

**DETEKSI DINI DAN KLASIFIKASI STAGING KANKER
PAYUDARA MENGGUNAKAN DNA METHYLATION
BERBASIS MULTI-LAYER PERCEPTRON**



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

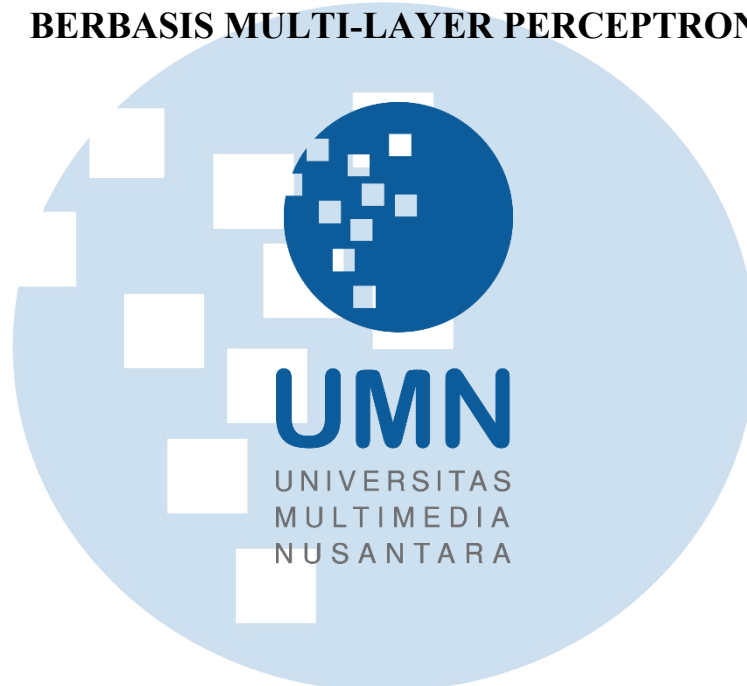
LAPORAN MBKM PENELITIAN

VINCENT KURNIAWAN

00000068404

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK & INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**DETEKSI DINI DAN KLASIFIKASI STAGING KANKER
PAYUDARA MENGGUNAKAN DNA METHYLATION
BERBASIS MULTI-LAYER PERCEPTRON**



LAPORAN MBKM PENELITIAN

UMN

VINCENT KURNIAWAN

00000068404

**UNIVERSITAS
MULTIMEDIA
NUSANTARA**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK & INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Vincent Kurniawan

Nomor Induk Mahasiswa : 00000068404

Program studi : Informatika

Laporan MBKM Penelitian dengan judul:

**DETEKSI DINI DAN KLASIFIKASI STAGING KANKER PAYUDARA
MENGUNAKAN DNA METHYLATION BERBASIS MULTI-LAYER
PERCEPTRON**

merupakan hasil karya saya sendiri bukan plagiat dari karya 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 MBKM, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk laporan MBKM yang telah saya tempuh.

Tangerang, 20 Juni 2025



Vincent Kurniawan

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertandatangan di bawah ini:

Nama : Vincent Kurniawan
NIM : 00000068404
Program Studi : Informatika
Fakultas : Teknik & Informatika
Jenis Karya : Laporan MBKM Penelitian

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Multimedia Nusantara Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul:

DETEKSI DINI DAN KLASIFIKASI STAGING KANKER PAYUDARA MENGUNAKAN DNA METHYLATION BERBASIS MULTI-LAYER PERCEPTRON

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 20 Juni 2025

Yang menyatakan,

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Vincent Kurniawan

KATA PENGANTAR

Puji Syukur atas selesainya penulisan laporan penelitian ini dengan judul: Deteksi Dini Dan Klasifikasi Staging Kanker Payudara Menggunakan DNA Methylation Berbasis Multi-Layer Perceptron dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik & Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo S.T, M.Sc., selaku Dekan Fakultas Teknik & Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc, OCA., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Vincentius Kurniawan, S.Kom., M. Eng.sc, sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Bapak David Agustriawan Ph.D., sebagai Pembimbing Lapangan yang telah memberikan bimbingan, dan arahan, atas terselesainya laporan MBKM Penelitian ini.

Semoga karya ilmiah ini dapat memberikan dampak yang positif dan signifikan di dunia medis terutama untuk klasifikasi stadium pada kanker payudara

Tangerang, 20 Juni 2025



Vincent Kurniawan

DETEKSI DINI DAN KLASIFIKASI STAGING KANKER PAYUDARA MENGUNAKAN DNA METHYLATION BERBASIS MULTI-LAYER PERCEPTRON

Vincent Kurniawan

ABSTRAK

Kanker payudara merupakan salah satu penyebab kematian tertinggi pada wanita, dan diagnosis dini sangat penting untuk meningkatkan tingkat kelangsungan hidup. Penelitian ini bertujuan untuk membangun model deteksi dan klasifikasi *stadium* kanker payudara berdasarkan pola *metilasi DNA* menggunakan algoritma *Multi-Layer Perceptron* (MLP). Data yang digunakan berasal dari basis data *TCGA-BRCA* dan mencakup informasi *DNA methylation Illumina 450K* serta data klinis pasien. Fitur dipilih melalui kombinasi metode *Limma*, *MethylSuite*, dan *Recursive Feature Elimination* (RFE), dengan fokus pada identifikasi *biomarker* yang paling signifikan. Arsitektur MLP dirancang dalam dua konfigurasi: sederhana dan kompleks, dan model dievaluasi menggunakan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa kombinasi *MethylSuite* dan *RFE* dengan 100 *CpG sites* menghasilkan performa terbaik (akurasi 96,5%) dibandingkan metode lain. Penerapan *SMOTE* juga terbukti meningkatkan kinerja model pada *dataset* tidak seimbang. Namun, evaluasi pada populasi kulit hitam menunjukkan adanya *bias*, dengan penurunan drastis akurasi hingga 22%. Hal ini menekankan pentingnya representasi data yang adil dalam pembangunan model medis. Penelitian ini menunjukkan potensi besar penggunaan analisis *metilasi DNA* dan *deep learning* untuk personalisasi diagnosis kanker payudara.

Kata kunci: kanker payudara, *DNA methylation*, *Multi-Layer Perceptron*, klasifikasi stadium

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

EARLY DETECTION AND STAGING CLASSIFICATION OF BREAST CANCER BASED ON DNA METHYLATION USING MULTI-LAYER PERCEPTRON

Vincent Kurniawan

ABSTRACT

Breast cancer is one of the leading causes of death among women, and early diagnosis is crucial for improving survival rates. This study aims to develop detection and staging classification models for breast cancer based on DNA methylation patterns using the Multi-Layer Perceptron (MLP) algorithm. The data used were obtained from the TCGA-BRCA database, including Illumina 450K DNA methylation profiles and clinical patient records. Features were selected using a combination of Limma, MethylSuite, and Recursive Feature Elimination (RFE), with a focus on identifying the most significant biomarkers. The MLP architecture was designed in two configurations: simple and complex, and model performance was evaluated using metrics such as accuracy, precision, recall, and F1-score. Results showed that the combination of MethylSuite and RFE with 100 CpG sites yielded the best performance (96.5% accuracy) compared to other methods. The application of SMOTE further improved the model's performance on imbalanced datasets. However, evaluation on a Black patient population revealed potential bias, with a drastic drop in accuracy to 22%. This highlights the importance of fair data representation in developing medical models. This study demonstrates the significant potential of DNA methylation analysis and deep learning for personalized breast cancer diagnosis.

Keywords: *Breast cancer, DNA methylation, Multi-Layer Perceptron, staging classification*

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

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERNYATAAN TIDAK PLAGIAT.....	ii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iii
KATA PENGANTAR	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	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 II TINJAUAN PUSTAKA.....	5
2.1 Kanker Payudara	5
2.2 Next Generation Sequencing (NGS).....	5
2.3 DNA Methylation	6
2.4 Limma & Methysuite.....	7
2.5 Recursive Feature Elimination (RFE).....	8
2.6 Artificial Neural Network	8
2.7 Multi-Layer Perceptron.....	9
2.8 Evaluation Metrics	11
2.8.1 Confusion Matrix.....	11
2.8.2 Precision.....	12
2.8.3 Recall	12
2.8.4 F1-Score.....	13
BAB III METODE PENELITIAN.....	14

3.1 Analisis Kebutuhan	14
3.2 Pengumpulan Data	15
3.3 Preprocessing Data	16
3.4 Data Labelling	18
3.5 Feature Selection	20
3.6 Proses Training Model	23
3.7 Evaluasi Model.....	25
BAB IV HASIL DAN PEMBAHASAN	27
4.1 Spesifikasi Sistem	27
4.1.1 Hardware & Software	27
4.2 Pemilihan Gene/Fitur	27
4.2.1 Pengecekan Mutasi pada Gen	31
4.3 Skema Pembangunan Model.....	32
4.4 Hasil Training.....	33
4.5 Evaluasi Model MLP	35
4.5.1 Evaluasi Model Menggunakan SMOTE.....	36
4.5.2 Evaluasi Model menggunakan populasi pasien kulit hitam	39
4.6 Diskusi & Pembahasan	40
BAB V SIMPULAN SARAN.....	42
5.1 Simpulan	42
5.2 Saran.....	42
DAFTAR PUSTAKA	44
LAMPIRAN	48

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

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion Matrix	12
Tabel 3. 1 Sampel Data TCGA-BRCA.methylation450.....	15
Tabel 3. 2 Penjelasan dari Fitur T	19
Tabel 3. 3 Penjelasan dari Fitur N.....	20
Tabel 3. 4 Skema Feature Selection Penelitian ini.....	22
Tabel 3. 5 Arsitektur model Multi-Layer Perceptron pada penelitian ini	24
Tabel 3. 6 Arsitektur model Multi-Layer Perceptron ke-2 pada penelitian ini.....	25
Tabel 4. 1 Tabel Hasil Feature Selection	28
Tabel 4. 2 List nama gen hasil konversi probes CpG	29
Tabel 4. 3 Hasil tes GSEA dan list set gen	30
Tabel 4. 4 Skenario Pembangunan model pada penelitian ini	33
Tabel 4. 5 Hasil Evaluasi dari tiap skenario Pembangunan model	36
Tabel 4. 6 Hasil Evaluasi model dengan SMOTE	37
Tabel 4. 7 Hasil Evaluasi model pada Ras Kulit Hitam.....	39



DAFTAR GAMBAR

Gambar 2. 1 Mekanisme dari DNA Methylation.....	7
Gambar 2. 2 Struktur Artificial Neural Network (ANN).....	9
Gambar 2. 3 Arsitektur Multi-Layer Perceptron.....	11
Gambar 3. 1 Alur Penelitian yang digarap.....	14
Gambar 3. 2 Tahapan pemecahan data TCGA-BRCA-Methylation450k	17
Gambar 3. 3 Labelling Stage Kanker Payudara berdasarkan AJCC Manual.....	19
Gambar 3. 4 Flowchart pemilihan fitur.....	21
Gambar 4. 1 Hasil Pengecekan Mutasi pada Gen	32
Gambar 4. 2 Hasil Pengecekan Mutasi pada Gen	34
Gambar 4. 3 Grafik accuracy dan loss over epochs pada skenario 2	34
Gambar 4. 4 Grafik accuracy dan loss over epochs pada skenario 4.....	35



DAFTAR LAMPIRAN

Lampiran A. Surat Pengantar MBKM – MBKM 01.....	48
Lampiran B. Kartu MBKM (MBKM 02)	49
Lampiran C. Daily Task Penelitian (MBKM 03).....	50
Lampiran D. Form Bimbingan.....	60
Lampiran E. Lembar Verifikasi Laporan MBKM Penelitian (MBKM 04).....	61
Lampiran F. Surat Penerimaan MBKM (LoA)	62
Lampiran G. Pengecekan hasil Turnitin	63

