

Comparison of Light Gradient Boosting Machine, eXtreme Gradient Boosting, and CatBoost with Balancing and Hyperparameter Tuning for Hypertension Risk Prediction on Clinical Dataset

Dewi Ayu Murtiningsih ^{1*}, Bety Wulan Sari ², Ika Nur Fajri ³

* Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta, Indonesia

dewiayu@students.amikom.ac.id ¹, bety@amikom.ac.id ², fajri@amikom.ac.id ³

Article Info

Article history:

Received 2025-07-22

Revised 2025-09-09

Accepted 2025-09-20

Keyword:

Hypertension,

Prediction,

LGBM,

SMOTE,

Feature selection.

ABSTRACT

Hypertension is a long-lasting condition that is highly prevalent and significantly contributes to cardiovascular issues, making early identification a crucial preventive action. This research evaluates the efficacy of three boosting algorithms, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LGBM), and CatBoost in forecasting hypertension risk. A publicly accessible dataset consisting of 4,363 samples was employed, followed by data preprocessing, feature selection through a voting method that integrates Boruta, Recursive Feature Elimination (RFE), and SelectKBest, as well as addressing class imbalance using the Synthetic Minority Over-sampling Technique (SMOTE) and ADASYN (Adaptive Synthetic Sampling Approach). The models were additionally fine-tuned through hyperparameter optimization using GridSearchCV and Repeated Stratified K-Fold Cross Validation. The evaluation results demonstrate that all three algorithms exhibited strong predictive capabilities, with CatBoost leading the way, achieving an accuracy of 0.992, precision of 0.992, recall of 0.992, F1-score of 0.992, and ROC-AUC of 0.9987. Analyzing the confusion matrix further validated that CatBoost had the lowest number of misclassifications when compared to XGBoost and LGBM. Additionally, the use of SHapley Additive exPlanations (SHAP) for model interpretability highlighted that the key factors influencing the prediction of hypertension risk are blood pressure, body mass index (BMI), overall physical activity, waist circumference, triglyceride levels, age, and LDL cholesterol levels, aligning with established medical knowledge. To facilitate real-world use, the top-performing model was implemented into a user-friendly website interface, allowing users to predict their hypertension risk interactively. These findings illustrate that boosting algorithms, especially CatBoost, offer an accurate, dependable, and interpretable machine learning method for creating hypertension risk prediction systems.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Hipertensi merupakan kondisi saat tekanan darah terhadap dinding arteri meningkat secara berkelanjutan, di mana tekanan darah sebesar 130/80 mmHg atau lebih maka sudah tergolong dalam kategori tekanan darah tinggi [1]. Apabila kondisi ini tidak ditangani dengan serius, akan menyebabkan berbagai komplikasi penyakit serius lainnya seperti ginjal, penyakit jantung, dan stroke. Menurut laporan World Health

Organization (WHO) terdapat sekitar 1,28 miliar orang dewasa direntang umur 30-79 di seluruh dunia mengalami hipertensi, namun 46% orang dewasa ini tidak menyadari bahwa mereka terkena kondisi penyakit tersebut [2]. Hipertensi dikenal dengan julukan “the silent killer” karena tidak menimbulkan gejala pada tahap awal, namun menjadi salah satu faktor risiko utama penyakit kardiovaskular kronis seperti stroke dan gagal jantung [3].

Menurut Survei Kesehatan Indonesia (SKI) 2023, prevalensi hipertensi mencapai 10,7% pada kelompok usia 18–24 tahun dan meningkat menjadi 17,4% pada kelompok usia 25–34 tahun [4]. Kondisi ini tentu mengkhawatirkan, karena hipertensi tidak hanya menyerang kelompok usia lanjut tetapi juga mulai ditemukan pada usia produktif. Hal ini menunjukkan bahwa hipertensi menjadi salah satu penyakit tidak menular yang perlu mendapatkan perhatian dalam deteksi dini dan pencegahan khususnya di Indonesia [5]. Sayangnya rendahnya kesadaran Masyarakat untuk pemeriksaan tekanan darah secara rutin dan keterlambatan diagnosis karena minimnya gejala yang tampak menjadi tantangan utama dalam pengendalian hipertensi [6].

Dalam beberapa tahun terakhir, penggunaan teknologi seperti Artificial Intelligence (AI) dan khususnya Machine Learning (ML) sangat berkembang pesat dalam dunia medis. Teknologi ini telah diterapkan dalam berbagai aspek penanganan medis, terutama dalam hal prediksi dan klasifikasi penyakit seperti hipertensi [7]. Dengan penggunaan machine learning, data medis dapat di analisis dalam jumlah data besar dan dapat menemukan pola tersembunyi yang tidak mudah dikenali oleh manusia [8]. Algoritma machine learning yang digunakan dalam bidang medis cukup beragam, seperti Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR) dan K-Nearest Neighbor (KNN) telah digunakan dalam klasifikasi dan diagnosis penyakit di bidang medis di Indonesia [9].

Algoritma boosting sudah banyak digunakan karena keunggulannya dalam menghasilkan akurasi yang tinggi dan akurat, menangani dataset yang berskala besar serta dapat menangani variabel numerik maupun kategorikal [10]. Pada penelitian ini menggunakan algoritma machine learning berbasis boosting seperti XGBoost, LGBM dan CatBoost. XGBoost dikenal sebagai model yang efektif dalam peningkatan akurasi prediksi dengan efisiensi komputasi yang baik [11], sedangkan model LGBM memiliki kecepatan pelatihan dan efisiensi memori pada dataset yang kompleks dan besar [12]. CatBoost memiliki kemampuan dalam menangani variabel kategorikal secara langsung dan memperkecil kemungkinan adanya bias menggunakan teknik ordered boosting [13]. Dengan pemilihan ketiga algoritma ini dalam perbandingan, diharapkan dapat menghasilkan model yang akurat dalam memberikan prediksi hipertensi.

Beberapa penelitian terdahulu telah mengevaluasi efektivitas algoritma machine learning berbasis boosting yang dibandingkan dengan metode yang lain. Seperti penelitian oleh Ahmad Rafi, dkk melakukan perbandingan algoritma LGBM dan Artificial Neural Networks (ANN) dalam memprediksi kualitas anggur merah. Memberikan hasil bahwa LGBM memiliki akurasi yang lebih tinggi sebesar r 82% dan F1-Score 0,80, serta mampu mencapai keseimbangan yang baik antara presisi dan recall pada kedua kelas kualitas anggur dan komputasi waktu yang lebih efisien dibandingkan ANN yang hanya mendapatkan akurasi sebesar 75,31% dan F1-Score 0,74 [14]. Pada tahun yang sama penelitian oleh Susan, dkk yang membandingkan XGBoost

dan Decision Tree (DT) dalam prediksi penyakit jantung, menunjukkan keunggulan model XGBoost yang mencapai akurasi 93% daripada Decison Tree yang hanya memperoleh 90%[15]. Penelitian lainnya oleh Abdullahi melakukan perbandingan beberapa algoritma meliputi Logistic Regression, Random Forest, Adaboost, XGBoost, Neural Networks, Gradient Boosting, CatBoost dan DT, melaporkan bahwa algoritma CatBoost menghasilkan skor terbaik 94% serta memberikan prediksi lebih akurat untuk persetujuan pinjaman dan promosi pegawai dibandingkan algoritma lainnya[16]. Di bidang lainnya, memprediksi produksi dan tarikan (trip generation) di Ilorin City dengan memanfaatkan data karakteristik rumah tangga, perjalanan, populasi dan peta kota. Hasil penelitian ini menunjukkan bahwa CatBoost mampu memprediksi lebih tepat jumlah perjalanan dibandingkan ANN, dengan R-squared 0,99 untuk produksi perjalanan dan akurasi 0,994 untuk tarikan perjalanan, menegaskan kemampuannya menghasilkan prediksi yang sangat akurat dan stabil[17].

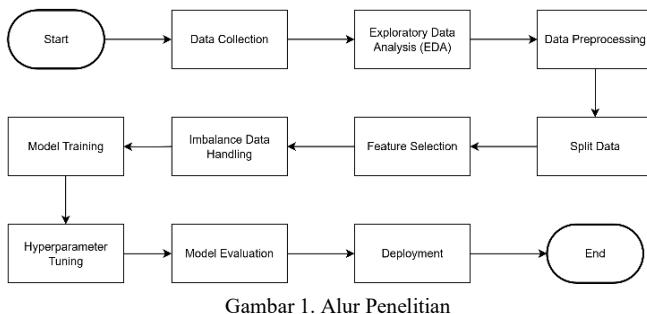
Terdapat penelitian lainnya yang mengkaji topik yang serupa mengenai prediksi hipertensi. Penelitian tahun 2022 oleh Rico Kurniawan, dkk membandingkan algoritma DT, RF, Gradient Boosting, dan LR dengan metrik AUC, akurasi, F1-score, presisi, dan recall. Algoritma LR menjadi model terbaik dengan performa AUC 0,829, akurasi 89,6%, presisi 0,878, recall 0,896, dan F1-score 0,877 [18]. Penelitian oleh Franscicus dan Caecilla mengimplementasikan model Gradient Boosting Decision Tree (GDPT) pada dataset sebanyak 70.693 dengan 17 perilaku dan riwayat kesehatan, serta dioptimalkan menggunakan Random Search, Grid Search, Bayesian Optimization, Grading Search dan Tree Parzen Estimation. Validasi tertinggi dicapai dengan Tree Parzen Estimation mencapai akurasi 74,43% [19]. Penelitian lain oleh Chai, dkk menggunakan sampel dengan 30,1% subjek hipertensi pertama-tama dibagi menjadi 90% untuk pelatihan dan 10% untuk validasi. Dataset diseimbangkan menggunakan SMOTE dan random under-sampling, dengan hasil bahwa model boosting modern seperti LightGBM, XGBoost, dan CatBoost mampu memberikan performa terbaik pada akurasi dan AUC dibandingkan model lama [20]. Pada tahun 2024 terdapat penelitian oleh Thomas Mroz, dkk melatih tiga algoritma boosting modern (XGBoost, LGBM, dan CatBoost) pada 350.008 pasien, menghasilkan akurasi AUC sebanyak 76%. Pada penelitian ini juga menegaskan perlunya eksplorasi lebih lanjut pada integrasi algoritma boosting dengan teknik balancing data dan feature selection untuk meningkatkan akurasi prediksi [21].

Berdasarkan penelitian-penelitian sebelumnya, terlihat bahwa algoritma boosting modern telah menunjukkan potensi yang kuat dalam memprediksi hipertensi. Namun, sebagian besar penelitian hanya berfokus pada penerapan satu metode seleksi fitur atau satu algoritma boosting saja, serta masih terbatas dalam mengeksplorasi antara. Oleh karena itu, penelitian ini menghadirkan kebaruan dengan membandingkan secara sistematis tiga algoritma boosting modern (XGBoost, LGBM, dan CatBoost) yang dipadukan

dengan teknik seleksi fitur (RFE, Boruta, dan SelectKBest), balancing data (SMOTE/ADASYN), validasi silang, serta analisis interpretabilitas menggunakan SHAP. Sehingga penelitian ini tidak hanya bertujuan menghasilkan model prediksi hipertensi yang akurat dan stabil, tetapi juga menyediakan hasil yang lebih mudah dijelaskan serta diimplementasikan dalam bentuk website sederhana sebagai aplikasi praktis.

II. METODE

Penelitian ini bertujuan untuk membangun model prediksi risiko hipertensi dengan membandingkan kinerja tiga algoritma antara XGBoost, LGBM dan CatBoost. Untuk mencapai tujuan tersebut terdapat beberapa tahapan penelitian yang perlu dilakukan, mulai dari pengumpulan data, preprocessing hingga deployment. Alur penelitian ini digambarkan pada Gambar 1.



A. Data Collection

Pada penelitian ini menggunakan dataset publik yang tersedia pada situs Kaggle dengan judul “Hypertension_Arterial_Mexico” [22]. Dataset tersebut terdiri dari 4.363 data dengan 36 atribut. Informasi lebih lengkap mengenai dataset dapat dilihat pada Tabel 1.

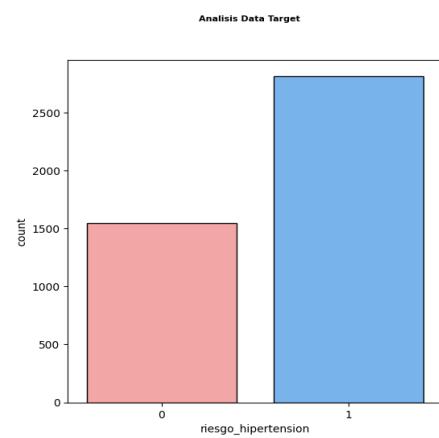
TABEL 1
INFORMASI DATASET

| Judul | Keterangan |
|------------------|--|
| Dataset | Hypertension Arterial Mexico |
| Atribut | 36 (Seluruhnya numerik) |
| Total Data | 4.363 |
| Duplikasi | 0 |
| Distribusi Kelas | Risiko: 2.816 (64.5%), Tidak Risiko: 1.547 (35.5%) |
| Target | Hypertension (1 = Risiko, 0 = Tidak Risiko) |

B. Exploratory Data Analysis (EDA)

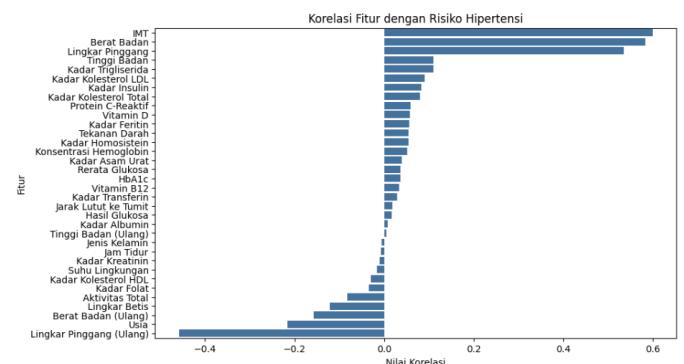
Exploratory Data Analysis (EDA) digunakan untuk memahami struktur data, mendeteksi pola, mengidentifikasi

anomali, serta menguji asumsi sebelum proses pemodelan dilakukan [23]. Pada tahap ini dilakukan analisis karakteristik data secara menyeluruh mulai dari analisis distribusi hingga korelasi antar fitur dengan target fitur. Target variabel yang digunakan adalah riesgo_hipertension, yaitu variabel biner yang terdiri dari dua kelas yaitu kelas 0 (tidak berisiko) dan 1 (berisiko). Distribusi kelas target ditunjukkan pada Gambar 3.



Gambar 3. Bar Chart Distribusi Kelas riesgo_hipertension

Pada gambar 3 menunjukkan ketidakseimbangan distribusi kelas yang cukup signifikan. Dari total 4.363 data terdapat sebanyak 2.816 (64,5%) pasien yang dikategorikan berisiko hipertensi pada kelas 1, sedangkan 1.547 (35,5%) pasien yang tidak berisiko pada kelas 0. Ketidakseimbangan ini perlu ditangani lebih lanjut untuk mencegah bias model terhadap kelas mayoritas. Oleh karena itu, penggunaan teknik penyeimbangan data seperti SMOTE atau ADASYN dilakukan untuk memastikan performa model yang lebih adil dan akurat dalam klasifikasi kedua kelas.



Gambar 4. Korelasi Fitur dengan Kelas Target

Selain itu pada proses EDA dilakukan analisis korelasi untuk melihat keterkaitan antar fitur dengan target hipertensi. Pada gambar 4 menunjukkan hasil analisis korelasi dilakukan untuk melihat keterikatan fitur-fitur dengan target hipertensi. Hasil analisis menunjukkan fitur Indeks Masa Tubuh (IMT), berat badan, lingkar pinggang hingga tinggi badan memiliki korelasi positif yang kuat. Sedangkan beberapa variabel

lingkar pinggang ulang dan usia menunjukkan korelasi negatif yang relatif lemah.

C. Data Preprocessing

Tahapan ini merupakan langkah penting untuk memastikan data bersih dan siap digunakan dalam pengembangan model. Proses ini dilakukan penghapusan duplikasi data, penanganan missing values dan penghapusan fitur yang tidak relevan seperti "FOLIO_I". Proses encoding categorical variabel tidak dilakukan pada proses ini karena semua fitur tidak ada yang memiliki tipe kategorikal. Selain itu, nama-nama fitur yang sebelumnya dalam bahasa Meksiko kemudian diterjemahkan ke dalam bahasa Indonesia, sehingga dataset dapat dianalisis dengan mudah tanpa kebingungan dalam pemahaman.

D. Split Data

Setelah tahapan preprocessing, data akan dibagi menjadi data latih dan data uji dengan rasio 70:30. Data latih akan digunakan untuk melatih model supaya bisa mengenali pola dan karakteristik data, sedangkan data uji akan digunakan untuk mengevaluasi kinerja model. Proses pembagian menggunakan metode stratified sampling, distribusi kelas pada kedua subset tetap konsisten dengan dataset asli. Pada data latih yang terdiri dari 3.054 sampel, terdapat 35,46% pasien tidak berisiko hipertensi dan 64,54% pasien berisiko hipertensi. Sementara itu, pada data uji yang terdiri dari 1.309 sampel, distribusi kelas juga seimbang, yaitu 35,45% untuk tidak berisiko dan 64,55% untuk berisiko. Proses pemisahan dilakukan sebelum proses pelatihan untuk menghindari data leakage.

E. Feature Selection

Feature selection atau seleksi fitur berguna untuk menyederhanakan model, mempercepat proses pelatihan, serta meningkatkan akurasi dan generalisasi. Dalam penelitian ini menggunakan tiga teknik yaitu RFE, SelectKBest, dan Boruta. RFE merupakan metode seleksi fitur yang berbasis wrapper dimana fitur-fitur yang memiliki kontribusi rendah akan dihapus secara bertahap hingga tersisa fitur yang relevan [24]. Boruta merupakan metode seleksi fitur berbasis random forest yang membandingkan fitur asli dengan fitur bayangan untuk memilih fitur yang memiliki signifikansi besar terhadap target. Sedangkan SelectKBest akan memilih beberapa fitur terbaik berdasarkan skor statistik tertentu, seperti nilai F-score (ANOVA) terhadap variabel target[25].

Dari ketiga metode feature selection tersebut menghasilkan jumlah dan nama fitur yang berbeda karena perbedaan cara evaluasi pentingnya atribut. Untuk memperoleh fitur-fitur yang lebih stabil dan tangguh, digunakan pendekatan Voting Feature Selection [26]. Dengan cara ini, akan memilih fitur yang muncul minimal dua kali dari tiga metode tersebut, sehingga dapat mempertahankan fitur yang konsisten dan dianggap penting oleh mayoritas metode. Metode ini

diharapkan mampu memberikan kontribusi signifikan pada tahap pelatihan model.

F. Handling Imbalance Data

Dataset yang digunakan memiliki distribusi kelas yang tidak seimbang, dimana jumlah pasien yang berisiko hipertensi lebih banyak dibandingkan dengan yang tidak berisiko. Data yang tidak seimbang seperti ini akan menyebabkan model cenderung bias terhadap kelas mayoritas sehingga mengurangi kemampuan model dalam mengenali kelas minoritas. Oleh karena itu, diperlukan penanganan imbalance data pada proses pelatihan model.

Dalam penelitian ini digunakan teknik SMOTE dan ADASYN. SMOTE bekerja dengan membuat sampel sintesis baru pada kelas minoritas menggunakan teknik interpolasi antar tetangga terdekat. Dengan cara ini, SMOTE tidak hanya dapat menduplikasi data yang ada, tetapi dapat menghasilkan data baru menyerupai distribusi kelas minoritas. Dengan metode ini terbukti dapat meningkatkan oversampling konvensional [27]. Sementara itu, ADASYN merupakan pengembangan dari SMOTE yang lebih adaptif. Metode ini menghasilkan lebih banyak data sintesis pada area di sekitar secison boundary atau titik minoritas yang sulit dipelajari model, sehingga tidak hanya menyeimbangkan distribusi data namun juga dapat membantu model belajar lebih baik pada kasus-kasus sulit dan dapat meningkatkan kemampuan generasi [28].

Dengan adanya penerapan metode balancing ini model dapat belajar mengenali pola dari kedua kelas secara adil. Hal ini diharapkan mampu meningkatkan kinerja model khususnya pada metrik recall dan F1-Score yang lebih relevan dalam konteks medis. Serta proses ini dilakukan setelah proses feature selection agar sintesis data hanya dilakukan pada fitur yang relevan sehingga dapat mengurangi potensi noise dan menjaga kualitas data dalam proses pelatihan model.

G. Model Training

Tahapan model training merupakan inti dari proses penelitian ini, bagaimana model dilatih untuk mengenali pola pada data dan melakukan prediksi risiko hipertensi. Pada penelitian ini menggunakan tiga algoritma utama, yaitu LGBM, XGBoost dan CatBoost. Ketiga algoritma tersebut merupakan bagian dari algoritma ensamble learning berbasis gradient boosting yang dikenal mampu memberikan performa tinggi, efisiensi komputasi serta penanganan berbagai tipe dataset dengan baik.

Algoritma XGBoost merupakan salah satu algoritma boosting yang banyak digunakan karena kemampuannya memberikan hasil akurat dengan efisiensi komputasi yang tinggi. Studi terbaru dalam bidang kesehatan menunjukkan bahwa XGBoost mampu mengidentifikasi faktor risiko penyakit kronis dengan performa lebih baik dibandingkan metode konvensional [11]. LightGBM dirancang untuk mempercepat proses pelatihan dan efisiensi penggunaan

memori tanpa mengorbankan akurasi. Algoritma ini menggunakan pendekatan leaf-wise growth strategy serta histogram-based learning yang membuatnya unggul dalam menangani dataset berukuran besar. Sebuah penelitian yang memprediksi komplikasi hipertensi menunjukkan bahwa LightGBM memberikan hasil yang kompetitif dengan waktu pelatihan yang lebih cepat dibandingkan metode lain berbasis boosting [12]. Sedangkan algoritma CatBoost memiliki keunggulan dalam menangani fitur kategorikal secara langsung tanpa memerlukan proses encoding tambahan. Hal ini menjadikannya lebih efisien dan mengurangi risiko prediction shift. Pada domain kesehatan, CatBoost telah diaplikasikan untuk mendeteksi penyakit kardiovaskular dengan akurasi tinggi, menunjukkan potensi besar algoritma ini dalam pemodelan prediktif berbasis data klinis [13].

H. Hyperparameter Tuning

Pada penelitian ini terdapat tahapan Hyperparameter Tuning yang digunakan untuk memperoleh konfigurasi parameter terbaik dan model dapat meningkatkan kinerjanya dalam memprediksi risiko hipertensi. Hiperparameter merupakan komponen yang ditentukan sebelum proses pelatihan model dan sangat memengaruhi performa akhir. Beberapa hiperparameter penting pada algoritme berbasis boosting antara lain jumlah pohon (n_estimators), kedalaman pohon (max_depth), serta laju pembelajaran (learning_rate) yang akan diatur kedalam masing-masing model yang digunakan [29].

Metode tuning yang digunakan adalah GridSearchCV, yaitu teknik optimisasi hiperparameter dengan cara membangun model untuk setiap kombinasi hiperparameter yang ada dalam grid kemudian mengevaluasi performanya menggunakan k-fold cross-validation. Dengan pendekatan ini, diperoleh kombinasi parameter yang paling optimal dan memiliki kemampuan generalisasi yang baik, sekaligus membantu mencegah masalah overfitting [30].

I. Model Evaluation

Evaluasi model dilakukan untuk mengukur performa algoritma dalam memprediksi risiko hipertensi. Metode yang digunakan salah satunya dengan Confusion Matrix, yang menyajikan informasi mengenai prediksi benar dan salah pada kelas aktual maupun hasil prediksi model seperti pada tabel 2.

TABEL 2
CONFUSION MATRIX

| Actual/Predicted | Positive | Negative |
|------------------|---------------------|---------------------|
| Positive | True Positive (TP) | False Positive (FP) |
| Negative | False Negative (FN) | True Negative (TN) |

Pada tabel 1 menunjukkan empat komponen yang True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Komponen-komponen ini merepresentasikan hasil prediksi model dibandingkan dengan

label sebenarnya. Komponen-komponen ini menjadi dasar dalam perhitungan metrik evaluasi seperti accuracy, precision, recall, f1-score, dan AUC. Berikut adalah rumus yang digunakan untuk menghitung matriks evaluasi :

1) Precision

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

2) Recall

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

3) Accuracy

$$Accuracy = \frac{(TP + TN)}{(TP + FN + FP + FN)} \quad (3)$$

4) F1-score

$$F1-score = 2 \times \frac{(precision \times recall)}{(precision + recall)} \quad (4)$$

5) AUC

$$AUC = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (5)$$

III. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi dan evaluasi dari tiga algoritma, yaitu LGBM, XGBoost dan CatBoost, dalam melakukan prediksi risiko hipertensi. Penelitian ini dilakukan dengan pendekatan sistematis yang meliputi preprocessing data, seleksi fitur dengan metode voting dari tiga metode RFE, SelectKBest dan Boruta kemudian dataset akan dilakukan penanganan ketidakseimbangan, pelatihan model, serta evaluasi performa menggunakan metrik klasifikasi.

A. Preprocessing data

Pada tahapan preprocessing dilakukan beberapa langkah untuk memastikan kualitas data sebelum masuk ke tahap analisis lebih lanjut. Hasil pengecekan terhadap data kosong dan duplikat menunjukkan bahwa dataset yang digunakan sudah bersih sehingga tidak diperlukan proses imputasi maupun penghapusan data ganda. Selanjutnya, dilakukan penghapusan fitur yang tidak berkontribusi, yaitu kolom FOLIO_I, karena tidak memiliki relevansi dengan target prediksi. Sehingga jumlah fitur yang akan diolah sebanyak 35 fitur.

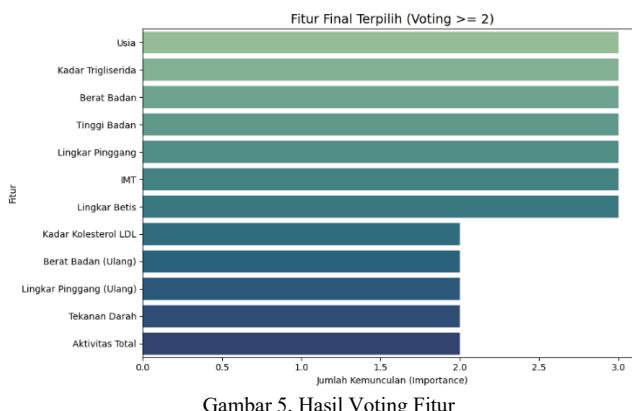
B. Feature Selection

Setelah data diproses, dilakukan proses seleksi fitur guna mengidentifikasi fitur yang paling relevan dalam memprediksi risiko hipertensi dan meningkatkan perfoma model. Tiga metode yang diterapkan yaitu Boruta, RFE dan SeleckBest. Hasil seleksi fitur menunjukkan bahwa tidak semua fitur memiliki kontribusi yang signifikan terhadap prediksi.

TABEL 3
HASIL SELEKSI FITUR

| Metode | Fitur |
|-------------|---|
| Boruta | Usia, Kadar Kolesterol LDL, Kadar Trigliserida, Berat Badan, Tinggi Badan, Lingkar Pinggang, Berat Badan (Ulang), Lingkar Betis, Lingkar Pinggang (Ulang), Tekanan Darah, IMT, dan Aktivitas Total. |
| RFE | Usia, Kadar Kolesterol LDL, Kadar Trigliserida, Berat Badan, Tinggi Badan, Lingkar Pinggang, Lingkar Betis, Tekanan Darah, IMT, dan Aktivitas Total. |
| SelectKBest | Usia, Kadar Trigliserida, Protein C-Reaktif, Berat Badan, Tinggi Badan, Lingkar Pinggang, Berat Badan (Ulang), Lingkar Betis, Lingkar Pinggang (Ulang), dan IMT. |

Dari ketiga metode di atas digunakan cara untuk memperoleh hasil yang lebih stabil dan kuat yaitu dengan cara Voting. Metode voting akan memilah dari ketiga hasil dengan syarat minimal fitur yang muncul dua kali dari ketiga metode.



Gambar 5. Hasil Voting Fitur

Pada gambar di atas menunjukkan hasil voting fitur unggulan setiap metode seleksi fitur sebanyak 12 fitur. Dengan melakukan cara ini, dapat membuat model menjadi lebih fokus pada variabel yang paling relevan, sehingga meningkatkan efisiensi dan potensi akurasi prediksi.

C. Imbalance data

Analisis distribusi awal data menunjukkan adanya ketidakseimbangan kelas target 0 dan 1. Sebelum dilakukan penyeimbangan, jumlah data kelas Tidak Berisiko lebih dominan dibandingkan kelas Berisiko, yaitu 1971 sampel berbanding 1083 sampel dengan total 3054 data. Kondisi ini

berpotensi menimbulkan bias pada model, di mana model cenderung melakukan prediksi pada kelas mayoritas.. Oleh karena itu diatasi dengan teknik imbalance data, seperti SMOTE dan ADASYN.

Berikut adalah perbandingan distribusi kelas sebelum dan sesudah imbalance menggunakan SMOTE dan ADASYN.

TABEL 4
PERBANDINGAN DISTRIBUSI DATA

| Data | Jumlah data | | Total Data |
|----------------|-------------|---------|------------|
| | Kelas 0 | Kelas 1 | |
| Sebelum SMOTE | 1083 | 1971 | 3054 |
| Setelah SMOTE | 1971 | 1971 | 3942 |
| Sebelum ADASYN | 1083 | 1971 | 3054 |
| Setelah ADASYN | 2000 | 1971 | 3971 |

Berdasarkan hasil pada tabel 1, SMOTE menghasilkan distribusi kelas yang seimbang dengan jumlah sampel yang sama antara kelas tidak berisiko dan berisiko dengan perbandingan 1971 : 1971. Sementara itu, ADASYN menghasilkan distribusi yang relatif seimbang, namun dengan jumlah kelas minoritas yang sedikit lebih besar dibandingkan kelas mayoritas sebanyak 2000 : 1971. Dalam perbandingan kedua metode tersebut, SMOTE memiliki hasil yang lebih stabil dan tidak terlalu banyak data sintesis.

Meskipun SMOTE efektif dalam menyeimbangkan distribusi kelas, metode ini berpotensi untuk bias karena sampel sintetis dibuat dari interpolasi fitur-fitur kelas minoritas. Oleh karena itu, SMOTE hanya diterapkan pada data latih, sedangkan data uji tetap asli. Selain itu, penerapan cross-validation membantu memastikan model tidak terlalu bergantung pada sampel sintetis tertentu, sehingga kemampuan generalisasi tetap terjaga.

D. Model Training

Pada tahapan ini algoritma XGBoost, LGBM dan CatBoost akan dilatih dengan untuk mengenali pola pada data dan melakukan prediksi risiko hipertensi. Sebelum tahap pelatihan utama, dilakukan perbandingan performa model berdasarkan dua teknik penyeimbangan data, yaitu SMOTE dan ADASYN, untuk menentukan metode yang paling efektif dalam meningkatkan akurasi prediksi. Berikut adalah tabel perbandingan antara SMOTE dan ADASYN.

TABEL 5
PERBANDINGAN AKURASI SMOTE

| Model | Accuracy | Precision | Recall | F1-Score |
|----------|----------|-----------|--------|----------|
| XGBoost | 0.9878 | 0.9964 | 0.9846 | 0.9905 |
| LGBM | 0.9893 | 0.9976 | 0.9858 | 0.9917 |
| CatBoost | 0.9924 | 0.9976 | 0.9905 | 0.9941 |

TABEL 6
PERBANDINGAN AKURASI ADASYN

| Model | Accuracy | Precision | Recall | F1-Score |
|----------|----------|-----------|--------|----------|
| XGBoost | 0.9862 | 0.9964 | 0.9822 | 0.9893 |
| LGBM | 0.9893 | 0.9976 | 0.9858 | 0.9917 |
| CatBoost | 0.9916 | 0.9976 | 0.9893 | 0.9893 |

Pada tabel 5 dan 6 Setelah dilakukan evaluasi, terlihat bahwa penggunaan SMOTE memberikan performa sedikit lebih unggul dibandingkan ADASYN pada sebagian besar metrik evaluasi, terutama pada akurasi dan F1-Score. Terlihat pada tabel 4 dan tabel 5, dimana algoritma terbaik pada akurasi yaitu CatBoost dengan SMOTE menunjukkan nilai akurasi 0.9924 dan F1-Score 0.9941, lebih tinggi dibandingkan CatBoost dengan ADASYN yang memiliki akurasi 0.9916 dan F1-Score 0.9935.

Berdasarkan hasil perbandingan antara SMOTE dan ADASYN, metode SMOTE dipilih sebagai metode penyeimbangan data pada tahapan pelatihan model utama. Pemilihan ini bertujuan untuk menjaga kestabilan prediksi, meminimalkan bias terhadap kelas mayoritas, serta memastikan model mampu mengenali pola pada kelas minoritas dengan baik pada saat proses training maupun testing.

E. Hyperparameter Tuning

Setelah data sudah diseimbangkan menggunakan balancing data, tahap berikutnya adalah melakukan hyperparameter tuning kepada masing-masing algoritma agar perbandingan perfoma pada ketiga model dapat adil. Dengan metode hyperparameter, parameter yang akan ditentukan akan sangat berpengaruh pada kemampuan model dalam mempelajari pola pada prediksi data.

TABEL 7
PARAMETER GRID SEARCH

| Metode | Parameter | Hasil Grid Search |
|----------|--|-------------------|
| XGBoost | colsample_bytree: 1.0, learning_rate: 0.05, max_depth: 6, n_estimators: 200, subsample: 1.0 | 0.9977 |
| LGBM | colsample_bytree: 0.8, learning_rate: 0.05, max_depth: -1, n_estimators: 200, subsample: 0.8 | 0.9974 |
| CatBoost | depth: 10, iterations: 200, learning_rate: 0.1 | 0.9977 |

Pada penelitian ini menggunakan teknik GridSearchCV dengan metode Repeated Stratified K-Fold Cross-Validation. Penggunaan Repeated Stratified K-Fold Cross-Validation dilakukan pada data latih yang sudah diseimbangkan menjadi beberapa fold/subset secara stratifikasi, sehingga proporsi pada dua kelas target tetap terjaga di setiap fold. Setiap fold digunakan secara bergantian sebagai data validasi sementara sisanya digunakan untuk pelatihan, dan proses ini diulang beberapa kali untuk meningkatkan stabilitas evaluasi. Penggunaan lipatan atau K-fold yang digunakan sebanyak 5

fold dan akan diulang sebanyak 3 kali, sehingga setiap kombinasi parameter akan dievaluasi secara menyeluruh pada semua berbagai macam data latih.

F. Evaluasi Model

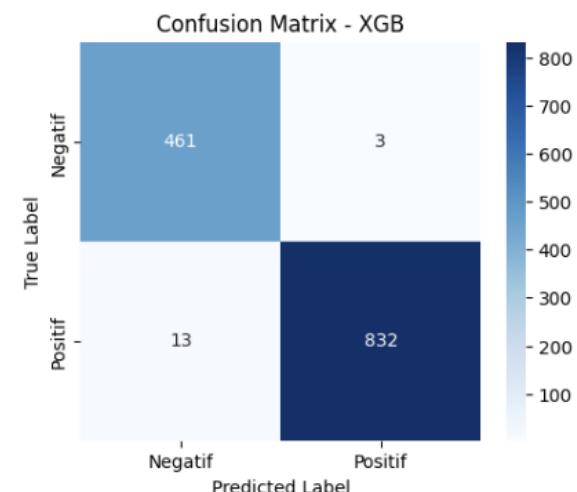
Setelah hyperparameter tuing dilakukan pada ketiga algoritma yang dilatih pada data latih. Evaluasi performa dilakukan menggunakan metrik klasifikasi, meliputi accuracy, precision, recall, dan F1-score, serta melalui confusion matrix untuk menilai kemampuan model dalam membedakan kelas "Tidak Berisiko" dan "Berisiko"

Pada tabel berikut menunjukkan hasil performa masing-masing algoritma setelah dilakukan hyperparameter tuning seperti tabel 8,

TABEL 8
EVALUASI MODEL SEBELUM TUNING

| Model | Accuracy | Precision | Recall | F1-score |
|----------|----------|-----------|--------|----------|
| XGBoost | 0.9878 | 0.9964 | 0.9846 | 0.9905 |
| LGBM | 0.9893 | 0.9976 | 0.9858 | 0.9917 |
| CatBoost | 0.9924 | 0.9976 | 0.9905 | 0.9941 |

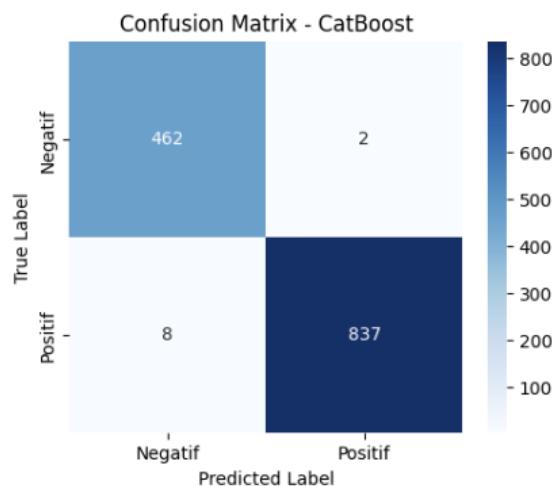
Dari tabel diatas menunjukkan algoritma CatBoost memiliki akurasi 99,77%, precision 99.76%, recall 99.05%, dan F1-Score sebanyak 99.41%, menjadikan model terbaik dibandingkan model XGBoost dan LGBM walaupun hanya memiliki selisih yang sedikit. Hal ini menunjukkan bahwa ketiga algoritma boosting ini mampu memprediksi risiko hipertensi dengan sangat baik.



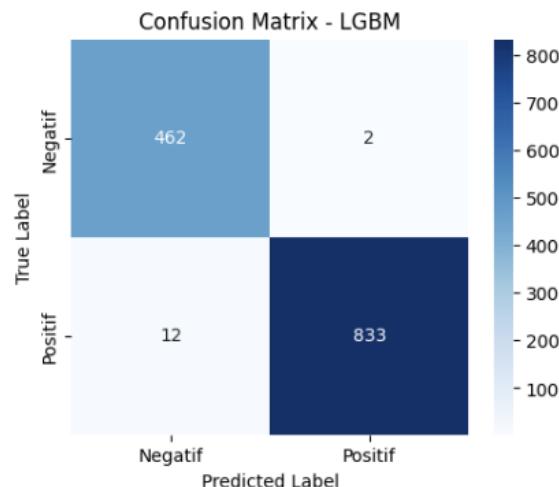
Gambar 6. Confusion Matrix Algoritma XGBoost

Berdasarkan hasil evaluasi menggunakan confusion matrix, terlihat bahwa ketiga algoritma yang diuji, yaitu XGBoost, CatBoost dan LGBM mampu memberikan performa klasifikasi yang baik setelah dilakukan penanganan data tidak seimbang dengan metode SMOTE. Pada gambar 6 algoritma XGBoost, model berhasil mengklasifikasikan 461 data negatif dan 832 data positif dengan benar, meskipun

masih terdapat 3 data negatif yang salah diprediksi sebagai positif dan 13 data positif yang salah terklasifikasi sebagai negatif. Hal ini menunjukkan bahwa XGBoost cukup baik dalam mendeteksi kelas positif, namun masih menghasilkan kesalahan klasifikasi pada sebagian data positif sehingga recall sedikit lebih rendah dibanding algoritma lainnya.



Gambar 7. Confusion Matrix Algoritma CatBoost



Gambar 8. Confusion Matrix Algoritma LGBM

Gambar 7 menunjukkan performa CatBoost paling baik di antara ketiga algoritma. Model ini mampu mengklasifikasikan 462 data negatif dan 837 data positif dengan benar hanya 2 data negatif dan 8 data positif yang salah diprediksi. Kesalahan klasifikasi yang sangat minim ini menjadikan CatBoost unggul dalam hal recall maupun precision, sehingga memberikan keseimbangan yang optimal pada nilai F1-score.

Gambar 8 algoritma LGBM juga menunjukkan performa yang mendekati CatBoost, dengan hasil klasifikasi benar sebanyak 462 data negatif dan 833 data positif, serta kesalahan pada 2 data negatif dan 12 data positif. Performa LGBM tetap cukup tinggi, meskipun recall sedikit lebih

rendah dibanding CatBoost akibat jumlah kesalahan pada data positif yang lebih banyak.

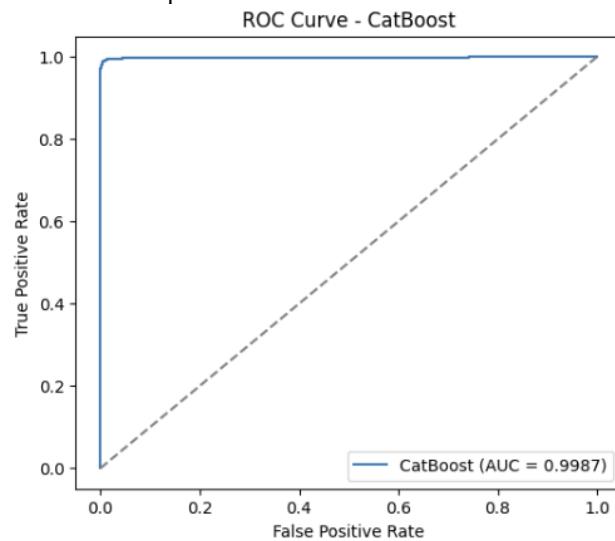
Pada ketiga algoritma boosting yang digunakan mampu memberikan akurasi dan konsistensi hasil prediksi yang baik. Namun, CatBoost menempati posisi terbaik karena menghasilkan jumlah kesalahan klasifikasi paling sedikit, diikuti oleh LGBM yang memiliki performa hampir setara dan XGBoost yang meskipun kuat namun masih sedikit tertinggal dalam hal recall. Hasil ini mengindikasikan bahwa CatBoost merupakan algoritma yang paling optimal untuk digunakan dalam sistem prediksi risiko hipertensi pada penelitian ini.

TABEL IX
CLASSIFICATION REPORT

| | Precision | Recall | F1-score | Support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.983 | 0.996 | 0.989 | 464 |
| 1 | 0.998 | 0.991 | 0.994 | 845 |
| Accuracy | | | 0.992 | 1309 |
| Macro avg | 0.990 | 0.993 | 0.992 | 1309 |
| Weighted avg | 0.992 | 0.992 | 0.992 | 1309 |

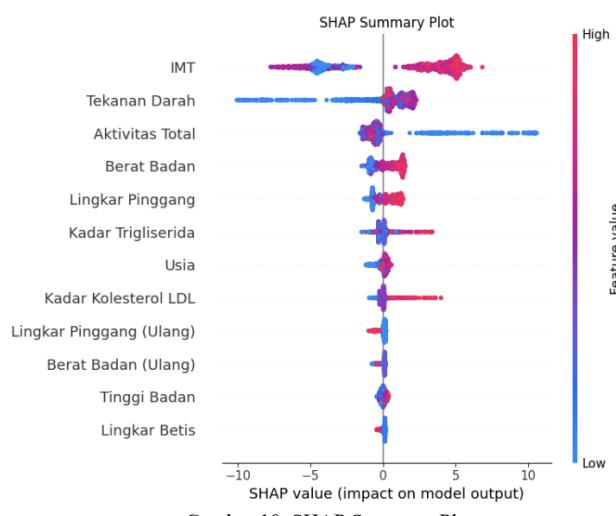
Berdasarkan Tabel IV Classification Report, algoritma CatBoost menunjukkan performa yang sangat baik dalam klasifikasi risiko hipertensi. Pada kelas 0 (negatif), model memperoleh precision sebesar 0,983, recall sebesar 0,996 dan F1-score sebesar 0,989 dengan jumlah data uji sebanyak 464. Nilai recall yang tinggi pada kelas ini menunjukkan bahwa hampir seluruh data negatif dapat diprediksi dengan benar, sehingga risiko false positive relatif kecil.

Untuk kelas 1 (positif), CatBoost mencapai precision sebesar 0,998, recall sebesar 0,991, dan F1-score sebesar 0,994 dengan jumlah data uji sebanyak 845. Nilai precision yang sangat tinggi menandakan bahwa hampir seluruh data yang diprediksi positif memang benar-benar positif, sementara nilai recall yang mendekati sempurna menunjukkan bahwa model mampu mendeteksi hampir semua kasus hipertensi.



Gambar 9. Curva ROC-AUC Algoritma CatBoost

Berdasarkan hasil classification report menunjukkan model kuat dengan nilai accuracy sebesar 0,992 precision 0,992 recall 0,992, dan F1-score 0,992. Hal ini menunjukkan bahwa CatBoost mampu menjaga keseimbangan antara kemampuan mengidentifikasi data positif (recall) dan ketepatan prediksi (precision). Selain itu, nilai ROC-AUC sebesar 0,9987 memperlihatkan bahwa model memiliki kemampuan diskriminasi yang hampir sempurna dalam membedakan antara individu berisiko hipertensi dan yang tidak. Sehingga performa algoritma CatBoost menunjukkan kinerja yang sangat baik dalam memprediksi risiko hipertensi.

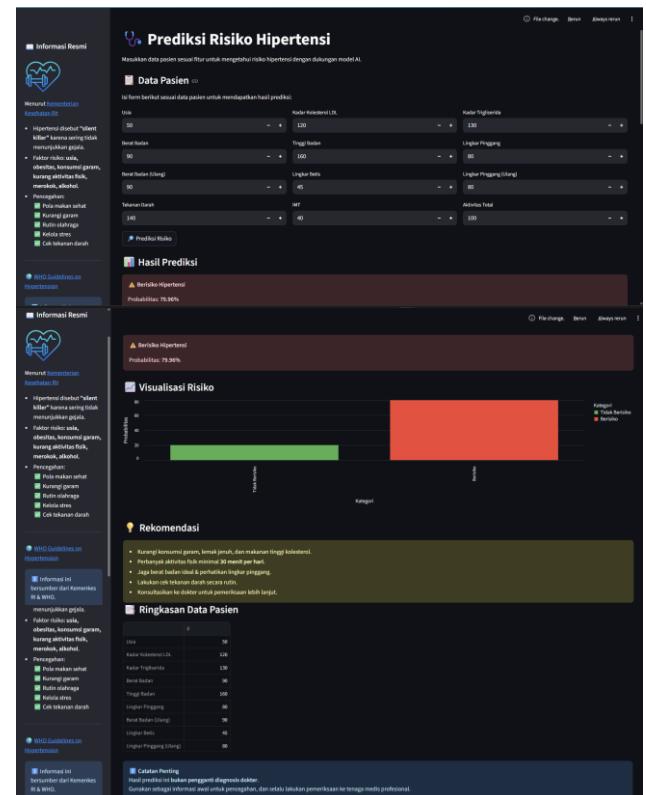


Gambar 10. SHAP Summary Plot

Gambar 10 menunjukkan hasil plot ringkasan SHAP yang menunjukkan gambaran interpretatif mengenai kontribusi setiap fitur terhadap prediksi risiko hipertensi. Fitur yang sebelumnya dihasilkan dari seleksi fitur menggunakan teknik voting kemudian pada tahapan ini memberikan gambaran interpretatif mengenai kontribusi setiap fitur terhadap prediksi. Hasil menunjukkan bahwa faktor yang paling berpengaruh adalah Indeks Masa Tubuh (IMT). Tekanan darah juga muncul sebagai faktor dominan, dengan nilai yang semakin tinggi mendorong prediksi menuju hipertensi, sedangkan nilai minus atau rendah menurunkan risiko. Aktivitas total memiliki peran penting karena tingkat aktivitas yang rendah berkaitan dengan risiko lebih tinggi, sehingga menegaskan pentingnya faktor gaya hidup. Faktor metabolismik seperti kadar trigliserida dan kolesterol LDL juga menunjukkan hubungan positif terhadap risiko hipertensi, khususnya pada nilai yang tinggi. Fitur-fitur seperti usia, berat badan, lingkar pinggang, tinggi badan, dan lingkar betis tetap berkontribusi terhadap prediksi, meskipun pengaruhnya relatif lebih kecil dibandingkan dengan variabel utama. Temuan ini memperkuat bukti bahwa faktor yang mempengaruhi perkembangan penyakit hipertensi yaitu obesitas, tekanan darah, pola hidup, dan kondisi metabolismik.

G. Deploy

Pada gambar 11 di bawah ditampilkan antarmuka website hasil deploy dari sistem prediksi risiko hipertensi. Website ini dirancang sebagai bentuk aplikasi dengan inputan data pasien seperti, usia, jenis kelamin, tekanan darah, indeks massa tubuh, serta lingkar perut dan pinggang yang berasal dari hasil voting feature selection. Setelah data dimasukkan, sistem akan memproses input tersebut menggunakan model catboost yang telah dilatih, kemudian menghasilkan output berupa probabilitas risiko hipertensi.



Gambar 11. Antarmuka Website

Website ini juga menampilkan hasil prediksi dan grafik risiko yang membedakan kategori normal dan berisiko hipertensi. Selain itu, sistem ini memberikan saran tentang cara menjaga pola makan, mengurangi konsumsi garam, berolahraga lebih banyak, dan sering melakukan pemeriksaan kesehatan. Selain itu, ringkasan data pasien ditampilkan, sehingga pengguna dapat dengan mudah memverifikasi kembali data yang dimasukkan. Sistem ini diharapkan dapat membantu pengguna, baik tenaga medis maupun masyarakat umum, lebih memahami faktor risiko hipertensi dan melakukan pencegahan secara lebih dini dengan memiliki antarmuka yang sederhana dan informatif.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem prediksi risiko hipertensi dengan memanfaatkan algoritma machine learning berbasis boosting, yaitu XGBoost, LGBM, dan CatBoost. Dengan tahapan preprocessing, penanganan data tidak seimbang menggunakan SMOTE dan ADASYN, serta pemilihan fitur dengan metode voting (RFE, Boruta, dan SelectKBest). Proses hyperparameter tuning dengan GridSearchCV dan Repeated Stratified K-Fold Cross-Validation menghasilkan peningkatan performa pada ketiga algoritma.

Evaluasi model menunjukkan bahwa ketiga algoritma boosting mampu memberikan prediksi dengan akurasi tinggi, namun CatBoost menempati posisi terbaik dengan akurasi sebesar 99,2%, precision 99,8%, recall 99,1%, F1-score 99,4%, serta nilai ROC-AUC 0,9987. Analisis SHAP juga menegaskan bahwa faktor dominan dalam prediksi hipertensi adalah indeks massa tubuh, tekanan darah, aktivitas fisik, serta faktor metabolik seperti triglycerida dan kolesterol LDL.

Implementasi sistem dalam bentuk website menggunakan model CatBoost memungkinkan pengguna untuk memasukkan data pasien secara langsung dan memperoleh hasil prediksi risiko hipertensi, disertai visualisasi risiko, ringkasan data pasien, serta rekomendasi gaya hidup. Dengan demikian, sistem ini diharapkan dapat mendukung upaya deteksi dini hipertensi serta membantu masyarakat dan tenaga medis dalam melakukan pencegahan dan pengelolaan risiko penyakit secara lebih efektif.

DAFTAR PUSTAKA

- [1] Mayo Clinic, "High blood pressure (hypertension)." Accessed: Jul. 17, 2025. [Online]. Available: <https://www.mayoclinic.org/diseases-conditions/high-blood-pressure/symptoms-causes/syc-20373410>
- [2] World Health Organization, "Hypertension." Accessed: Jul. 17, 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/hypertension>
- [3] K. Sawicka, M. Szczęska, I. Jastrzębska, M. Prasał, A. Zwolak, and J. Daniłuk, "Hypertension-The Silent Killer," 2011. [Online]. Available: www.jpcer.eu
- [4] A. Silvianah, P. Studi Ilmu Keperawatan, and S. Tinggi Ilmu Kesehatan Dian Husada Mojokerto, "HUBUNGAN KEPATUHAN MINUM OBAT HIPERTENSI DENGAN PERUBAHAN TEKANAN DARAH PADA LANSIA DI POSYANDU LANSIA," 2024.
- [5] B. Zhou *et al.*, "Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants," *The Lancet*, vol. 398, no. 10304, pp. 957–980, Sep. 2021, doi: 10.1016/S0140-6736(21)01330-1.
- [6] T. Inoue, "Unawareness and untreated hypertension: a public health problem needs to be solved," Apr. 01, 2025, *Springer Nature*. doi: 10.1038/s41440-025-02118-x.
- [7] J. S. Cho and J. H. Park, "Application of artificial intelligence in hypertension," Dec. 01, 2024, *BioMed Central Ltd*. doi: 10.1186/s40885-024-00266-9.
- [8] A. Az'zahra Tarimana, M. Ryan, S. Fajar, M. A. Saktiawan, and R. A. Saputra, "PREDIKSI PENYAKIT HIPERTENSI
- [9] G. Almuzadid and R. Subhiyakto, "Stroke Risk Classification Using the Ensemble Learning Method of XGBoost and Random Forest," 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [10] A. Hadianto and W. H. Utomo, "CatBoost Optimization Using Recursive Feature Elimination," vol. 9, pp. 169–178, 2024, doi: 10.15575/join.v9i1.1324.
- [11] W. Chang *et al.*, "Prediction of hypertension outcomes based on gain sequence forward tabu search feature selection and xgboost," *Diagnostics*, vol. 11, no. 5, May 2021, doi: 10.3390/diagnostics11050792.
- [12] J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," *J Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00369-8.
- [13] X. Ji *et al.*, "Prediction Model of Hypertension Complications Based on GBDT and LightGBM," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Feb. 2021. doi: 10.1088/1742-6596/1813/1/012008.
- [14] A. R. Rafi *et al.*, "PERBANDINGAN ALGORITMA LIGHTGBM DAN ANN UNTUK MENENTUKAN KUALITAS ANGGUR MERAH," 2025.
- [15] D. Mayang Pratiwi and L. Mufidah, *Perbandingan Metode Decision Tree Classifier dan XGBoost Classifier Dalam Memprediksi Penyakit Jantung*, vol. 4, no. 1. 2024.
- [16] A. A. Ibrahim, R. L. Ridwan, M. M. Muhammed, R. O. Abdulaziz, and G. A. Saheed, "Comparison of the CatBoost Classifier with other Machine Learning Methods," 2020. [Online]. Available: www.ijacs.thesai.org
- [17] O. T. Biala, "A COMPARATIVE STUDY OF CATBOOST AND ARTIFICIAL NEURAL NETWORKS IN ENHANCING TRIP GENERATION MODELLING FOR ILORIN CITY," *Journal of Civil Engineering, Science and Technology*, vol. 15, no. 1, pp. 18–29, Apr. 2024, doi: 10.33736/jcest.6196.2024.
- [18] R. Kurniawan *et al.*, "Hypertension prediction using machine learning algorithm among Indonesian adults," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 2, pp. 776–784, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp776-784.
- [19] F. V. Ongkosianbhara and C. C. Lestari, "Pengembangan Model Prediksi Risiko Hipertensi Menggunakan Algoritma Gradient Boosting Decision Tree Yang Dioptimalkan," *Jurnal Informatika dan Sistem Informasi*, vol. 9, no. 2, pp. 90–99, Dec. 2023, doi: 10.37715/juisi.v9i2.4403.
- [20] S. S. Chai, K. L. Goh, W. L. Cheah, Y. H. R. Chang, and G. W. Ng, "Hypertension Prediction in Adolescents Using Anthropometric Measurements: Do Machine Learning Models Perform Equally Well?," *Applied Sciences (Switzerland)*, vol. 12, no. 3, Feb. 2022, doi: 10.3390/app12031600.
- [21] T. Mroz *et al.*, "Predicting hypertension control using machine learning," *PLoS One*, vol. 19, no. 3 March, Mar. 2024, doi: 10.1371/journal.pone.0299932.
- [22] Axel Frederick Félix Jiménez and Vania Stephany Sánchez Lee, "Hipertension Arterial Mexico Data Set." Accessed: Jul. 19, 2025. [Online]. Available: <https://www.kaggle.com/datasets/frederickfelix/hipertension-arterial-mxico>
- [23] IBM, "What is exploratory data analysis (EDA)?," 2021. Accessed: Jul. 21, 2025. [Online]. Available: <https://www.ibm.com/think/topics/exploratory-data-analysis>
- [24] R. Febrian and A. Mudya Yolanda, "Comparison of Recursive Feature Elimination and Boruta as Feature Selection in Greenhouse Gas Emission Data Classification," 2024.
- [25] M. A. Tariq, "A Study on Comparative Analysis of Feature Selection Algorithms for Students Grades Prediction," *Journal of Information and Organizational Sciences*, vol. 48, no. 1, pp. 133–147, Jun. 2024, doi: 10.31341/jios.48.1.7.
- [26] M. Hasan *et al.*, "Enhancing stroke disease classification through machine learning models via a novel voting system by feature

- selection techniques,” Jan. 01, 2025, *Public Library of Science*. doi: 10.1371/journal.pone.0312914.
- [27] A. Fernández, S. García, F. Herrera, and N. V. Chawla, “SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary,” Apr. 01, 2018, *AI Access Foundation*. doi: 10.1613/jair.1.11192.
- [28] H. He, Y. Bai, E. A. Garcia, and S. Li, “ADASYN: Adaptive synthetic sampling approach for imbalanced learning,” in *Proceedings of the International Joint Conference on Neural Networks*, 2008, pp. 1322–1328. doi: 10.1109/IJCNN.2008.4633969.
- [29] A. Damayanti and A. Baita, “Comparison of Support Vector Machine (SVM) and Random Forest (RF) Algorithm Performance with Random Undersampling Technique to Predict Gestational Diabetes Mellitus Risk,” 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [30] S. Sah, B. Surendiran, R. Dhanalakshmi, and M. Yamin, “Covid-19 cases prediction using SARIMAX Model by tuning hyperparameter through grid search cross-validation approach,” *Expert Syst*, vol. 40, no. 5, Jun. 2023, doi: 10.1111/exsy.13086.