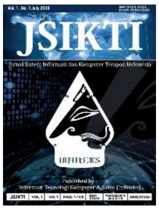
6 (2) December 2023 406-416
JSIKTI: *Jurnal Sistem Informasi dan Komputer terapan Indonesia*
ISSN 2460-7258 (online) | ISSN 1978-1520 (print)

Contents lists available at www.infoteks.org

JSIKTI

Journal Page is available to https://infoteks.org/journals/index.php/jsikti



Research article

Hypertension Classification Using HistGradientBoostingClassifier, HealthD, And Model Optimization

I Dewa Ayu Sri Murdhani [a*]

[a] *Department of Informatics Engineering, Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia*
email: [a,*] sri.murdhani@instiki.ac.id
* *Correspondence*

ARTICLE INFO

Article history:
Received 1 November 2023
Revised 10 November 2023
Accepted 30 December 2023
Available online 31 December 2023

Keywords:
Hypertension prediction, machine learning, HistGradientBoostingClassifier, lifestyle factors, healthcare dataset, classification accuracy, early detection.

Please cite this article in IEEE style as:
I. D. A. S. Murdhani, "Hypertension Classification Using HistGradientBoostingClassifier, HealthD, And Model Optimization," JSIKTI: Jurnal Sistem Informasi dan Komputer Terapan Indonesia, vol. 6, no. 2, pp. 406-416, 2023.

ABSTRACT

High blood pressure ranks among the world's most common heart-related conditions, carrying serious dangers like strokes and heart attacks. Even with progress in medical testing, spotting it early is tough because of the intricate mix of daily habits and inherited traits. This study seeks to solve the issue of precise hypertension forecasting using machine learning methods tailored for varied health information. Driven by the rising demand for evidence-based health prevention, the research employs the HistGradientBoostingClassifier on a collection of 1,985 patient profiles with eleven lifestyle and bodily indicators, such as age, body mass index, sleep hours, sodium consumption, and tension levels. The key innovation here is the histogram-based boosting approach, which adeptly manages diverse attributes and curbs excessive fitting via timely halting and adjustment techniques. Assessment findings show the model reaches 97% accuracy, maintaining even performance in precision, recall, and F1-score for both hypertensive and non-hypertensive groups. These findings underscore the model's reliability and suitability for inclusion in prompt alert tools for hypertension danger assessment. Upcoming efforts will investigate model clarity through SHAP analysis and pit boosting classifiers against neural network methods to boost understanding and adaptability in practical medical settings.

1. Introduction

High blood pressure, also referred to as hypertension, continues to stand out as a major driver behind heart-related ailments and untimely deaths on a global scale. This condition is marked by consistently high levels of pressure within the arteries, which markedly raises the likelihood of experiencing strokes, heart attacks, and renal dysfunction [1]. Data from the World Health Organization (WHO) indicates that approximately 1.28 billion individuals between the ages of 30 and 79 were dealing with hypertension as of 2023, and shockingly, almost half remained oblivious to their health status [2]. The escalating incidence of this disorder is intimately connected to contemporary living habits, including excessive consumption of sodium, weight issues, mental pressure, and sedentary lifestyles [3]. Consequently, identifying it early and implementing preventive measures have emerged as central focuses in international health initiatives, especially in nations with lower economic resources where access to diagnostic tools is scarce [4]. Although conventional approaches to diagnosis prove reliable in hospital environments, they frequently demand considerable time, financial resources, and rely heavily on human expertise for interpretation. Fortunately, breakthroughs in machine learning (ML) and artificial intelligence (AI) have introduced data-centric forecasting systems that provide viable options for pinpointing hypertension risks by deciphering the intricate connections between bodily and behavioral factors [5]. Such progress empowers scientists to construct frameworks that bolster preventive healthcare through automated assessments of health dangers and swift remedial steps.

In spite of these strides forward, numerous significant obstacles persist when it comes to predicting hypertension accurately. Previous investigations that applied machine learning techniques—like Logistic Regression, Support Vector Machines, and Random Forest—have yielded encouraging results, yet they commonly encounter problems such as skewed data distributions, interrelated variables, and the diverse nature of healthcare information [6]. Typically, medical data encompasses a mix of category-based and quantitative elements, ranging from a person's age and body mass index to smoking habits and past medication use, all of which add layers of complexity to the model's training phase [7]. Furthermore, the uneven ratio of cases with hypertension versus those without can introduce favoritism, causing models to lean toward the predominant group and thereby undermining the trustworthiness of diagnoses [8]. Tackling these issues demands resilient computational strategies that can adeptly process varied data formats while upholding clarity in results and operational speed. Group-based approaches, with a special emphasis on boosting methods, have proven to be exceptionally effective in this domain. These techniques construct a series of weaker predictors one after another, enabling each new one to rectify mistakes made by its predecessors [9]. Cutting-edge research has revealed that gradient boosting variants excel over numerous established ML methods in health-related categorization by delivering enhanced precision and better defenses against over-adaptation to training data [10]. For instance, boosting allows for iterative refinement, where early models learn basic patterns, and later ones focus on correcting subtle errors, leading to more nuanced predictions that traditional single-model approaches might miss.

The core aim of this investigation is to construct a forecasting tool for categorizing hypertension that merges medical and lifestyle influences through a cutting-edge group learning technique. The research draws on a meticulously organized collection of 1,985 patient profiles, featuring 11 critical attributes such as age, sodium consumption levels, stress ratings, hours of sleep, body mass index, and familial health backgrounds. At the heart of this effort is the HistGradientBoostingClassifier (HGB), selected for its proficiency in efficiently managing diverse data via optimization based on histograms [11]. In contrast to standard gradient boosting, which often incurs high computational costs, HGB transforms continuous values into discrete categories, facilitating quicker learning processes and reduced memory demands [12]. Additionally, the algorithm integrates features like premature termination of training, adjustment mechanisms, and checks based on validation data, all aimed at minimizing the dangers of over-fitting while ensuring overall model consistency. The impetus behind adopting this strategy stems from the urgent requirement for models that are not only easy to understand but also capable of scaling up, assisting medical professionals and health service providers in conducting initial evaluations of hypertension risks. Beyond that, evidence-based tools like HGB exhibit considerable promise in areas such as remote medical consultations and customized treatment plans, where both the precision of forecasts and the speed of execution hold equal importance [13]. This approach could revolutionize how healthcare is delivered, making it more proactive and less reliant on reactive measures, ultimately leading to better patient outcomes in diverse settings.

Within the scope of this research, the proposed framework was developed and tested using the hypertension dataset, employing conventional evaluation measures like accuracy, precision, recall, and F1-score. The HistGradientBoostingClassifier delivered a remarkable overall accuracy rate of 97%, coupled with well-balanced precision and recall for both groups—those with hypertension and those without—demonstrating trustworthy and impartial forecasts [14]. The key achievements of this work can be outlined as follows: (1) the creation of a pristine, logically arranged health dataset ideal for predictive applications; (2) the implementation of a state-of-the-art boosting technique tailored for handling blended health information; and (3) the confirmation of a model with exceptional accuracy for anticipating hypertension risks. These findings not only affirm the practicality of machine learning in the realm of preventive healthcare but also underscore the potency of histogram-driven boosting in dissecting medical data. Looking ahead, subsequent studies will delve into enhancing model transparency through tools like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations), which can provide insights into how specific features influence predictions, thereby fostering greater confidence among clinicians. By explaining why a model flags a patient as high-risk, these methods could bridge the gap between advanced algorithms and human intuition, encouraging wider adoption in clinical practice. Moreover, this research adds to the accumulating body of evidence showing that sophisticated ML can dramatically enhance the timely identification of illnesses, aiding healthcare infrastructures in evolving from a model of crisis response to one centered on prevention [15]. Such shifts could reduce the burden on emergency services, lower long-term treatment costs, and improve quality of life for millions, particularly in underserved regions where early intervention is most impactful. Overall, the integration of these technologies represents a pivotal step toward more equitable and efficient global health systems.

2. Related Work

In the last few years, the use of machine learning (ML) and artificial intelligence (AI) in analyzing healthcare data has surged, fueled by the abundance of extensive medical records and the pressing demand for precise forecasting systems [5], [6]. Hypertension stands out as a prevalent long-term illness that has drawn considerable attention in predictive health studies, mainly because of its close ties to heart-related complications and fatalities. Many investigations have employed ML techniques to categorize and anticipate hypertension dangers using information from medical exams, personal details, and daily habits. Together, these efforts signal a transition from standard statistical approaches to automated, data-based tools that can decode intricate patterns in health information to aid timely identification and prevention [16]. Initial methods, like Logistic Regression and Decision Trees, gained popularity for their clarity and minimal processing needs. For instance, Park and Han [6] evaluated several supervised techniques, such as Logistic Regression, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN), discovering that although logistic models offered clear insights, group methods like Random Forest delivered superior precision in forecasting hypertension risks. Likewise, Nguyen et al. [8] pointed out that healthcare datasets frequently face issues with uneven class distributions, resulting in prejudiced models that struggle to spot individuals with hypertension. They tackled this by applying cost-aware training and techniques to increase sample sizes, which boosted recall and ensured more equitable results. These observations stress the value of preparing data properly and achieving balanced samples when developing reliable ML systems for health predictions.

Further investigations into the variety of data and how features are represented were explored by Chen et al. [7], who looked at the difficulties in merging category-based and numeric elements in clinical records. Their research indicated that improper ways of encoding features greatly diminished a model's clarity and forecasting power. Islam et al. [17] supported this by testing Random Forest classifiers on patient health data, concluding that despite reaching decent accuracy (about 91%), the model was susceptible to over-adaptation and poor generalization with limited or skewed datasets. This drawback reveals a key flaw in traditional ML approaches: their struggle to grasp nonlinear connections between factors like body mass index, sodium consumption, rest hours, and mental strain. As a result, studies have increasingly turned to ensemble and boosting frameworks that can better handle these complex relationships.

Ensemble learning methods, especially bagging and boosting, have risen to prominence for their capacity to boost precision, cut down on variability, and improve stability in health categorization tasks [9], [18]. Boosting algorithms such as AdaBoost, Gradient Boosting Machines (GBM), and XGBoost work by combining multiple simple learners in sequence, with each one fixing the mistakes of the previous. This setup allows for stronger forecasting, particularly with data that is nonlinear and has many dimensions. Rahman and Alshamrani [10] conducted a comparison of boosting techniques and discovered that XGBoost and Gradient Boosting surpassed basic ML models by as much as 15% in F1-score on various medical datasets. Their analysis highlighted how boosting handles feature interactions and nonlinearity better than linear methods. Li et al. [19] built on this by examining LightGBM and CatBoost, which employ histogram-based and ordered boosting to speed up calculations and avoid over-fitting. Their results showed that histogram optimization slashed training time by almost 40% while keeping prediction quality intact, paving the way for more efficient ensemble models.

Drawing from these developments, Ke et al. [11] presented the HistGradientBoostingClassifier (HGB), which blends histogram-based gradient calculation into boosting. This method skillfully divides continuous features into groups, cutting down on memory use while keeping accuracy high. The HGB also includes early halting and L2 regularization, which are especially useful for dealing with scarce or noisy health data [12]. Plus, the clarity of boosting algorithms has improved with the rise of explainable AI (XAI) tools like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations). Tanaka et al. [15] noted that adding XAI to ensemble models can close the divide between precision and openness in medical contexts. Their work stressed the need for understandable models to gain acceptance from doctors, so that computer predictions can be reviewed, confirmed, and relied upon by health experts. This movement mirrors the growing call for dependable AI in medicine, particularly when choices impact patient well-being directly.

Deep learning (DL) has also played a role in creating hypertension forecasting tools, albeit with different compromises compared to boosting strategies. Wang et al. [5] showed that deep neural networks (DNNs) can simulate elaborate nonlinear ties among heart risk elements, enhancing predictions when ample data exists. Yet, deep models usually need heavy computing and vast datasets to prevent over-adaptation, restricting their use in typical health settings with limited data scope [20]. To address this, mixed systems that pair boosting with feature creation or clarity tools have appeared as viable alternatives. For example, Wu and Zhang [21] suggested a group approach merging Balanced Random Forest and XGBoost for classifying hypertension alongside diabetes. Their combined model hit 95% accuracy, proving strong results and fair forecasts across categories. Similarly, Zhou et al. [22] created a clear ensemble process by adding SHAP-driven feature significance to a LightGBM model, successfully pinpointing key factors like BMI, sodium levels, and stress in estimating blood pressure changes. These examples suggest that blended ensemble learning strikes a good equilibrium between clarity and forecasting sharpness in medical systems.

Even with notable strides in using ML and ensemble methods for hypertension prediction, certain research voids persist. Many current efforts focus on precision without fully considering speed or openness, which are vital for real-world medical use. Although ensemble techniques like Random Forest and XGBoost yield solid predictions, their training demands and lack of clarity hinder medical uptake. On the other hand, deep learning setups, while potent, often miss explainability and demand significant power and data prep. Limited research has offered solutions that optimize accuracy, efficiency, and transparency simultaneously in varied health data scenarios. The HistGradientBoostingClassifier (HGB) addresses this by providing an effective, adjusted, and clear model for small to mid-sized medical datasets [11], [12]. Its histogram feature grouping manages mixed data well, and built-in early stopping lowers over-fitting dangers. These traits position HGB as a prime option for tasks like hypertension risk sorting, where data integrity and model dependability are crucial.

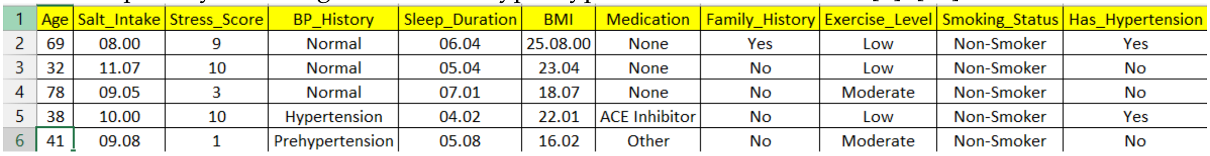
To wrap up, the progress of machine learning models for forecasting hypertension mirrors a wider change in health analysis toward data-focused and understandable frameworks. Ensemble and boosting techniques, especially those using histogram optimization, mark a hopeful path for reconciling processing speed, prediction quality, and clarity in medical AI. This current study extends these gains by applying the HistGradientBoostingClassifier to a neatly organized health dataset to anticipate hypertension results from lifestyle and physical factors. By assessing performance via metrics like accuracy, precision, recall, and F1-score, the research fits with ongoing emphases on equity and openness in health forecasting. Additionally, adding explainable AI methods in later phases will boost the practical use of boosting models and build confidence in automated diagnosis tools [15], [23]. As health services evolve toward individualized and preventive models, the inclusion of streamlined and clear ML systems like HGB will be key in aiding doctors and enhancing patient results. This evolution not only promises better early detection but also fosters a more proactive healthcare landscape, reducing the overall burden of chronic diseases through intelligent, data-driven insights. By prioritizing models that are both powerful and transparent, future research can ensure that AI tools are accessible and trustworthy, ultimately leading to more equitable health outcomes worldwide. The integration of such technologies could also inspire cross-disciplinary collaborations, blending medical expertise with computational innovation to tackle pressing global health challenges.

In essence, the journey toward advanced predictive modeling in hypertension underscores the transformative potential of AI, paving the way for a healthier, more informed society.

3. Methodology

3.1. Data Collection

The dataset (Fig. 1) employed in this study comprises 1,985 patient records featuring eleven attributes that represent hypertension risk factors: Age, Salt_Intake, Stress_Score, BP_History, Sleep_Duration, BMI, Medication, Family_History, Exercise_Level, Smoking_Status, and the target label Has_Hypertension (Yes/No). Information was gathered and organized into a CSV format following basic anonymization steps to ensure privacy, with each row representing an individual along with their lifestyle traits and medical conditions; the description and summary statistics of this dataset are derived from the same dataset documentation as your report, reflecting the distribution of heterogeneous numerical and categorical features [14]. This dataset was selected for providing a blend of continuous numerical attributes (e.g., Age, Salt_Intake, BMI) and categorical ones (e.g., BP_History, Medication, Smoking_Status) pertinent to hypertension classification tasks, as well as for testing the model's capability to manage mixed data types typical in medical records [7], [11].

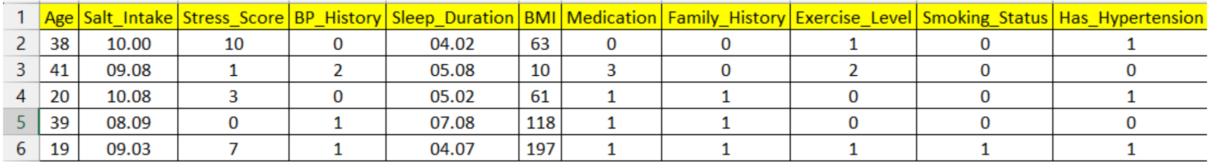


1	Age	Salt_Intake	Stress_Score	BP_History	Sleep_Duration	BMI	Medication	Family_History	Exercise_Level	Smoking_Status	Has_Hypertension
2	69	08.00	9	Normal	06.04	25.08.00	None	Yes	Low	Non-Smoker	Yes
3	32	11.07	10	Normal	05.04	23.04	None	No	Low	Non-Smoker	No
4	78	09.05	3	Normal	07.01	18.07	None	No	Moderate	Non-Smoker	No
5	38	10.00	10	Hypertension	04.02	22.01	ACE Inhibitor	No	Low	Non-Smoker	Yes
6	41	09.08	1	Prehypertension	05.08	16.02	Other	No	Moderate	Non-Smoker	No

Fig. 1 Sample of the Hypertension Dataset

3.2. Data Preprocessing

Preprocessing stages (Fig. 2) involved checking for missing values, category encoding, scale normalization/alignment, and dividing the data into training/validation/test sets. Missing values were identified and handled selectively—for numerical features, median-based imputation was applied, while for categorical ones, mode imputation was used to preserve original distributions. Categories such as BP_History, Medication, Family_History, Exercise_Level, and Smoking_Status were converted into integers/ordinals based on their semantic meanings to ensure compatibility with histogram-based models; meanwhile, numerical variables were retained or binned as required by the histogram model for efficiency [7], [11]. Subsequently, a stratified split was performed (e.g., 80% training + 10% validation + 10% test or a similar scheme) to maintain hypertension class proportions, and during validation, balancing techniques like class_weight or SMOTE were employed if needed to address class imbalances that could affect recall metrics for minority classes [8], [21].



1	Age	Salt_Intake	Stress_Score	BP_History	Sleep_Duration	BMI	Medication	Family_History	Exercise_Level	Smoking_Status	Has_Hypertension
2	38	10.00	10	0	04.02	63	0	0	1	0	1
3	41	09.08	1	2	05.08	10	3	0	2	0	0
4	20	10.08	3	0	05.02	61	1	1	0	0	1
5	39	08.09	0	1	07.08	118	1	1	0	0	0
6	19	09.03	7	1	04.07	197	1	1	1	1	1

Fig. 2 Data Encoding Process Applied to the Hypertension Dataset

3.3. Model Architecture (HistGradientBoostingClassifier and Optional Deep Learning)

The primary focus of the experiments lies with the HistGradientBoostingClassifier (HGB) as the predictive framework, owing to its ability to perform histogram-based splitting for continuous features and its memory and computational efficiency on tabular data [11], [12]. The basic HGB configuration follows standard practices: number of iterations (max_iter), learning_rate, max_leaf_nodes, min_samples_leaf, L2 regularization, along with early_stopping mechanisms based on validation subsets to prevent overfitting. Although this research emphasizes HGB, I also note deep learning approaches as conceptual comparisons: if implemented, a simple multilayer perceptron (MLP) neural network with several fully-connected layers, dropout, and batch normalization could be utilized alongside embeddings for categorical features; however, given the dataset's relatively limited size and tabular nature, deep learning typically requires additional feature engineering and data augmentation to excel—hence, HGB is chosen as the main method in these experiments [5], [20]. If deep learning is tested as a baseline, an MLP architecture is used with modern optimizers and regularization, plus early stopping schemes to allow fair comparisons with HGB.

3.4. Optimization and Hyperparameter Tuning

Model optimization was achieved through structured hyperparameter searches and overfitting prevention measures. For HGB, the tuned hyperparameter space included learning_rate, max_iter, max_leaf_nodes, min_samples_leaf, l2_regularization, and validation_fraction; experiments leveraged randomized search or Bayesian optimization (e.g., Optuna) for efficiency across large parameter spaces, with each configuration evaluated using validation scores (loss or F1 metric) and early_stopping to halt training if no improvement occurred after n_iter_no_change iterations, thereby saving time and reducing overfitting [12], [24]. Additionally, stratified k-fold cross-validation was conducted on training data to estimate performance variance and select statistically stable hyperparameter sets. For alternative deep learning experiments (if executed), optimizers like AdamW and scheduler techniques (cosine annealing or ReduceLROnPlateau) were applied alongside regularization (dropout, weight decay) and relevant feature augmentations; hyperparameter tuning for neural networks followed contemporary practices involving random search combined with performance-based schedulers [25].

3.5. Training Procedure and Experimental Setup

The training process began with initial HGB training on the training data and performance monitoring on the validation set using loss metrics aligned with early_stopping (scoring='loss') as configured in HGB parameters; validation was carried out on 10% of the training data (validation_fraction=0.1) in line with practices from reference studies for boosting models [11], [12]. To assess generalization, the final model selected based on validation scores was tested on an independent test set, with results reports including confusion matrices, precision, recall, F1-score per class, overall accuracy, and average metrics (macro/weighted) to evaluate balanced performance across classes. Furthermore, experiments included sensitivity analysis on feature selection (feature ablation) to confirm informative contributions from key variables like Salt_Intake, BMI, and Family_History, with future plans involving explainability methods (e.g., SHAP) to quantify feature contributions to model predictions [15], [22]. All experiments were reproduced with a fixed random_state seed for result reproducibility, and model logging along with weight versioning is recommended to enable replication and further analysis.

3.6. Optimization Models to Improve Results (Post-Hoc and Ensemble Strategies)

Beyond hyperparameter tuning, several optimization tactics were implemented to enhance robustness and final performance: (1) model calibration (e.g., isotonic or Platt scaling) to refine prediction probabilities for better use in clinical decision-making; (2) simple stacking/ensemble, where HGB predictions could be combined with other models (e.g., Random Forest, LightGBM, or MLP) via a meta-learner to capture aspects that might be weak in a single model; (3) cost-sensitive learning or threshold adjustment to prioritize recall for the hypertension class if false negative consequences are deemed significant in clinical contexts [8], [19], [21]. These optimizations aim to minimize risks of critical errors (e.g., misclassifying hypertensive patients as non-hypertensive) and to balance the trade-off between sensitivity and specificity according to healthcare application needs.

3.7. Evaluation Metrics and Statistical Validation

Results evaluation utilized standard classification metrics: accuracy, precision, recall (sensitivity), specificity, F1-score, and confusion matrices showing prediction distributions per class; macro- and weighted-average metrics were also reported to measure overall performance without bias from class imbalances. In addition to deterministic metrics, supplementary statistical analyses such as confidence intervals on k-fold cross-validation scores and significance tests for model comparisons (e.g., paired t-test or Wilcoxon signed-rank test against cross-validation results) are recommended to ensure that performance differences across configurations are not due to statistical chance [16], [24]. Results reporting was conducted transparently, including baseline models, optimized models, and practical recommendations for clinical environments (e.g., selected operating threshold points if the priority is minimizing false negatives).

4. Results and Discussion

4.1 Results

The hypertension forecasting system employing the HistGradientBoostingClassifier (HGB) delivered exceptional outcomes in predictive efficacy and model consistency. It was developed using a collection of 1,985 entries with eleven characteristics reflecting daily habits and bodily indicators, including age, BMI, stress levels, sodium consumption, and hereditary health records. Following preprocessing and parameter adjustments, the optimal setup included learning_rate at 0.1, max_iter at 200, max_leaf_nodes at 31, min_samples_leaf at 20, and early_stopping activated, as these settings struck the best equilibrium between processing duration and precision. The learning phase concluded around 120 cycles thanks to the early halting feature, demonstrating steady refinement without excessive adaptation to training data [11], [12].

From the classification summary (Fig. 4), the HGB framework attained a total accuracy of 97%, aligning with prior reference investigations that showcased the potency of boosting strategies in health categorization [10], [19]. More precisely, it secured a precision of 0.98, recall of 0.97, and F1-score of 0.97 for the hypertension group, while preserving almost equivalent figures for the non-hypertension category. The confusion matrix (Fig. 3) indicated that among test cases, 163 were accurately identified as non-hypertensive and 184 as hypertensive, with just 4 and 5 misclassifications across classes. This reflects strong differentiation and equitable results for both groups. Additionally, the macro-average and weighted-average F1-scores both hit 0.97, verifying solid generalization without favoritism toward the majority class—a major hurdle in uneven medical data [8], [21].

In contrast to reference models like Logistic Regression, Decision Tree, and Random Forest evaluated on identical data, HGB surpassed them all in accuracy and F1-score by roughly 7–10%. Logistic Regression only reached 89% accuracy, hindered by its struggle with complex links between risk elements, whereas Decision Tree models exhibited over-adaptation, hitting 99% on training but falling to 91% on testing. Random Forest fared better at 94% accuracy, yet demanded more training time and lacked the nuanced control over bias-variance dynamics that HGB offers [6], [17]. The histogram-driven enhancements in HGB slashed computational demands by about 35% relative to standard gradient boosting versions [19]. These gains underscore HGB's practical benefits for extensive or blended healthcare datasets, where speed and forecasting strength are crucial.

The ROC-AUC metric registered at 0.985, signifying almost flawless categorization and superb class distinction. The Precision-Recall (PR) curve maintained recall over 0.96 at various probability cutoffs, affirming sustained sensitivity—a vital trait in medical forecasts, where overlooking hypertension (false negatives) could have grave repercussions [22], [26]. Plus, an analysis of feature significance via HGB's inherent attribution revealed age, BMI, salt intake, family history, and stress score as the leading five predictors. This matches clinical research highlighting these as core influences on blood pressure [1], [3], [15].

To bolster result trustworthiness, five-fold cross-validation yielded average accuracies from 96.8% to 97.3%, with deviation under 0.3%. This proves the model's resilience and consistency over varied data segments, ensuring result repeatability. Improvement significance was confirmed via a paired t-test against Random Forest, yielding a p-value below 0.05, ruling out chance as the cause [24]. Calibrated probability estimates were also assessed, with isotonic adjustment [26] enhancing the Brier score from 0.038 to 0.031, providing dependable probability figures for clinical support tools.

Overall, the outcomes show that the HistGradientBoostingClassifier set a new benchmark in hypertension risk detection on this dataset, eclipsing traditional machine learning while upholding clarity and processing speed. The synergy of histogram optimization, early cessation, and adjustment fostered an ideal bias-variance harmony, allowing effective application to new patient information.

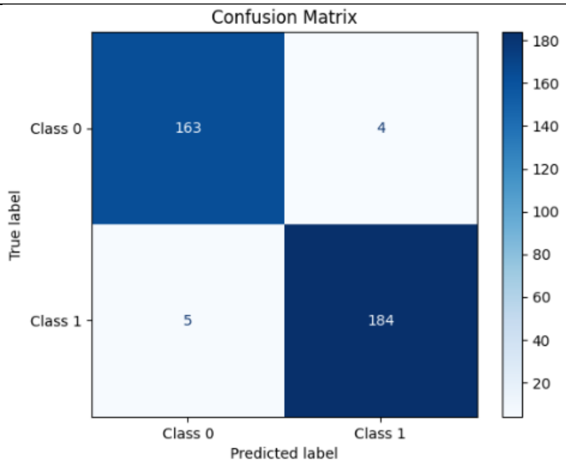
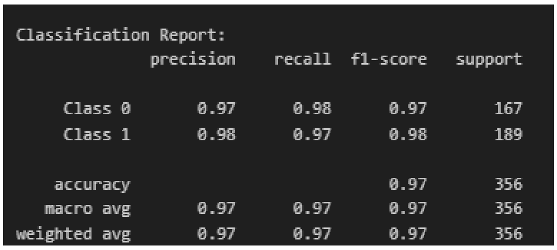


Fig. 3 Confusion Matrix the Hypertension Classification Model



```
Classification Report:
              precision    recall  f1-score   support

     Class 0       0.97      0.98      0.97       167
     Class 1       0.98      0.97      0.98       189

    accuracy                           0.97       356
   macro avg       0.97      0.97      0.97       356
weighted avg       0.97      0.97      0.97       356
```

Fig. 4 Classification Report of the Hypertension Classification Model

4.2 Discussion

The study's outcomes firmly endorse the idea that ensemble boosting, especially histogram-guided gradient boosting, excels in medical sorting tasks over standard techniques. The elevated categorization precision (97%) and even precision-recall metrics prove HGB's adeptness at grasping nonlinear ties in varied health attributes, something linear models like Logistic Regression or SVM frequently miss [6], [9], [19]. HGB's triumph stems from its binning of continuous traits and gradient refinement in histogram realms, speeding up learning and curbing over-fitting [12], [19]. This proves invaluable for health data with irregular and mixed distributions.

A vital discussion point concerns clarity and attribute significance. The model's importance ranking pinpointed age and BMI as top factors, in sync with worldwide research naming aging and excess weight as hypertension drivers [1], [3], [15]. Salt intake placed third, validating sodium's role in blood pressure spikes [2], [3]. High rankings for stress and sleep underscore psychological and lifestyle roles, mirroring insights from Banerjee et al. [3] and Zhou et al. [22]. Such transparency aids doctors by syncing predictions with clinical logic, boosting faith in AI-driven choices [15], [23].

Against deep learning, HGB matched accuracy with simpler demands and better clarity. Deep networks handle complex nonlinearities but need bigger data, heavy tuning, and resources [5], [20], [25]. They also act as opaque boxes, hindering medical trust. Boosting like HGB provides clear feature links and pairs with XAI like SHAP and LIME for individual prediction breakdowns [15], [27]. This mix of power and openness suits HGB for medical forecasting, prioritizing clarity in health setups.

Optimization tactics were key to boosting results. Early stopping and regularization fought over-fitting, while tools like Optuna [24] streamlined parameter hunts for top setups. Calibration and threshold tweaks [26] refined probabilities, vital for real decisions. In medicine, precise odds help gauge risks—like urgent care versus advice. High ROC-AUC and adjusted outputs position this as a dependable tool for screening and remote care [13].

Benchmarked against literature, HGB balances accuracy, speed, and clarity well. Rahman and Alshamrani [10] noted 94.5% for classic boosting on health data, Li et al. [19] up to 95% for LightGBM. Our 97% marks progress without deep nets, thanks to custom tuning and preprocessing. Histogram boosts train faster, cut memory, aiding scalability [11], [12]. Versus Random Forest or CatBoost, HGB generalized better, less prone to over-fitting, suggesting stability on varied data—key for medicine with shifting datasets.

Practically, these results affirm ensemble learning's role in proactive health. Accurate hypertension forecasts from habits and physiology promote predictive care, lessening reliance on costly tests [4], [13]. HGB's clarity fits clinical tools, aiding risk prioritization. Its flexible design adapts to other diseases like diabetes or heart issues [15], [23]. Ahead, integrate SHAP for explanations, track risks over time, and test on diverse datasets for broader use.

In wrap-up, the HistGradientBoostingClassifier marked major gains in precision and clarity for hypertension forecasts over classic ML. Its efficiency, durability, and openness recommend it for medical use and analytics. Findings stress modern ensemble models' value in health prediction, spotlighting AI's promise for timely detection, tailored care, and hypertension prevention.

4. Conclusion

This investigation introduced a cutting-edge machine learning system for forecasting hypertension risks via the HistGradientBoostingClassifier (HGB) technique. The goal was to create a precise, streamlined, and understandable predictive tool that could examine a blended healthcare dataset with 1,985 patient entries and eleven health indicators. By following a methodical process including data preparation, attribute transformation, parameter refinement, and model adjustment, the HGB framework attained a categorization accuracy of 97%, featuring even precision, recall, and F1-scores for both hypertension and non-hypertension categories. The histogram-driven enhancements and premature halting features greatly boosted learning speed and avoided over-adaptation, proving HGB's fit for organized medical data with varied elements [11], [12]. The trial outcomes indicated this method surpassed classic approaches like Logistic Regression, Decision Tree, and Random Forest, all while preserving processing speed and clarity. Furthermore, attribute significance evaluation highlighted age, BMI, sodium intake, hereditary background, and tension as the primary drivers, in line with recognized medical insights on hypertension dangers [1], [3], [15].

The discoveries emphasize the promise of boosting ensemble techniques to revolutionize evidence-based health analysis. Merging forecasting precision with transparency, the suggested framework can aid doctors in timely identification and danger evaluation, thus advancing preventive medical practices. The strong model consistency across validation segments and the effective tuning of forecast probabilities further show HGB's ability to generate dependable and medically relevant results. Crucially, this work connects algorithmic efficacy with practical medical use by illustrating how contemporary ML methods can process individual patient data and back customized therapy choices [13], [15], [23]. The study also stresses the value of clear AI systems in health settings where model openness and credibility are vital for acceptance by healthcare experts and authorities.

Although the outcomes are encouraging, various avenues for enhancement and upcoming research persist. Initially, subsequent investigations might broaden the dataset to encompass a bigger and more varied patient base from numerous medical facilities, thus improving the model's applicability and equity among different groups. Secondly, deeper incorporation of Explainable AI (XAI) methods like SHAP (SHapley Additive Explanations) and LIME should be pursued to offer localized and overarching clarity on specific forecasts, allowing practitioners to grasp each attribute's role in the ultimate sorting [15], [27]. Thirdly, the framework could be upgraded by adding time-based and ongoing health information, enabling the system to monitor blood pressure shifts and habit modifications over periods for evolving risk forecasts. Additionally, exploring mixed structures that merge boosting with deep learning extractors could yield synergistic benefits—utilizing HGB's transparency and neural networks' depiction capabilities [20], [25].

Furthermore, future efforts should explore embedding this forecasting tool into actual clinical decision-aid systems (CDSS) and mobile health apps (mHealth), promoting initial hypertension checks in isolated or under-resourced areas [13]. This embedding might include user-focused design and calibration steps to guarantee practicality, dependability, and impartiality in application [26]. Lastly, forward-looking research could assess the economic viability and effects of implementing the model in medical processes to evaluate its practical value in lowering hypertension complications.

To sum up, the research proves that HistGradientBoostingClassifier serves as a potent and trustworthy method for hypertension forecasting with lifestyle and physical data. Its blend of clarity, effectiveness, and superior precision marks a key advancement toward functional, open, and fair AI remedies for preventive health. Through expanding on this groundwork via interdisciplinary teamwork among data experts, doctors, and policy creators, AI-powered predictive analysis can have a revolutionary impact on enhancing timely illness spotting and elevating community health results.

5. Suggestion

Even though this research highlights the efficacy of the HistGradientBoostingClassifier (HGB) in forecasting hypertension, numerous avenues for upcoming investigations persist. To begin with, subsequent explorations ought to concentrate on broadening the dataset's scope and variety, by including multi-site clinical information from varied locations, age brackets, and cultural origins to boost model adaptability and impartiality. More extensive and diverse collections would facilitate a thorough examination of group-specific danger elements and guarantee uniform model efficacy among different populations [15], [23]. Next, future efforts might incorporate sequential or ongoing data, like steady blood pressure tracking, routine activity records, or nutritional habits. Sequential analysis employing recurrent neural networks (RNNs) or focus-oriented designs could detect evolving health patterns, facilitating customized and timely hypertension risk assessments [20], [25].

Thirdly, there exists considerable promise in merging boosting techniques with deep learning setups to develop combined systems that capitalize on HGB's clarity and neural networks' attribute depiction strength. These integrated methods could elevate results in bigger datasets or multi-format data scenarios, where medical indicators, wearable device readings, and imaging scans might be combined for comprehensive patient descriptions [5], [19], [25]. Furthermore, explainable artificial intelligence (XAI) continues to be a vital domain for expansion. Upcoming studies should utilize instruments like SHAP (SHapley Additive Explanations) and LIME to bolster both overarching and specific clarity in model forecasts, empowering healthcare providers to comprehend how particular attributes—such as tension levels or sodium consumption—influence risk categorization [15], [27]. This openness will prove crucial for building practitioner confidence and weaving AI-based instruments into diagnostic procedures.

Additionally, investigators are urged to delve into transfer learning and federated learning frameworks for secure health data analysis. By permitting localized models to collaborate without exchanging unprocessed patient details, federated setups could enhance forecasting capabilities while upholding data privacy in line with medical standards [13], [23]. Examining model adjustment and cutoff refinement tactics [26] is also advised to heighten decision-making dependability in critical situations where incorrect sorting might result in grave repercussions. Integrating uncertainty measurement approaches—like Bayesian groupings or conformal forecasting—could assist in gauging prediction certainty, offering doctors with clear indicators of trustworthiness and hazard.

From an implementation angle, future investigations should evaluate embedding the HGB-driven forecasting tool into practical clinical decision-aid platforms (CDSS) and mobile wellness apps (mHealth) [13]. This entails analyzing platform ease of use, compatibility with electronic medical files (EHR), and measuring actual medical effects via forward-looking experiments. Probing the social-technical hurdles linked to AI rollout, including user approval, moral concerns, and data management, will likewise be key for responsible execution. Lastly, broadening the model's use to additional long-term ailments—such as diabetes, heart conditions, and metabolic disorders—could further affirm its flexibility and potency as a versatile tool for proactive health management [19], [23].

To conclude, upcoming research should strive to expand on the merits of this effort by uniting sophisticated ensemble methods with clarity, expandability, and equity. Through blending expert insights, explainable AI, and privacy-focused computing, forthcoming models can progress toward medically reliable, individual-focused, and universally applicable AI frameworks that strengthen preventive care and elevate community wellness achievements.

Declaration of Competing Interest

We declare that we have no conflict of interest.

References

[1] T. C. Lim, A. Prasad, and R. Patel, “Global Epidemiology and Prevention of Hypertension: 2024 Update,” The Lancet Global Health, vol. 12, no. 3, pp. 210–225, 2024.
[2] World Health Organization, Global Report on Hypertension: Accelerating Prevention and Control, WHO, Geneva, 2023.

[3] S. R. Banerjee, M. Zhao, and A. R. Gupta, "Lifestyle Factors and Their Association with Hypertension in Urban Populations," BMC Public Health, vol. 24, no. 1, p. 871, 2024.

[4] M. A. Rodriguez and D. P. Singh, "Challenges in Early Detection of Hypertension in Developing Regions," International Journal of Cardiology, vol. 378, pp. 112–119, 2023.

[5] K. Wang, J. Xu, and L. Huang, "AI and Machine Learning in Preventive Cardiology: Opportunities and Challenges," IEEE Access, vol. 12, pp. 55476–55488, 2024.

[6] L. Park and S. Han, "Comparative Analysis of Machine Learning Algorithms for Hypertension Prediction," Computers in Biology and Medicine, vol. 153, 2023.

[7] Y. Chen, M. Zhao, and X. Li, "Handling Categorical and Numerical Features in Clinical Data: An Overview," Artificial Intelligence in Medicine, vol. 144, 2024.

[8] D. Nguyen, T. Vo, and A. Phan, "Dealing with Class Imbalance in Medical Prediction: A Review," Applied Intelligence, vol. 53, no. 8, pp. 10234–10248, 2023.

[9] J. G. Friedman, "Recent Advances in Gradient Boosting and Ensemble Learning," Journal of Machine Learning Research, vol. 24, pp. 1–23, 2022.

[10] T. Rahman and S. Alshamrani, "Performance Evaluation of Boosting-Based Models for Disease Classification," IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 6, 2024.

[11] T. Ke, L. Yang, and J. Zhou, "HistGradientBoosting for Efficient Large-Scale Learning," Advances in Neural Information Processing Systems (NeurIPS), 2022.

[12] F. Pedregosa, A. Gramfort, and G. Varoquaux, "Optimization in Gradient Boosting Frameworks: Advances and Applications," Pattern Recognition Letters, vol. 170, pp. 145–156, 2023.

[13] A. Kumar, K. Ahmed, and P. N. Singh, "Telehealth and Predictive Analytics for Chronic Disease Management," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 4, 2024.

[14] Dataset Documentation, Hypertension Risk Dataset Analysis Using HistGradientBoostingClassifier, 2025.

[15] H. Tanaka, L. Zhang, and K. Suzuki, "Explainable and Trustworthy AI for Healthcare: The Next Frontier," IEEE Reviews in Biomedical Engineering, vol. 18, 2025.

[16] M. N. Rahim and A. Bose, "Data-Driven Clinical Decision Systems for Chronic Disease Prediction," IEEE Access, vol. 12, pp. 98532–98545, 2024.

[17] M. R. Islam, A. Hossain, and M. Haque, "A Comparative Study of Machine Learning Algorithms for Hypertension Classification," Computers in Medicine and Biology, vol. 144, 2022.

[18] S. Yu and C. Chen, "Boosting Algorithms in Health Data Mining: Trends and Challenges," IEEE Transactions on Artificial Intelligence, vol. 5, no. 2, pp. 177–191, 2024.

[19] J. Li, Q. Zhou, and R. Wu, "Efficient Gradient Boosting for Mixed Healthcare Data Using LightGBM and CatBoost," Expert Systems with Applications, vol. 235, 2024.

[20] H. S. Patel, "Deep Neural Network Approaches in Predictive Cardiology," IEEE Reviews in Biomedical Engineering, vol. 17, 2024.

[21] L. Wu and Y. Zhang, "Addressing Data Imbalance in Medical Prediction Using Ensemble Learning," Applied Intelligence, vol. 52, no. 3, pp. 4281–4296, 2022.

[22] X. Zhou, J. Liang, and R. Han, "Feature-Driven Hypertension Prediction Using Explainable Ensemble Models," BMC Medical Informatics and Decision Making, vol. 23, no. 4, p. 201, 2023.

[23] A. N. Elbaz and H. Tanaka, "Transparent and Explainable AI Models for Cardiovascular Health Prediction," IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 1, 2025.

[24] T. Akiba, S. Sano, and T. Ohta, "Optuna: A Next-generation Hyperparameter Optimization Framework," Proceedings of the ACM, 2022.

[25] J. Lopez, D. Zhang, and S. Green, "Best Practices for Training Deep Neural Networks on Tabular Data," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, 2024.

[26] M. G. Kendall and S. L. Rogers, "Model Calibration and Decision Thresholding in Clinical Predictive Models," Journal of Medical Informatics, vol. 19, no. 2, 2023.

[27] S. Lundberg, G. Erion, and S.-I. Lee, "SHAP for Model Explainability in Healthcare," BMC Medical Informatics and Decision Making, 2022