanggotaan (Membership Function) dari tiga gen hasil seleksi.	40
Gambar 4.3	Visualisasi confusion matrix skenario 90:10	43
Gambar 4.4	Visualisasi ROC skenario 90:10	44
Gambar 4.5	Visualisasi ROC skenario 90:10	46
Gambar 4.6	Visualisasi ROC skenario 90:10	47



DAFTAR KODE

Kode 4.1	Fungsi calculate_membership_params	31
Kode 4.2	Fungsi diskritisasi	32
Kode 4.3	Fungsi mencari discriminant fuzzy pattern	35



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

Rumus 2.1	Transcripts Per Million (TPM)	8
Rumus 2.2	Fungsi Karakteristik (Crisp)	8
Rumus 2.3	Triangular Membership Function	9
Rumus 2.4	Trapezoidal Membership Function	10
Rumus 2.5	Gaussian Membership Function	11
Rumus 2.6	Fuzzy Intersection (AND)	11
Rumus 2.7	Fuzzy Union (OR)	12
Rumus 2.8	Fuzzy Complement (NOT)	12
Rumus 2.9	Defuzzification (Centroid)	12
Rumus 2.10	Gaussian Fuzzification Function	13
Rumus 2.11	Fuzzy Pattern	14
Rumus 2.12	Discriminative Fuzzy Pattern Definition	14
Rumus 2.13	SVM Hyperplane Equation	15
Rumus 2.14	SVM Objective Function	15
Rumus 2.15	Optimization Constraints	15
Rumus 2.16	Lagrange Dual Problem	16
Rumus 2.17	Dual Problem Constraints	16
Rumus 2.18	SVM Decision Function	16
Rumus 2.19	Linear Kernel	16
Rumus 2.20	Polynomial Kernel	16
Rumus 2.21	RBF (Gaussian) Kernel	17
Rumus 2.22	Accuracy	17
Rumus 2.23	Sensitivity (Recall)	18
Rumus 2.24	Specificity	18
Rumus 2.25	Precision	18
Rumus 2.26	F1-Score	18
Rumus 2.27	TPR and FPR for ROC	18
Rumus 4.1	Pearson Correlation Coefficient	38



DAFTAR LAMPIRAN

Lampiran 1	MBKM-01 Cover Letter MBKM Research & Technology Program	55
Lampiran 2	MBKM-02 Research & Technology Program Card	56
Lampiran 3	MBKM-03 Daily Task Research & Technology Program	57
Lampiran 4	MBKM-04 Verification Form of Research & Technology Program	68
Lampiran 5	MBKM-05 Letter of Acceptance Research & Technology Program	69
Lampiran 6	Form Bimbingan	70
Lampiran 7	Pengecekan Hasil Similarity Turnitin	71
Lampiran 8	Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)	80
Lampiran 9	Research Paper	81
Lampiran 10	Surat Kerja Sama MoU FTI dan FKUB 2024-2027	91

