Journal of Golden Generation Multidisciplinary
Volume: 2, Nomor: 3, (2026), Juni: 15 – 33
Link Publish: https://ejournal.lppnusantara.com/index.php/jggmultidisciplinary/article/view/754
E-ISSN: 3123-6278

COMPARATIVE ANALYSIS OF EXPLAINABLE AI USING LIME AND SHAP FOR DIABETES PREDICTION BASED ON LIFESTYLE FACTORS

Ricky Salim [1], **Agung Mulyo Widodo** [2]
Universitas Esa Unggul, Jakarta, Indonesia [1,2]
Corresponding Author: rickysalim81@student.esaunggul.ac.id [1*]

Info Artikel
Submitted: *19 Maret 2026*
Revised : *31 Maret 2026*
Accepted: *25 April 2026*
Published: *30 April 2026*

Keywords: *Artificial Neural Network, Explainable AI, LIME, SHAP*

Abstract

The rapid advancement of artificial intelligence (AI) has significantly impacted the healthcare sector, particularly in supporting the early detection of diabetes; however, many AI models still face challenges due to their black-box nature, where decision-making processes are not easily understood. This study aims to compare two Explainable Artificial Intelligence (XAI) methods, namely Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive Explanations (SHAP), in interpreting the prediction results of an Artificial Neural Network (ANN) model using the Diabetes Health Indicator dataset. Prior to modeling, the data were preprocessed through cleaning and normalization to ensure quality and consistency. The trained ANN model was then analyzed using LIME and SHAP to evaluate the contribution of each feature to the prediction outcomes. The results show that both methods are capable of providing meaningful and interpretable explanations, although SHAP demonstrates more consistent and stable interpretations across the dataset. These findings highlight the importance of integrating XAI techniques to enhance model transparency, thereby increasing trust and supporting more reliable decision-making in clinical settings, particularly for diabetes diagnosis.

Abstrak

Kemajuan pesat kecerdasan buatan (AI) telah berdampak signifikan pada sektor kesehatan, khususnya dalam mendukung deteksi dini diabetes; namun, banyak model AI masih menghadapi tantangan karena sifatnya yang seperti kotak hitam, di mana proses pengambilan keputusan tidak mudah dipahami. Studi ini bertujuan untuk membandingkan dua metode Kecerdasan Buatan yang Dapat Dijelaskan (XAI), yaitu Penjelasan Model-Agnostik yang Dapat Diinterpretasikan Secara Lokal (LIME) dan Penjelasan Aditif SHapley (SHAP), dalam menginterpretasikan hasil prediksi model Jaringan Saraf Buatan (ANN) menggunakan dataset Indikator Kesehatan Diabetes. Sebelum pemodelan, data diproses terlebih dahulu melalui pembersihan dan normalisasi untuk memastikan kualitas dan konsistensi. Model ANN yang telah dilatih kemudian dianalisis menggunakan LIME dan SHAP untuk mengevaluasi kontribusi setiap fitur terhadap hasil prediksi. Hasil menunjukkan bahwa kedua metode mampu memberikan penjelasan yang bermakna dan dapat diinterpretasikan, meskipun SHAP menunjukkan interpretasi yang lebih konsisten dan stabil di seluruh dataset. Temuan ini menyoroti pentingnya mengintegrasikan teknik XAI untuk meningkatkan transparansi model, sehingga meningkatkan kepercayaan dan mendukung pengambilan keputusan yang lebih andal dalam pengaturan klinis, khususnya untuk diagnosis diabetes.

Kata Kunci: Jaringan Saraf Buatan, AI yang Dapat Dijelaskan, LIME, SHAP

Publisher: Lembaga Penerbit Penelitian Nusantara

Introduction

The accelerated advancement of Artificial Intelligence (AI) has driven significant transformative changes across multiple sectors, particularly in healthcare, finance, education, and governance [1][2]. In the medical domain, AI is increasingly deployed to facilitate disease diagnosis, risk prediction, and clinical decision support, offering speed and accuracy that augment human expertise [3][4].

Among the wide range of AI algorithms, Artificial Neural Networks (ANN) have become a leading option owing to their capacity to capture intricate, non-linear interactions among variables, resulting in highly accurate predictive outputs [5][6]. In particular, ANN is widely utilized for the classification of chronic conditions such as diabetes a disease with a high global prevalence [7].

The selection of ANN in this research is driven by its capacity to capture intricate patterns between lifestyle factors and health risks that traditional linear or tree-based models often fail to map optimally [8] Furthermore, ANN exhibits strong scalability in processing large-scale datasets, such as the Diabetes Health Indicators dataset, where the model conducts automatic feature extraction via hidden layers to identify underlying risk patterns [9].

Despite its predictive capabilities, ANN is often described as a "black box" due to its internal decision-making mechanisms being difficult for human users to interpret [10][11]. This opacity represents more than a technical hurdle; it constitutes a considerable risk to patient safety and medical accountability. In clinical settings, physicians and researchers must comprehend the reasoning underlying a prediction to ensure decisions are scientifically and ethically defensible [12][13]. The lack of explainability in model outputs can undermine trust and potentially result in fatal medical errors [14].

This "responsibility gap" complicates the clinical justification of AI-driven predictions, ultimately hindering the integration of AI technology within healthcare. Consequently, there is an urgent need for approaches that make AI models more transparent and interpretable [15].

Explainable AI (XAI) has developed into a dedicated domain aimed at overcoming these challenges [16][17]. XAI seeks to offer a more transparent interpretation of how AI produces predictions, both at the local and global levels [18]. Two widely recognized XAI methods are Local Interpretable Model-agnostic Explanations (LIME) and Shapley Additive Explanations (SHAP) [19][20].

LIME functions by constructing a simplified local surrogate model to interpret individual predictions, whereas SHAP applies principles from game theory to determine the exact contribution of each feature to the final output [21]. These approaches enable medical practitioners to better understand the reasoning underlying ANN-based decisions [15][22].

While previous studies have highlighted the predictive accuracy of ANN in diabetes detection, comparative evaluations concerning the effectiveness of LIME and SHAP in interpreting these outcomes remain limited [23][21]. Some researchers, such as [24], have developed ensemble approaches with XAI components; however, many lack a granular focus on the specific behavior of ANN architectures.

This study addresses this gap by concentrating specifically on the application and comparative evaluation of LIME and SHAP within ANN models for diabetes prediction. By utilizing the Diabetes Health Indicator dataset, which includes essential features such as glucose levels, BMI, age, and lifestyle factors, this research seeks to convert the ANN from an opaque system into a transparent and clinically accountable tool. Specifically, this study aims to address the following research questions:

1. In what ways can the "black box" characteristics of ANN be reduced to ensure transparency and medical accountability for healthcare practitioners?
2. What is the level of consistency in feature interpretation between LIME and SHAP when explaining lifestyle factors in ANN-based diabetes predictions?
3. Which method LIME or SHAP provides superior computational performance and reliability for clinical decision-making?

Ultimately, this research contributes to the advancement of XAI in the medical domain, highlighting the necessity of transparency and accountability in the implementation of artificial intelligence [25][16].

Method

A. Proposed Method

This study adopts a comparative research design to evaluate two Explainable Artificial Intelligence (XAI) techniques, namely LIME and SHAP, in interpreting predictions generated by an Artificial Neural Network (ANN) model for diabetes detection. Comparative research is used to identify differences between methods and assess their impact on research outcomes [83].

A Planned Approach is applied to ensure a structured research process, covering stages from problem identification to evaluation, allowing systematic and logical analysis.

The research workflow is adapted to a data science context and aims to estimate diabetes risk based on clinical and lifestyle factors using an ANN model, followed by interpretation using XAI techniques. The stages of the research are illustrated in FIGURE 1, which represents the overall process from initial observation to evaluation and further development.

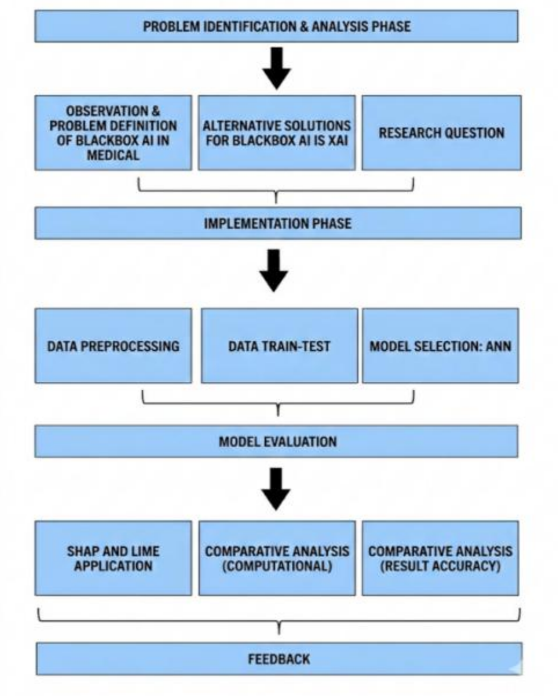


Figure 1. Planned Approach Stages (Thierauf & Klecamp, 1985**)**

B. Research Stages

The study begins with problem identification, where the increasing prevalence of diabetes is recognized as a major health concern. The objective is to develop a prediction model based on lifestyle variables such as physical activity, diet, smoking, and alcohol consumption [50][53][54][56].

Data preparation includes cleaning, feature selection, and dataset splitting (70% training, 30% testing). Min-Max normalization is applied to standardize feature values, while class imbalance is handled using SMOTE to improve minority class representation [22][32][34].

Model development involves constructing an ANN architecture with two hidden layers (8 and 4 neurons, ReLU activation) and a sigmoid output layer [5][9][36]. The model is trained using the Adam optimizer and evaluated using accuracy, precision, recall, F1-score, and ROC-AUC to ensure performance reliability [66].

C. Dataset Description

This study uses the Diabetes Health Indicators dataset from Kaggle, derived from the 2015 BRFSS survey. After preprocessing, 253,680 observations with 22 variables are retained. The analysis focuses on lifestyle-related factors influencing diabetes risk, including physical activity, diet, smoking, and alcohol consumption [50][53][54][57][56].

D. Data Preprocessing

The preprocessing stage includes statistical analysis to describe the dataset distribution. This involves calculating mean, median, standard deviation, and other descriptive measures.

In addition, Kendall's Tau correlation is used to analyze relationships between variables, especially for ordinal and binary data [93].

E. Data Training And Testing

The dataset is split into training (70%) and testing (30%) sets. The target variable is binary, representing diabetes (1) and non-diabetes (0).

Min-Max normalization is applied to scale the data into a uniform range, improving training efficiency and model stability [32][34][35].

To handle class imbalance, SMOTE is applied to generate synthetic samples for the minority class [22]. This improves the model's ability to recognize underrepresented cases. Cross-validation is also used to evaluate model generalization and reduce overfitting.

F. Model Selection: Ann

The ANN model consists of an input layer, hidden layers, and an output layer [5][9]. This model is chosen due to its ability to capture complex non-linear relationships between variables [36].

The architecture includes:

- Input layer: 5 lifestyle features
- Hidden layer 1: 8 neurons (ReLU)
- Hidden layer 2: 4 neurons (ReLU)
- Output layer: 1 neuron (Sigmoid)

The model uses the Adam optimizer, 20 epochs, batch size of 128, and Binary Cross-Entropy loss. This configuration is designed to balance predictive performance and interpretability analysis.

G. Model Evalutation

Model performance is evaluated using accuracy, precision, recall, F1-score, and ROC-AUC [66]. These metrics measure classification performance and prediction reliability.

The formulas remain unchanged as in your original version and should be preserved exactly. The equations for the aforementioned metrics are defined as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

Where:

1. TP (True Positive): Individuals with diabetes/pre-diabetes correctly identified as diabetes/pre-diabetes.
2. FP (False Positive): Individuals without diabetes incorrectly classified as diabetes/pre-diabetes.
3. TN (True Negative): Individuals without diabetes correctly classified as non-diabetic.
4. FN (False Negative): Individuals with diabetes/pre-diabetes incorrectly classified as non-diabetic.

H. Implementation of Shap and Lime

SHAP is applied to measure both global and local feature importance, providing a comprehensive understanding of model behavior. Computational time is also recorded to assess efficiency.

LIME is then applied to generate local explanations for individual predictions. Similar to SHAP, runtime performance is evaluated.

I. Comparative Analysis of Shap and Lime

The comparison focuses on:

1. Computational Performance: An evaluation of the processing speed and efficiency of the interpretation process for both methods.
2. Consistency of Feature Rankings: A comparison of feature importance orderings produced through SHAP and LIME for determining the degree of convergence or divergence between the two approaches.

This analysis provides insight into the strengths and limitations of each method in explaining ANN predictions.

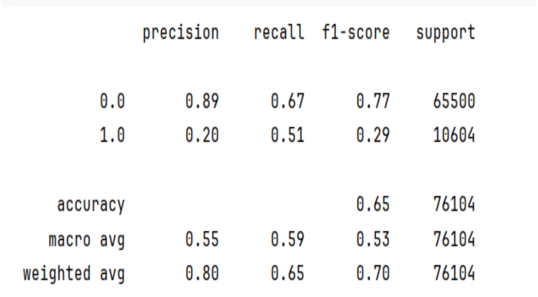
J. Feedback and Conclusion

The final stage evaluates overall findings and provides recommendations for selecting appropriate XAI methods. Future work may include dataset expansion and integration into decision support systems

Result and Disccusion

A. Performance Evaluation of The Ann Model

Prior to the interpretation stage using Explainable Artificial Intelligence (XAI), the performance of the Artificial Neural Network (ANN) model was evaluated to ensure that it achieved a reliable level of predictive capability. Based on the confusion matrix presented in FIGURE 2, the model achieved an overall accuracy of 65% on the testing dataset consisting of 76,104 instances. This level of performance can be categorized as moderate and is commonly observed in classification tasks involving imbalanced datasets and lifestyle-related health variables, which tend to exhibit complex and non-linear relationships [27][29][32].



```
              precision    recall  f1-score   support

         0.0       0.89      0.67      0.77     65500
         1.0       0.20      0.51      0.29     10604

    accuracy                           0.65     76104
   macro avg       0.55      0.59      0.53     76104
weighted avg       0.80      0.65      0.70     76104
```

Figure 2. Confusion Matrix Analysis

A more detailed analysis shows that the model performs better in identifying non-diabetic individuals, achieving a precision of 0.89 and a recall of 0.67. In contrast, for the diabetes class, the model obtained a recall of 0.51 but a relatively low precision of 0.20. This indicates that while the model is able to identify more than half of the actual diabetic cases, it also produces a considerable number of false positives. From a healthcare perspective, such a trade-off is often acceptable in early-stage screening systems, where sensitivity is prioritized to minimize missed diagnoses [50][53][54].

However, the low precision highlights a critical limitation, emphasizing the need for interpretability techniques to better understand the model's decision-making process. This is particularly important given the black-box nature of neural networks, which often lack transparency in explaining how predictions are generated [10][12][18].

The training dynamics illustrated in FIGURE 10 further confirm the stability of the model. The loss curve shows a clear convergence pattern after approximately 10 epochs, while the ROC-

AUC remains stable at around 0.62, indicating moderate discriminative capability. The absence of significant fluctuations in performance metrics suggests that the model does not suffer from severe overfitting, a common challenge in ANN-based modeling [36][46].

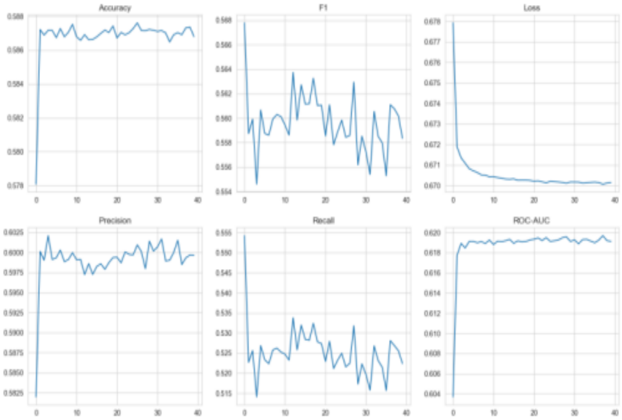


Figure 3. Model Training Dynamics

B. Interpretability Using Lime

LIME provides local explanations by LIME (Local Interpretable Model-Agnostic Explanations) was applied to provide local interpretations of the ANN model's predictions. As shown in FIGURE 4, the explanation highlights that physical inactivity, smoking, and alcohol-related variables contribute positively to the prediction of diabetes risk, while fruit consumption exhibits a protective effect.

These findings are consistent with existing epidemiological evidence indicating that sedentary behavior and smoking significantly increase the risk of Type 2 diabetes, whereas healthy dietary patterns contribute to risk reduction [50][53][54]. However, it is important to note that LIME explanations are instance-specific and rely on local approximations of the model, meaning that the results cannot be generalized across the entire dataset [22][24].

The prominence of certain features, particularly alcohol consumption, reflects LIME's sensitivity to local variations. This characteristic enables LIME to capture sharp changes in feature influence for individual cases, but it may also lead to inconsistencies when compared to global interpretation methods.

Figure 4. LIME tabular explainer

C. Interpretability Using Shap

To complement the local explanations provided by LIME, SHAP (SHapley Additive exPlanations) was employed to analyze both local and global feature contributions. The force plot in FIGURE 5 demonstrates how individual features shift the prediction from a baseline value toward a higher probability of diabetes. Features such as physical inactivity and smoking increase the prediction value, while vegetable consumption contributes negatively, indicating a protective effect.

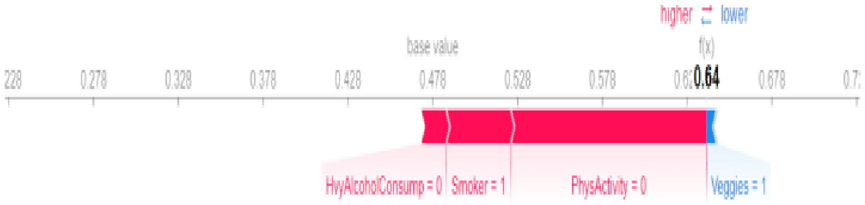


Figure 5. SHAP Explanation in Force Plot

The SHAP waterfall plot in FIGURE 6 further quantifies these contributions, confirming that physical inactivity is the most influential factor, followed by smoking. These results are consistent with prior research highlighting the role of lifestyle behaviors as primary determinants of diabetes risk [50][54].

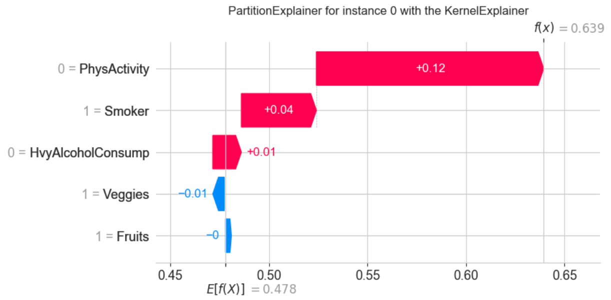


Figure 6. SHAP Explanation using Waterfall Plot

At a broader level, the SHAP clustering visualization in FIGURE 7 reveals heterogeneous patterns across different groups of individuals. The variation in feature contributions suggests that the model captures complex, non-linear interactions among lifestyle variables, rather than relying on uniform relationships. This capability is one of the key strengths of ANN models in handling real-world health data [36].

Additionally, the SHAP violin plot in FIGURE 7 provides a global perspective on feature impact distribution. The visualization shows a clear separation between low and high feature values, particularly for physical activity and smoking, indicating strong and consistent effects on prediction outcomes. Such patterns reinforce the reliability of SHAP in identifying dominant predictors across the population [44].

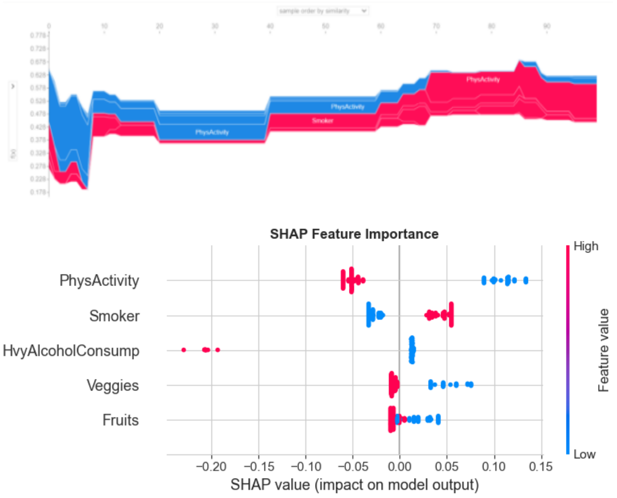


Figure 7. SHAP Explanation in Supervised Clustering

D. Comparative Analysis of Shap and Lime

The comparative results between SHAP and LIME, as illustrated in FIGURE 16, highlight both similarities and differences in their interpretability approaches. Both methods consistently identify physical activity and smoking as the most influential features, demonstrating strong agreement and reinforcing the robustness of these variables as key predictors of diabetes risk.

However, differences emerge in feature ranking. SHAP, which is based on Shapley values, provides a more stable and consistent global interpretation by distributing feature contributions fairly across all variables [22][24]. In contrast, LIME focuses on local approximations, resulting in higher sensitivity to specific instances. This explains why alcohol consumption appears more dominant in LIME, despite having a relatively moderate impact in SHAP analysis.

From a computational perspective, SHAP demonstrates significantly higher efficiency, requiring only 21.92 seconds compared to 442.2 seconds for LIME. This difference reflects the optimization of SHAP in handling large datasets, whereas LIME relies on repeated local sampling, which Increases Computational Cost.

E. Discussion of Xai Results

The comparative evaluation summarized in TABLE 6 indicates that SHAP outperforms LIME in terms of computational efficiency, stability, and interpretability scope. SHAP provides deterministic results based on a strong mathematical foundation, ensuring consistency across repeated analyses. In contrast, LIME exhibits stochastic behavior due to its reliance on sampling techniques, which may lead to variability in results [22][66].

Despite these differences, both methods play complementary roles. SHAP is more suitable for global model interpretation and large-scale clinical applications, while LIME is valuable for analyzing individual cases and detecting local anomalies.

Importantly, both approaches consistently highlight physical inactivity and smoking as the most influential risk factors. This alignment with established medical literature strengthens the validity of the ANN model and demonstrates that, despite its black-box nature, the model captures clinically meaningful patterns [50][53][54].

The observed variation in alcohol-related features further indicates the presence of complex, non-linear interactions within the dataset. Such findings confirm that ANN models are capable of uncovering hidden relationships that may not be easily identified using traditional statistical methods [36].

TABLE 6. COMPARISON MATRIX OF LIME AND SHAP PARAMETERS

PARAMETER	LIME	SHAP	KEY FINDINGS
Computational Time	442.2 seconds (44.22 ms/instance)	21.92 seconds (2.19 ms/instance)	SHAP operates approximately 20 times faster than LIME in processing this dataset.
Feature Consistency	Sensitive to local anomalies	Stable across population trends (PhysActivity)	SHAP is more reliable in capturing the

	(Heavy Alcohol)		broader landscape of risk predictors.
Result Stability	Stochastic (Results vary based on sampling)	Deterministic (Results remain constant based on mathematical foundation)	SHAP provides higher interpretational certainty for medical audits.
Interpretation Scope	Focuses on Local Explanation	Covers both Local and Global Explanations	SHAP provides comprehensive insights from the individual to the population level.
Visual Complexity	Simple (Contribution Bar Plots)	Comprehensive (Waterfall, Force, Violin, Clustering)	SHAP is capable of presenting feature impact distributions in greater depth and detail.

CONCLUSION

This study evaluates the interpretability of a diabetes prediction model by applying Explainable Artificial Intelligence (XAI) methods, namely SHAP and LIME. The results show that lifestyle factors, particularly physical activity and smoking, are the most influential predictors of diabetes risk, which is consistent with previous epidemiological findings [72][75][77]. In terms of technical performance, SHAP demonstrates better efficiency and stability, requiring significantly less computation time and providing consistent interpretations due to its Shapley value foundation [22][44]. Meanwhile, LIME offers valuable local insights by capturing specific patterns, such as the strong influence of alcohol consumption in certain cases, which may not be fully reflected in global analysis. Overall, the findings indicate that combining SHAP and LIME provides a more comprehensive understanding of model behavior, improving transparency and supporting the development of reliable AI-based systems in healthcare.

REFERENCES

[1] Y. Suh, A. Chandramowlishwaran, and Y. Won, "Recent progress of artificial intelligence for liquid-vapor phase change heat transfer," npj Computational Materials, vol. 10, no. 1, pp. 1–14, Mar. 2024, doi: https://doi.org/10.1038/s41524-024-01223-8.

[2] S. Das, A. Dey, A. Pal, and N. Roy, "Applications of Artificial Intelligence in Machine Learning: Review and Prospect," International Journal of Computer Applications, vol. 115, no. 9, pp. 31–41, Apr. 2015, doi: https://doi.org/10.5120/20182-2402.

[3] M. Nguyen, T. He, L. An, D. C. Alexander, J. Feng, and B. T. T. Yeo, "Predicting Alzheimer's disease progression using deep recurrent neural networks," NeuroImage, vol. 222, no. 9, p. 117203, Nov. 2020, doi: https://doi.org/10.1016/j.neuroimage.2020.117203.

[4] Q. An, S. Rahman, J. Zhou, and J. J. Kang, "A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges," Sensors, vol. 23, no. 9, p. 4178, Jan. 2023, doi: https://doi.org/10.3390/s23094178.

[5] A. P. Engelbrecht, Computational Intelligence. John Wiley & Sons, 2007.

[6] S. S. Fiyadh et al., "A comprehensive review on modelling the adsorption process for heavy metal removal from waste water using artificial neural network technique," Heliyon, vol. 9, no. 4, pp. e15455–e15455, Apr. 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e15455.

[7] Parvathaneni Naga Srinivasu, S. Ahmed, Mahmoud Hassaballah, and Naif Almusallam, "An Explainable Artificial Intelligence Software System for Predicting Diabetes," Heliyon, vol. 10, no. 16, pp. e36112–e36112, Aug. 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e36112.

[8] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep Learning for healthcare: review, Opportunities and Challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, 2018, doi: https://doi.org/10.1093/bib/bbx044.

[9] [9] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," Heliyon, vol. 4, no. 11, p. e00938, Nov. 2018, doi: https://doi.org/10.1016/j.heliyon.2018.e00938.

[10] W. J. von Eschenbach, "Transparency and the Black Box Problem: Why We Do Not Trust AI," Philosophy & Technology, vol. 34, no. 4, pp. 1607–1622, Sep. 2021, doi: https://doi.org/10.1007/s13347-021-00477-0.V. L. Raposo, "T

[11] he fifty shades of black: about black box AI and explainability in healthcare," Medical Law Review, vol. 33, no. 1, Feb. 2025, doi: https://doi.org/10.1093/medlaw/fwaf005.

[12] D. S. Watson et al., "Clinical applications of machine learning algorithms: beyond the black box," BMJ, vol. 364, p. l886, Mar. 2019, doi: https://doi.org/10.1136/bmj.l886.

[13] A. J. London, "Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability," Hastings Center Report, vol. 49, no. 1, pp. 15–21, Jan. 2019, doi: https://doi.org/10.1002/hast.973.

[14] D. Heaven, "Why deep-learning AIs are so easy to fool," Nature, vol. 574, no. 7777, pp. 163–166, Oct. 2019, doi: https://doi.org/10.1038/d41586-019-03013-5.

[15] A. A. Adeniran, A. P. Onebunne, and P. William, "Explainable AI (XAI) in healthcare: Enhancing trust and transparency in critical decision-making," World Journal of Advanced Research and Reviews, vol. 23, no. 3, pp. 2447–2658, Sep. 2024, doi: https://doi.org/10.30574/wjarr.2024.23.3.2936.

[16] D. Gunning, E. Vorm, J. Y. Wang, and M. Turek, "DARPA's Explainable AI (XAI) program: a Retrospective," Applied AI Letters, vol. 2, no. 4, Dec. 2021, doi: https://doi.org/10.1002/ail2.61.

[17] F. Doshi-Velez and B. Kim, "Towards A Rigorous Science of Interpretable Machine Learning," arXiv:1702.08608 , vol. 2, no. 2, Mar. 2017, Available: https://arxiv.org/abs/1702.08608

[18] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," Artificial Intelligence, vol. 267, pp. 1–38, Feb. 2019, doi: https://doi.org/10.1016/j.artint.2018.07.007.

[19] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," arXiv.org, Feb. 16, 2016. https://arxiv.org/abs/1602.04938

[20] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," arXiv:1705.07874 , vol. abs/1705.07874, Nov. 2017, Available: https://arxiv.org/abs/1705.07874

[21] A. Salih et al., "Commentary on explainable artificial intelligence methods: SHAP and LIME," arXiv.org, May 08, 2023. https://arxiv.org/abs/2305.02012

[22] P. B. Khokhar, V. Pentangelo, F. Palomba, and C. Gravino, "Towards Transparent and Accurate Diabetes Prediction Using Machine Learning and Explainable Artificial Intelligence," arXiv.org, 2025. https://arxiv.org/abs/2501.18071

[23] Y. Aydın, C. Cakiroglu, G. Bekdaş, and Z. W. Geem, “Explainable Ensemble Learning and Multilayer Perceptron Modeling for Compressive Strength Prediction of Ultra-High-Performance Concrete,” Biomimetics, vol. 9, no. 9, pp. 544–544, Sep. 2024, doi: https://doi.org/10.3390/biomimetics9090544.

[24] H. B. Kibria, M. Nahiduzzaman, Md. O. F. Goni, M. Ahsan, and J. Haider, “An Ensemble Approach for the Prediction of Diabetes Mellitus Using a Soft Voting Classifier with an Explainable AI,” Sensors, vol. 22, no. 19, p. 7268, Sep. 2022, doi: https://doi.org/10.3390/s22197268.

[25] S. Wachter, B. Mittelstadt, and L. Floridi, “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,” International Data Privacy Law, vol. 7, no. 2, pp. 76–99, May 2017, doi: https://doi.org/10.1093/idpl/ipx005.

[26] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, “Explaining Explanations: An Overview of Interpretability of Machine Learning,” arXiv:1806.00069 , Feb. 2019, Available: https://arxiv.org/abs/1806.00069

[27] V. Chang, M. A. Ganatra, K. Hall, L. Golightly, and Q. A. Xu, “An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators,” Healthcare Analytics, vol. 2, p. 100118, Oct. 2022, doi: https://doi.org/10.1016/j.health.2022.100118.

[28] R. Ahmad, L. A. Maghrabi, Ishfaq Ahmad Khaja, L. A. Maghrabi, and M. Ahmad, “SMOTE-Based Automated PCOS Prediction Using Lightweight Deep Learning Models,” Diagnostics, vol. 14, no. 19, pp. 2225–2225, Oct. 2024, doi: https://doi.org/10.3390/diagnostics14192225.

[29] X. Luo, X. Li, Z. Wang, and J. Liang, “Discriminant autoencoder for feature extraction in fault diagnosis,” Chemometrics and Intelligent Laboratory Systems, vol. 192, pp. 103814–103814, Jul. 2019, doi: https://doi.org/10.1016/j.chemolab.2019.103814.

[30] J. Cai, J. Luo, S. Wang, and S. Yang, “Feature selection in machine learning: A new perspective,” Neurocomputing, vol. 300, pp. 70–79, Jul. 2018, doi: https://doi.org/10.1016/j.neucom.2017.11.077.

[31] S. M. Nagarajan, V. Muthukumaran, R. Murugesan, R. B. Joseph, M. Meram, and A. Prathik, “Innovative feature selection and classification model for heart disease prediction,” Journal of Reliable Intelligent Environments, vol. 8, Aug. 2021, doi: https://doi.org/10.1007/s40860-021-00152-3.

[32] D. Swift, K. Cresswell, R. E. Johnson, S. C. Stilianoudakis, and X. Wei, “A review of normalization and differential abundance methods for microbiome counts data,” A review of normalization and differential abundance methods for microbiome counts data, vol. 15, no. 1, May 2022, doi: https://doi.org/10.1002/wics.1586.

[33] A. Elkhouly, A. M. Andrew, H. A. Rahim, N. Abdulaziz, M. F. Abd Malek , and S. Siddique, “Data-driven audiogram classifier using data normalization and multi-stage feature selection,” Scientific Reports, vol. 13, no. 1, Feb. 2023, doi: https://doi.org/10.1038/s41598-022-25411-y.

[34] Y.-S. Kim, Moon Keun Kim, N. Fu, J. Liu, J. Wang, and Jelena Srebric, “Investigating the Impact of Data Normalization Methods on Predicting Electricity Consumption in a Building Using different Artificial Neural Network Models.,” Sustainable Cities and Society, vol. 118, pp. 105570–105570, Jun. 2024, doi: https://doi.org/10.1016/j.scs.2024.105570.

[35] P. J. Muhammad Ali, “Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements,” ARO-THE SCIENTIFIC JOURNAL OF KOYA UNIVERSITY, vol. 10, no. 1, pp. 85–91, Jun. 2022, doi: https://doi.org/10.14500/aro.10955.

[36] L. Alzubaidi et al., “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” Journal of Big Data, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: https://doi.org/10.1186/s40537-021-00444-8.

[37] A. Tello, V. Degeler, and A. Lazovik, “Too Good To Be True: accuracy overestimation in (re)current practices for Human Activity Recognition,” 2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), pp. 511–517, Mar. 2024, doi: https://doi.org/10.1109/percomworkshops59983.2024.10503465.

[38] S. Ahmed, M. Shamim Kaiser, Mohammad Shahadat Hossain, and K. Andersson, “A Comparative Analysis of LIME and SHAP Interpreters with Explainable ML-Based Diabetes Predictions,” IEEE access, vol. 13, pp. 37370–37388, Jun. 2024, doi: https://doi.org/10.1109/access.2024.3422319.

[39] X. Qi, Y. Lu, Y. Shi, H. Qi, and L. Ren, “A deep neural network prediction method for diabetes based on Kendall’s correlation coefficient and attention mechanism,” PloS one, vol. 19, no. 7, pp. e0306090–e0306090, Jul. 2024, doi: https://doi.org/10.1371/journal.pone.0306090.

[40] H. V. Nguyen, Y. Choi, and H. Byeon, “An explainable hybrid deep learning model for prediabetes prediction in men aged 30 and above,” Journal of Men’s Health, vol. 20, no. 10, p. 52, 2024, doi: https://doi.org/10.22514/jomh.2024.166.

[41] S. A. Tanim, A. R. Aurnob, T. E. Shrestha, M. R. I. Emon, M. F. Mridha, and M. S. U. Miah, “Explainable deep learning for diabetes diagnosis with DeepNetX2,” Openreview.net, 2025. https://openreview.net/forum?id=fqGwL5mcKU

[42] I. Tasin, T. U. Nabil, S. Islam, and R. Khan, “Diabetes prediction using machine learning and explainable AI techniques,” Healthcare Technology Letters, vol. 10, no. 1–2, Dec. 2022, doi: https://doi.org/10.1049/htl2.12039.

[43] P. F. Khan and K. Meehan, “Diabetes prognosis using white-box machine learning framework for interpretability of results,” 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), Jan. 2021, doi: https://doi.org/10.1109/ccwc51732.2021.9375927.

[44] Y. Du, A. R. Rafferty, F. M. McAuliffe, L. Wei, and C. Mooney, “An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus,” Scientific Reports, vol. 12, no. 1, Jan. 2022, doi: https://doi.org/10.1038/s41598-022-05112-2.

[45] J. Liu and J. Cuadros, “Eye-Net: An Interpretable Machine Learning Ensemble for Feature Engineering, Classification, and Lesion Localization of Diabetic Retinopathy,” 2022 IEEE MIT Undergraduate Research Technology Conference (URTC), pp. 1–7, Sep. 2022, doi: https://doi.org/10.1109/urtc56832.2022.10002209.

[46] E. Shakeri, T. Crump, E. Weis, R. Souza, and B. Far, “Using SHAP Analysis to Detect Areas Contributing to Diabetic Retinopathy Detection,” IEEE Xplore, Aug. 01, 2022. https://ieeexplore.ieee.org/document/9874222

[47] Abdussamad, H. Daud, R. Sokkalingam, M. Zubair, I. K. Khan, and Z. Mahmood, “A deep learning framework with hybrid stacked sparse autoencoder for type 2 diabetes prediction,” Scientific Reports, vol. 15, no. 1, pp. 36678–36678, Oct. 2025, doi: https://doi.org/10.1038/s41598-025-20534-4.

[48] María Herrero-Tudela, R. Romero-Oraá, R. Hornero, G. C. Gutiérrez, M. I. López, and María García, “An explainable deep-learning model reveals clinical clues in diabetic retinopathy through SHAP,” Biomedical Signal Processing and Control, vol. 102, pp. 107328–107328, Dec. 2024, doi: https://doi.org/10.1016/j.bspc.2024.107328.

[49] U. Das and B. Ahmed, “An Explainable ML Approach for Diabetes Detection Using the Influential Features,” International Conference on Computer and Information Technology (ICCIT), p. pp. 1-5, Dec. 2023, doi: https://doi.org/10.1109/iccit60459.2023.10441134.

[50] F. B. Hu et al., “Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women,” New England Journal of Medicine, vol. 345, no. 11, pp. 790–797, Sep. 2001, doi: https://doi.org/10.1056/nejmoa010492.

[51] D. Mozaffarian, L. J. Appel, and L. Van Horn, “Components of a Cardioprotective Diet,” Circulation, vol. 123, no. 24, pp. 2870–2891, Jun. 2011, doi: https://doi.org/10.1161/circulationaha.110.968735.

[52] D. L. Katz, E. P. Frates, J. P. Bonnet, S. K. Gupta, E. Vartiainen, and R. H. Carmona, “Lifestyle as Medicine: The Case for a True Health Initiative,” American Journal of Health Promotion, vol. 32, no. 6, pp. 1452–1458, May 2017, doi: https://doi.org/10.1177/0890117117705949.

[53] W. C. Knowler et al., “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” The New England Journal of Medicine, vol. 346, no. 6, pp. 393–403, Feb. 2002, doi: https://doi.org/10.1056/NEJMoa012512.

[54] T. A. Khan et al., “Combination of Multiple Low-Risk Lifestyle Behaviors and Incident Type 2 Diabetes: A Systematic Review and Dose-Response Meta-analysis of Prospective Cohort Studies,” Diabetes Care, vol. 46, no. 3, pp. 643–656, Feb. 2023, doi: https://doi.org/10.2337/dc22-1024.

[55] L. Hakim, “Penelitian Komparatif: Pengertian, Jenis dan Contoh,” Deepublish Store, Apr. 15, 2025. https://deepublishstore.com/blog/penelitian-komparatif/

[56] L. Schwingshackl, B. Missbach, J. König, and G. Hoffmann, “Adherence to a Mediterranean Diet and Risk of diabetes: a Systematic Review and meta-analysis,” Public Health Nutrition, vol. 18, no. 07, pp. 1292–1299, Aug. 2014, doi: https://doi.org/10.1017/s1368980014001542.

[57] I. Kyrou et al., “Sociodemographic and lifestyle-related risk factors for identifying vulnerable groups for type 2 diabetes: a narrative review with emphasis on data from Europe,” BMC Endocrine Disorders, vol. 20, no. S1, Mar. 2020, doi: https://doi.org/10.1186/s12902-019-0463-3.

[58] A. Pandey and D. K. Vishwakarma, “VABDC-Net: A framework for Visual-Caption Sentiment Recognition via spatio-depth visual attention and bi-directional caption processing,”

Knowledge-Based Systems, vol. 269, p. 110515, Jun. 2023, doi: https://doi.org/10.1016/j.knosys.2023.110515.