

**SELEKSI MIRNA BERBASIS FUZZY MUTUAL  
INFORMATION UNTUK KLASIFIKASI  
STADIUM KANKER LAMBUNG**



**LAPORAN RESEARCH TECHNOLOGY PROGRAM**

**WENCYVONNEYWIJAYA  
00000071833**

**PROGRAM STUDI INFORMATIKA  
FAKULTASTEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA  
NUSANTARA TANGERANG  
2026**

**SELEKSI MIRNA BERBASIS FUZZY MUTUAL  
INFORMATION UNTUK KLASIFIKASI  
STADIUM KANKER LAMBUNG**



**LAPORAN RESEARCH TECHNOLOGY PROGRAM**

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

**WENCYVONNEYWIJAYA  
00000071833**

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

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Wency Yvonney Wijaya  
Nomor Induk Mahasiswa : 00000071833  
Program Studi : Informatika

Skripsi dengan judul:

**Seleksi miRNA Berbasis *Fuzzy Mutual Information* untuk Klasifikasi Stadium Kanker Lambung**

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, 01 Januari 2026



Wency Yvonney Wijaya

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 : Wency Yvonney Wijaya  
NIM : 00000071833  
Program Studi : Informatika  
Judul Laporan : Seleksi miRNA Berbasis *Fuzzy*  
*Mutual Information* untuk Klasifikasi  
Stadium Kanker Lambung

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



Wency Yvonney Wijaya



## HALAMAN PERNYATAAN KEABSAHAN PERUSAHAAN

Yang bertanda tangan dibawah ini:

Nama : Wency Yvonney Wijaya  
NIM : 00000071833  
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, 01 Januari 2026

Menyatakan



Wency Yvonney Wijaya

## HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Wency Yvonney Wijaya  
NIM : 00000071833  
Program Studi : Informatika  
Jenjang : S1  
Judul Karya Ilmiah : Seleksi miRNA Berbasis *Fuzzy Mutual Information* untuk Klasifikasi Stadium Kanker Lambung

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



Wency Yvonney Wijaya

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 seleksi fitur berbasis *fuzzy logic mutual information* untuk identifikasi biomarker miRNA dalam klasifikasi stadium kanker lambung. 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., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Aditiyawan, S.Kom., M.Si., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya laporan penelitian 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 laporan MBKM ini.

Semoga penelitian ini dapat memberikan kontribusi bagi perkembangan ilmu pengetahuan, khususnya dalam bidang informatika dan onkologi, serta bermanfaat bagi upaya peningkatan diagnosis dan pengobatan kanker lambung.

Tangerang, 02 Januari 2026



Wency Yvonne Wijaya

# SELEKSI MIRNA BERBASIS FUZZY MUTUAL INFORMATION UNTUK KLASIFIKASI STADIUM KANKER LAMBUNG

Wency Yvonne Wijaya

## ABSTRAK

Kanker lambung merupakan salah satu penyebab utama mortalitas global, dengan tingkat prognosis yang sangat bergantung pada stadium penyakit saat diagnosis. Penelitian ini bertujuan mengembangkan model klasifikasi stadium kanker lambung berbasis data ekspresi microRNA (miRNA) dengan mengintegrasikan seleksi fitur berbasis *fuzzy information theory* dan algoritma *supervised machine learning*. Dataset diperoleh dari The Cancer Genome Atlas (TCGA) dan dikelompokkan menjadi stadium awal dan stadium lanjut berdasarkan *AJCC pathologic T-stage*. Mengingat karakteristik data miRNA yang berdimensi tinggi dan memiliki tingkat redundansi yang signifikan, diterapkan metode seleksi fitur berbasis *True Fuzzy Mutual Information* untuk mengidentifikasi 15 miRNA paling informatif pada setiap *fold*. Evaluasi model dilakukan menggunakan skema *10-fold stratified cross-validation* dengan tiga algoritma klasifikasi, yaitu *Support Vector Machine* (SVM), *Random Forest* (RF), dan *K-Nearest Neighbors* (KNN), serta teknik penyeimbangan data SMOTETomek. Hasil eksperimen menunjukkan performa klasifikasi yang relatif stabil, dengan nilai *accuracy* berkisar antara 0,62–0,65, F1-score antara 0,69–0,76, dan AUC antara 0,65–0,67. Analisis *confusion matrix* mengindikasikan sensitivitas yang lebih tinggi dalam mendeteksi stadium lanjut, yang secara klinis memiliki prioritas lebih besar. Selain itu, beberapa miRNA terpilih secara konsisten pada seluruh *fold* dan lintas model, menunjukkan potensi sebagai kandidat *biomarker* stadium kanker lambung. Temuan ini menegaskan bahwa integrasi seleksi fitur berbasis *fuzzy* dan *machine learning* efektif untuk analisis data miRNA berdimensi tinggi.

**Kata kunci:** *Fuzzy Mutual Information*, Kanker Lambung, *Machine Learning*, MicroRNA, Seleksi Fitur

U M N  
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

# **FUZZY MUTUAL INFORMATION-BASED MIRNA SELECTION FOR GASTRIC CANCER STAGE CLASSIFICATION**

Wency Yvonney Wijaya

## **ABSTRACT**

*Gastric cancer is one of the leading causes of cancer-related mortality worldwide, with patient prognosis strongly influenced by the stage of the disease at diagnosis. This study aims to develop a gastric cancer stage classification model based on microRNA (miRNA) expression data by integrating fuzzy information theory-based feature selection with supervised machine learning algorithms. The dataset was obtained from The Cancer Genome Atlas (TCGA) and categorized into early-stage and late-stage groups according to the AJCC pathological T-stage. Given the high dimensionality and redundancy inherent in miRNA expression data, a True Fuzzy Mutual Information-based feature selection method was applied to identify the 15 most informative miRNAs in each fold. Model evaluation was conducted using a 10-fold stratified cross-validation scheme with three classification algorithms such as, Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN) in conjunction with the SMOTETomek data balancing technique. Experimental results demonstrated relatively stable classification performance, with accuracy ranging from 0.62 to 0.65, F1-scores between 0.69 and 0.76, and AUC values from 0.65 to 0.67. Confusion matrix analysis indicated higher sensitivity toward late stage detection, which is clinically prioritized. Moreover, several miRNAs were consistently selected across folds and models, suggesting their potential as candidate biomarkers for gastric cancer staging. These findings indicate that the integration of fuzzy-based feature selection and machine learning is effective for analyzing high-dimensional miRNA data.*

**Keywords:** Feature Selection, Fuzzy Mutual Information, Gastric Cancer, Machine Learning, MicroRNA



## DAFTAR ISI

|                                                                          |           |
|--------------------------------------------------------------------------|-----------|
| HALAMAN JUDUL . . . . .                                                  | i         |
| PERNYATAAN TIDAK MELAKUKAN PLAGIAT . . . . .                             | ii        |
| PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (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      |
| <b>BAB 1 PENDAHULUAN . . . . .</b>                                       | <b>1</b>  |
| 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 . . . . .                                         | 3         |
| 1.6 Luaran Penelitian . . . . .                                          | 3         |
| 1.7 Manfaat Penelitian . . . . .                                         | 4         |
| 1.8 Sistematika Penulisan . . . . .                                      | 4         |
| <b>BAB 2 LANDASAN TEORI . . . . .</b>                                    | <b>5</b>  |
| 2.1 MicroRNA sebagai Biomarker Molekuler Kanker Lambung . . . . .        | 5         |
| 2.2 Stage Stadium Kanker . . . . .                                       | 5         |
| 2.3 Seleksi Fitur pada Data Berdimensi Tinggi . . . . .                  | 6         |
| 2.4 Teori Fuzzy dalam Analisis Data Biomedis . . . . .                   | 6         |
| 2.5 Fuzzy Mutual Information dan True Fuzzy Mutual Information . . . . . | 7         |
| 2.6 Penanganan Ketidakseimbangan Data dengan SMOTETomek . . . . .        | 7         |
| 2.7 Machine Learning untuk Klasifikasi Stadium Kanker . . . . .          | 8         |
| 2.8 Evaluasi Model dan Validasi . . . . .                                | 8         |
| <b>BAB 3 METODOLOGI PENELITIAN . . . . .</b>                             | <b>10</b> |
| 3.1 Metodologi Penelitian . . . . .                                      | 10        |
| 3.2 Pengumpulan dan Persipan Dataset . . . . .                           | 11        |
| 3.3 Seleksi Fitur Berbasis True Fuzzy Mutual Information . . . . .       | 13        |
| 3.3.1 Normalisasi Data Min-Max Scaling . . . . .                         | 14        |
| 3.3.2 Relasi Ekuivalensi Fuzzy (Fuzzy Equivalence Relation) . . . . .    | 15        |
| 3.3.3 Entropi Fuzzy (Fuzzy Entropy) . . . . .                            | 16        |
| 3.3.4 Entropi Fuzzy Label (Fuzzy Entropy of Class Label) . . . . .       | 16        |
| 3.3.5 Entropi Gabungan Fuzzy (Fuzzy Joint Entropy) . . . . .             | 16        |
| 3.3.6 True Fuzzy Mutual Information (TFMI) . . . . .                     | 17        |
| 3.4 Penanganan Ketidakseimbangan Kelas . . . . .                         | 17        |
| 3.4.1 Synthetic Minority Over-sampling Technique (SMOTE) . . . . .       | 17        |
| 3.4.2 Tomek Links . . . . .                                              | 18        |
| 3.4.3 Kombinasi SMOTETomek . . . . .                                     | 18        |
| 3.4.4 Penerapan pada Pipeline Penelitian . . . . .                       | 18        |
| 3.5 Machine Learning Model . . . . .                                     | 19        |
| 3.5.1 Support Vector Machine (SVM) . . . . .                             | 19        |
| 3.5.2 Random Forest (RF) . . . . .                                       | 20        |
| 3.5.3 K-Nearest Neighbors (KNN) . . . . .                                | 20        |
| 3.6 Metrik Evaluasi Kinerja Model . . . . .                              | 20        |
| 3.7 Analisis Confusion Matrix dan Kurva ROC . . . . .                    | 21        |
| 3.7.1 Confusion Matrix . . . . .                                         | 21        |
| 3.7.2 Kurva ROC . . . . .                                                | 21        |
| 3.8 Dokumentasi dan Analisis Stabilitas Fitur . . . . .                  | 21        |
| <b>BAB 4 HASIL DAN DISKUSI . . . . .</b>                                 | <b>22</b> |
| 4.1 Hasil Eksperimen Klasifikasi . . . . .                               | 22        |
| 4.1.1 Performa Model Klasifikasi . . . . .                               | 22        |

|                |                                   |    |
|----------------|-----------------------------------|----|
| 4.1.2          | Analisis Confusion Matrix dan ROC | 23 |
| 4.1.3          | Confusion Matrix                  | 23 |
| 4.1.4          | ROC Curve dan AUC                 | 24 |
| 4.1.5          | Analisis Stabilitas Fitur miRNA   | 26 |
| 4.1.6          | Analisis Korelasi Fitur           | 27 |
| 4.1.7          | Relevansi Biologis                | 28 |
| 4.1.8          | Ringkasan                         | 28 |
| 4.2            | Diskusi                           | 28 |
| BAB 5          | SIMPULAN DAN SARAN                | 30 |
| 5.1            | Simpulan                          | 30 |
| 5.2            | Saran                             | 30 |
| DAFTAR PUSTAKA |                                   | 32 |



## DAFTAR TABEL

|           |                                                                                               |    |
|-----------|-----------------------------------------------------------------------------------------------|----|
| Tabel 3.1 | Distribusi data berdasarkan AJCC pathologic T-stage pada dataset TCGA kanker lambung. . . . . | 12 |
| Tabel 3.2 | Distribusi dataset berdasarkan pengelompokan stadium awal dan stadium lanjut. . . . .         | 13 |
| Tabel 3.3 | Parameter default dan nilai yang digunakan untuk tuning model machine learning . . . . .      | 19 |
| Tabel 4.1 | Performa Model Klasifikasi (Mean $\pm$ Std) . . . . .                                         | 22 |
| Tabel 4.2 | Top 10 miRNA Paling Stabil . . . . .                                                          | 27 |



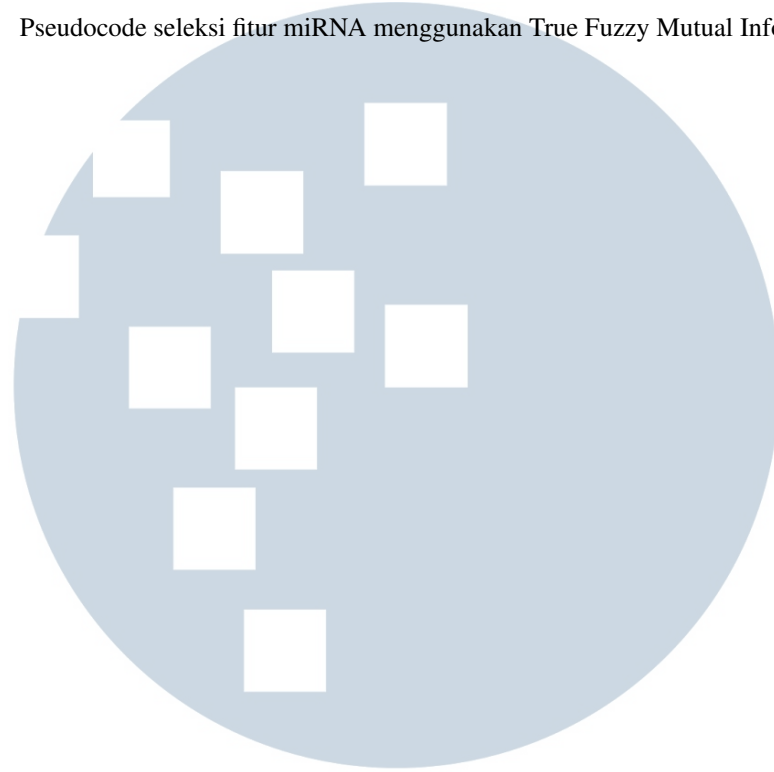
## DAFTAR GAMBAR

|            |                                                                                                                                                                                                                                                            |    |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Gambar 3.1 | Flowchart Metodologi Penelitian . . . . .                                                                                                                                                                                                                  | 11 |
| Gambar 3.2 | T-stage Distribution . . . . .                                                                                                                                                                                                                             | 12 |
| Gambar 4.1 | Confusion matrix untuk model SVM: (a) aggregated confusion matrix hasil penggabungan seluruh fold 10-fold cross-validation, dan (b) normalized confusion matrix yang merepresentasikan proporsi prediksi pada masing-masing kelas. . . . .                 | 23 |
| Gambar 4.2 | Confusion matrix untuk model Random Forest: (a) aggregated confusion matrix hasil penggabungan seluruh fold 10-fold cross-validation, dan (b) normalized confusion matrix yang merepresentasikan proporsi prediksi pada masing-masing kelas. . . . .       | 23 |
| Gambar 4.3 | Confusion matrix untuk model K-Nearest Neighbors: (a) aggregated confusion matrix hasil penggabungan seluruh fold 10-fold cross-validation, dan (b) normalized confusion matrix yang merepresentasikan proporsi prediksi pada masing-masing kelas. . . . . | 24 |
| Gambar 4.4 | Mean ROC curve untuk SVM dengan fitur miRNA terpilih. . . . .                                                                                                                                                                                              | 25 |
| Gambar 4.5 | Mean ROC curve untuk RF dengan fitur miRNA terpilih. . . . .                                                                                                                                                                                               | 25 |
| Gambar 4.6 | Mean ROC curve untuk KNN dengan fitur miRNA terpilih. . . . .                                                                                                                                                                                              | 26 |
| Gambar 4.7 | Matriks korelasi miRNA terpilih. . . . .                                                                                                                                                                                                                   | 27 |



## DAFTAR KODE

Kode 3.1      Pseudocode seleksi fitur miRNA menggunakan True Fuzzy Mutual Information    14



UMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR RUMUS

|            |                                         |    |
|------------|-----------------------------------------|----|
| Rumus 3.1  | Min–Max Normalization . . . . .         | 15 |
| Rumus 3.2  | Fuzzy Equivalence Relation . . . . .    | 15 |
| Rumus 3.3  | Fuzzy Entropy . . . . .                 | 16 |
| Rumus 3.4  | Fuzzy Entropy of Class Label . . . . .  | 16 |
| Rumus 3.5  | Fuzzy Joint Entropy . . . . .           | 16 |
| Rumus 3.6  | True Fuzzy Mutual Information . . . . . | 17 |
| Rumus 3.7  | SMOTE Sample Generation . . . . .       | 17 |
| Rumus 3.8  | Accuracy . . . . .                      | 20 |
| Rumus 3.9  | Precision . . . . .                     | 20 |
| Rumus 3.10 | Recall / Sensitivity . . . . .          | 20 |
| Rumus 3.11 | Specificity . . . . .                   | 20 |
| Rumus 3.12 | F1-Score . . . . .                      | 20 |



## DAFTAR LAMPIRAN

|             |                                                                      |    |
|-------------|----------------------------------------------------------------------|----|
| Lampiran 1  | MBKM-01 Cover Letter MBKM Research & Technology Program . . . . .    | 34 |
| Lampiran 2  | MBKM-02 Research & Technology Program Card . . . . .                 | 35 |
| Lampiran 3  | MBKM-03 Daily Task Research & Technology Program . . . . .           | 36 |
| Lampiran 4  | MBKM-04 Verification Form of Research & Technology Program . . . . . | 67 |
| Lampiran 5  | MBKM-05 Letter of Acceptance Research & Technology Program . . . . . | 68 |
| Lampiran 6  | Form Bimbingan . . . . .                                             | 69 |
| Lampiran 7  | Pengecekan Hasil <i>Similarity</i> Turnitin . . . . .                | 70 |
| Lampiran 8  | Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI) . . . . .   | 75 |
| Lampiran 9  | Research Paper . . . . .                                             | 76 |
| Lampiran 10 | Surat Kerja Sama MoU FTI dan FKUB 2024-2027 . . . . .                | 87 |

