

## PERNYATAAN

Saya menyatakan bahwa skripsi ini masih mengandung plagiat dibawah batas yang diterapkan , dan apabila di kemudian hari terbukti terdapat plagiat , dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang – undangan .



Bojonegoro, 13 Agustus 2023



Elly Khusniawati

# UNUGIRI


## HALAMAN PERSETUJUAN

Nama : Elly Khusniawati  
NIM : 2520190037  
Judul : Implementasi Algoritma XGBoost untuk Klasifikasi Kasus  
Stroke di RSUD Dr. R. Sosodoro Djatikoesoemo.Bojonegoro.

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam ujian skripsi .

Bojonegoro, 5 Agustus 2023

Pembimbing I

  
Nita Cahyani, M. Stat.

NIDN. 0704038906

Pembimbing II

  
Denny Nurdiansyah, M. Si.

NIDN. 07605870

# HALAMAN PENGESAHAN

Nama : Elly Khusniawati

Nim : 2520190037

Judul Skripsi : Implementasi Algoritama XGBoost untuk Klasifikasi Kasus Stroke di RSUD Dr. R. Sosodoro Djatikoesoemo Bojonegoro.

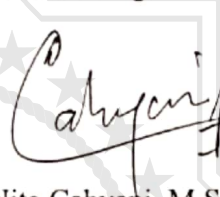
Telah diujikan dalam ujian skripsi pada tanggal 28 Agustus 2023

Dewan Penguji

Tim Pembimbing

Penguji I

Pembimbing I



Alif Yuanita Kartini, M.Si.

Nita Cahyani, M.Stat

NIDN. 0721048606

NIDN. 0704038906

Penguji II

Pembimbing II



Dr. H. M. Ridlwan Hambali, Lc., M.A.

Denny Mardiansyah, M.Si.

NIDN. 2117056803

NIDN. 0726058702

Mengetahui,

Mengetahui,

Dekan Fakultas Sains & Teknologi

Ketua Program Studi



Siti Wahyudhi, M.Pd

Nita Cahyani, M.Stat

NIDN. 0709058902

NIDN. 0704038906

## MOTTO

إِنَّ مَعَ الْعُسْرِ يُسْرًا فَإِذَا فَرَغْتَ فَانصَبْ وَإِلَىٰ رَبِّكَ فَارْغَبْ

“Maka sesungguhnya Bersama kesulitan ada kemudahan. Maka apabila engkau telah selesai(dari suatu urusan) tetaplah bekerja keras (untuk urusan yang lain). Dan hanya kepada TUHAN mu lah engkau berharap”

(QS. Al-Insyirah,6-8)

“Success and happiness lies in yourself. Stay happy because of your happiness and you will form a strong character to fight difficulties.”

(Helen Keller)

## PERSEMBAHAN

Tiada lembar yang paling inti dalam laporan skripsi ini kecuali persembahan.laporan skripsi saya ini saya persembahkan sebagai tanda bukti kepada orang tua, adik,sahabat,serta orang orang terdekat yang selalu memberi support untuk menyelesaikain skripsi ini . Ku persembahkan skripsi untuk orang tercinta dan tersayang atas kasihnya yang berlimpah.

**Teristimewa Ayahanda (Sarjono) dan Ibunda (Musyarofah) tercinta,  
tersayang, terkasih, dan yang terhormat**

Kupersembahkan skripsi ini kepada kalian atas kasih sayang dan bimbingan selama ini sehingga saya dapat menyelesaikan skripsi ini dengan baik. Banyak sekali hal yang ingin saya ungkapkan, tetapi tidak dapat dituliskan satu persatu. Somaga hasil dari perjuangan saya selama ini dapat berbuah hasil yang manis. Semangat yang terus berkobar dalam diri saya agar sanggup menghadapi dunia luar yang sebenarnya. Selama 14 tahun telah menempuh pendidikan telah tiba saatnya saya akan membuktikan kepada kalian dan dunia saya telah siap untuk membuka lembaran baru sebagai seseorang yang bertanggung jawab dan berkarya bagi kepentingan banyak orang.

## KATA PENGANTAR

Puji dan sukur penulis ucapkan kepada ALLAH SWT yang telah memberikan berupa kesehatan, kesempatan kepada penulis sehingga mampu menyelesaikan proposal skripsi ini yang berjudul “Implementasi Algoritama XGBoost untuk klasifikasi kasus stroke di RSUD Dr.R.Sosodoro Djatikoesoemo Bojonegoro, pada waktu yang tepat. Shalawat teriring salam tercurahkan kepada baginda Nabi Muhammad Saw, yang telah membawa umatnya dari zaman kebodohan ke zaman yang penuh dengan ilmu dan teknologi ini, serta atas segala keteladanan dan pengorbanan beliau dalam mendidik para umatnya agar menjadi manusia yang berakhlak mulia..

Dalam proses pembuatan laporan ini tak lupa saya menghaturkan sujud kepada orangtua saya yang telah banyak memberikan dorongan semangat dari awal hingga selesainya laporan ini. Tak lupa juga saya mengucapkan terima kasih pada teman teman dikampus yang telah memberikan dorongan moril dan material serta informasi. Juga dengan segala hormat saya ucapkan banyak terima kasih pada bapak bapak ibu dosen di UNUGIRI sehingga kami dapat menerapkan ilmu yang diberikan pada kami. Ucapan terima kasih ini juga saya ucapkan kepada:

1. Bapak KH.M.Jauharul Ma'arif, M. Pd.I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Sunu Wahyudhi, M.Pd sebagai Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Ibu Nita Cahyani, M.Stat sebagai Ketua Program Studi Statistika.
4. Pembimbing I Ibu Nita Cahyani, M.Stat. yang senantiasa membimbing, mengarahkan, dan memberikan motivasi kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini tepat pada waktunya.
5. Pembimbing II Bapak Denny Nurdiansyah, M.Si yang senantiasa membimbing, mengarahkan, dan memberikan motivasi kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini tepat

pada waktunya.

6. Dosen pembimbing akademik Ibu Nita Cahyani, M. Stat., yang senantiasa membimbing, mengarahkan, dan memberikan motivasi kepada penulis.
7. Kepala RSUD Dr. R. Sosodoro Djatikoesoemo yang telah mengizinkan peneliti melakukan penelitian serta memberikan informasi yang peneliti butuhkan.
8. Bapak Sarjono serta Ibu Musyarofah selaku orang tua penulis yang selalu memberikan dukungan dan memberi semangat yang tidak hentinya mendoakan.
9. Teman-teman seperjuangan angkatan 2019 prodi statistika yang tidak dapat disebutkan, karena kalian menjadi tempat bertukar pikiran dalam penulisan proposal skripsi ini, dan juga pengalaman dengan kalian tidak akan terlupakan.

Serta semua pihak yang namanya tidak dapat penulis sebutkan satu persatu, yang turut membantu dan memotivasi penulis dalam penyusunan proposal skripsi ini. Penulis mengucapkan terima kasih yang sebanyak-banyaknya, semoga Allah membalas dengan rahmat dan karuniaNya yang tak terhingga.

Bojonegoro, 12 Januari 2023

UNUGIRI  
Elly Khusniawati

## **ABSTRAC**

*Stroke is a brain disease or disorder characterized by nerve paralysis caused by blockage of blood flow to the brain. Simply put, acute stroke is a brain disorder caused by the cessation of blood flow to the brain due to blockage (ischemic stroke) or bleeding (hemorrhagic stroke). Stroke risk factors are factors that cannot be changed, such as age, gender, race, and family history. Stroke is a disease that is estimated to kill 5 million people worldwide by 2030. The incidence of stroke in Indonesia in 2018 is cases per 1,000 population, with a predominance of men aged 75 years and over. Stroke is caused by impaired blood flow to the brain which is characterized by loss of movement and communication abilities as well as cognitive impairment (Sulaeman, 2022). The XGBoost method is an ensemble algorithm-based gradient tree enhancement development algorithm that can deal with large machine learning cases effectively. The XGBoost method was chosen because it has several additional features that are useful for speeding up computing systems and preventing overconfiguration. Based on the results of the research that has been done, it can be concluded that the stroke data is processed by applying data mining techniques using the Extreme Gradient Boosting (XGBOOST) algorithm, obtaining a precision of 100%, a recall of 100%, an AUC of 1,000 and an accuracy of 100%. Based on the classification results, it can be said that the classification is excellent (very good accuracy).*

**Keywords:** *Stroke, Extreme Gradient Boosting.*

## ABSTRAK

Stroke adalah penyakit atau gangguan otak yang ditandai dengan kelumpuhan saraf yang disebabkan oleh tersumbatnya aliran darah ke otak. Sederhananya, stroke akut adalah gangguan otak yang disebabkan oleh terhentinya aliran darah ke otak akibat penyumbatan (ischemic stroke) atau pendarahan (hemorrhagic stroke). Faktor risiko stroke adalah faktor yang tidak dapat diubah, seperti usia, jenis kelamin, ras, dan riwayat keluarga. Stroke merupakan penyakit yang diperkirakan membunuh 5 juta orang di seluruh dunia pada tahun 2030. Angka kejadian stroke di Indonesia pada tahun 2018 adalah kasus per 1.000 penduduk, dengan dominasi pria berusia 75 tahun ke atas. Stroke disebabkan oleh gangguan aliran darah ke otak yang ditandai dengan hilangnya kemampuan gerak dan komunikasi serta gangguan kognitif (Sulaeman, 2022). Metode XGBoost adalah algoritme pengembangan peningkatan pohon gradien berbasis algoritme ansambel yang dapat menangani kasus pembelajaran mesin besar secara efektif. Metode XGBoost dipilih karena memiliki beberapa fitur tambahan yang berguna untuk mempercepat sistem komputasi dan mencegah overconfiguration. Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa data stroke diolah dengan menerapkan teknik data mining menggunakan algoritma Extreme Gradient Boosting (XGBOOST), memperoleh hasil precision sebesar 100%, recall sebesar 100 %, AUC sebesar 1.000 dan akurasi sebesar 100%. Berdasarkan hasil klasifikasi dapat dikatakan excellent classification (akurasi sangat baik).

# UNUGIRI

**Kata Kunci:** *Stroke, Extreme Gradient Boosting.*



## DAFTAR ISI

	Halaman
<b>COVER .....</b>	<b>i</b>
<b>HALAMAN SAMPUL DALAM.....</b>	<b>ii</b>
<b>HALAMAN PERNYATAAN.....</b>	<b>iii</b>
<b>HALAMAN PERSETUJUAN .....</b>	<b>iv</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>v</b>
<b>HALAMAN MOTTO DAN PERSEMBAHAN .....</b>	<b>vi</b>
<b>KATA PENGANTAR.....</b>	<b>vii</b>
<b>ABSTRAC.....</b>	<b>ix</b>
<b>ABSTRAK .....</b>	<b>x</b>
<b>DAFTAR ISI .....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xiv</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....</b>	<b>6</b>
2.1 Tinjauan Pustaka .....	6
2.2 Dasar Teori .....	11
2.2.1 Statistika Deskriptif .....	11
2.2.2 Data Mining .....	13
2.2.3 Klasifikasi .....	16
2.2.4 Pre- Processing Data.....	16
2.2.5 Ensemble Learning .....	17
2.2.6 Boosting .....	18
2.2.7 Algoritma XGBOOST .....	19
2.2.8 Confusion matric.....	23
2.2.9 Stroke .....	25
2.2.10 Faktor faktor Resiko Penyakit Stroke.....	27

<b>BAB III METODE PENELITIAN .....</b>	<b>29</b>
3.1 Sumber Data .....	29
3.2 Variabel Penelitian .....	29
3.3 Struktur Data .....	31
3.4 Langkah- Langkah Analisis.....	31
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>34</b>
4.1 Analisis Deskriptif.....	34
4.1.1 Jenis Stroke (Y) .....	34
4.2 Tabel kontingensi .....	42
4.2.1 Tabel kontingensi Jenis stroke berdasarkan jenis kelamin .....	42
4.2.2 Tabel kontingensi Jenis stroke berdasarkan status hipertensi.....	43
4.2.3 Tabel kontingensi Jenis stroke berdasarkan status diabetes militus .....	43
4.2.4 Tabel kontingensi Jenis stroke berdasarkan status hiperkolestrol .....	43
4.2.5 Tabel kontingensi Jenis stroke berdasarkan indeks masa tubuh.....	44
4.2.6 Tabel kontingensi Jenis stroke berdasarkan status merokok .....	44
4.3 Normalisasi Data .....	45
4.4 Pembagian Data Training dan Data Testing.....	46
4.5 Confusion Matrik .....	46
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>49</b>
5.1 Kesimpulan.....	49
<b>5.2 Saran .....</b>	<b>49</b>
<b>DAFTAR PUSTAKA .....</b>	<b>50</b>
<b>LAMPIRAN.....</b>	<b>52</b>



# UNUGIRI

## DAFTAR TABEL

Tabel	Halaman
Tabel 2. 1 Penelitian Terdahulu .....	6
Tabel 2. 2 Confusion Matric .....	23
Tabel 3. 1 Variabel Penelitian.....	29
Tabel 3. 2 Struktur Data.....	31
Tabel 4. 1 Statistika Deskriptif Usia.....	36
Tabel 4. 2 kontingensi Jenis stroke berdasarkan jenis kelamin.....	42
Tabel 4. 3 kontingensi Jenis stroke berdasarkan status hipertensi .....	43
Tabel 4. 4 kontingensi Jenis stroke berdasarkan status diabetes militus.....	43
Tabel 4. 5 kontingensi Jenis stroke berdasarkan status hiperkolestrol.....	44
Tabel 4. 6 kontingensi Jenis stroke berdasarkan indeks masa tubuh .....	44
Tabel 4. 7 kontingensi Jenis stroke berdasarkan status merokok.....	45
Tabel 4. 8 Normalisasi Data pada Data Kasus Stroke .....	45
Tabel 4. 9 Pembagian Data Training dan Data Testing pada Data Kasus Stroke .....	46
Tabel 4. 10 Confusion Matric Data Testing .....	46

UNUGIRI

## DAFTAR GAMBAR

Gambar	Halaman
Gambar 2. 1 Tahapan Knowledge discovery .....	14
Gambar 2. 2 Metode Boosting .....	18
Gambar 2. 3 Stroke Iskemik .....	26
Gambar 2. 4 Stroke Hemoragik .....	26
Gambar 3. 1 Diagram alir Metode XGBOOST.....	33
Gambar 4. 1 Diagram Lingkaran Jenis Stroke.....	34
Gambar 4. 2 Diagram Lingkaran Jenis Kelamin.....	35
Gambar 4. 3 Diagram Lingkaran Status hipertensi .....	37
Gambar 4. 4 Diagram lingkaran Status Diabetes Millitus .....	38
Gambar 4. 5 Diagram Lingkaran Status Hiperkolestrol.....	39
Gambar 4. 6 Diagram Lingkaran Status Riwayat Stroke Keluarga .....	40
Gambar 4. 7 Diagram Lingkaran IMT .....	41
Gambar 4. 8 Kurva ROC.....	48



UNUGIRI

## DAFTAR LAMPIRAN

Lampiran 1 Output Crostabulation Statistika deskriptif .....	52
Lampiran 2 Output Diagram lingkaran di SPSS .....	53
Lampiran 3 syntax Metode xgboost di software pyhton .....	57



# UNUGIRI