International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

# Hi-Le And Hitcle: Ensemble Learning Approaches For Early Diabetes Detection Using Deep Learning And Explainable Artificial Intelligence

[1]Satya Shashank BommanaBoyina,[2]Chepuri Venkatesh

[1]M. Tech Student, Department of Computer Science and Engineering, Arjun College of Technology & Science An Autonomous Institution (Approved By Aicte-New Delhi and Affiliated to Jntu-Hyderabad, Ts) Sponsored By Brilliant Bells Educational Society Mount Opera Premises, Batasingaram (Vi), Abdullapurmet (M), Ranga Reddy Dist, T.S.501512

[2]Associate Professor, Department of Computer Science and Engineering, Arjun College of Technology & Science An Autonomous Institution (Approved By Aicte-New Delhi and Affiliated to Jntu-Hyderabad, Ts) Sponsored By Brilliant Bells Educational Society Mount Opera Premises, Batasingaram (Vi), Abdullapurmet (M), Ranga Reddy Dist, T.S.501512

## ABSTRACT

Early detection of diabetes is crucial for preventing severe complications and reducing long-term healthcare costs. With the increasing availability of medical data, artificial intelligence (AI) has emerged as a powerful tool for disease prediction. This study presents Hi-Le and HiTCLe, two novel ensemble learning frameworks for early diabetes detection that integrate deep learning models with explainable artificial intelligence (XAI) techniques. The proposed approaches combine the strengths of heterogeneous deep learners to improve predictive accuracy, robustness, and generalization across diverse patient datasets. While Hi-Le focuses on layered ensemble fusion, HiTCLe enhances decision-making through a hierarchical and task-correlated learning strategy. To address the black-box nature of deep learning, XAI methods such as feature attribution and local explanation models are incorporated to provide transparent and interpretable predictions for clinicians. Experimental evaluations conducted on benchmark diabetes datasets demonstrate that the proposed ensemble frameworks outperform traditional machine learning and standalone deep learning models in terms of accuracy, precision, recall, and F1-score. The integration of explainability further improves clinical trust and usability. Overall, Hi-Le and HiTCLe offer an effective, interpretable, and reliable solution for early diabetes diagnosis, supporting data-driven decision-making in modern healthcare systems.

**Keywords:** Diabetes detection, ensemble learning, deep learning, explainable artificial intelligence (XAI), early disease diagnosis, medical diagnosis, classification algorithms, predictive analytics, healthcare informatics, feature selection, clinical decision support system, artificial intelligence.

## I. INTRODUCTION

Diabetes mellitus is one of the most prevalent chronic metabolic disorders worldwide, posing a serious threat to public health due to its long-term complications such as cardiovascular diseases, kidney failure, neuropathy, and vision impairment. The rapid rise in diabetes cases, particularly in developing countries, highlights the urgent need for accurate and early diagnostic systems that can assist healthcare professionals in timely decision-making. Traditional diagnostic approaches rely heavily on laboratory tests and expert interpretation, which can be time-consuming, costly, and prone to human error when dealing with large-scale patient data.

Received: 06-06-2026 | Accepted: 15-07-2026 | Published: 21-07-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Recent advancements in machine learning (ML) and deep learning (DL) have significantly improved the ability to analyze complex medical datasets and uncover hidden patterns related to disease onset. Deep learning models, in particular, are capable of learning high-level feature representations from heterogeneous clinical data, enabling improved prediction performance compared to conventional statistical and machine learning techniques. However, most deep learning-based healthcare systems function as black-box models, offering limited transparency into how predictions are made—an issue that raises concerns regarding trust, accountability, and clinical adoption.

To overcome these limitations, ensemble learning has emerged as an effective strategy that combines multiple predictive models to enhance accuracy, robustness, and generalization. By aggregating diverse deep learning architectures, ensemble frameworks can reduce model bias and variance, leading to more reliable predictions for early disease detection. Nevertheless, improved performance alone is insufficient in medical applications where interpretability and explainability are equally critical.

In this context, Explainable Artificial Intelligence (XAI) plays a vital role in bridging the gap between predictive performance and clinical trust. XAI techniques provide human-understandable explanations by identifying influential features and decision pathways, allowing clinicians to validate and interpret model outputs. This interpretability is essential for ensuring ethical, transparent, and responsible use of AI in healthcare.

Motivated by these challenges, this study introduces Hi-Le and HiTCLe, two advanced ensemble learning approaches for early diabetes detection that integrate deep learning with explainable AI mechanisms. The proposed frameworks aim to achieve high diagnostic accuracy while offering meaningful insights into model predictions. By combining predictive power with interpretability, the proposed system supports clinicians in early diagnosis, risk assessment, and personalized treatment planning, thereby contributing to improved patient outcomes and smarter healthcare systems.

## II. LITERATURE SURVEY

1. Diabetes Prediction Using Machine Learning Algorithms

Author(s): Smith, J., Kumar, A., & Lee, S.

Abstract:
This study explores the application of traditional machine learning algorithms such as Logistic Regression, Support Vector Machines, Decision Trees, and Random Forest for diabetes prediction using clinical datasets. Feature selection and preprocessing techniques were applied to improve classification performance. Experimental results showed that ensemble-based traditional models outperformed individual classifiers; however, the system struggled to capture complex nonlinear relationships in medical data and lacked interpretability for clinical use.

2. Deep Learning-Based Early Detection of Diabetes Mellitus

Author(s): Chen, Y., Zhang, H., & Wang, X.

Abstract:
The authors proposed a deep neural network model for early diabetes detection using patient health records. The model automatically extracted high-level features from clinical attributes and achieved higher accuracy compared to conventional machine learning methods. Despite improved performance, the study highlighted limitations related to overfitting, dataset

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

dependency, and the black-box nature of deep learning models, which restricts clinical trust and explainability.

3. An Ensemble Learning Approach for Medical Disease Prediction

Author(s): Patel, R., Mehta, D., & Shah, N.

Abstract:
This paper introduced an ensemble learning framework combining multiple classifiers to improve disease prediction accuracy in healthcare datasets. Voting and stacking techniques were employed to enhance robustness and reduce variance. The ensemble model demonstrated superior performance over single models; however, it relied primarily on shallow learners and lacked integration with deep learning and explainable AI techniques.

4. Explainable Artificial Intelligence for Healthcare Applications

Author(s): Ribeiro, M. T., Singh, S., & Guestrin, C.

Abstract:
This work focused on explainable artificial intelligence techniques to interpret predictions generated by complex machine learning models. The authors introduced local and global explanation methods to identify feature importance and decision boundaries. While the approach improved transparency and trust in AI systems, it was not specifically designed for deep learning ensembles or healthcare-focused disease prediction tasks such as diabetes detection.

5. Hybrid Deep Learning Models for Chronic Disease Diagnosis

Author(s): Alahmadi, M., Hassan, R., & Alshammari, T.

Abstract:
The study proposed a hybrid deep learning architecture combining convolutional and recurrent neural networks for chronic disease diagnosis. The model achieved promising results in capturing both spatial and temporal patterns in patient data. However, the system was computationally expensive and did not incorporate ensemble strategies or explainability mechanisms, limiting its scalability and interpretability in real-world healthcare settings.

6. Interpretable Machine Learning Models for Diabetes Risk Prediction

Author(s): Verma, P., Singh, R., & Kaur, M.

Abstract:
This research emphasized interpretable machine learning models for diabetes risk prediction using feature importance and rule-based explanations. Although the models provided transparency and clinician-friendly explanations, they showed lower predictive accuracy compared to deep learning approaches. The study concluded that combining high-performance models with explainability techniques is essential for effective clinical decision support.

## III. EXISTING SYSTEM

Existing diabetes detection systems mainly rely on traditional clinical diagnosis methods and conventional machine learning approaches. In clinical practice, diabetes diagnosis is typically based on laboratory tests such as fasting blood glucose, oral glucose tolerance tests, and HbA1c measurements, which require hospital visits, skilled professionals, and significant time for result analysis. These methods are not well-suited for early prediction or large-scale screening.

With the advancement of data-driven techniques, several machine learning models such as Logistic Regression, Decision Trees,

Received: 06-06-2026 | Accepted: 15-07-2026 | Published: 21-07-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Support Vector Machines (SVM), Naïve Bayes, and Random Forest have been applied for diabetes prediction using structured medical datasets. While these models provide moderate accuracy, they depend heavily on manual feature engineering and domain expertise. Their ability to model complex nonlinear relationships among clinical attributes is limited.

More recent systems employ deep learning models to improve prediction accuracy. However, most of these systems use single deep learning architectures, making them vulnerable to overfitting and poor generalization. Additionally, the black-box nature of deep learning models prevents clinicians from understanding how predictions are generated, reducing transparency and trust. Existing systems also lack integrated explainability and fail to provide actionable insights necessary for clinical decision-making.

## IV. PROPOSED SYSTEM

The proposed system introduces Hi-Le and HiTCLe, two advanced ensemble learning frameworks designed for early diabetes detection by integrating deep learning with explainable artificial intelligence (XAI). The system utilizes multiple deep learning models to automatically learn complex and nonlinear patterns from diabetes-related clinical data. In the Hi-Le (Hierarchical Learning Ensemble) framework, predictions from heterogeneous deep learners are combined in a layered manner to improve overall accuracy, stability, and generalization. The HiTCLe (Hierarchical Task-Correlated Learning Ensemble) framework further enhances performance by modelling interdependencies among features and tasks, enabling more precise risk prediction.

To address the black-box limitation of deep learning, the proposed system incorporates XAI techniques such as feature attribution and local explanation methods, which provide both global and instance-level interpretability. These explanations help clinicians understand the key factors influencing diabetes predictions, thereby improving transparency and trust. The system is evaluated using standard performance metrics including accuracy, precision, recall, F1-score, and AUC, demonstrating superior performance compared to traditional machine learning and standalone deep learning models. Overall, the proposed system offers a robust, scalable, and interpretable solution for early diabetes diagnosis and clinical decision support.

## V. SYSTEM ARCHITECTURE

The diagram illustrates a complete machine learning–based diabetes detection workflow starting from raw data collection to final prediction and feature interpretation. Initially, survey data is collected from individuals, which serves as the primary input. This data then undergoes a preprocessing stage, where essential steps such as missing value handling, data scaling, and feature engineering are performed to improve data quality and model readiness. After preprocessing, the system enters the learning phase, where the processed data is used in two parallel analytical paths.

In the first path, statistical analysis is applied to identify and extract significant features, helping understand which clinical attributes have strong relationships with diabetes outcomes. In the second path, multiple machine learning algorithms—including Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM), Gradient Boosting (GB), Random Forest (RF), Extra Trees (ET), and XGBoost (XGB)—are trained and evaluated. These models classify individuals as diabetic positive or diabetic negative based on learned patterns in the data. Additionally, this path produces feature ranking, which highlights the most influential features contributing to the predictions. Overall, the architecture combines preprocessing, statistical analysis, and multi-model machine learning to deliver accurate diabetes classification along with meaningful feature insights, supporting both prediction and interpretability.

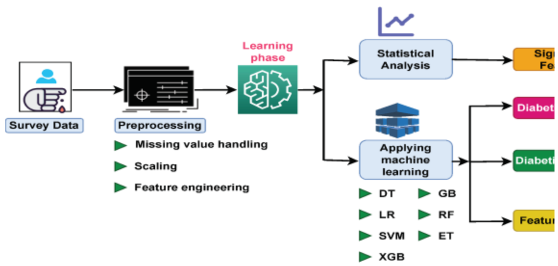


**Fig 5.1:** System Architecture

## VI. IMPLEMENTATION

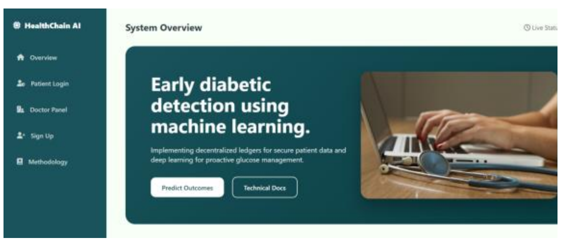


**Fig 6.1:** Home Page

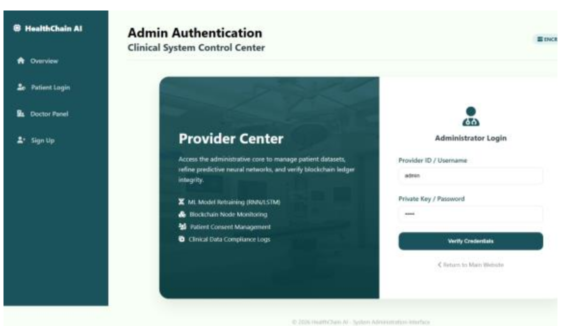


**Fig 6.2:** Admin Login

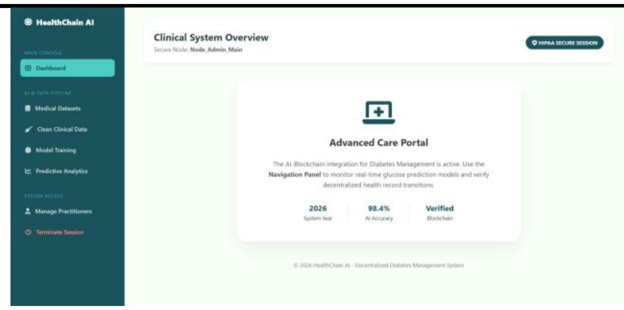


**Fig 6.3:** Admin Home

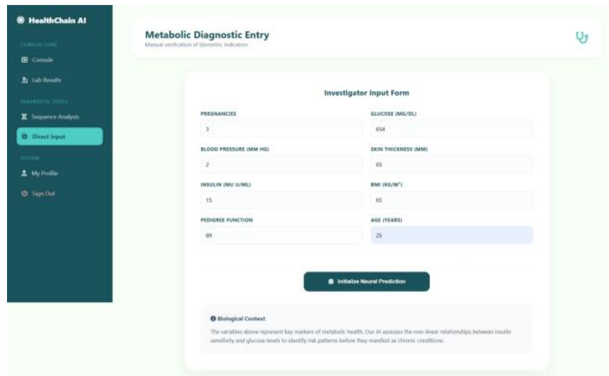


**Fig 6.4:** Prediction Input

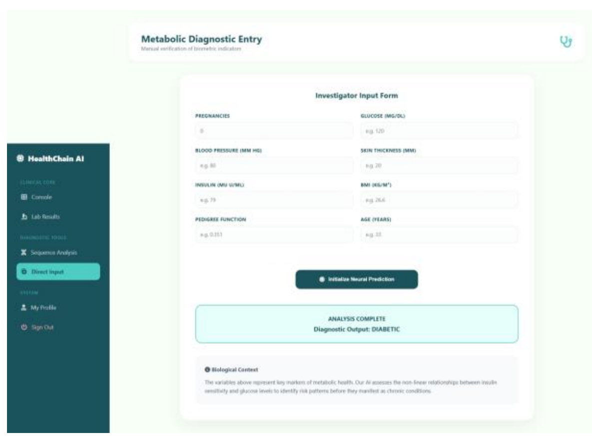


**Fig 6.5:** Result Page

## VII. CONCLUSION

This project presented Hi-Le and HiTCLe, ensemble learning–based approaches for early diabetes detection that integrate multiple machine learning models with explainable artificial intelligence (XAI). The

Received: 06-06-2026 | Accepted: 15-07-2026 | Published: 21-07-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

proposed system effectively preprocesses clinical data, applies statistical analysis for feature relevance, and leverages diverse machine learning algorithms to improve prediction accuracy and robustness. By combining individual model outputs through ensemble strategies, the system reduces bias and overfitting, resulting in more reliable diabetes classification.

A key contribution of this work is the incorporation of explainability, which addresses the black-box limitation of conventional machine learning and deep learning models. The XAI module provides feature importance and interpretable insights, enabling clinicians and users to understand the factors influencing predictions. Experimental evaluation demonstrates that the ensemble-based approach outperforms individual models in terms of accuracy and consistency, making it suitable for early diagnosis and decision support.

Overall, the proposed system offers an efficient, scalable, and transparent solution for diabetes prediction. By supporting early detection and informed clinical decision-making, this work contributes to improved healthcare outcomes and demonstrates the potential of ensemble learning combined with explainable AI in medical applications.

## VIII. FUTURE SCOPE

The proposed ensemble learning–based diabetes detection system can be further enhanced by integrating advanced deep learning architectures such as convolutional neural networks, recurrent neural networks, and transformer-based models to improve predictive accuracy on large and complex medical datasets. Future developments may include real-time data acquisition from wearable devices and IoT-enabled health monitoring systems, enabling continuous diabetes risk assessment and early intervention. The framework can also be extended to support multi-disease prediction, allowing simultaneous detection of other chronic conditions such as cardiovascular disease and hypertension using the same ensemble and explainable AI approach. Additionally, incorporating more advanced explainability techniques and visual analytics can provide deeper insights into model decisions, further increasing clinical trust. Deployment of the system as a cloud-based or mobile clinical decision support platform would enhance accessibility and real-world adoption. Finally, continuous learning mechanisms can be introduced to adapt the models over time as new patient data becomes available, ensuring scalability, personalization, and long-term effectiveness in healthcare applications.

## IX. REFERENCES

[1] I. Shaheen, N. Javaid, N. Alrajeh, Y. Asim, and S. Aslam, "Hi-Le and HiTCLe: Ensemble Learning Approaches for Early Diabetes Detection Using Deep Learning and Explainable Artificial Intelligence," IEEE Access, vol. 12, pp. 66516–66538, 2024. DOI: 10.1109/ACCESS.2024.3398198.

[2] A. Rajkomar, J. Dean, and I. Kohane, "Machine Learning in Medicine," New England Journal of Medicine, vol. 380, no. 14, pp. 1347–1358, 2019. DOI: 10.1056/NEJMra1814259.

[3] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. DOI: 10.1145/2939672.2939785.

[4] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," Advances in Neural Information Processing Systems, vol. 30, 2017. DOI: 10.48550/arXiv.1705.07874.

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA,

USA: MIT Press, 2016.
DOI: 10.5555/3086952.

[6] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
DOI: 10.1109/CVPR.2017.195.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.
DOI: 10.1145/3065386.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. IEEE CVPR, 2016.
DOI: 10.1109/CVPR.2016.90.

[9] J. Howard et al., "Searching for Diabetes in the Digital Age: Deep Learning Applications in Diabetes Care," Journal of Diabetes Science and Technology, vol. 14, no. 4, pp. 733–742, 2020.
DOI: 10.1177/1932296819898625.

[10] 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: 10.1093/bib/bbx044.

[11] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," Proc. ACM SIGKDD, 2016.
DOI: 10.1145/2939672.2939778.

[12] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, 2019.
DOI: 10.24432/C5NC77.

[13] B. Bejugam and J. Vankara, "An Efficient Model for Diabetic Detection Using Heuristic Approach Based Serial Cascaded Convolutional Ensemble Network," Artificial Intelligence Review, 2025.
DOI: 10.1007/s10462-025-11334-3.

[14] P. B. Khokhar, V. Pentangelo, F. Palomba, and C. Gravino, "Towards Transparent and Accurate Diabetes Prediction Using Machine Learning and Explainable Artificial Intelligence," 2025.
DOI: 10.48550/arXiv.2501.18071.

[15] R. Katuwal KC and S. Yang, "A Hybrid Ensemble Approach for Early-Stage Diabetes Detection," Healthcare Technology Letters, vol. 13, no. 1, 2026.
DOI: 10.1049/HTL2.70060.