

**PREDIKSI RISIKO HIPERTENSI MENGGUNAKAN
PENGGABUNGAN CATBOOST, XGBOOST, DAN
LIGHTGBM: PENDEKATAN PEMBELAJARAN MESIN**

**LAPORAN NON-REGULER
SKEMA SCIENTIST**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

ABISAKHA SAIF ALFATH

22.11.5285

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**PREDIKSI RISIKO HIPERTENSI MENGGUNAKAN
PENGGABUNGAN CATBOOST, XGBOOST, DAN
LIGHTGBM: PENDEKATAN PEMBELAJARAN MESIN**

**LAPORAN NON-REGULER
SKEMA SCIENTIST**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

ABISAKHA SAIF ALFATH

22.11.5285

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALOR NON-REGULER

**PREDIKSI RISIKO HIPERTENSI MENGGUNAKAN
PENGGABUNGAN CATBOOST, XGBOOST, DAN
LIGHTGBM: PENDEKATAN PEMBELAJARAN MESIN**

yang disusun dan diajukan oleh

Abisakha Salf Alfath
22.11.5285

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Desember 2025

Dosen Pembimbing,



Rumli, S.Kom., M.Kom.
NIK. 190302246

HALAMAN PENGESAHAN

JALUR NON-REGULER

**PREDIKSI RISIKO HIPERTENSI MENGGUNAKAN
PENGGABUNGAN CATBOOST, XGBOOST, DAN
LIGHTGBM: PENDEKATAN PEMBELAJARAN MESIN**

yang disusun dan diajukan oleh

Abisakha Saif Alfath

22.11.5285

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Desember 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Subekti Ningsih, S.Kom., M.Kom.
NIK. 190302413



Bayu Setiaji, M.Kom.
NIK. 190302216



Rumini, S.Kom., M.Kom.
NIK. 190302246



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Desember 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Abisakha Saif Alfath

NIM : 22.11.5285

Menyatakan bahwa Laporan dengan judul berikut:

Hypertension Risk Prediction Using Stacking Ensemble of CatBoost, XGBoost, and LightGBM: A Machine Learning Approach

Dosen Pembimbing : Rumini, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 Desember 2025

Yang Menyatakan,

Abisakha Saif Alfath

HALAMAN PERSEMBAHAN

Dengan memanjatkan puji syukur ke hadirat Allah SWT atas segala limpahan rahmat, petunjuk, dan kemudahan-Nya, penulis akhirnya dapat menyelesaikan Tugas Akhir ini sebagai bagian dari pemenuhan proses akademik pada jenjang sarjana. Karya ini terwujud melalui perjalanan panjang yang disertai usaha, doa, ketekunan, serta dukungan moral dari berbagai pihak. Sebagai bentuk penghargaan dan rasa terima kasih yang tulus, dengan segala kerendahan hati penulis mempersembahkan karya ini kepada:

1. Kepada Bapak tercinta Moch. Ali dan Ibunda Santi Sugiharti, yang dengan ketulusan hati, kasih sayang tanpa batas, serta doa yang tak pernah putus telah menjadi pilar utama dalam setiap langkah perjalanan hidup penulis, sekaligus membentuk karakter, semangat, dan keteguhan dalam menyelesaikan tugas akhir ini, serta kepada seorang wanita tercinta yang senantiasa memberikan ketenangan, pengertian, dan dukungan emosional di tengah berbagai tekanan dan keraguan, dengan doa dan perhatian yang selalu menguatkan. Semoga karya ini menjadi persembahan kecil yang mampu membalas cinta, pengorbanan, dan keikhlasan yang begitu besar dari orang-orang terbaik dalam hidup penulis.
2. Ibu Rumini, S.Kom., M.Kom., selaku dosen pembimbing, yang dengan penuh ketelitian dan ketulusan telah memberikan arahan, bimbingan, serta dukungan selama proses penyusunan laporan ini. Setiap masukan dan penjelasan beliau sangat berarti dan membantu penulis dalam menyempurnakan karya ini.
3. Bapak Ajie Kusuma Wardhana, S.Kom., M.Eng., selaku dosen yang telah membantu dalam proses riset, serta memberikan arahan, wawasan, dan motivasi yang sangat berarti selama penyusunan laporan ini. Bimbingan beliau menjadi bagian penting dalam memperkaya pemahaman penulis dan menyempurnakan hasil penelitian ini.
4. Bapak Kamarudin, S.Kom., M.Kom., selaku dosen wali, yang telah

memberikan bimbingan, arahan, serta dukungan akademik selama penulis menempuh pendidikan. Perhatian dan bantuan beliau turut menjadi bagian penting dalam kelancaran proses studi hingga terselesaikannya laporan ini.

5. Kedua adik penulis, Sulthan Azka Assidqon dan Rachmalia Aisyah Putri, yang merupakan cahaya dan kebahagiaan dalam kehidupan penulis. Senyum, canda, dan kehadiran kalian selalu menjadi penguat di tengah rasa lelah dan tekanan selama proses penyusunan karya ini. Semoga keberhasilan kecil ini dapat menjadi contoh dan motivasi bagi kalian untuk terus bermimpi dan berjuang meraih masa depan yang terbaik.
6. Teman-teman tongkrongan Rambe, yang selalu memberikan semangat, kebersamaan, dan tawa yang menghidupkan setiap proses perjalanan ini. Dukungan dan suasana positif yang kalian hadirkan menjadi pelengkap penting dalam melewati masa studi hingga terselesaikannya laporan ini.
7. Teman-teman tongkrongan Rambe dan IF 12, yang telah memberikan semangat, kebersamaan, serta dukungan yang berarti sepanjang perjalanan akademik ini. Terima kasih atas tawa, diskusi, dan motivasi yang membantu penulis menyelesaikan laporan ini dengan baik.
8. Untuk diri penulis sendiri, Abisakha Saif Alfath, yang telah bertahan, berusaha, dan tidak menyerah meskipun melalui berbagai tekanan, kegagalan, dan proses panjang yang penuh tantangan. Terima kasih telah tetap melangkah, belajar, dan mempercayai bahwa setiap usaha akan membawa hasil. Semoga pencapaian ini menjadi pengingat bahwa kerja keras dan keteguhan hati selalu membuka jalan menuju kesempatan dan keberhasilan berikutnya.

KATA PENGANTAR

Dengan menyebut nama Allah SWT yang Maha Pengasih dan Maha Penyayang, penulis menyampaikan rasa syukur yang mendalam atas limpahan rahmat, kemudahan, dan petunjuk-Nya sehingga laporan kelulusan ini dapat diselesaikan dengan baik. Penyusunan laporan ini merupakan proses panjang yang tidak hanya membutuhkan kemampuan akademik, tetapi juga dukungan, bimbingan, dan bantuan dari berbagai pihak yang telah memberikan kontribusi berarti bagi penulis. Pada kesempatan ini, penulis menyampaikan ucapan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta,
2. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta,
3. Ibu Eli Pujiastuti, M.Kom., selaku Ketua Program Studi S1 Informatika Universitas AMIKOM Yogyakarta,
4. Ibu Rumini, S.Kom., M.Kom., selaku Dosen Pembimbing, atas kesediaan beliau membimbing dengan teliti dan konsisten. Nasihat, koreksi, serta arahan yang diberikan sangat membantu penulis dalam memahami substansi penelitian dan menyempurnakan laporan ini.
5. Bapak Ajie Kusuma Wardhana, S.Kom., M.Eng., selaku dosen yang turut mendampingi proses riset, atas penjelasan, ide, serta dukungan akademik yang memperkaya pemahaman penulis pengembangan penelitian.

Yogyakarta, 8 Desember 2025

Penulis

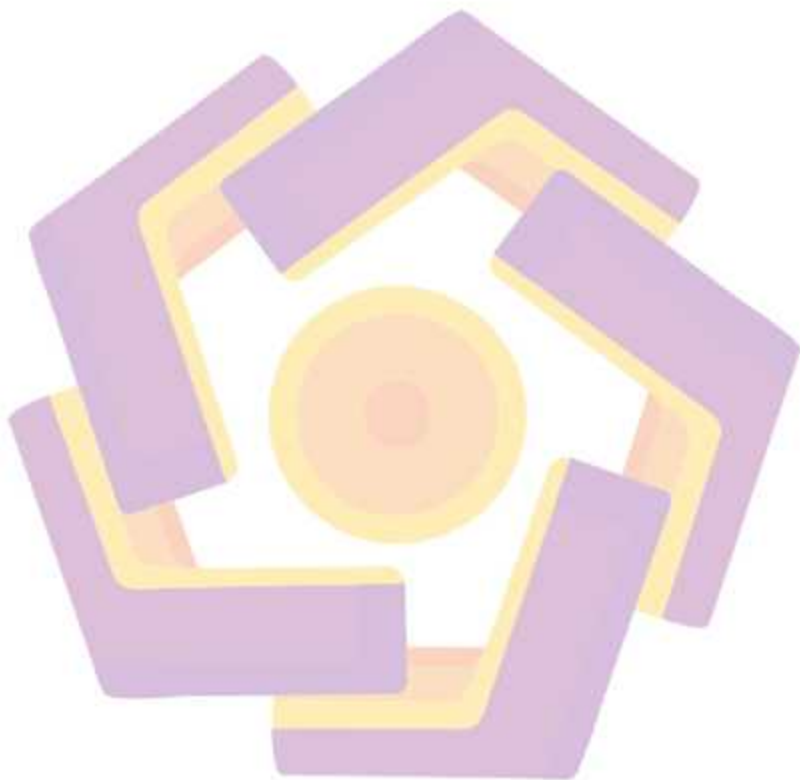
DAFTAR ISI

Halaman judul	i
Halaman persetujuan	ii
Halaman Pengesahan	iii
Halaman Pernyataan Keaslian Karya	iv
Halaman Persembahan	v
Kata Pengantar	vii
Daftar Isi	viii
Daftar Tabel	x
Daftar Gambar	xi
Daftar Lampiran	xii
Daftar Lambang dan Singkatan	xiii
Daftar Istilah	xiv
Intisari	xv
Abstract	xvi
Bab I Pendahuluan	1
1.1. Gambaran Umum	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	3
1.4. Tujuan	3
Bab II Tinjauan Pustaka	4
2.1. Studi Literatur	4
2.2. Landasan Teori	6
2.2.1. Teori Hipertensi	6
2.2.2. Machine Learning	6
2.2.3. Ensemble Learning	6
2.2.4. Stacking Ensemble	7
2.2.5. Algoritma Base Learner	7
2.2.6. Algoritma Meta Learner	8
Bab III Metode Penelitian	9
3.1. Metode	9

3.1.1. Dataset.....	9
3.1.2. Preprocessing Data.....	10
3.1.3. Balancing Data SMOTE.....	11
3.1.4. Train dan Test Split.....	11
3.1.5. Algoritma Model.....	12
3.1.6. Evaluasi.....	16
Bab IV Pembahasan.....	17
4.1. Sub Pembahasan.....	17
4.1.1. Preprocessing.....	17
4.1.2. Balancing Data SMOTE.....	17
4.1.3. Train dan Test Split.....	19
4.1.4. Evaluasi.....	20
4.2. Gambar Website.....	30
Bab V Kesimpulan.....	32
5.1. Kesimpulan.....	32
5.2. Saran.....	33
Referensi.....	34
Curriculum Vitae.....	37
Lampiran dan Bukti Pendukung.....	38
a. Letter of Acceptance (LOA).....	38
b. Lembar Review.....	39
c. Bukti terbit/Terindex.....	41
d. Bukti Pembayaran.....	42

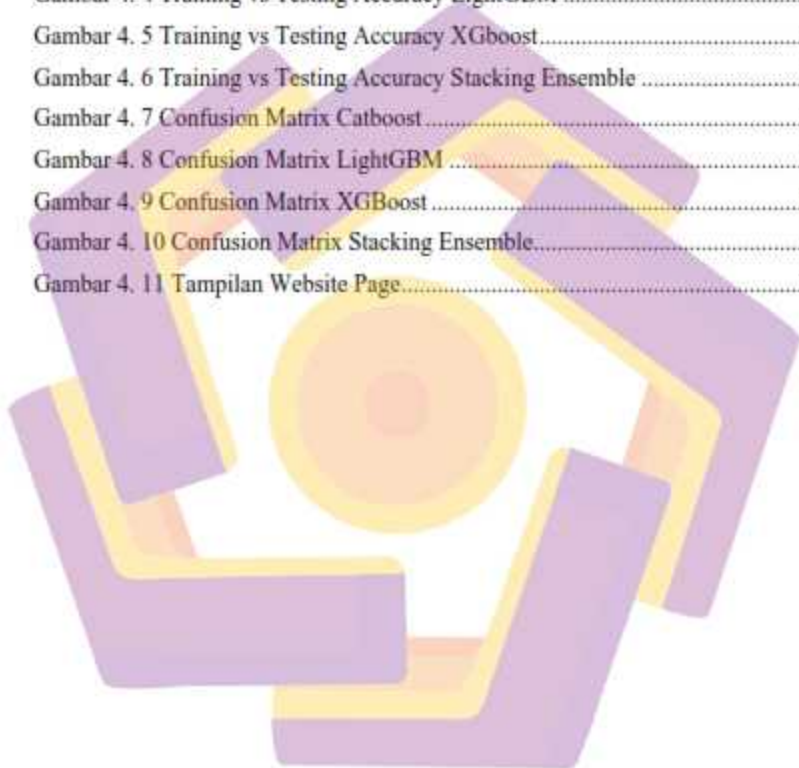
DAFTAR TABEL

Tabel 3. 1 Tipe Dataset	10
Tabel 4. 1 Pembagian Data Training dan Testing.....	20
Tabel 4. 2 Classification Report.....	20



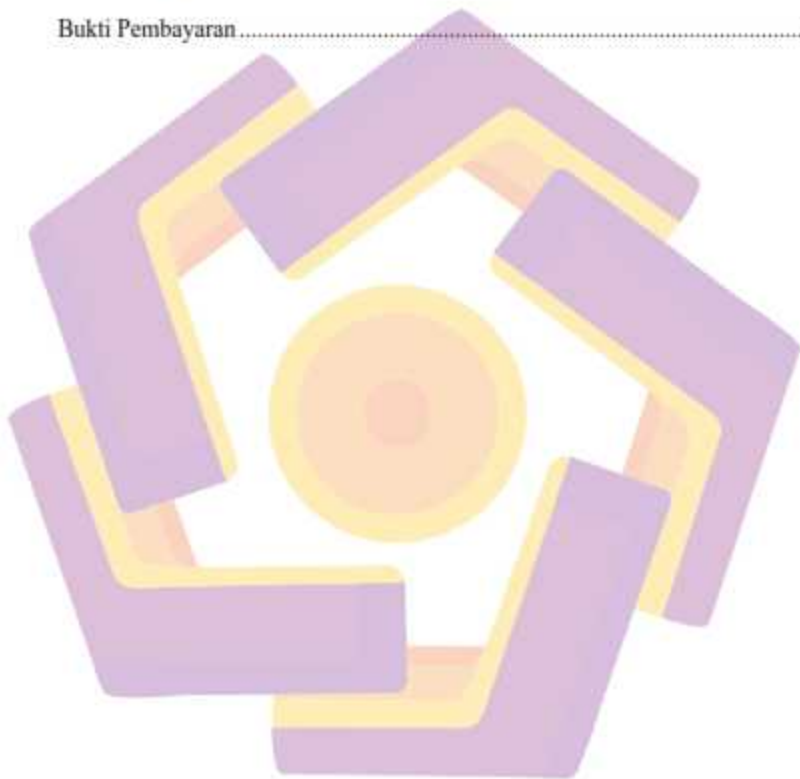
DAFTAR GAMBAR

Gambar 3. 1 Metode Penelitian.....	9
Gambar 4. 1 Perbandingan Data.....	18
Gambar 4. 2 Hasil Balancing data SMOTE.....	19
Gambar 4. 3 Training vs Testing Accuracy Catboost.....	22
Gambar 4. 4 Training vs Testing Accuracy LightGBM	23
Gambar 4. 5 Training vs Testing Accuracy XGboost.....	24
Gambar 4. 6 Training vs Testing Accuracy Stacking Ensemble	25
Gambar 4. 7 Confusion Matrix Catboost	26
Gambar 4. 8 Confusion Matrix LightGBM	27
Gambar 4. 9 Confusion Matrix XGBoost	28
Gambar 4. 10 Confusion Matrix Stacking Ensemble.....	29
Gambar 4. 11 Tampilan Website Page.....	30

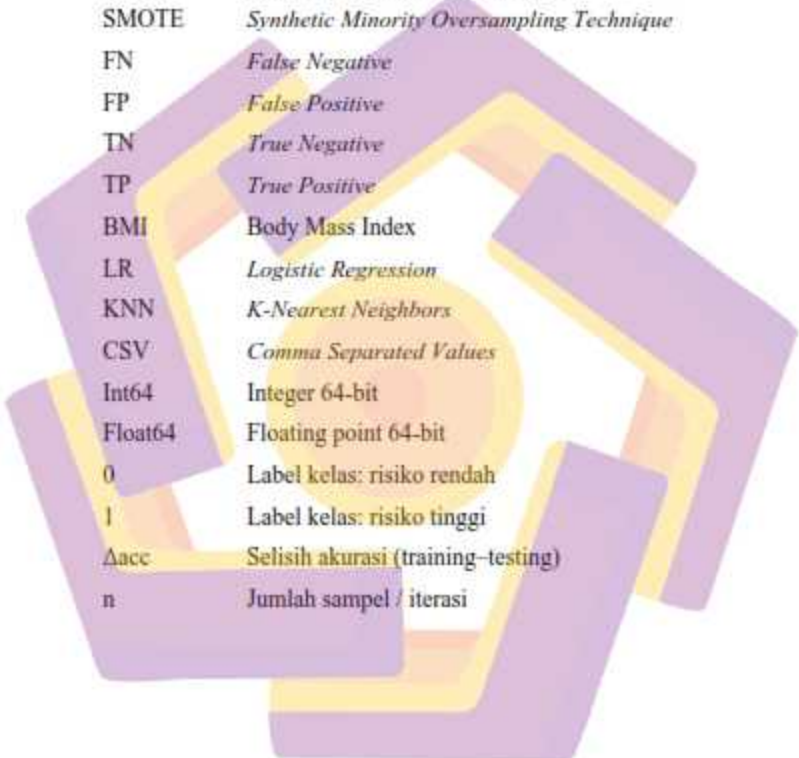


DAFTAR LAMPIRAN

Letter of Acceptance (LOA)	38
Lembar Review	39
Bukti terbit/Terindex	41
Bukti Pembayaran	42



DAFTAR LAMBANG DAN SINGKATAN



ML	<i>Machine Learning</i>
AI	<i>Artificial Intelligence</i>
ANN	<i>Artificial Neural Network</i>
GBDT	<i>Gradient Boosting Decision Trees</i>
SMOTE	<i>Synthetic Minority Oversampling Technique</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
BMI	Body Mass Index
LR	<i>Logistic Regression</i>
KNN	<i>K-Nearest Neighbors</i>
CSV	<i>Comma Separated Values</i>
Int64	Integer 64-bit
Float64	Floating point 64-bit
0	Label kelas: risiko rendah
1	Label kelas: risiko tinggi
Δacc	Selisih akurasi (training-testing)
n	Jumlah sampel / iterasi

DAFTAR ISTILAH



<i>Overfitting</i>	Kecocokan berlebih model
<i>Dataset</i>	Kumpulan data terstruktur
<i>Confusion Matrix</i>	Matriks evaluasi klasifikasi
<i>Precision</i>	Ketepatan prediksi positif
<i>Recall</i>	Cakupan prediksi positif
<i>F1-score</i>	Rata-rata precision-recall
<i>Accuracy</i>	Proporsi prediksi benar
<i>Training Set</i>	Data untuk pelatihan
<i>Testing Set</i>	Data untuk pengujian
<i>Feature Engineering</i>	Rekayasa fitur data
<i>Feature Selection</i>	Seleksi fitur relevan
<i>Imputasi Median</i>	Pengisian nilai median
<i>Oversampling</i>	Penambahan sampel minoritas
<i>Hypertuning</i>	Penyetelan hiperparameter

INTISARI

Hipertensi merupakan penyebab utama penyakit kardiovaskular, gagal ginjal kronis, dan stroke, yang memengaruhi jutaan orang di seluruh dunia. Deteksi dini dan prediksi risiko yang akurat sangat penting untuk manajemen dan pencegahan yang efektif. Studi ini bertujuan untuk mengevaluasi dan membandingkan kinerja berbagai algoritma untuk memprediksi risiko hipertensi menggunakan pendekatan *stacking ensemble*. Model ini menggabungkan tiga algoritma penguat gradien XGBoost, LightGBM, dan CatBoost sebagai pembelajar dasar, dengan Regresi Logistik sebagai pembelajar meta. Dataset, yang bersumber dari Kaggle, berisi 4.240 instans dengan atribut demografis dan klinis yang relevan dengan hipertensi. Langkah-langkah *praproses* meliputi imputasi nilai yang hilang menggunakan median, menghapus entri nol residual, dan menangani ketidakseimbangan kelas melalui algoritma SMOTE. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Evaluasi menunjukkan bahwa model *stacking ensemble* mencapai akurasi keseluruhan sebesar 92,65%, dengan presisi, recall, dan skor F1 secara konsisten mencapai 0,92 untuk kedua kelas. Matriks *konfusi* menunjukkan kesalahan klasifikasi yang minimal, menunjukkan kemampuan model yang kuat untuk membedakan antara individu berisiko rendah dan tinggi. Hasil ini menekankan bahwa tujuan utama penelitian ini adalah untuk mengidentifikasi algoritma mana yang memberikan kinerja terbaik untuk prediksi risiko hipertensi. Dengan mengevaluasi dan membandingkan berbagai model, studi ini menawarkan wawasan dalam memilih algoritma yang paling efektif untuk pengambilan keputusan klinis dan strategi deteksi dini.

Kata kunci: Prediksi Hipertensi, Stacking Ensemble, Pembelajaran Mesin, Penanganan Data Tidak Seimbang, Metrik Klasifikasi

ABSTRACT

Hypertension is a leading cause of cardiovascular diseases, chronic kidney failure, and strokes, affecting millions worldwide. Early detection and accurate risk prediction are crucial for effective management and prevention. This study aims to evaluate and compare the performance of different algorithms for predicting hypertension risk using a stacking ensemble approach. The model combines three gradient boosting algorithms XGBoost, LightGBM, and CatBoost as base learners, with Logistic Regression as the meta learner. The dataset, sourced from Kaggle, contains 4,240 instances with demographic and clinical attributes relevant to hypertension. The preprocessing steps included imputing missing values using the median, removing residual null entries, and addressing class imbalance through the SMOTE algorithm. Data were divided into 80% for training and 20% for testing. The evaluation showed that the stacking ensemble model achieved an overall accuracy of 92.65%, with precision, recall, and F1-scores consistently reaching 0.92 for both classes. The confusion matrix revealed minimal misclassification, indicating the model's strong ability to differentiate between low and high risk individuals. These results emphasize that the primary goal of this research is to identify which algorithm provides the best performance for hypertension risk prediction. By evaluating and comparing different models, this study offers insights into choosing the most effective algorithm for clinical decision-making and early detection strategies.

Keyword: Hypertension Prediction, Stacking Ensemble, Machine Learning, Imbalanced Data Handling, Classification Metrics