



**KOMPARASI ALGORITMA *BOOSTING* PADA KLASIFIKASI
PENYAKIT TUMOR GLIOMA DENGAN PENYEIMBANGAN
DATA MENGGUNAKAN ADASYN**

SKRIPSI

**Untuk memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

**Oleh
ARIFAH ANNISA RAHMI
NIM 2211016220034**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MEI 2026**

SKRIPSI

KOMPARASI ALGORITMA *BOOSTING* PADA KLASIFIKASI PENYAKIT TUMOR GLIOMA DENGAN PENYEIMBANGAN DATA MENGUNAKAN ADASYN

Oleh:

ARIFAH ANNISA RAHMI

2211016220034

Telah dipertahankan di depan Dosen Penguji pada tanggal 17 Juni 2026

Susunan Dosen Penguji:

Pembimbing Utama



Triando Hamonangan Saragih, S.Kom., M.Kom.
NIP. 199308242019031012

Penguji

1.



Muliadi, S.Kom., M.Sc.
NIP. 197804222010121002

Pembimbing Pendamping



Muhammad Itqan Mazdadi, S.Kom., M.Kom.
NIP. 199006122019031013

2.



Rudy Hertono, S.Kom., M.Kom.
NIP. 198809252022031003

Banjarbaru, 25 Juni 2026

Koordinator Program Studi Ilmu Komputer



Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 01 Mei 2026



Arifah Annisa Rahmi

NIM. 2211016220034

ABSTRAK

KOMPARASI ALGORITMA *BOOSTING* PADA KLASIFIKASI PENYAKIT TUMOR GLIOMA DENGAN PENYEIMBANGAN DATA MENGGUNAKAN ADASYN

(Oleh: Arifah Annisa Rahmi; Pembimbing: Triando Hamonangan Saragih, S.Kom., M.Kom. dan Muhammad Itqan Mazdadi, S.Kom, M.Kom.; 2026; 73 halaman)

Glioma merupakan salah satu jenis tumor otak yang memiliki tingkat keganasan berbeda sehingga diperlukan proses klasifikasi yang tepat untuk membantu diagnosis. Namun, pada data klasifikasi glioma sering ditemukan kondisi data yang tidak seimbang, di mana jumlah data pada salah satu kelas lebih dominan dibandingkan kelas lainnya. Penelitian ini bertujuan untuk membandingkan performa algoritma *boosting* serta menganalisis pengaruh penerapan metode ADASYN pada klasifikasi glioma. Dataset yang digunakan yaitu *Glioma Grading Clinical and Mutation Features Dataset* dari *UCI Machine Learning Repository* dengan jumlah data sebanyak 839 data. Penelitian menggunakan algoritma Gradient Boosting, XGBoost, dan CatBoost dengan metode Stratified 10-Fold Cross Validation. Hasil penelitian menunjukkan bahwa model terbaik diperoleh oleh Gradient Boosting dengan ADASYN dengan nilai *accuracy* sebesar 0,8725, *recall* sebesar 0,9319, *F1-score* sebesar 0,8606, dan *AUC* sebesar 0,9156. Penerapan ADASYN juga meningkatkan performa model terbaik, ditunjukkan dengan kenaikan nilai *recall* sebesar 3,45% dan *F1-score* sebesar 1,70% dibandingkan tanpa ADASYN. Berdasarkan hasil penelitian, algoritma *boosting* dengan penerapan ADASYN dapat digunakan dengan cukup baik untuk melakukan klasifikasi glioma pada data yang tidak seimbang.

Kata Kunci: klasifikasi glioma, ADASYN, data tidak seimbang, *boosting*

ABSTRACT

COMPARISON OF BOOSTING ALGORITHMS FOR GLIOMA CLASSIFICATION USING ADASYN-BASED DATA BALANCING

(By: Arifah Annisa Rahmi; Supervisor: Triando Hamonangan Saragih, S.Kom., M.Kom. and Muhammad Itqan Mazdadi, S.Kom, M.Kom.; 2026; 73 pages)

Glioma is a type of brain tumor with varying degrees of malignancy, necessitating an accurate classification process to aid in diagnosis. However, glioma classification datasets often suffer from data imbalance, where the number of data points in one class significantly outnumbers those in other classes. This study aims to compare the performance of boosting algorithms and analyze the impact of applying the ADASYN method on glioma classification. The dataset used is the Glioma Grading Clinical and Mutation Features Dataset from the UCI Machine Learning Repository, containing 839 data points. The study employed the Gradient Boosting, XGBoost, and CatBoost algorithms using the Stratified 10-Fold Cross-Validation method. The results showed that the best model was achieved by Gradient Boosting with ADASYN, yielding an accuracy of 0.8725, a recall of 0.9319, an F1-score of 0.8606, and an AUC of 0.9156. The application of ADASYN also improved the performance of the best model, as evidenced by an increase in recall of 3.45% and an increase in the F1-score of 1.70% compared to without ADASYN. Based on the results of this study, boosting algorithms with the application of ADASYN can be used quite effectively for glioma classification on imbalanced data.

Keywords: glioma classification, ADASYN, imbalanced data, boosting

PRAKATA

Puji Syukur penulis panjatkan ke Tuhan Yang Maha Esa karena atas berkat Rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul **“KOMPARASI ALGORITMA *BOOSTING* PADA KLASIFIKASI PENYAKIT TUMOR GLIOMA DENGAN PENYEIMBANGAN DATA MENGGUNAKAN ADASYN”** untuk memenuhi syarat dalam menyelesaikan Pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, Adapun yang dimaksud Adalah sebagai berikut:

1. Orang tua, kakak, serta seluruh keluarga besar yang selalu memberikan dukungan, doa, motivasi, dan kasih sayang dalam proses penyelesaian skripsi ini.
2. Bapak Triando Hamonangan Saragih S.Kom., M.Kom selaku dosen pembimbing utama yang telah memberikan bimbingan, membantu dan meluangkan waktu dalam proses penyelesaian skripsi ini.
3. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom., selaku dosen pembimbing pendamping yang telah memberikan bimbingan, membantu dan meluangkan waktu dalam proses penyelesaian skripsi ini.
4. Bapak Muliadi, S.Kom., M.Cs. dan Bapak Rudy Herteno, S.Kom., M.Kom. selaku dosen penguji yang telah memberikan masukan-masukan untuk penyajian skripsi.
5. Bapak Rudy Herteno, S.Kom., M.Kom. selaku Dosen Pembimbing Akademik dan Ibu Dwi Kartini, S.Kom., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
6. Seluruh dosen dan staf Program Studi Ilmu Komputer FMIPA ULM yang telah memberikan ilmu yang sangat bermanfaat selama penulis menempuh pendidikan.
7. Seluruh sahabat dan teman dan rekan mahasiswa yang telah memberikan

banyak dukungan, bantuan, dan doa selama masa studi hingga penyelesaian skripsi ini.

8. Semua pihak yang tidak bisa disebutkan satu persatu, yang turut berperan dan membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penyusunan skripsi ini masih terdapat berbagai kekurangan dan belum mencapai kesempurnaan yang diharapkan. Oleh karena itu, penulis mengharapkan bantuan berupa kritik dan saran yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini.

Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 01 Mei 2026

A handwritten signature in black ink, appearing to read 'arifah', followed by a horizontal line and a double hash mark '=='. The signature is stylized and cursive.

Arifah Annisa Rahmi

DAFTAR ISI

HALAMAN COVER.....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK.....	iv
<i>ABSTRACT</i>	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
BAB I PENDAHULUAN.....	12
1.1 Latar Belakang.....	12
1.2 Rumusan Masalah.....	14
1.3 Tujuan Penelitian.....	15
1.4 Manfaat Penelitian.....	15
1.5 Batasan Masalah.....	15
BAB II TINJAUAN PUSTAKA.....	17
2.1 Kajian Terdahulu.....	17
2.2 Landasan Teori.....	27
2.2.1 Glioma.....	27
2.2.2 <i>Machine Learning</i> untuk Klasifikasi Medis.....	27
2.2.3 Data Tidak Seimbang (Imbalanced Data).....	28
2.2.4 Teknik Penanganan Data Tidak Seimbang (<i>Oversampling</i> /ADASYN).....	28
2.2.5 Algoritma <i>Boosting</i>	30
2.2.6 Gradient Boosting.....	30
2.2.7 Extreme Gradient Boosting (XGBoost).....	31
2.2.8 Categorical Boosting (CatBoost).....	33
2.2.9 Metode Evaluasi Kerja.....	34
BAB III METODE PENELITIAN.....	37
3.1 Alat Penelitian.....	37

3.2	Bahan Penelitian.....	37
3.3	Prosedur Penelitian.....	38
3.3.1	Pengumpulan Dataset	39
3.3.2	Prapemrosesan Data.....	39
3.3.3	Pembagian dan Validasi Data.....	39
3.3.4	Penanganan Ketidakseimbangan Data (ADASYN)	40
3.3.5	Pelatihan Model <i>Boosting</i> dan Optimasi Parameter	41
3.3.6	Evaluasi dan Analisis Akhir.....	42
BAB IV HASIL DAN PEMBAHASAN.....		44
4.1	Hasil.....	44
4.1.1	Deskripsi Data.....	44
4.1.2	Prapemrosesan Data.....	47
4.1.3	Penerapan ADASYN	48
4.1.4	Mekanisme Penerapan ADASYN.....	49
4.1.5	Hasil Pengujian Gradient Boosting.....	51
4.1.6	Hasil Pengujian XGBoost.....	53
4.1.7	Hasil Pengujian CatBoost	55
4.1.8	Perbandingan Performa Model	58
4.1.9	Hasil Model Terbaik	61
4.2	Pembahasan	65
BAB V PENUTUP		69
5.1	Kesimpulan.....	69
5.2	Saran	69
DAFTAR PUSTAKA		71

DAFTAR TABEL

Tabel 1 Keaslian Penelitian.....	21
Tabel 2 Perancangan Penelitian	26
Tabel 3 Daftar Fitur Dataset.....	44
Tabel 4 Distribusi Dataset Glioma	46
Tabel 5 Hasil Prapemrosesan Data.....	47
Tabel 6 Distribusi Data Sebelum dan Sesudah Penerapan ADASYN	48
Tabel 7 Hasil Pengujian Gradient Boosting.....	52
Tabel 8 Hasil Pengujian XGBoost	55
Tabel 9 Hasil Pengujian CatBoost.....	57
Tabel 10 Perbandingan Performa Seluruh Model	58
Tabel 11 Hasil Model Terbaik	61
Tabel 12 Parameter Terbaik Model Gradient Boosting.....	64
Tabel 13 Peringkat Performa Model	65
Tabel 14 Confusion Matrix Model Terbaik.....	66
Tabel 15 Top 5 Feature Importance.....	68

DAFTAR GAMBAR

Gambar 1 Diagram Alir ADASYN	29
Gambar 2 Diagram Alir Gradient Boosting	31
Gambar 3 Diagram Alir XGBoost	32
Gambar 4 Diagram Alir CatBoost.....	34
Gambar 5 Alur penelitian	38
Gambar 6 Distribusi Kelas Sebelum ADASYN	46
Gambar 7 Distribusi Kelas Setelah ADASYN.....	49
Gambar 8 Contoh Pohon Keputusan pada Model Gradient Boosting	52
Gambar 9 Contoh Pohon Keputusan pada Model XGBoost.....	54
Gambar 10 Contoh Pohon Keputusan pada Model CatBoost.....	57
Gambar 11 Perbandingan F1-Score Model	60
Gambar 12 Perbandingan Recall Model	60
Gambar 13 Confusion Matrix Model Terbaik.....	62
Gambar 14 ROC Curve Model Terbaik	63
Gambar 15 Top 5 Feature Importance.....	64
Gambar 16 Ringkasan Metrik Model Terbaik.....	67