EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

Explainable Diabetes Classification Using Machine Learning Algorithm and SHAP Analysis on Clinical Laboratory Data

Siti Agus Kartini[1*], Muhammad Fauzi[2], Riki Winanjaya[3]
[1*,2] Department of Information Systems, Universitas Tjuk Nyak Dhien, Medan, Indonesia
[3] Department of Information Systems, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia
[1*]sitiaguskartini11@gmail.com, [2]mfauzi@utnd.ac.id, [3]riki@amiktunasbangsa.ac.id
[*)] sitiaguskartini11@gmail.com

Abstract−Diabetes mellitus is a chronic metabolic disorder that continues to pose significant health challenges worldwide due to its increasing prevalence and potential complications. Early and accurate identification of diabetes is essential to support timely intervention and effective disease management. Recent advances in machine learning have enabled the development of intelligent classification systems; however, many predictive models still suffer from limited interpretability, reducing their applicability in clinical environments. Therefore, this study proposes an explainable diabetes classification framework using machine learning algorithms and SHapley Additive exPlanations (SHAP) analysis on clinical laboratory data. The dataset consists of 1,000 patient records containing demographic and laboratory attributes, including age, gender, glycated hemoglobin (HbA1c), cholesterol, triglycerides, lipoprotein levels, creatinine, urea, and body mass index (BMI). Four machine learning algorithms, namely Decision Tree, Random Forest, Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost), were developed and evaluated. The performance of each model was assessed using Accuracy, Precision, Recall, F1-Score, Classification Report, and Confusion Matrix. Experimental results indicate that Random Forest and XGBoost achieved the highest classification accuracy of 98.5%, outperforming Decision Tree (98.0%) and SVM (94.5%). Due to its strong predictive capability and compatibility with explainable artificial intelligence techniques, XGBoost was selected for further SHAP analysis. The SHAP results successfully identified the most influential features contributing to diabetes classification and provided transparent explanations regarding model predictions. The proposed framework demonstrates that combining machine learning algorithms with explainable artificial intelligence improves predictive accuracy and interpretability, supporting reliable clinical decision-making for diabetes diagnosis.

Keywords: Diabetes Mellitus, Machine Learning, XGBoost, SHAP, Clinical Laboratory Data

1. INTRODUCTION

Diabetes mellitus has become one of the most serious global health challenges because its prevalence continues to rise across both developed and developing countries [1]–[5]. Recent global evidence shows that more than 800 million adults were living with diabetes in 2022, and more than half of adults over 30 years old with diabetes were not receiving treatment [6]–[8]. This condition is concerning because diabetes is not only related to high blood glucose, but also to long-term complications involving the heart, kidneys, nerves, blood vessels, and eyes [9]–[12]. The increasing burden of diabetes creates pressure on healthcare systems, especially in countries with limited access to early screening, laboratory testing, and continuous disease monitoring [13]–[15]. Early diagnosis is therefore essential to prevent severe complications and reduce treatment costs [16]. In clinical practice, laboratory indicators such as glycated hemoglobin (HbA1c), lipid profile, kidney function markers, and body mass index are commonly used to support diabetes assessment [17]–[20]. However, these variables often interact in complex and non-linear ways, making manual interpretation difficult. Therefore, the use of machine learning has become increasingly relevant to support diabetes classification and improve data-driven clinical decision-making.

The main problem in diabetes classification is not only how to build a model with high predictive accuracy, but also how to make the prediction understandable for clinical users [21]–[23]. Many machine learning algorithms can identify hidden patterns in clinical laboratory data, yet complex models such as ensemble methods and boosting algorithms are often considered black-box systems [24]. This creates a challenge because healthcare professionals need to understand why a patient is classified as diabetic, non-diabetic, or pre-diabetic before trusting the model output [25]. In addition, clinical datasets may contain heterogeneous variables, different measurement scales, label inconsistencies, and imbalanced class distributions. These issues can affect model performance and reduce the reliability of prediction results [26]. A model that only reports accuracy is insufficient for medical research because it does not explain which features contribute most strongly to the classification [27]. For example, HbA1c, BMI, cholesterol, triglycerides, creatinine, and urea may have different effects on each prediction [28]. Therefore, diabetes classification research requires a combined approach that compares several machine learning algorithms and applies explainable artificial intelligence to interpret the best-performing model.

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

Several recent studies have investigated the application of machine learning and explainable artificial intelligence (XAI) for diabetes prediction. Hasan et al. (2025) integrated Automated Machine Learning and Explainable Artificial Intelligence to improve diabetes risk prediction and model interpretability, demonstrating the growing importance of transparent predictive systems in healthcare [29]. Kutlu et al. (2024) employed XGBoost and SHAP to enhance interpretability in diabetes risk assessment and highlighted the importance of understanding model decision-making processes beyond predictive performance alone [30]. Islam et al. (2024) proposed an explainable machine learning framework for type 2 diabetes classification and demonstrated that machine learning models can achieve high performance while maintaining explainability through SHAP-based analysis [31]. Netayawijit et al. (2025) emphasized the growing trend of interpretable machine learning frameworks for diabetes prediction and reported that balancing predictive accuracy and explainability remains a challenging research problem [32]. Rafie et al. (2025) utilized XGBoost combined with explainable AI techniques and identified important predictive factors through SHAP analysis, showing the effectiveness of interpretable ensemble learning approaches for diabetes prediction [33]. Agarwal et al. (2025) conducted a comparative study integrating XGBoost with explainable AI techniques and evaluated several machine learning algorithms using performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC [34]. Ahmad et al. (2026) developed a stacked ensemble model enhanced with SHAP explainability for early-stage diabetes prediction and demonstrated that explainable models can provide valuable insights into disease diagnosis [35]. Although these studies have contributed significantly to the field, most of them focus on public benchmark datasets such as the Pima Indians Diabetes Dataset or demographic health datasets. Furthermore, limited studies have simultaneously compared Decision Tree, Random Forest, Support Vector Machine, and XGBoost using clinical laboratory data while incorporating SHAP-based explainability to interpret model predictions. Therefore, a research gap remains in developing a comprehensive explainable diabetes classification framework that combines comparative machine learning analysis with SHAP interpretation on clinical laboratory datasets, providing both predictive performance and clinical interpretability.

To address this problem, this study proposes an explainable diabetes classification approach using machine learning algorithms and SHAP analysis on clinical laboratory data. The workflow begins with data preprocessing, including removing irrelevant identity columns, encoding categorical variables, encoding the target class, standardizing numerical features, and dividing the dataset into training and testing sets. Several machine learning algorithms are then trained and compared, namely Decision Tree, Random Forest, Support Vector Machine, and XGBoost. These algorithms are selected because they represent different classification mechanisms, ranging from rule-based tree structures to ensemble learning, margin-based classification, and gradient boosting. Model performance is evaluated using accuracy, precision, recall, F1-score, classification reports, and confusion matrices. The best-performing model is then interpreted using SHAP to identify which clinical features contribute most strongly to diabetes classification. Through SHAP summary and beeswarm plots, the model becomes more transparent because each feature contribution can be visually examined. This approach is expected to produce a predictive model that is not only accurate but also explainable and clinically meaningful.

The main objective of this study is to develop an explainable diabetes classification model using machine learning algorithms and SHAP analysis based on clinical laboratory data. Specifically, this study aims to compare the performance of Decision Tree, Random Forest, SVM, and XGBoost in classifying diabetes status, determine the best-performing algorithm, and interpret the contribution of each clinical feature to the prediction result. This research also aims to demonstrate that explainability is an important component in medical machine learning, especially when the model is intended to support clinical decision-making. The expected contribution of this study is a classification framework that can help identify diabetes status more accurately while providing understandable explanations for each prediction. The findings are expected to support early screening, improve awareness of influential laboratory indicators, and strengthen trust in data-driven healthcare systems. In addition, this study can serve as a foundation for future research that uses larger datasets, external validation, and more advanced explainable artificial intelligence methods for diabetes diagnosis and risk prediction.

2. RESEARCH METHODS

This study employed a quantitative experimental approach to develop an explainable machine learning framework for diabetes classification using clinical laboratory data. The proposed framework consists of several stages, including data collection, data preprocessing, machine learning model development, model evaluation, and explainability analysis using SHapley Additive exPlanations (SHAP). Four machine learning algorithms, namely Decision Tree, Random Forest, Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost), were utilized and compared to identify the most effective model for diabetes classification. The overall research framework is illustrated in Figure 1.

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer
DOI: https://doi.org/10.47065/explorer.v6i2.2987

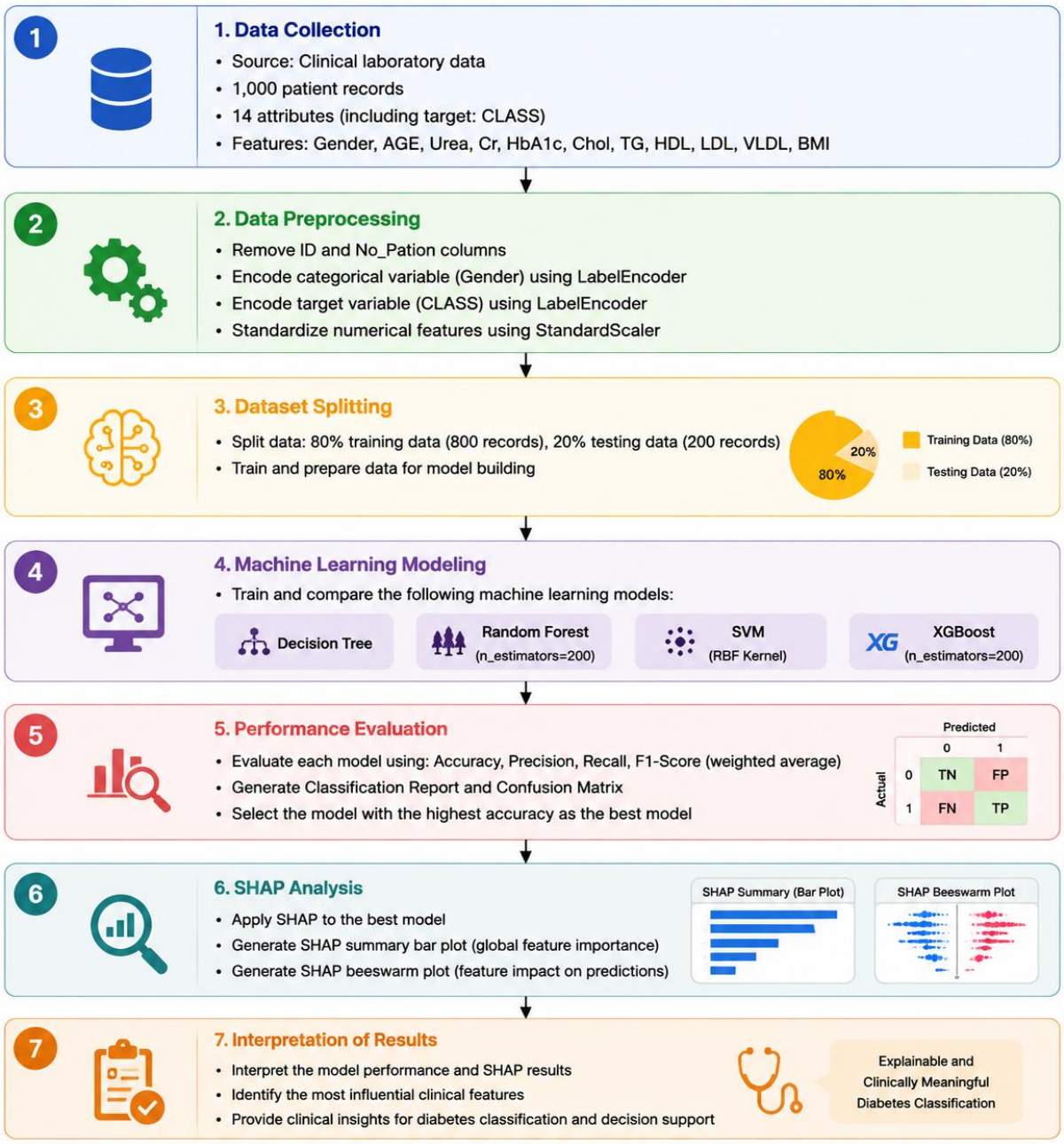

Figure 1. Research methodology workflow

Figure 1 illustrates the research methodology workflow used in this study. The process begins with data collection, where clinical laboratory data consisting of 1,000 patient records and 14 attributes were obtained. The data were then processed through data preprocessing, including removing non-relevant identification columns, encoding categorical variables, encoding the target class, and standardizing numerical features. After preprocessing, the dataset was divided into 80% training data and 20% testing data to support model development and evaluation. The next stage involved machine learning modeling, where four classification algorithms were trained and compared, namely Decision Tree, Random Forest, Support Vector Machine, and XGBoost. Each model was evaluated using accuracy, precision, recall, F1-score, classification report, and confusion matrix. The model with the highest performance was then selected as the best model for further analysis. Finally, SHAP analysis was applied to explain the contribution of each clinical feature to the prediction results, followed by the interpretation of results to identify the most influential variables and provide clinically meaningful insights for diabetes classification.

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

2.1 Data Collection

This study used a clinical laboratory diabetes dataset obtained in Excel format. The dataset consisted of 1,000 records and 14 initial attributes, including patient identification variables, demographic information, and laboratory measurement variables. The clinical features included Gender, AGE, Urea, Creatinine, HbA1c, Cholesterol, Triglycerides, HDL, LDL, VLDL, BMI, and CLASS as the target label. The target variable represented the diabetes classification status of each patient. Since the study focused on predictive modeling and explainability, only clinically relevant variables were retained for further analysis, while identification attributes were excluded from the modeling process.

2.2 Data Preprocessing

Data preprocessing was conducted to improve data quality and ensure optimal model performance. Initially, irrelevant attributes such as ID and No_Pation were removed from the dataset. Categorical variables such as Gender were transformed into numerical values using Label Encoding. Similarly, the target variable CLASS was encoded into numerical form to support machine learning algorithms. Furthermore, feature scaling was applied using StandardScaler to normalize numerical features. The standardization process is defined as:

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

where: X represents the original value, μ represents the mean, σ represents the standard deviation.

This step ensures that all features contribute equally to the model and improves the performance of distance-based and gradient-based algorithms.

2.3 Dataset Splitting

The dataset was divided into training and testing sets using an 80:20 ratio. Specifically, 80% of the data (800 samples) were used for model training, while 20% (200 samples) were used for testing. A random state parameter was applied to ensure reproducibility of the experimental results. The dataset splitting can be expressed as:

$$D_{train} = 80\%, D_{test} = 20\%$$

2.4 Machine Learning Modeling

This study employs multiple machine learning algorithms to perform diabetes classification, including Decision Tree, Random Forest, Support Vector Machine (SVM), and XGBoost. Decision Tree is used as a baseline model due to its interpretability. Random Forest is an ensemble learning method that combines multiple decision trees to improve prediction stability, with 200 estimators used in this study. Support Vector Machine (SVM) utilizes a Radial Basis Function (RBF) kernel to handle nonlinear relationships in the data. Meanwhile, XGBoost is applied as a gradient boosting algorithm with the following parameters: 200 estimators, maximum depth of 5, learning rate of 0.1, subsample ratio of 0.8, and column sampling rate of 0.8. All models were trained using the same training dataset to ensure a fair comparison of performance.

2.5 Performance Evaluation

Model performance was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score [36]–[39].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - Score = 2\,x\,\frac{Precision\ x\ Recall}{Precision + Recall} \quad (5)$$

Where: True Positive (TP) refers to the number of correctly predicted positive cases, while True Negative (TN) represents the number of correctly predicted negative cases. False Positive (FP) indicates cases that are incorrectly predicted as positive, and False Negative (FN) represents cases that are incorrectly predicted as negative. Precision measures the proportion of correctly predicted positive observations among all predicted positive cases. Recall measures the proportion of correctly identified positive cases among all actual positive cases. Accuracy represents

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

the overall proportion of correct predictions, while F1-score is the harmonic mean of precision and recall, providing a balance between both metrics, especially in imbalanced datasets.

2.6 SHAP Analysis

SHAP (SHapley Additive exPlanations) was used to interpret the predictions of the machine learning model. SHAP is based on cooperative game theory and assigns each feature an importance value for a particular prediction. The SHAP value is mathematically defined as:

$$\emptyset i = \sum_{S \subseteq N\{i\}} \frac{|S|!\,(n - |S| - 1)!}{n!} [f(S \cup \{i\}) - f(S)] \quad (6)$$

where:
$\emptyset i$ is the contribution of feature i,
N is the set of all features,
S is a subset of features,
$f(S)$ is the model prediction based on subset S.
SHAP summary plots were used to evaluate global feature importance, while SHAP beeswarm plots were used to analyze the impact of each feature on individual predictions.

2.7 Interpretation of Results

The final stage involves interpreting both the predictive performance and the explainability results. The best-performing model was selected based on evaluation metrics, particularly accuracy. The SHAP analysis was then used to identify the most influential clinical features affecting diabetes prediction. These findings provide insights into the relationship between clinical variables and diabetes risk. This interpretability enhances the transparency of the machine learning model and supports the application of Explainable Artificial Intelligence (XAI) in clinical decision-making systems.

3. RESULTS AND DISCUSSION

3.1 Dataset Characteristics

The clinical laboratory dataset used in this study consisted of 1,000 patient records containing demographic and laboratory measurement attributes associated with diabetes mellitus. After removing non-relevant identification attributes, the dataset contained clinical features including Gender, Age, Urea, Creatinine (Cr), HbA1c, Cholesterol (Chol), Triglycerides (TG), High-Density Lipoprotein (HDL), Low-Density Lipoprotein (LDL), Very Low-Density Lipoprotein (VLDL), and Body Mass Index (BMI). These variables are commonly utilized in clinical practice to evaluate metabolic conditions and diabetes risk.

Prior to model development, preprocessing procedures were applied to improve data quality and ensure compatibility with machine learning algorithms. Categorical variables were transformed into numerical representations using label encoding, while numerical attributes were standardized using StandardScaler. This normalization process reduced the influence of differing measurement scales and facilitated more effective learning by the classification algorithms. The dataset was subsequently divided into training and testing subsets using an 80:20 ratio. The training subset contained 800 records and was used to develop the machine learning models, while the testing subset consisted of 200 records used to evaluate model performance. This experimental design ensured a fair and unbiased assessment of each classification model. The processed dataset provided a reliable basis for comparing machine learning algorithms and conducting explainability analysis through SHAP.

3.2 Comparative Performance of Machine Learning Models

To identify the most effective classifier for diabetes prediction, four machine learning algorithms were evaluated, namely Decision Tree, Random Forest, Support Vector Machine (SVM), and XGBoost. The performance of each model was measured using Accuracy, Precision, Recall, and F1-Score.

Table 1. Performance Comparison of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.980	0.980	0.980	0.980
Random Forest	0.985	0.985	0.985	0.985
SVM	0.945	0.945	0.945	0.944
XGBoost	0.985	0.985	0.985	0.985

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

Based on Table 2, Random Forest and XGBoost achieved the highest classification performance, both obtaining an accuracy of 98.5%. These results indicate that ensemble learning techniques outperform conventional classification approaches when applied to clinical laboratory data. The ability of ensemble methods to aggregate multiple decision structures allows them to capture complex nonlinear relationships among clinical variables more effectively.

Decision Tree also produced excellent performance with an accuracy of 98.0%, demonstrating that the clinical features contained strong discriminative information. However, its performance remained slightly lower than that of Random Forest and XGBoost due to its susceptibility to overfitting. In contrast, SVM achieved an accuracy of 94.5%, which was the lowest among the evaluated models. Although SVM is widely recognized as an effective classifier, its performance may be influenced by the characteristics of the dataset and parameter settings. The results demonstrate that XGBoost and Random Forest provide the most reliable predictive performance for diabetes classification based on clinical laboratory measurements.

3.3 Confusion Matrix Analysis

To further evaluate the classification performance of each model, confusion matrix analysis was conducted. The confusion matrix provides detailed information regarding correctly and incorrectly classified observations.

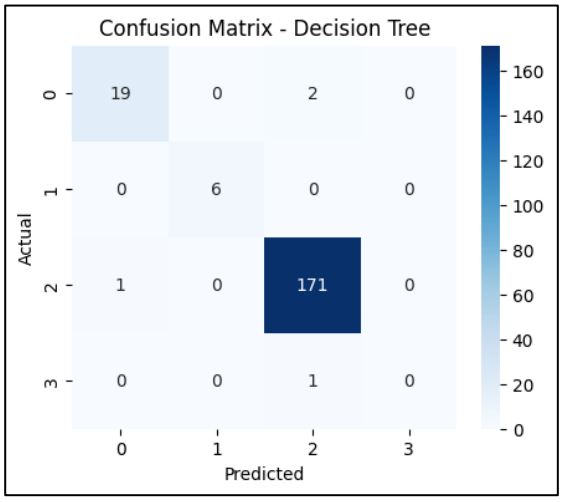

Figure 2. Confusion Matrix – Decision Tree

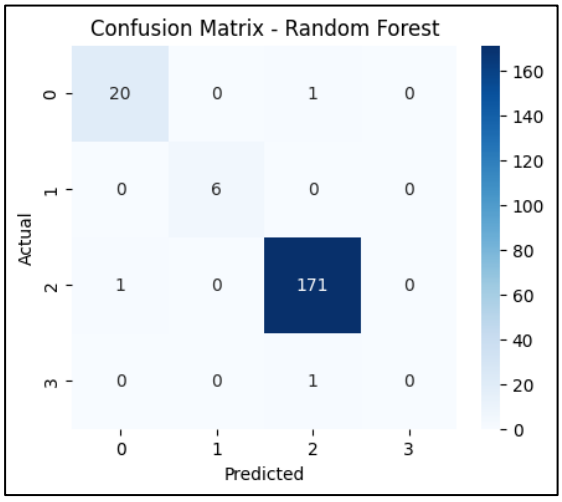

Figure 3. Confusion Matrix – Random Forest

The Decision Tree confusion matrix (Figure 2) demonstrates that the model successfully classified the majority of instances across all classes. Only a small number of misclassifications were observed, contributing to the high overall accuracy of 98.0%. Random Forest (Figure 3) achieved superior classification performance with very few classification errors. The confusion matrix indicates strong predictive consistency across all diabetes categories. The ensemble mechanism effectively reduced variance and improved classification stability.

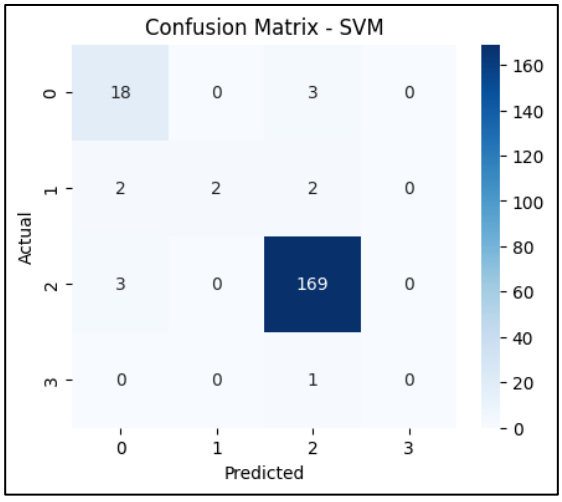

Figure 4. Confusion Matrix – SVM

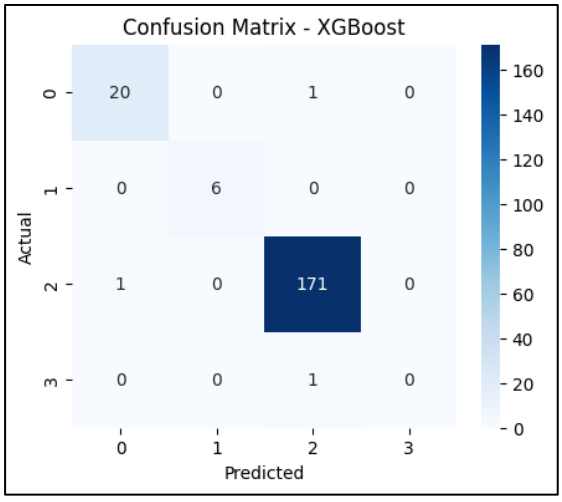

Figure 5. Confusion Matrix – XGBoost

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer **DOI:** https://doi.org/10.47065/explorer.v6i2.2987

The confusion matrix of the SVM model (Figure 4) reveals a higher number of misclassifications compared to the other models. Although the overall accuracy remained acceptable, the model showed limitations in distinguishing certain diabetes categories, which contributed to its lower performance. The XGBoost confusion matrix (Figure 5) exhibited excellent classification capability with only a minimal number of errors. The model successfully identified the majority of diabetic, pre-diabetic, and non-diabetic instances. These findings confirm the robustness of XGBoost in handling clinical laboratory data and justify its selection for subsequent explainability analysis.

Overall, the confusion matrix results are consistent with the quantitative evaluation metrics and further demonstrate the effectiveness of ensemble learning approaches for diabetes classification.

3.4 SHAP-Based Explainability Analysis

Although predictive performance is important, model interpretability is equally critical in healthcare applications. Therefore, SHapley Additive exPlanations (SHAP) were employed to provide transparent explanations regarding the predictions generated by the XGBoost model.

3.4.1 SHAP Summary Plot

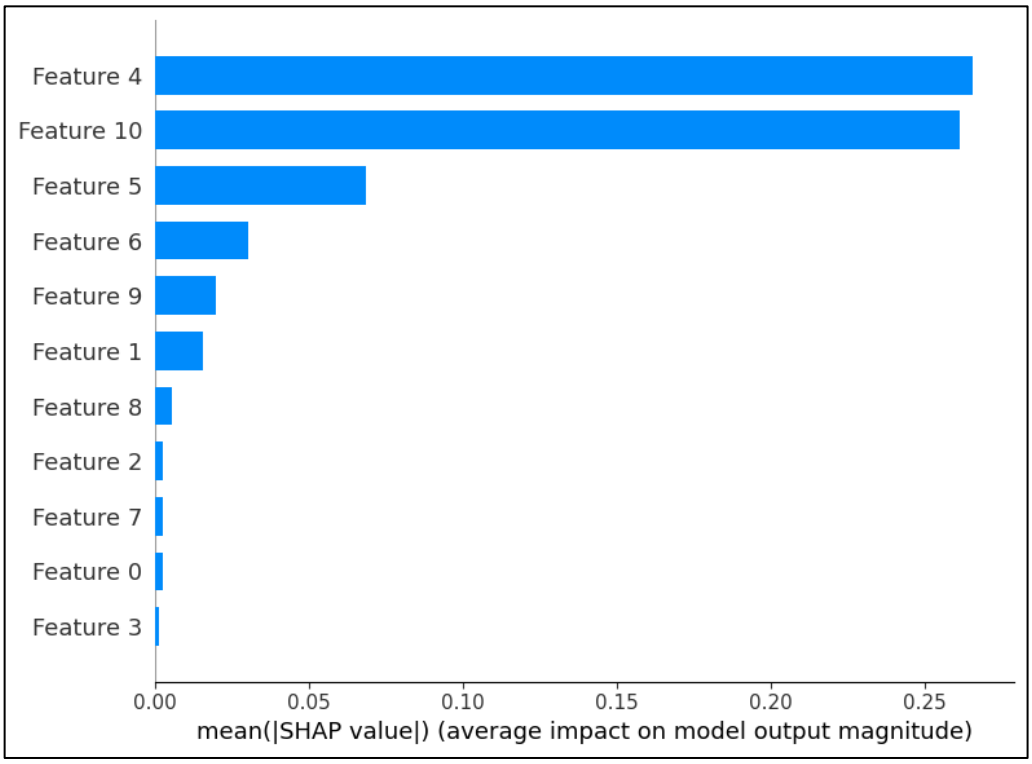

Figure 6. Confusion Matrix – SVM

The SHAP Summary Plot provides a global interpretation of feature importance. The results indicate that Feature 4 contributed most significantly to diabetes classification, followed by Feature 10, Feature 5, Feature 6, and Feature 9. These features consistently exhibited larger SHAP values, indicating stronger influence on prediction outcomes. The distribution of SHAP values further demonstrates that both high and low feature values contribute differently to the classification process. Features with larger absolute SHAP values exert greater influence on the model's predictions.

3.4.2 SHAP Beeswarm Plot

While the SHAP Summary Plot provides a global overview of feature importance across the entire dataset, a more detailed understanding of feature behavior can be obtained through the SHAP Beeswarm Plot. This visualization not only ranks features according to their importance but also illustrates how individual feature values influence model predictions. By analyzing the distribution of SHAP values for each feature, it becomes possible to identify whether higher or lower feature values contribute positively or negatively to diabetes classification. Consequently, the SHAP Beeswarm Plot offers a comprehensive interpretation of the interaction between clinical laboratory variables and prediction outcomes, thereby improving the transparency and reliability of the XGBoost classification model.

EXPLORER
Journal of Computer Science and Information Technology
E-ISSN 2774-4647
Website: https://journal.fkpt.org/index.php/Explorer DOI: https://doi.org/10.47065/explorer.v6i2.2987

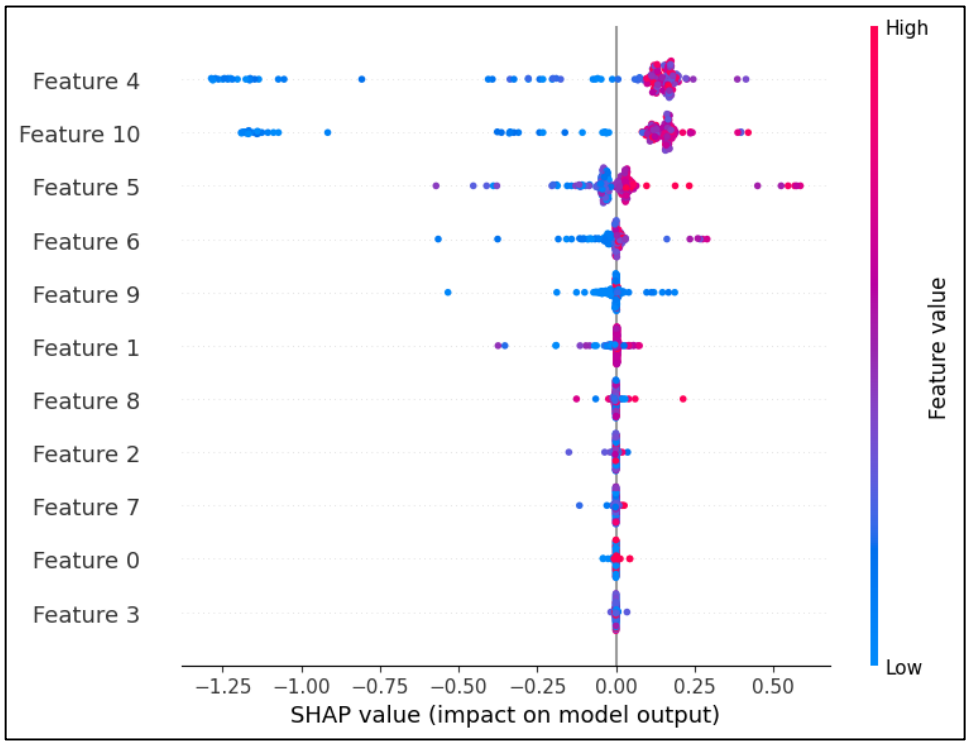

Figure 7. SHAP Beeswarm Plot

The SHAP Beeswarm Plot illustrates the relationship between feature values and their contributions to model predictions. Each point represents an individual observation, while the color gradient indicates the magnitude of the feature value. The plot reveals substantial variability in feature contributions across observations, indicating that diabetes prediction is influenced by complex interactions among clinical variables. Features located near the top of the plot consistently exhibit higher impact on classification outcomes. The SHAP feature importance analysis confirms the dominant role of several laboratory variables in diabetes classification. The ranking obtained from SHAP aligns with medical knowledge, as metabolic indicators typically play a central role in diabetes diagnosis and monitoring. The explainability results demonstrate that the proposed framework not only achieves high predictive performance but also provides meaningful insights into the factors influencing diabetes classification.

3.5 Discussion

The experimental results demonstrate that ensemble learning approaches provide superior performance for diabetes classification using clinical laboratory data. Both Random Forest and XGBoost achieved an accuracy of 98.5%, outperforming Decision Tree and Support Vector Machine. These findings are consistent with previous studies that reported the effectiveness of ensemble learning techniques in healthcare prediction tasks due to their ability to capture complex nonlinear relationships and reduce model variance.

The strong performance of XGBoost may be attributed to its gradient boosting mechanism, which sequentially improves prediction performance by minimizing classification errors. Similarly, Random Forest benefits from the aggregation of multiple decision trees, resulting in improved robustness and generalization capability. In contrast, SVM produced comparatively lower performance, suggesting that kernel-based methods may be less effective than ensemble approaches for the analyzed dataset. An important contribution of this study lies in the integration of explainable artificial intelligence through SHAP analysis. Unlike conventional black-box models, the proposed framework enables the interpretation of feature contributions and provides transparency regarding the decision-making process. This capability is particularly important in healthcare applications, where trust and interpretability are essential requirements for clinical adoption.

The SHAP results revealed that several features exerted substantial influence on diabetes classification. These findings indicate that machine learning models can successfully identify meaningful patterns within clinical laboratory measurements while simultaneously providing interpretable explanations. Consequently, healthcare practitioners may gain valuable insights into the variables associated with diabetes status and utilize these findings to support clinical decision-making.

Overall, the proposed framework successfully combines predictive performance and interpretability, demonstrating the potential of explainable machine learning for diabetes classification based on clinical laboratory

data. The integration of XGBoost and SHAP provides a promising approach for developing transparent artificial intelligence systems capable of supporting healthcare professionals in disease diagnosis and risk assessment.

4. CONCLUSION

This study proposed an explainable diabetes classification framework using machine learning algorithms and SHapley Additive exPlanations (SHAP) analysis on clinical laboratory data. Four machine learning algorithms, namely Decision Tree, Random Forest, Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost), were evaluated and compared using a dataset consisting of 1,000 patient records with demographic and laboratory attributes. The experimental results demonstrated that Random Forest and XGBoost achieved the highest classification performance with an accuracy of 98.5%, outperforming Decision Tree (98.0%) and SVM (94.5%). These findings indicate that ensemble learning approaches provide superior predictive capabilities for diabetes classification based on clinical laboratory data. To enhance model transparency and interpretability, SHAP analysis was applied to the XGBoost model. The explainability results successfully identified the most influential features contributing to diabetes classification and provided insights into how individual variables affect model predictions. The integration of machine learning and explainable artificial intelligence enables not only accurate classification but also meaningful interpretation of prediction outcomes, which is essential in healthcare applications. Overall, the proposed framework demonstrates that combining high-performance machine learning algorithms with SHAP-based explainability can support reliable and interpretable diabetes diagnosis. The findings contribute to the development of transparent artificial intelligence systems for healthcare and may assist medical practitioners in understanding the factors associated with diabetes risk. Future work may focus on addressing class imbalance, incorporating larger multi-center datasets, and exploring advanced explainable deep learning models to further improve classification performance and interpretability.

One limitation of this study is the imbalance in class distribution, which may affect the performance of minority class prediction. Future studies may employ data balancing techniques such as SMOTE and evaluate additional explainable machine learning approaches.

REFERENCRE

[1] S. Chen *et al.*, “The global macroeconomic burden of diabetes mellitus,” *Nature Medicine*, vol. 32, no. January, pp. 126–138, 2026, doi: 10.1038/s41591-025-04027-5.

[2] I. Genitsaridi *et al.*, “11th edition of the IDF Diabetes Atlas : global , regional , and national diabetes prevalence estimates for 2024 and projections for 2050,” *The Lancet Diabetes & Endocrinology*, vol. 14, no. 2, pp. 149–156, 2026, doi: 10.1016/S2213-8587(25)00299-2.

[3] J. Hossain, Al-Mamun, and R. Islam, “Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused,” *Health Science Reports*, vol. 7, no. e2004, pp. 5–9, 2024, doi: 10.1002/hsr2.2004.

[4] S. A. Alinaghian *et al.*, “Burden of type 2 diabetes and its relationship with human development index in Asian countries: Global Burden of Disease Study in 2019,” *BMC Public Health*, vol. 25, no. 402, pp. 1–12, 2025, doi: 10.1186/s12889-025-21608-8.

[5] C. Pan *et al.*, “Global burden of diabetes mellitus 1990–2021: epidemiological trends, geospatial disparities, and risk factor dynamics,” *Frontiers in Endocrinology*, vol. 16, no. July, pp. 1–8, 2025, doi: 10.3389/fendo.2025.1596127.

[6] B. Zhou *et al.*, “Worldwide trends in diabetes prevalence and treatment from 1990 to 2022 : a pooled analysis of 1108 population- representative studies with 141 million participants,” *The Lancet*, vol. 404, no. 10467, pp. 2077–2093, 2024, doi: 10.1016/S0140-6736(24)02317-1.

[7] A. Bawden, “More than 800 million people around the world have diabetes, study finds,” *The Guardian*, 2024. [Online]. Available: https://www.theguardian.com/society/2024/nov/13/diabetes-rates-increase-world-study?

[8] W. M. Team, “Urgent action needed as global diabetes cases increase four-fold over past decades,” *World Health Organization*, 2024. [Online]. Available: https://www.who.int/news/item/13-11-2024-urgent-action-needed-as-global-diabetes-cases-increase-four-fold-over-past-decades

[9] E. Ru, O. Ec, and M. Gi, “Complications Associated with Type 2 Diabetes Mellitus, Pathophysiology, Diagnosis and Management: A Concise Review of Current Literature,” *Saudi Journal of Biomedical Research Abbreviated*, vol. 10, no. 7, pp. 201–244, 2025, doi: 10.36348/sjbr.2025.v10i07.001.

[10] K. Islam, R. Islam, I. Nguyen, H. Malik, and H. Pirzadah, “Diabetes Mellitus and Associated Vascular Disease : Pathogenesis , Complications , and Evolving Treatments,” *Advances in Therapy*, vol. 42, no. 6,

pp. 2659–2678, 2025, doi: 10.1007/s12325-025-03185-9.
[11] Z. Shariat-madar and F. Mahdi, "Potential Molecular Biomarkers for Predicting and Monitoring Complications in Type 2 Diabetes Mellitus," *Molecules*, vol. 30, no. 22, pp. 1–48, 2025, doi: 10.3390/molecules30224448.
[12] G. Manoukian *et al.*, "Post-COVID Syndrome in Patients With Comorbid Hypertension or Diabetes: A Narrative Review of Long-Term Outcomes," *Cureus*, vol. 18, no. 1, pp. 1–12, 2026, doi: 10.7759/cureus.102117.
[13] O. B. Ayoade, S. Shahrestani, and C. Ruan, "Comparative Epidemiology of Machine and Deep Learning Diagnostics in Diabetes and Sickle Cell Disease: Africa's Challenges, Global Non-Communicable Disease Opportunities," *Electronics*, vol. 15, no. 2, pp. 1–41, 2026, doi: 10.3390/electronics15020394.
[14] S. Cassambai *et al.*, "Strengthening diabetes care in sub-Saharan Africa: health systems, digital innovations, lifestyle interventions, and socioeconomic dimensions," *The Lancet Diabetes and Endocrinology*, vol. 14, no. 6, pp. 523–532, 2026, doi: 10.1016/S2213-8587(26)00035-5.
[15] D. Ikhile, S. Seidu, D. Omodara, J. C. Katte, K. Ramaiya, and K. Khunti, "Diabetes-related complications and multiple long-term conditions in sub-Saharan Africa: determinants and management strategies," *The Lancet Diabetes and Endocrinology*, vol. 14, no. 6, pp. 512–522, 2026, doi: 10.1016/S2213-8587(26)00036-7.
[16] A. Mannucci and A. Goel, "Advances in pancreatic cancer early diagnosis, prevention, and treatment: The past, the present, and the future," *Ca-A Cancer Journal for Clinicians*, vol. 76, no. 1, pp. 1–30, 2026, doi: 10.3322/caac.70035.
[17] X. Yao, D. Huang, X. Shi, and F. Xiang, "Clinical ef fi cacy , safety , and predictors of treatment response to SGLT2 versus DPP- 4 inhibitors in type 2 diabetes : a retrospective comparative study," *Frontiers in Endocrinology*, vol. 17, no. May, pp. 1–15, 2026, doi: 10.3389/fendo.2026.1818019.
[18] I. Z. Sadiq *et al.*, "Data-driven diabetes mellitus prediction and management: a comparative evaluation of decision tree classifier and artificial neural network models along with statistical analysis," *Scientific Reports*, vol. 15, no. 19339, pp. 1–16, 2025, doi: 10.1038/s41598-025-03718-w.
[19] Y. Zhuang *et al.*, "The effect of multidisciplinary team and experience-based co-design on the care of older adult patients with type 2 diabetes: A randomized controlled trial," *Diabetes Research and Clinical Practice*, vol. 221, no. March, p. 112028, 2025, doi: 10.1016/j.diabres.2025.112028.
[20] L. N. Adam, "Association between dyslipidemia and glycated hemoglobin in diabetic patients from Zakho General Hospital, Iraq," *The Egyptian Journal of Internal Medicine*, vol. 37, no. 70, pp. 1–6, 2025, doi: 10.1186/s43162-025-00460-7.
[21] S. Singh, N. A. Wani, R. Kumar, and J. Bedi, "DiaXplain: A transparent and interpretable artificial intelligence approach for Type-2 diabetes diagnosis through deep learning," *Computers and Electrical Engineering*, vol. 126, no. August, p. 110470, 2025, doi: 10.1016/j.compeleceng.2025.110470.
[22] S. Mukherjee and D. De, "XAI-IoT based real-time diabetes prediction using TabTransformer," *Medicine in Novel Technology and Devices*, vol. 30, no. May, p. 100435, 2026, doi: 10.1016/j.medntd.2026.100435.
[23] M. Elfayoumi, M. AbouElazm, O. Mohamed, T. Abuhmed, and S. El-Sappagh, "Knowledge Augmented Significant Language Model-Based Chatbot for Explainable Diabetes Mellitus Prediction," in *2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, 2025, pp. 1–8. doi: 10.1109/IMCOM64595.2025.10857525.
[24] O. O. Awe, P. N. Mwangi, S. K. Goudoungou, R. V. Esho, and O. S. Oyejide, "Explainable AI for enhanced accuracy in malaria diagnosis using ensemble machine learning models," *BMC Medical Informatics and Decision Making*, vol. 25, no. 162, pp. 1–26, 2025, doi: 10.1186/s12911-025-02874-3.
[25] H. Fan, Z. Li, N. Yan, and H. Liu, "Scientific Reports Article in Press An interpretable progressive residual network for automated multiclass diabetes diagnosis IN AR IN," *Scientific Reports*, vol. May, pp. 1–50, 2026, doi: 10.1038/s41598-026-51603-x.
[26] A. Ashraf *et al.*, "A preprocessing-enhanced stacking classifier for generalized cardiovascular disease detection across diverse datasets," *Scientific Reports*, vol. 16, no. 16206, pp. 1–16, 2026, doi: 10.1038/s41598-026-41042-z.
[27] Z. Ding *et al.*, "Trade-offs between machine learning and deep learning for mental illness detection on social media," *Scientific Reports*, vol. 15, no. 14497, pp. 1–14, 2025, doi: 10.1038/s41598-025-99167-6.
[28] L. Jiang *et al.*, "Assessment of six insulin resistance surrogate indexes for predicting stroke incidence in Chinese middle-aged and elderly populations with abnormal glucose metabolism: a nationwide prospective cohort study," *Cardiovascular Diabetology*, vol. 24, no. 56, pp. 1–13, 2025, doi: 10.1186/s12933-025-02618-7.
[29] R. Hasan, V. Dattana, S. Mahmood, and S. Hussain, "Towards Transparent Diabetes Prediction : Combining AutoML and Explainable AI for Improved Clinical Insights," *Information*, vol. 16, no. 7, pp. 1–31, 2025, doi: 10.3390/info16010007.

[30] M. Kutlu, T. B. Donmez, and C. Freeman, "Machine Learning Interpretability in Diabetes Risk Assessment: A SHAP Analysis," *Computers and Electronics in Medicine*, vol. 1, no. 1, pp. 34–44, 2024, doi: 10.69882/adba.cem.2024075.

[31] M. Islam, H. R. Rifat, S. Bin Shahid, A. Akhter, A. Uddin, and K. M. M. Uddin, "Explainable Machine Learning for Efficient Diabetes Prediction Using Hyperparameter Tuning, SHAP Analysis, Partial Dependency, and LIME," *Engineering Reports*, vol. 7, no. e13080, pp. 1–22, 2024, doi: 10.1002/eng2.13080.

[32] P. Netayawijit, W. Chansanam, and K. Sorn-In, "Interpretable Machine Learning Framework for Diabetes Prediction: Integrating SMOTE Balancing with SHAP Explainability for Clinical Decision Support," *Healthcare*, vol. 13, no. 20, pp. 1–26, 2025, doi: 10.3390/healthcare13202588.

[33] Z. Rafie *et al.*, "Leveraging XGBoost and explainable AI for accurate prediction of type 2 diabetes," *BMC Public Health*, vol. 25, no. 3688, pp. 1–17, 2025, doi: 10.1186/s12889-025-24953-w.

[34] R. Agarwal, K. K. Gangwar, A. Pandey, V. Sharma, and A. Sanghi, "Integrating XGBoost with Explainable AI Techniques for Diabetes Prediction: A Comparative Study," in *2025 14th International Conference on System Modeling & Advancement in Research Trends (SMART)*, IEEE, 2025, pp. 513–519. doi: 10.1109/SMART66937.2025.11389515.

[35] S. Ahmad, S. Hussain, M. Arif, and M. A. Ansari, "Early-Stage Diabetes Prediction Using a Stacked Ensemble Model Enhanced with SHAP Explainability," *Biomedical and Pharmacology Journal*, vol. 19, no. 1, pp. 246–263, 2026, doi: 10.13005/bpj/3350.

[36] A. Wanto, Y. Yuhandri, and O. Okfalisa, "RetMobileNet: A New Deep Learning Approach for Multi-Class Eye Disease Identification," *Revue d'Intelligence Artificielle*, vol. 38, no. 4, pp. 1055–1067, 2024, doi: 10.18280/ria.380401.

[37] S. P. Siregar, I. Akbari, Poningsih, A. Wanto, and Solikhun, "Enhancing Tomato Leaf Disease Detection via Optimized VGG16 and Transfer Learning Techniques," *Jurnal RESTI*, vol. 9, no. 3, pp. 570–580, 2025, doi: 10.29207/resti.v9i3.6410.

[38] A. I. R. H. Barat, W. S. Astuti, D. Hartama, A. P. Windarto, and A. Wanto, "Enhanced Flower Image Classification Using Mobilenetv2 With Optimized Performance," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 11, no. 1, pp. 180–191, 2025, doi: 10.33480/jitk.v11i1.6497.

[39] H. R. M. W. Jasa, A. Wanto, and R. K. Sormin, "Optimization of ShuffleNetV2 Using Self-Knowledge Distillation for Cocoa Fruit Disease Classification," *Jurnal Teknik Informatika (JUTIF)*, vol. 7, no. 3, pp. 2670–2689, 2026, doi: 10.52436/1.jutif.2026.7.3.5649.