

**PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI
CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR
GLCM**



**KARUNIA AGUSTIANI
NIM. 2110511002**

**S-1 INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
JAKARTA
2025**

**PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI
CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR
GLCM**

KARUNIA AGUSTIANI

NIM. 2110511002

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer

S1 INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

JAKARTA

2025

PERNYATAAN ORISINALITAS

Tugas akhir ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Karunia Agustiani

NIM : 2110511002

Tanggal : 26 Juni 2025

Judul Skripsi : **PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR GLCM**

Bilamana pada kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 26 Juni 2025

Yang menyatakan,



Karunia Agustiani

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Karunia Agustiani

NIM : 2110511002

Fakultas : Ilmu Komputer

Program Studi : S-1 Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan karya ilmiah saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exchange Royalty Free Right*) untuk dipublikasi dengan judul:

PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR GLCM

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media atau memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat. Dan mempublikasi artikel ilmiah selama tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.

Dengan pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 26 Juni 2025

Yang menyatakan,



Karunia Agustiani

LEMBAR PENGESAHAN

Judul : Penerapan Extreme Gradient Boosting untuk Klasifikasi Citra MRI
Prostat Berdasarkan Ekstraksi Fitur GLCM
Nama : Karunia Agustiani
NIM : 2110511002
Program Studi : SI Informatika

Disetujui oleh :

Penguji 1:

Dr. Widya Cholil, M.I.T.

Penguji 2:

Muhammad Adrezo, S.Kom., M.Sc.

Pembimbing 1:

Neny Rosmawarni, S.Kom., M.Kom.

Pembimbing 2:

Hamonangan Kinantan Prabu, S.T, M.T.



Diketahui oleh:

Koordinator Program Studi:

Dr. Widya Cholil, M.I.T

NIP. 221112080

Dekan Fakultas Ilmu Komputer:

Prof. Dr. Ir. Supriyanto, M.Sc., IPM.

NIP. 197605082003121002



Tanggal Ujian Tugas Akhir:

04 Juni 2025

PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR GLCM

KARUNIA AGUSTIANI

ABSTRAK

Kanker prostat merupakan salah satu jenis kanker yang umum terjadi pada pria dan menjadi salah satu penyebab utama kematian akibat kanker di dunia. Di Indonesia, jumlah kasus kanker prostat terus meningkat, sehingga dibutuhkan metode deteksi dini melalui pemeriksaan *Magnetic Resonance Imaging* (MRI). Citra MRI memiliki peran penting dalam diagnosis dini dan perencanaan pengobatan kanker prostat. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model *machine learning* yang dapat melakukan klasifikasi terhadap citra MRI prostat. Dataset yang digunakan terdiri atas 894 citra MRI prostat yang diklasifikasikan dalam dua kelas, yaitu positif dan negatif. Model klasifikasi dibangun menggunakan metode *Gray Level Co-Occurrence Matrix* (GLCM) untuk ekstraksi fitur tekstur pada citra MRI, serta algoritma *Extreme Gradient Boosting* (XGBoost) sebagai algoritma klasifikasi. Penelitian ini dilakukan dalam enam percobaan skema model berdasarkan kombinasi antara penggunaan *hyperparameter tuning Optuna* dan metode *resampling*, yaitu *Random Under-Sampling* (RUS) dan *Synthetic Minority Over-Sampling Technique* (SMOTE) untuk mengatasi ketidakseimbangan data. Hasil penelitian menunjukkan bahwa model yang dibangun dengan metode data *sampling* RUS memberikan kinerja evaluasi terbaik dengan tingkat akurasi pada data uji sebesar 0,82, presisi sebesar 0,82, *recall* sebesar 0,82, dan *F1-score* sebesar 0,82. Selain itu, model berhasil diterapkan dalam sistem antarmuka GUI yang dapat mempermudah proses klasifikasi dengan menggunakan model yang dibangun. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam mendukung proses deteksi dini kanker prostat serta menunjukkan potensi penggunaan algoritma *XGBoost* dalam klasifikasi citra medis.

Kata Kunci: Kanker Prostat, Citra MRI, GLCM, *XGBoost*

PENERAPAN *EXTREME GRADIENT BOOSTING* UNTUK KLASIFIKASI CITRA MRI KANKER PROSTAT BERDASARKAN EKSTRAKSI FITUR GLCM

KARUNIA AGUSTIANI

ABSTRACT

Prostate cancer is one of the most common cancers in men and one of the leading causes of cancer deaths in the world. In Indonesia, the number of prostate cancer cases continues to increase, so an early detection method through Magnetic Resonance Imaging (MRI) examination is needed. MRI images have an important role in early diagnosis and treatment planning of prostate cancer. Therefore, this study aims to develop a machine learning model that can classify prostate MRI images. The dataset used consists of 894 prostate MRI images classified into two classes, namely positive and negative. The classification model was built using the Gray Level Co-Occurrence Matrix (GLCM) method for extracting texture features in MRI images, and the Extreme Gradient Boosting (XGBoost) algorithm as the classification algorithm. This research was conducted in six model scheme experiments based on a combination of using Optuna hyperparameter tuning and resampling methods, namely Random Under-Sampling (RUS) and Synthetic Minority Over-Sampling Technique (SMOTE) to overcome data imbalance. The results show that the model built with the RUS method provides the best evaluation performance with an accuracy rate on test data of 0.82, precision of 0.82, recall of 0.82, and F1-score of 0.82. In addition, the model was successfully implemented in a GUI interface system that can facilitate the classification process using the built model. This research is expected to contribute in supporting the process of early detection of prostate cancer and demonstrate the potential use of the XGBoost algorithm in medical image classification.

Keyword: Prostate Cancer, MRI Image, GLCM, XGBoost

KATA PENGANTAR

Segala puji dan syukur peneliti panjatkan pada Tuhan Yang Maha Esa berkat rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan skripsi dengan judul “Penerapan *Extreme Gradient Boosting* untuk Klasifikasi Citra MRI Prostat Berdasarkan Ekstraksi Fitur GLCM”. Skripsi ini disusun sebagai syarat dalam penyelesaian mata kuliah Proposal studi S1-Informatika di Universitas Pembangunan Nasional Veteran Jakarta.

Dalam penyusunan skripsi ini, peneliti mendapatkan banyak dukungan, motivasi, serta bimbingan dari berbagai pihak yang berperan penting dalam penyelesaian penelitian ini. Setiap masukan dan arahan yang diberikan selama proses penyusunan sangat berharga dalam menghadapi tantangan penelitian. Oleh karena itu, pada kesempatan ini, peneliti ingin mengucapkan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta.
2. Ibu Dr. Widya Cholil, M.I.T. selaku Kepala Program Studi S-1 Informatika Universitas Pembangunan Nasional “Veteran” Jakarta.
3. Bapak Bayu Hananto, S.Kom., M.Kom. selaku Dosen Pembimbing Akademik.
4. Ibu Neny Rosmawarni, M. Kom. selaku Dosen Pembimbing 1 Proposal Skripsi.
5. Bapak Hamonangan Kinantan Prabu, S.T, M.T. selaku Dosen Pembimbing 2 Proposal Skripsi.
6. Orang tua dan keluarga, yang selalu memberikan dukungan, semangat, dan doa yang berharga sebagai pendorong dalam penyelesaian Proposal Skripsi.
7. Sahabat dan teman-teman seperjuangan mahasiswa atas semangat belajar dan dukungan yang menginspirasi selama perkuliahan.

Peneliti berharap penelitian ini dapat menjadi landasan awal dan referensi yang bermanfaat dalam pengembangan skripsi yang akan datang dan dapat memberi wawasan lebih terkait penerapan teknologi *machine learning*, khususnya dalam konteks pengembangan klasifikasi citra medis. Peneliti juga berharap skripsi ini dapat menghasilkan dan memberikan kontribusi positif terhadap bidang yang berkaitan.

Peneliti menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan dan keterbatasan, baik dari segi data, metode, maupun ruang lingkup pembahasan yang belum sepenuhnya optimal. Oleh karena itu, peneliti mengharapkan dengan tulus dan terbuka atas segala bentuk kritik dan saran yang membangun dalam perbaikan dan penyempurnaan penelitian ini di masa mendatang. Besar harapan

peneliti agar hasil karya ini dapat memberikan manfaat yang luas, tidak hanya bagi kalangan akademisi, tetapi juga bagi para praktisi, peneliti, dan pihak-pihak lain yang memiliki perhatian terhadap penerapan teknologi kecerdasan buatan dalam dunia medis. Semoga ke depan, penelitian ini dapat terus berkembang dan menjadi kontribusi berarti dalam pengembangan ilmu pengetahuan dan inovasi teknologi yang berkelanjutan.

Jakarta, 2 September 2024

Karunia Agustiani

2110511002

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	iii
DAFTAR TABEL	v
DAFTAR GAMBAR	vii
DAFTAR RUMUS	ix
DAFTAR LAMPIRAN	x
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan dan Manfaat Penelitian.....	4
1.5 Sistematika Penulisan	4
BAB 2. TINJAUAN PUSTAKA	6
2.1 Kajian Teoretis	6
2.1.1 Prostat.....	6
2.1.2 Kanker Prostat	6
2.1.3 <i>Magnetic Resonance Imaging (MRI)</i>	7
2.1.4 Pengolahan Citra Digital	8
2.1.5 Ekstraksi Ciri.....	9
2.1.6 <i>Label Encoding</i>	11
2.1.7 <i>Data Imbalance</i>	12
2.1.8 Normalisasi Data	12
2.1.9 <i>Ensemble Learning</i>	13
2.1.10 <i>Gradient Boosting Decision Tree</i>	14
2.1.11 <i>Extreme Gradient Boosting</i>	15
2.1.12 <i>Hyperparameter Tuning</i>	17
2.1.13 Evaluasi Model Klasifikasi	17
2.2 Penelitian Terdahulu.....	19
BAB 3. METODOLOGI.....	22
3.1 Metode Penelitian	22

3.1.1	Identifikasi Masalah.....	22
3.1.2	Studi Literatur	23
3.1.3	Pengumpulan Data	23
3.1.4	Praproses Citra	23
3.1.5	Ekstraksi Ciri <i>Gray Level Co-Occurrence Matrix</i> (GLCM).....	23
3.1.6	Praproses Data Tabular	23
3.1.7	Pembagian Data	24
3.1.8	<i>Resampling</i>	24
3.1.9	Klasifikasi <i>Extreme Gradient Boosting</i>	24
3.1.10	Evaluasi <i>Confusion Matrix</i>	25
3.1.11	Implementasi GUI.....	25
3.2	Rancangan Solusi / Metode yang Diusulkan.....	26
3.3	Metode Analisis.....	33
3.4	Alat dan Bahan Penelitian	45
3.5	Jadwal Penelitian	45
BAB 4. HASIL DAN PEMBAHASAN		47
4.1	Analisis Hasil	47
4.1.1	Pengumpulan Data	47
4.1.2	Praproses Citra	49
4.1.3	Ekstraksi Fitur <i>Gray Level Co-Occurrence Matrix</i> (GLCM)	52
4.1.4	Praproses Data.....	58
4.1.5	Pembagian Data	65
4.1.6	Hasil Model <i>Extreme Gradient Boosting</i>	67
4.1.8	Evaluasi Model.....	80
4.1.9	Implementasi GUI.....	95
4.2	Hasil dan Rekomendasi	101
BAB 5. PENUTUP		103
5.1	Kesimpulan.....	103
5.2	Saran	103
DAFTAR PUSTAKA		105
LAMPIRAN.....		109

DAFTAR TABEL

Tabel 2.1 Confusion Matrix	18
Tabel 2.2 Ringkasan penelitian terdahulu	19
Tabel 3.1 Contoh Confusion Matrix Dua Label	25
Tabel 3.2 Contoh Parameter XGBoost.....	31
Tabel 3.3 Contoh Hyperparameter Optimasi XGBoost	32
Tabel 3.4 Contoh Matriks Value Image.....	35
Tabel 3.5 Framework Matrix	35
Tabel 3.6 Hasil Co-Occurrence Matrix.....	35
Tabel 3.7 Contoh Hasil Perhitungan Fitur GLCM.....	38
Tabel 3.8 Contoh Data Dummy untuk Klasifikasi XGBoost.....	41
Tabel 3.9 Contoh Hasil Perhitungan Nilai Gradien dan Hessian.....	42
Tabel 3.10 Contoh Hasil Prediksi pada Iterasi Pertama.....	43
Tabel 3.11 Contoh Confusion Matrix Hasil Prediksi.....	44
Tabel 3.12 Jadwal Kegiatan	46
Tabel 4.1 Total Sumber Data Citra MRI.....	47
Tabel 4.2 Sampel Data Citra MRI.....	48
Tabel 4.3 Konfigurasi Ekstraksi Fitur GLCM	52
Tabel 4.4 Contoh plot dari Co-Occurrence Matrix	53
Tabel 4.5 Contoh Hasil Co-Occurrence Matrix	53
Tabel 4.6 Contoh Hasil Co-Occurrence Matrix setelah Normalisasi.....	54
Tabel 4.7 Contoh Hasil Perhitungan Fitur GLCM.....	56
Tabel 4.8 Hasil Nilai Fitur GLCM pada Data Citra MRI	57
Tabel 4.9 Detail Label Encoding pada Kelas Citra.....	58
Tabel 4.10 Hasil Fitur GLCM setelah Label Encoding	59
Tabel 4.11 Nilai Maksimum dan Minimum untuk Setiap Fitur GLCM	60
Tabel 4.12 Contoh Nilai GLCM pada Citra prostate_pos_1.jpg.	61
Tabel 4.13 Contoh Hasil Normalisasi Data pada Citra prostate_pos_1.jpg.....	63
Tabel 4.14 Hasil Normalisasi Data pada Data Citra MRI.....	64
Tabel 4.15 Hasil Pembagian Data.....	66
Tabel 4.16 Ringkasan Hasil Proses Resampling Jumlah Data Pelatihan.....	67
Tabel 4.17 Hasil Pelatihan Model tanpa Optuna	68
Tabel 4.18 Hasil Pencarian Hyperparameter Model tanpa Sampling	69
Tabel 4.19 Hasil Pelatihan Model dengan Optuna	71
Tabel 4.20 Hasil Pelatihan Model Dengan RUS dan Tanpa Optuna	72
Tabel 4.21 Hasil Pencarian Hyperparameter Model dengan RUS.....	74
Tabel 4.22 Hasil Pelatihan Model dengan RUS dan tanpa Optuna	75
Tabel 4.23 Hasil Pelatihan Model dengan SMOTE dan tanpa Optuna.....	76

Tabel 4.24 Hasil Pencarian Hyperparameter Model dengan SMOTE	78
Tabel 4.25 Hasil Pelatihan Model dengan SMOTE dan Optuna	79
Tabel 4.26 TP, TN, FP, FN untuk Model Klasifikasi Tanpa Optuna	82
Tabel 4.27 Hasil Evaluasi Model Klasifikasi Tanpa Optuna.....	83
Tabel 4.28 Hasil Evaluasi Model Klasifikasi Dengan Optuna	84
Tabel 4.29 Hasil Evaluasi Model Klasifikasi Dengan RUS dan Tanpa Optuna.....	86
Tabel 4.30 Hasil Evaluasi Model Klasifikasi Dengan RUS dan Optuna.....	88
Tabel 4.31 Hasil Evaluasi Model Klasifikasi Dengan SMOTE dan Tanpa Optuna ...	90
Tabel 4.32 Hasil Evaluasi Model Klasifikasi Dengan SMOTE Optuna	92
Tabel 4.33 Hasil Keseluruhan Skema Model Klasifikasi	93

DAFTAR GAMBAR

Gambar 1.1 Chart Angka Kasus Kanker Prostat di Indonesia Tahun 2022 (Ferlay et al. 2022)	1
Gambar 2.1 Ilustrasi Anatomi Kelenjar Prostat (dr. Pittara 2022).....	6
Gambar 2.2 Ilustrasi Prostat Normal dan Prostat dengan Kanker (Corbacserdar)	7
Gambar 2.3 Citra Magnetic Resonance Imaging (MRI).....	7
Gambar 2.4 Contoh Citra Grayscale	8
Gambar 2.5 Ilustrasi Arah Sudut GLCM (Yadav et al., 2022)	10
Gambar 2.6 Ilustrasi Alur Pohon Keputusan XGBoost (Sagemaker, Amazon Web Services...).....	15
Gambar 3.1 Tahapan Penelitian	22
Gambar 3.2 Rancangan Pengembangan Antarmuka Pengguna	26
Gambar 3.3 Alur Implementasi Model	27
Gambar 3.4 Rancangan Solusi Pengumpulan Data.....	27
Gambar 3.5 Rancangan Solusi Praproses Data Citra	28
Gambar 3.6 Rancangan Solusi Ekstraksi Fitur	28
Gambar 3.7 Rancangan Solusi Praproses Data	29
Gambar 3.8 Rancangan Solusi Pembuatan Model Klasifikasi.....	30
Gambar 4.1 Citra Hasil Image Resizing	49
Gambar 4.2 Contoh Nilai Piksel RGB Citra MRI.....	50
Gambar 4.3 Citra Hasil Grayscale	51
Gambar 4.4 Hasil Ukuran Citra setelah Grayscale	51
Gambar 4.5 Citra Hasil Praproses	52
Gambar 4.6 Source Code untuk Pembagian Data	66
Gambar 4.7 Grafik Hasil Learning Curve Pelatihan Model Tanpa Optuna.....	69
Gambar 4.8 Grafik Hasil Learning Curve Pelatihan Model Dengan Optuna	72
Gambar 4.9 Grafik Hasil Learning Curve Pelatihan Model Dengan RUS dan Tanpa Optuna	73
Gambar 4.10 Grafik Hasil Learning Curve Pelatihan Model Dengan RUS dan Optuna	76
Gambar 4.11 Grafik Hasil Learning Curve Pelatihan Model Dengan SMOTE dan Tanpa Optuna	77
Gambar 4.12 Grafik Hasil Learning Curve Pelatihan Model Dengan SMOTE dan Optuna	80
Gambar 4.13 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Tanpa Optuna	81

Gambar 4.14 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Dengan Optuna	84
Gambar 4.15 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Dengan <i>RUS</i> dan Tanpa Optuna	86
Gambar 4.16 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Dengan <i>RUS</i> dan Optuna.....	88
Gambar 4.17 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Dengan <i>SMOTE</i> dan Tanpa Optuna.....	90
Gambar 4.18 (a) Confusion Matrix pada Data Latih, (b) Confusion Matrix pada Data Validasi, dan (c) Confusion Matrix pada Data Uji untuk Model Klasifikasi Dengan <i>SMOTE</i> dan Optuna	92
Gambar 4.19 Source Code untuk Memuat Model dalam GUI.....	95
Gambar 4.20 Source Code untuk Melakukan Klasifikasi dalam GUI	95
Gambar 4.21 Tampilan awal GUI.....	96
Gambar 4.22 Tampilan Pop-up Jendela Folder.....	97
Gambar 4.23 Tampilan Hasil Praproses dan Segmentasi Citra MRI Prostat.....	98
Gambar 4.24 Tampilan Hasil Ekstraksi Fitur GLCM dari Citra MRI Prostat	99
Gambar 4.25 Tampilan Hasil Klasifikasi Citra MRI Prostat	100
Gambar 4.26 Tampilan ketika GUI dilakukan reset	101

DAFTAR RUMUS

(2.1) 9

(2.2) 11

(2.3) 11

(2.4) 11

(2.5) 11

(2.6) 13

(2.7) 14

(2.8) 14

(2.9) 16

(2.10) 16

(2.11) 16

(2.12) 18

(2.13) 19

(2.14) 19

(2.15) 19

DAFTAR LAMPIRAN

Lampiran 1. Bukti Proses Pengumpulan Data	109
Lampiran 2. Surat Keterangan Validasi	115
Lampiran 3. Bukti Validasi Citra MRI dengan Pakar	116