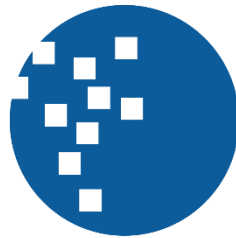


**KLASIFIKASI STADIUM KANKER PROSTAT
MENGUNAKAN MODEL XGBOOST
PADA DATA MIRNA**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

LAPORAN MBKM PENELITIAN

JHENO SYECHLO

00000067872

**PROGRAM STUDI INFORMATIKA
TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**KLASIFIKASI STADIUM KANKER PROSTAT
MENGUNAKAN MODEL XGBOOST
PADA DATA MIRNA**



LAPORAN MBKM PENELITIAN



JHENO SYECHLO

00000067872

**PROGRAM STUDI INFORMATIKA
TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Jheno Syechlo

Nomor Induk Mahasiswa : 00000067872

Program studi : Informatika

Laporan MBKM Penelitian dengan judul:

KLASIFIKASI STADIUM KANKER PROSTAT MENGGUNAKAN MODEL XGBOOST PADA DATA MIRNA

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

Jheno Syechlo

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

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

Nama : Jheno Syechlo
NIM : 00000067872
Program Studi : Informatika
Fakultas : Teknik dan 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:

KLASIFIKASI STADIUM KANKER PROSTAT MENGGUNAKAN MODEL XGBOOST PADA DATA MIRNA

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,



Jheno Syechlo

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Laporan MBKM Penelitian ini dengan judul: “Klasifikasi Stadium Kanker Prostat Menggunakan Model XGBoost pada Data miRNA” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Strata 1 Jurusan Informatika Pada Fakultas Teknik dan 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 Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Bapak Dr. Ivransa Zuhdi Pane, B.Eng., M.Eng., 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, arahan, dan motivasi atas terselesainya laporan MBKM Penelitian.
6. Keluarga yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan laporan MBKM ini.

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

Tangerang, 20, Juni 2025



Jheno Syechlo

KLASIFIKASI STADIUM KANKER PROSTAT MENGUNAKAN MODEL XGBOOST PADA DATA MIRNA

Jheno Syechlo

ABSTRAK

Penelitian ini bertujuan untuk membangun model klasifikasi stadium kanker prostat menggunakan algoritma XGBoost pada data ekspresi miRNA, dengan fokus pada analisis berdasarkan ras. Pemilihan topik ini didasari oleh tingginya angka kematian akibat kanker prostat serta ketimpangan diagnosis antar ras, khususnya antara ras kulit putih dan kulit hitam. Data yang digunakan diperoleh dari GDC TCGA XenaBrowser, mencakup ekspresi miRNA serta data klinis pasien. Penelitian ini menerapkan beberapa metode seleksi fitur seperti *Lasso* + RFE, EdgeR, dan ROC, serta teknik penyeimbangan data seperti RandomOversampler, SMOTE, SMOTEEN dan BorderlineSMOTE. Hasil penelitian menunjukkan bahwa model XGBoost mampu mencapai akurasi hingga 99% pada data ras kulit putih. Namun, saat diuji pada data ras kulit hitam, akurasi menurun menjadi 84–89%, yang mengindikasikan adanya keterbatasan model antar ras. Kendala utama dalam penelitian ini adalah keterbatasan jumlah data dan solusi yang diterapkan dalam membandingkan performa antar ras untuk mengidentifikasi dampak ras terhadap klasifikasi stadium kanker. Penelitian ini diharapkan dapat menjadi pijakan awal dalam pengembangan model klasifikasi kanker yang lebih inklusif dan adil.

Kata kunci: kanker prostat, miRNA, XGBoost, seleksi fitur, perbedaan ras



PROSTATE CANCER STAGE CLASSIFICATION USING XGBOOST MODEL ON MIRNA DATA

Jheno Syechlo

ABSTRACT

This study aims to develop a prostate cancer stage classification model using the XGBoost algorithm on miRNA expression data, with a focus on race-based analysis. The choice of this topic is based on the high mortality rate from prostate cancer and diagnostic disparities between races, particularly between White and Black populations. The data used were obtained from the GDC TCGA XenaBrowser, comprising miRNA expression and patient clinical data. This study applied several feature selection methods such as Lasso + RFE, edgeR, and ROC, as well as data balancing techniques including RandomOversampler, SMOTE, SMOTEEN, and BorderlineSMOTE. The results show that the XGBoost model achieved an accuracy of up to 99% on data from White patients. However, when tested on data from Black patients, accuracy decreased to 84–89%, indicating limitations in cross-race performance. The main challenge in this study was the limited amount of data and the approach used to compare performance across races to identify the impact of race on cancer stage classification. This study is expected to serve as an initial foundation for developing more inclusive and equitable cancer classification models.

Keywords: prostate cancer, miRNA, XGBoost, feature selection, racial disparity

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

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	2
1.3. Tujuan Penelitian	2
1.4. Urgensi Penelitian	2
1.5. Luaran Penelitian	3
1.6. Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Kanker Prostat.....	4
2.2 Data miRNA	5
2.3 Xgboost	6
2.4 Seleksi Fitur	8
2.4.1. Lasso + RFE.....	8
2.4.2. EdgeR.....	9
2.5 Matrik Evaluasi	10
BAB III METODE PENELITIAN.....	12
3.1. Pengumpulan Data	12
3.2 Preproses Data.....	13
3.3 Seleksi Fitur	14
3.4 Pembuatan Model.....	15

BAB IV HASIL DAN PEMBAHASAN	17
4.1 Pengumpulan dan Preproses Data	17
4.2 Seleksi Fitur	17
4.3 Pembangunan Model.....	20
4.3.1. 3 Fitur miRNA	20
4.3.2. 15 Fitur miRNA	21
4.3.3. 16 Fitur miRNA	22
4.3.4. 17 Fitur miRNA	23
4.3.5. 18 Fitur miRNA	24
4.3.6. 19 Fitur miRNA	25
4.3.7. 7 Fitur miRNA	26
4.4 Pembahasan.....	27
BAB V SIMPULAN SARAN.....	30
5.1 Simpulan	30
5.2 Saran.....	30
DAFTAR PUSTAKA	31
LAMPIRAN.....	34



DAFTAR TABEL

Tabel 3. 1 Syarat pemberian kanker tahap 1-4.....	13
Tabel 3. 2 Metode Pemilihan Fitur	14
Tabel 3. 3 Skenario Pembuatan Model Menggunakan Xgboost.....	15
Tabel 4. 1 Dataset miRNA	17
Tabel 4. 2 Hasil Seleksi Fitur	18
Tabel 4. 3 Biomarker miRNA pada Database.....	18
Tabel 4. 4 Hasil menggunakan 3 fitur miRNA	20
Tabel 4. 5 Hasil menggunakan 15 fitur miRNA	21
Tabel 4. 6 Hasil menggunakan 16 fitur miRNA	22
Tabel 4. 7 Hasil menggunakan 17 fitur miRNA	23
Tabel 4. 8 Hasil menggunakan 18 fitur miRNA	24
Tabel 4. 9 Hasil menggunakan 19 fitur miRNA	25
Tabel 4. 10 Hasil menggunakan 7 fitur miRNA	26
Tabel 4. 11 Akurasi terbaik dari setiap fitur miRNA yang digunakan	28



DAFTAR GAMBAR

Gambar 2. 1 Central Dogma miRNA.....	5
Gambar 3. 1 Alur Penelitian Prediksi Kanker Prostat.....	12
Gambar 3. 2 Pipeline Pembangunan Model.....	16



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Surat Pengantar MBKM (MBKM 01).....	34
Lampiran B Kartu MBKM (MBKM 02)	35
Lampiran C Daily Task (MBKM 03).....	36
Lampiran D Lembar Verifikasi Laporan MBKM Kewirausahaan (MBKM 04)..	47
Lampiran E Surat Penerimaan MBKM (LoA).....	48
Lampiran F Pengecekan Hasil Turn It In	49



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA