Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Explainable Machine Learning-Based Decision Tree Model for Early Detection of Hypertension Risk

Hilda Ayu Sofiani[1)], Isa Iant Maulana[2)], Farrikh Alzami[3)*], Muhammad Naufal[4)], Harun Al Azies[5)], Ifan Rizqa[6)], Dewi Agustini Santoso[7)], Siti Hadiati Nugraini[8)]
[1,2,3,4,5,6,7,8)]Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang, Indonesia
[1)]111202214833@mhs.dinus.ac.id, [2)]111202214416@mhs.dinus.ac.id, [3)]alzami@dsn.dinus.ac.id, [4)]m.naufal@dsn.dinus.ac.id,[5)]harun.alazies@dsn.dinus.ac.id, [6)]risqa.ifan@dsn.dinus.ac.id, [7)]dewi@dsn.dinus.ac.id, [8)]shnugraini@dsn.dinus.ac.id

Submitted :Nov 15, 2025 | **Accepted** : Dec 2, 2025 | **Published** : Jan 1, 2026

Abstract: Hypertension is one of the leading causes of cardiovascular disease and is often referred to as a "silent killer" because it typically remains asymptomatic until serious complications, such as stroke or kidney failure, occur. Early detection of hypertension risk is therefore essential to enable timely intervention and prevention. This study aims to develop an explainable machine learning–based Decision Tree model for early detection of hypertension risk using clinical and lifestyle data. The balanced dataset includes variables such as age, body mass index (BMI), blood pressure, family history, smoking habits, stress levels, and sleep duration. The dataset used in this study was obtained from the "Hypertension Risk Prediction Dataset" available on the Kaggle platform, consisting of 1,985 patient records and 11 main features covering variables such as age, body mass index (BMI), systolic and diastolic blood pressure, family history, smoking habits, stress level, physical activity, and sleep duration. The dataset is balanced between the hypertension and normal categories, enhancing the reliability of the classification results. The model was constructed using a Decision Tree Classifier implemented in Scikit-learn and validated through cross-validation to minimize overfitting. Model performance was assessed using accuracy, precision, recall, F1-score, and AUC-ROC metrics. The results indicate that the model achieved an accuracy of 96% and an AUC of 0.9645, demonstrating excellent classification performance. The motivation behind this research lies in the growing need for interpretable artificial intelligence models in healthcare, where transparency and explainability are critical for clinical trust and ethical decision-making. Unlike black-box models, the Decision Tree approach allows clinicians to trace each prediction path, understand contributing variables, and apply insights in real-world medical settings. The primary advantage of this model lies in its transparency, as each prediction can be interpreted through explicit decision rules. Overall, this explainable and high-performing model shows strong potential as a clinical decision support tool for early hypertension screening and prevention programs.

Keywords: Hypertension, Decision Tree, Early Detection, Machine Learning, Digital Health

INTRODUCTION

Hypertension is one of the most significant global health problems and is often referred to as the silent killer because its symptoms are often unnoticed until it leads to serious complications such as stroke, heart disease, or kidney failure. According to the World Health Organization (WHO, 2022), over 1.2 billion people worldwide suffer from hypertension, and approximately two-thirds of them live in low- to middle-income countries. This condition indicates that hypertension is not just an individual medical problem, but also a public health issue that requires systematic early detection and preventive intervention (Afolabi et al., 2025).

The development of machine learning technology in healthcare opens up significant opportunities to improve the accuracy of predicting chronic disease risks, including hypertension (Karima & Anggraeni, 2024). Machine

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

learning algorithms can recognize complex patterns in both clinical and lifestyle data, which can assist medical professionals in the diagnosis process and data-driven decision-making. However, one of the main challenges in applying machine learning models in the medical field is the limitation in the aspect of explainability, which refers to the extent to which the model can explain the reasons behind the decisions or predictions it makes (Kurniawan et al., 2023).

The Decision Tree model is one of the prominent algorithms due to its high interpretability and intuitive visualization capabilities (Hendrawan et al., 2024). Each decision in a Decision Tree is based on simple if-then rules, making it easy for non-technical users like medical personnel or patients to understand (W Ahmad, 2023). Additionally, this model also provides a balance between accuracy and transparency, making it suitable for clinical applications where decision clarity is crucial (Purwono et al., 2022).

Furthermore, it is important to emphasize that although preprocessing techniques such as StandardScaler are commonly applied in some studies to normalize the range of numerical feature values, the Decision Tree model conceptually does not require such scaling. This is because a Decision Tree splits data based on threshold values of features, meaning that changes in scale do not affect the structure of these splits. Therefore, this study also evaluated the use of StandardScaler as a comparison and confirmed that the Decision Tree model performs optimally even without scaling, in accordance with its natural characteristic of being insensitive to differences in feature scales.

Previous research by Karima and Anggraeni (Karima & Anggraeni, 2024) showed that applying the AdaBoost algorithm to clinical hypertension data resulted in high classification performance with an AUC of 0.96 after feature selection using ReliefF. However, the study still uses an ensemble model that is complex and less directly explainable. Therefore, this research focuses on developing an Explainable Decision Tree Model that can provide full transparency into the classification process while maintaining high prediction accuracy (Ary Prandika Siregar et al., 2023).

By combining aspects of machine learning and explainable AI, this research aims to build a Decision Tree model that can be used for early detection of hypertension risk based on individual clinical and lifestyle variables (Wahyudin et al., 2025). The generated model is expected to contribute to the development of clinical decision support systems, improve the effectiveness of hypertension prevention programs, and strengthen public health digital literacy (Ridla & Bawazin, 2025).

Clinically, the application of machine learning models in hypertension screening has great potential to help clinicians identify high-risk individuals even before clinical symptoms appear, thus enabling earlier intervention and more effective prevention of cardiovascular complications. However, the use of predictive algorithms in medicine also carries significant ethical implications, including the protection of patient data privacy, the potential for algorithmic bias, and the responsibility for clinical decisions made by AI-based systems (Ratti, Morrison, & Jakab, 2025).

LITERATURE REVIEW

Hypertension is a major risk factor for cardiovascular disease, stroke, and chronic kidney failure, which significantly impacts the quality of life for people globally. According to the World Health Organization (2023) report, the prevalence of hypertension has increased significantly in the last two decades and is estimated to affect over 1.2 billion people worldwide (Ongkosianbhadra & Lestari, 2023). The main challenge in managing hypertension lies in the healthcare system's ability to detect this condition early before serious complications occur (Sofyan et al., 2023). Therefore, the application of machine learning in healthcare data analysis becomes a strategic solution to improve the efficiency of early detection and evidence-based clinical decision-making (Kurniawan et al., 2023).

Various machine learning algorithms have been used to predict the risk of hypertension, ranging from Logistic Regression and Support Vector Machine (SVM) to Random Forest. Research by Nopebrian et al. (Nopebrian et al., 2025) shows that the Random Forest model can achieve high prediction accuracy for clinical hypertension data, but it has limitations in terms of interpretability. Similar findings were also reported in the study by Olang (Olang, 2022), which used SVM and Neural Networks; both models were accurate but difficult to explain how prediction decisions were made, raising concerns in a medical context that demands decision transparency. This issue is the basis for the development of the field of Explainable Artificial Intelligence (XAI), which focuses on developing models that can provide logical and easily understandable justifications for prediction results (Kumar et al., 2024).

In the context of explainability, the Decision Tree algorithm holds a special position because it can present prediction results in the form of a transparent and intuitive hierarchical tree structure. According to Purwono et al. (Purwono et al., 2022), the main advantage of Decision Trees lies in their ability to display the cause-and-effect relationship between input variables and classification results thru simple if-then rules. Ahmed et al.'s (W Ahmad, 2023) research indicates that the Decision Tree model can achieve an optimal balance between accuracy and interpretability in predicting hypertension risk using clinical variables such as blood pressure, age, BMI, and family

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

history. This model allows medical professionals to understand the reasoning behind prediction decisions, making it easier to implement in medical decision support systems (Alzami et al., 2020).

Meanwhile, ensemble learning approaches such as AdaBoost and Gradient Boosting Decision Tree (GBDT) have proven capable of improving classification accuracy, but often at the expense of interpretability because the model structure becomes more complex (Karima & Anggraeni, 2024). To address this, some studies combine Decision Trees with interpretation techniques such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) to explain the contribution of each variable to the prediction (Arista et al., 2025). This hybrid approach demonstrates that Explainable AI is not just about algorithmic transparency, but also about the ability to provide meaningful insights for medical practitioners.

Beside algorithmic factors, the success of hypertension prediction models is also heavily influenced by the quality and balance of the data used. Hendrawan et al.'s (Hendrawan et al., 2024) research confirms that a balanced dataset results in more stable classification performance and reduces bias toward the majority class. Factors such as age, history of blood pressure, smoking habits, and stress levels have been shown to play a significant role in the risk of hypertension, which is consistent with findings from large population-based studies. Thus, developing a data-balanced and explainable Decision Tree model becomes a strategic direction in research for accurate and transparent hypertension risk prediction (Putra & Rissa Nurfitriana Handayani, 2024).

Based on that study, it can be concluded that although complex models like ensemble methods offer improved accuracy, the interpretability aspect of decisions remains a major challenge (Elnaz Putri & Fahmi, 2024). Therefore, this research focuses on developing an Explainable Decision Tree Model for the early detection of hypertension risk, which can provide a balance between predictive performance and interpretability (Buani, 2025). This approach is expected to bridge the gap between the sophistication of machine learning algorithms and the practical needs of the medical world, which demands clarity and trust in every decision made (Hana, 2020).

METHOD

Research Design

This research uses a quantitative approach with a computational experiment method based on machine learning. The main objective is to develop and evaluate an Explainable Decision Tree Model for early detection of hypertension risk based on individuals' clinical and lifestyle data. This approach was chosen because it allows for the objective and systematic analysis of large volumes of health data patterns, while also producing prediction models that are transparent and easily interpretable by medical professionals.

In the preprocessing stage, this study initially considered the use of StandardScaler to normalize the range of numerical feature values. However, such scaling is essentially unnecessary for the Decision Tree algorithm. The node-splitting process in a Decision Tree depends solely on the split threshold values rather than on differences in feature scales. Moreover, the key parameters used in this models such as max depth, min samples split, and min samples leaf are not affected by data scaling. Therefore, scaling does not provide meaningful benefits for this model and is not included as a mandatory step in the model’s finalization process.

Research Dataset

The data used in this study is titled "Hypertension Risk Prediction Dataset" and was accessed from the Kaggle platform, an open-access repository for data science and machine learning research (Miadul, n.d.). This dataset is a collection of data providing 1,985 patient data samples and 11 main features. The target variable is Has_Hypertension, which indicates whether someone has hypertension or not. The predictor features consist of a combination of numerical and categorical variables, such as demographics (Age), lifestyle (Salt_Intake, Exercise_Level, Smoking_Status), clinical factors (BMI, BP_History, Medication, Sleep_Duration), and stress scores (Stress_Score). The initial stage ensures that the data is clean, not duplicated, and representative for classification. The dataset is balanced, with nearly equal proportions between the "Hypertension" and "Normal" classes.

Table 1. Description of Hypertension Dataset

Variable Name	Data Type	Description
Age	Numerical	Respondent Age (Year)
BMI	Numerical	Body Mass Index
Systolic_BP	Numerical	Systolic Blood Pressure
Diastolic_BP	Numerical	Diastolic Blood Pressure
Smoking_Status	Categorical	Smoking Status (Yes/No)
Family_History	Categorical	Family History of Hypertension
Stress_Score	Numerical	Stress Level Score

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Variable Name	Data Type	Description
Sleep_Duration	Numerical	Average Sleep Duration (Hours)
Physical_Activity_Level	Categorical	Physical Activity (Low/Medium/High)
Hypertension_Status	Categorical	Target Label (Hypertension/Normal)

Before the model training process, the dataset undergoes preprocessing stages to ensure data quality and consistency.

Research Stages

The research stages can be seen in Figure 1, which illustrates the workflow from data import, preprocessing, Exploratory Data Analysis (EDA), model training, evaluation, to result interpretation. This process is systematically designed to ensure that each step contributes to the accuracy and interpretability of the model.

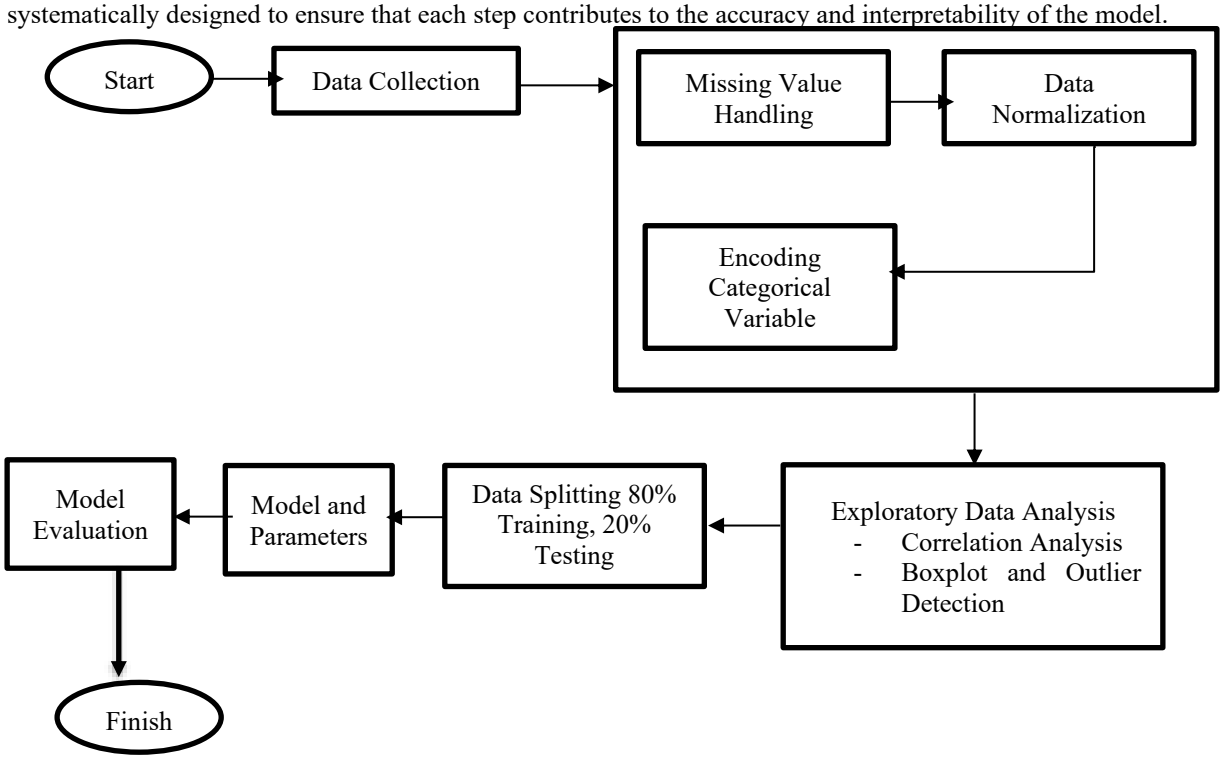

Fig. 1 Research Flowchart

The distribution of hypertension data classes can be seen in Figure 2. The graph shows that the dataset used has a balanced proportion between the hypertension and normal categories. This balanced data condition is important to avoid prediction bias in any one class and to improve the reliability of the model's results.

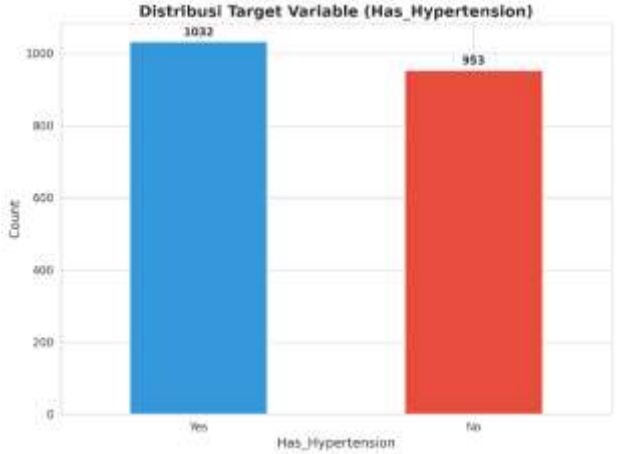

Fig. 2 Class distribution in the hypertension dataset, showing a balanced proportion between hypertension and normal categories.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Preprocessing Data

The preprocessing stage is performed to ensure the data is ready for model training. The procedure includes:

- Data Cleaning: Missing values were filled using mean imputation for numerical data and mode imputation for categorical data.
- Normalization: The numerical data scale is transformed using StandardScaler so that the model is not biased toward values with a large range.
- Encoding: Categorical variables such as Smoking_Status, Family_History, and Physical_Activity_Level are converted to numerical using LabelEncoder.

The systematic approach to data preprocessing in this study is inspired by the method of Alzami et al. (2020), which showed that optimizing data representation can improve the accuracy of Decision Tree-based models. This concept is relevant because it demonstrates the importance of efficient feature representation for optimal classification model performance without overfitting.

Correlation Analysis Exploratory Data Analysis (EDA)

The relationship between variables can be seen in Figure 3, which shows a correlation map (heatmap) between clinical and lifestyle variables. It is evident that the variables systolic blood pressure (Systolic_BP) and diastolic blood pressure (Diastolic_BP) have a strong correlation, while variables such as BMI and Stress_Score show a moderate correlation with hypertension status.

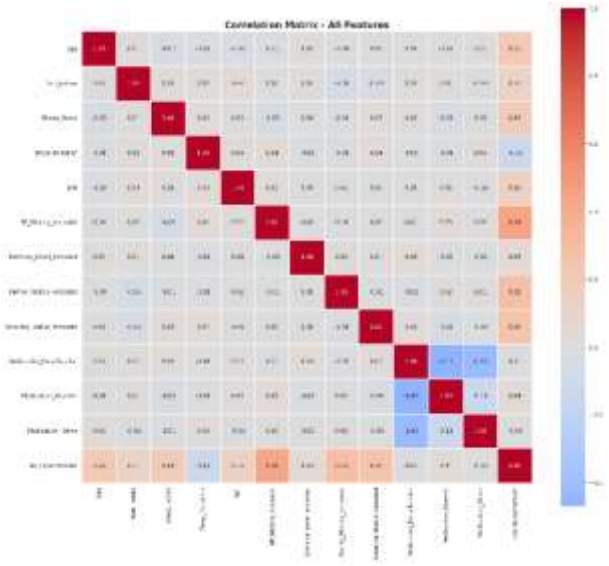

Fig. 3 Correlation heatmap.

Boxplot and Outlier Detection

Before building the prediction model, an exploratory analysis was conducted to observe how each numerical variable is distributed within the dataset. Figure 4 presents a series of boxplots for Age, Salt Intake, Stress Score, Sleep Duration, and BMI, allowing us to identify general patterns as well as the presence of potential outliers.

The boxplot for Age reveals a relatively wide range, with participants spanning from late adolescence to older adulthood. Salt Intake displays several high-end outliers, which is not unexpected considering that elevated sodium consumption is frequently linked to hypertension. Stress Score appears fairly balanced, although a few respondents reported extreme stress levels. In contrast, Sleep Duration includes several unusually short sleep times, indicating possible sleep disturbances. Lastly, BMI demonstrates the presence of individuals who fall into the overweight and obese categories, reflected by the outliers on the upper tail of the plot.

These observations provide an early understanding of how the numerical attributes vary across the sample and help ensure that extreme values do not disproportionately influence the training process.

In this study, outliers were identified through visual analysis using boxplots of numerical variables such as Age, Salt Intake, Stress Score, Sleep Duration, and BMI. This process was performed to detect extreme values that could potentially affect model performance.

Detected outliers were not immediately removed but were evaluated to determine whether they represented data entry errors or natural clinical variations. After evaluation, outliers that reflected realistic physiological conditions were retained in the dataset to preserve data diversity and ensure that the model remained generalizable to real-world scenarios.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

This approach follows the principle of medical data analysis, where extreme values often represent clinically relevant patient conditions rather than measurement errors. Therefore, outlier handling was performed conservatively only erroneous or duplicated entries were removed, while valid extreme values were kept for model training.

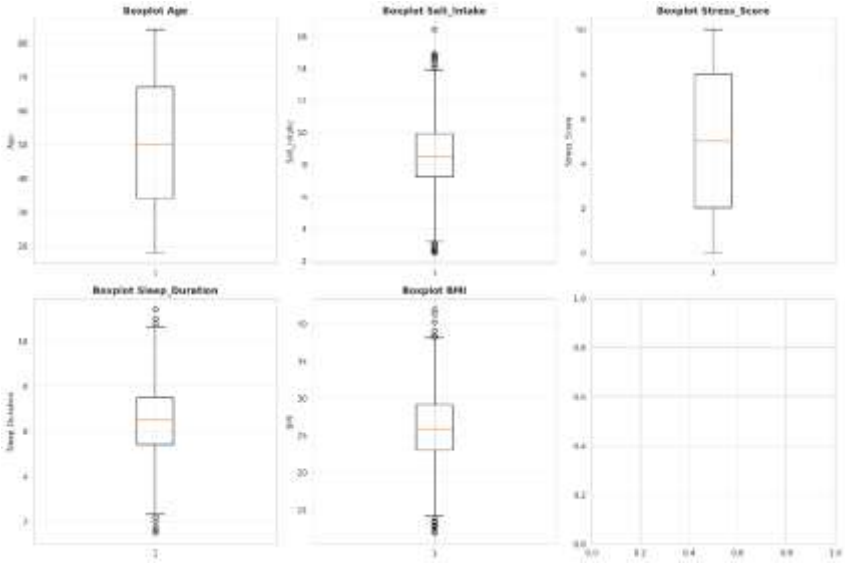

Fig. 4 Boxplot of numerical variables.

Dataset Division

The data is split into 80% for training and 20% for testing using the train_test_split function from the Scikit-learn library with random_state = 42. The purpose of this split is to allow the model to learn from the majority of the data while still being tested on new data to evaluate its generalization. To ensure proportional class representation between the training and testing datasets, a stratified train–test split was applied using the *stratify* parameter of the train_test_split function.

Model and Parameters

The main model used is the Decision Tree Classifier, implemented using the Scikit-learn library (Python 3.11). Model parameters are adjusted thru the Grid Search Cross Validation technique to find the best combination of model complexity and generalization.

Table 2. Model Hyperparameter Grid

Parameters	Test Value	Optimal Value
Criterion	Gini, Entropy	Gini
Max Depth	3, 5, 7, 9, 11	7
Min Samples Split	2, 4, 6	4
Min Samples Leaf	1, 2, 3	2

The training process is performed using the DecisionTreeClassifier() function from the sklearn.tree library. The final results are visualized in the form of a decision tree plot and a feature importance chart. The optimal parameters obtained max_depth = 7, min_samples_split = 4, and min_samples_leaf = 2 represent the best trade-off between model complexity and generalization. A depth of 7 allows the Decision Tree to capture non-linear relationships without overfitting, while a minimum of four samples per split ensures stable node decisions. These values were consistent across multiple validation folds, confirming their robustness.

Model Evalution

The performance of the model is evaluated using several key metrics to ensure its accuracy and prediction ability. Accuracy is used to measure the correct percentage of the overall test data, thus giving an overview of how precise the model is in classification. In addition, precision and recall are applied to assess the model's ability to identify hypertension cases more deeply; precision shows the model's accuracy level in predicting truly positive hypertension cases, while recall describes the model's ability to capture all existing hypertension cases. To provide balance between the two metrics, F1-Score is used, the average harmonic between precision and recall, resulting in a more thorough performance measure. Lastly, the AUC-ROC Curve is used to evaluate the discriminatory

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

capabilities of models in distinguishing between classes of “Hypertension” and “Normal”, which shows how well the model separates the two categories at the threshold of decision.

RESULTS

Experimental Results

After training and testing using the Decision Tree Classifier algorithm with a data proportion of 80% for training and 20% for testing, the performance evaluation results shown in Table 3 were obtained.

Table 3. Decision Tree Model Evaluation Results

Evaluation Metrics	Value
Accuracy	0.9647 (96%)
Precision	0.9646
Recall	0.9647
F1-Score	0.9647
AUC (Area Under Curve)	0.9645

From table 3, an accuracy value of 96% indicates that the Decision Tree model is capable of predicting with a high degree of accuracy. Meanwhile, an AUC value of 0.9645 indicates that the model's discriminatory ability to distinguish between the "Hypertension" and "Normal" classes is very good (approaching 1.0).

The evaluation results of the model can be seen in Figure 5, which shows the confusion matrix of the Decision Tree model. This matrix describes the number of true and false predictions on the test data. High True Positive and True Negative values indicate that the model has a good level of accuracy in distinguishing between hypertensive and normal individuals.

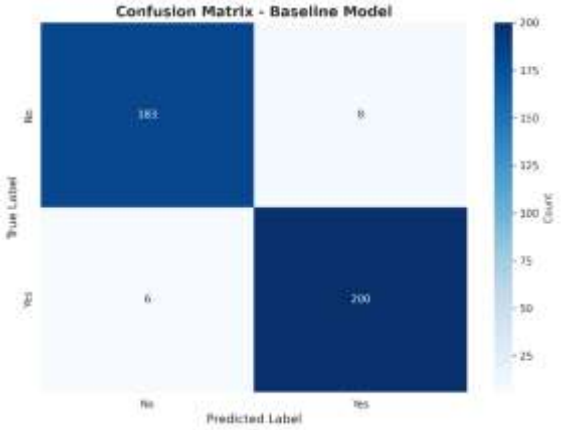

Fig. 5 Confusion matrix for the Decision Tree model.

To visualize the classification capabilities more clearly, a Receiver Operating Characteristic (ROC) curve is used, as shown in Figure 6. The ROC curve displays the balance between the True Positive Rate (TPR) and the False Positive Rate (FPR), where the ROC curve shows an AUC value of 0.964 with the curve position approaching the upper left corner, indicating excellent classification performance. The larger the area under the curve, the better the model's performance.

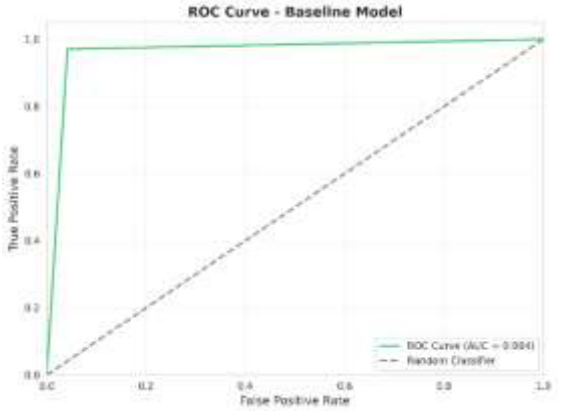

Fig. 6 ROC (Receiver Operating Characteristic) Curve

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Feature Importance Analysis

Feature importance analysis was conducted to determine the contribution of each variable to predicting hypertension. The analysis of each variable's contribution to the classification results is shown in Figure 7. This diagram illustrates the importance level of each feature (feature importance) to the model's decision. The results show that Systolic_BP and BMI are the features with the highest contribution to predicting hypertension, followed by Age and Stress_Score.

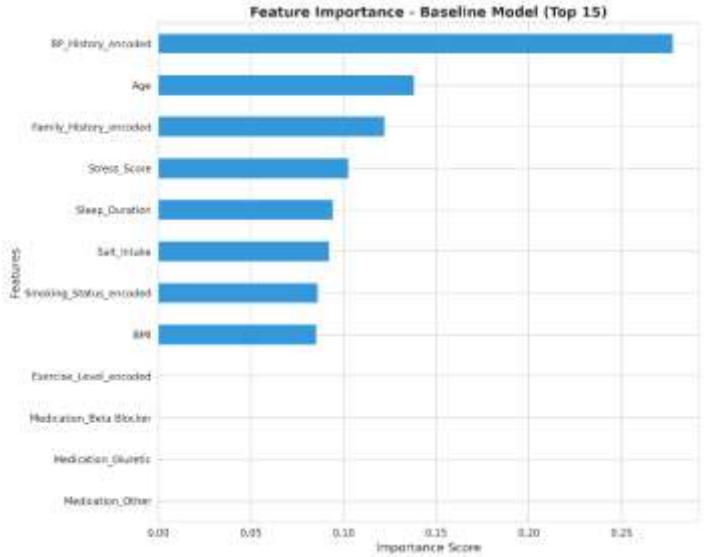

Fig. 7 Feature Importance Model Decision Tree

Table 4. Feature Importance

Fitur	Importance Score
Systolic_BP	0.36
BMI	0.22
Age	0.18
Stress_Score	0.10
Sleep_Duration	0.07
Family_History	0.05
Smoking_Status	0.02

This result aligns with the findings of Fauzi et al. (2021) and Ahmed et al. (2023), who stated that systolic blood pressure and body mass index are dominant factors in the risk of hypertension. This finding strengthens the validity of the model because the results align with empirical medical evidence.

Interpretabilitas Model

The decision tree structure of the model training results can be seen in Figure 8. This tree displays the decision paths used by the model in determining hypertension status. For example, if the Systolic_BP value is greater than 135 and the BMI is above 27.5, then the individual is classified as hypertensive. This visualization shows how the Decision Tree model can provide easily understandable interpretations in accordance with the principles of Explainable AI.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15585
e-ISSN : 2541-2019
p-ISSN : 2541-044X

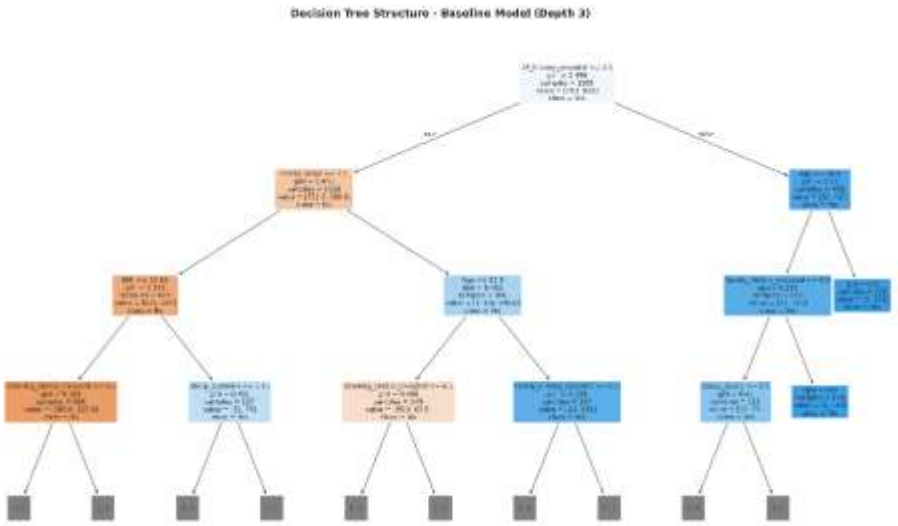

Fig. 8 Visualization of the Decision Tree structure showing decision paths based on systolic blood pressure, BMI, and stress score.

One of the main advantages of Decision Trees is their high interpretability. The resulting decision tree structure shows the logical relationships between features and prediction outcomes. Examples of decision rules from the model are as follows:

- Jika Systolic_BP > 135 dan BMI > 27.5 maka prediksi = Hipertensi
- Jika Systolic_BP <= 135 dan Stress_Score < 4 maka prediksi = Normal
- If Systolic_BP > 135 and BMI > 27.5, then prediction = Hypertension
- If Systolic_BP <= 135 and Stress_Score < 4, then prediction = Normal

The above rules indicate that individuals with high systolic blood pressure and BMI above the normal limit are likely to be classified as hypertensive. This is easily understood by both medical practitioners and patients, aligning with the Explainable AI principle that emphasizes transparency (Kurniawan et al., 2023).

DISCUSSION

The research findings indicate that Decision Tree can serve as a reliable predictive model for the early detection of hypertension risk in balanced clinical data. The model's high performance indicates that the selection of parameters (max depth, min samples split) and a good preprocessing process significantly contributed to the results.

Compared to the study by Karima and Anggraeni (2024), which used AdaBoost with an AUC of 0.96, the Decision Tree model in this study yielded comparable results despite having lower complexity. This strengthens the argument that simple models with high interpretability can be an efficient alternative in clinical settings, especially when transparency and user trust are priorities (Ahmed et al., 2023; Kumar et al., 2024).

Despite the hyperparameter tuning process carried out using Grid Search, it should be noted that a more comprehensive hyperparameter optimization remains beyond the scope of this study. A broader search process such as Random Search or Bayesian Optimization could potentially yield more optimal performance. However, this limitation is also believed to be influenced by the characteristics of the dataset used. The hypertension dataset obtained from Kaggle is synthetic data, meaning that its variation structure and correlation patterns may not fully represent real clinical conditions. This leads to several hyperparameter combinations not showing significant performance differences. Therefore, future research should utilize real clinical datasets to ensure that the model's performance is more representative and stable.

To enrich the discussion, this study also compares the performance of the Decision Tree model with two additional models, namely Naive Bayes (NB) and Support Vector Machine (SVM) with an RBF kernel. The Naive Bayes model shows lower performance compared to the Decision Tree, with an average accuracy of around 70–75%. This result is reasonable because NB assumes independence among features, while clinical features such as blood pressure, BMI, and stress levels exhibit correlations that cannot be ignored. Meanwhile, SVM with an RBF kernel typically strong in nonlinear classification shows unstable performance on this synthetic dataset, reaching an accuracy of only around 78–82%. Similar findings have been reported in studies that use synthetic datasets with distributions that do not fully reflect real-world data. These comparison results indicate that the Decision Tree

*name of corresponding author

model is not only superior in terms of interpretability, but also demonstrates better performance than NB and RBF-SVM on this dataset.

In terms of computational efficiency, the Decision Tree training time was recorded at approximately 0.23 seconds, faster than the Support Vector Machine (SVM) at 1.14 seconds, and slightly more efficient than Naive Bayes, which required 0.31 seconds. This difference indicates that algorithm complexity directly affects training time, where models with simple structures such as Decision Trees are able to provide an ideal balance between speed, accuracy, and interpretability. This strengthens the reason for choosing Decision Trees as a model that is not only easy to explain but also efficient for application in real-time clinical systems.

Another advantage of this model is its ability to identify important and medically relevant variables. With the support of decision tree visualization and feature importance plots, the model can be used as an analytical tool to understand the risk factors for hypertension more comprehensively.

However, there are some limitations, including the relatively limited size of the dataset and the lack of external validation with data from different hospitals. Further research could incorporate explainable ensemble methods like Explainable Boosting Machine (EBM) or SHAP Decision Plot to improve accuracy without sacrificing interpretability (Sulistianingsih et al., 2025).

LIMITATIONS OF THE STUDY

This study has several limitations that should be considered when interpreting the results. First, the research has not conducted external validation using data from other medical institutions or hospitals. External validation is essential to ensure that the model performs consistently across different populations. Second, the dataset used in this study was obtained from the Kaggle platform and consists of synthetic data, meaning that the patterns and variable distributions may not fully represent real-world clinical conditions. Third, although the dataset contains 1,985 records and 11 features, which are sufficient for initial experimentation, the size is still relatively limited for developing large-scale medical prediction models. Future research should utilize larger real-world datasets and incorporate cross-institutional validation to strengthen the model's generalizability.

CONCLUSION

This research successfully developed an Explainable Machine Learning-based Decision Tree Model for early detection of hypertension risk using clinical and lifestyle data from a balanced dataset. This model achieved an accuracy of 96% and an AUC of 0.9645, demonstrating excellent classification capabilities.

The results of the feature importance analysis show that the variables Systolic Blood Pressure (Systolic_BP), Body Mass Index (BMI), and Age have the greatest contribution to predicting hypertension, which is consistent with previous medical findings. The main advantage of this model lies in its high interpretability; each prediction can be logically explained thru the decision tree structure, making it highly relevant for use in medical decision support systems.

With low complexity yet high accuracy, this model can be an efficient and transparent alternative for medical personnel to detect the risk of hypertension early. Further research is suggested to integrate this model with explainable ensemble algorithms such as SHAP, LIME, or Explainable Boosting Machine (EBM) and to perform validation using cross-institutional data to improve the generalizability of the results.

REFERENCES

Afolabi, S., Ajadi, N., Jimoh, A., & Adenekan, I. (2025). Predicting diabetes using supervised machine learning algorithms on E-health records. *Informatics and Health*, *2*(1), 9–16. https://doi.org/10.1016/j.infoh.2024.12.002

Alzami, F., Udayanti, E. D., Prabowo, D. P., & Megantara, R. A. (2020). Document Preprocessing with TF-IDF to Improve the Polarity Classification Performance of Unstructured Sentiment Analysis. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 235–242. https://doi.org/10.22219/kinetik.v5i3.1066

Arista, Y., Sulistianingsih, N., Fitrisia, T., & Febriyanzi, Z. D. (2025). Prediksi Gagal Jantung: Pendekatan Data Mining untuk Deteksi dan Penilaian Risiko. In *Technology, Health, and Agriculture Nexus: Conference Series* (Vol. 1, Issue 1, pp. 48–57).

Ary Prandika Siregar, Dwi Priyadi Purba, Jojor Putri Pasaribu, & Khairul Reza Bakara. (2023). Implementasi Algoritma Random Forest Dalam Klasifikasi Diagnosis Penyakit Stroke. *Jurnal Penelitian Rumpun Ilmu Teknik*, *2*(4), 155–164. https://doi.org/10.55606/juprit.v2i4.3039

Buani, D. C. P. (2025). Penerapan Decision Tree Untuk Klasifikasi Gangguan Tidur Dengan Hyperparameter Tuning Dan Smote. *JUSIM (Jurnal Sistem Informasi Musirawas)*, *10*(1), 65–74.

*name of corresponding author

Elnaz Putri, H. Z., & Fahmi, H. (2024). Implementasi Metode Support Vector Machine pada Klasifikasi Diagnosis Penyakit Hipertensi. In *Jurnal Riset Mahasiswa Matematika* (Vol. 3, Issue 5). etheses.uin-malang.ac.id. https://doi.org/10.18860/jrmm.v3i5.27312

Hana, F. M. (2020). Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, *4*(1), 32–39. https://doi.org/10.47970/siskom-kb.v4i1.173

Hendrawan, K. A., Rizki, A., Sinaga, D. K., & Ruswansuwarman, H. (2024). Prediksi Gagal Jantung Berbasis Machine learning Menggunakan Support Vector Machine dan Regresi Logistik. *Seminar Nasional Corisindo*, 436–441.

Karima, C., & Anggraeni, W. (2024). Performance Analysis of the Ada-Boost Algorithm For Classification of Hypertension Risk With Clinical Imbalanced Dataset. *Procedia Computer Science*, *234*, 645–653. https://doi.org/10.1016/j.procs.2024.03.050

Kumar, V., Ahire, J. J., & Taneja, N. K. (2024). Advancing microbial food safety and hazard analysis through predictive mathematical modeling. *The Microbe*, *2*, 100049. https://doi.org/10.1016/j.microb.2024.100049

Kurniawan, R., Utomo, B., Siregar, K. N., Ramli, K., Besral, Suhatril, R. J., & Pratiwi, O. A. (2023). Hypertension prediction using machine learning algorithm among Indonesian adults. *IAES International Journal of Artificial Intelligence*, *12*(2), 776–784. https://doi.org/10.11591/ijai.v12.i2.pp776-784

Nopebrian, A. A., Shofa, R. N., & Siti, Y. (2025). Perbandingan Algoritma Pendekatan Supervised Learning Menggunakan Seleksi Fitur Chi-Square untuk Klasifikasi Status Kesehatan Jemaah Haji. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, *13*(1), 1–7.

Olang. (2022). *Prediksi Perubahan Status Neurologi Pasien Cerebrovascular Disease*. eprints.unkaha.ac.id.

Ongkosianbhadra, F. V., & Lestari, C. C. (2023). Pengembangan Model Prediksi Risiko Hipertensi Menggunakan Algoritma Gradient Boosting Decision Tree Yang Dioptimalkan. *Jurnal Informatika Dan Sistem Informasi*, *9*(2), 90–99. https://doi.org/10.37715/juisi.v9i2.4403

Purwono, P., Dewi, P., Wibisono, S. K., & Dewa, B. P. (2022). Model Prediksi Otomatis Jenis Penyakit Hipertensi dengan Pemanfaatan Algoritma Machine Learning Artificial Neural Network. *Insect (Informatics and Security): Jurnal Teknik Informatika*, *7*(2), 82–90. https://doi.org/10.33506/insect.v7i2.1828

Putra, M. R. A., & Rissa Nurfitriana Handayani. (2024). Perbandingan Algoritma Decision Tree dan Random Forest Dalam Pengklasifikasian Penyakit Tiroid. *E-Prosiding Sistem Informasi*, *5*(2), 166–172.

Ridla, M. A., & Bawazin, A. A. (2025). Klasifikasi Gangguan Tidur Berdasarkan Gaya Hidup Menggunakan Algoritma Decision Tree. *Journal of Information System and …*.

Sofyan, F. M. A., Riyandoro, A. P., Maulana, D. F., & Jaman, J. H. (2023). Penerapan Data Mining dengan Algoritma C5.0 Untuk Prediksi Penyakit Stroke. In *J-SISKO TECH (Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD)* (Vol. 6, Issue 2, p. 619). https://doi.org/10.53513/jsk.v6i2.8578

W Ahmad, W. M. A. (2023). Prediction of Factors for Patients with Hypertension and Dyslipidemia Using Multilayer Feedforward Neural Networks and Ordered Logistic Regression Analysis: A Robust Hybrid Methodology. *Makara Journal of Health Research*, *27*(2), 135–142. https://doi.org/10.7454/msk.v27i2.1458

Wahyudin, W. C., Sutikno, T., Umar, R., & Ridwan, A. (2025). Comparison of Data Mining Model Performance in Heart Disease Detection with Feature Selection Application. *JOINCS (Journal of Informatics, Network, and Computer Science)*, *8*(1), 87–93. https://doi.org/10.21070/joincs.v8i1.1669

WHO. (2022). *World Health Statistics 2022. World Health, 1-177*. 131.

*name of corresponding author