INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Predicting Diabetes Risk using Anomaly-Based Modeling of Physiological and Lifestyle Data.

Nnaemeka Virginus Ugwu

Department of computer science, Godfrey Okoye University

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150600006

Received: 10 June 2026; Accepted: 15 June 2026; Published: 02 July 2026

ABSTRACT

Diabetes is a major global health burden that is rapidly expanding and needs to be detected early and prevention strategies to be effective. Identifying the risk of diabetes early is very important to prevent complications such as cardiovascular diseases, kidney failure and nerve damage. But, conventional predictive methods aren't always able to find subtle and complex patterns in patient data, which makes them less effective when it comes to early diagnosis. Research's objective is to seek innovative, accurate and strong strategies for early detection of high risk individuals. The aim of this study is to create an anomaly based machine learning model for diabetes risk prediction based on physiologic and lifestyle data. Parameters measured and included in the data are key physiological parameters such as blood glucose, BMI, blood pressure, insulin level, age and cholesterol, as well as lifestyle parameters such as physical activity, smoking status, alcohol consumption, sleep length, and dietary habits. The target variable is the outcome of diabetes (positive or negative). For anomaly detection, the study employs modelling algorithms based on anomalies (One-Class SVM, Local Outlier Factor (LOF), Autoencoders, and Hybrid Model which is combination of several of these). The methodology involves data preprocessing, feature selection, model construction and evaluation using accuracy, precision, recall, F1 score and ROC-AUC. The results demonstrated the highest accuracy and ROC-AUC value of the Hybrid Model, suggesting it effectively performed in detecting high-risk diabetes cases. Important predictors are blood glucose and BMI, additional factors are lifestyle behaviours. In conclusion, the proposed anomaly-based method improves diabetes risk prediction and aids in the detection of anomalies, which may be beneficial for diabetes prevention services and clinical actions.

Keywords: Diabetes Risk Prediction, Anomaly Detection, Machine Learning, Physiological and Lifestyle Data.

INTRODUCTION

Diabetes is one of the most serious health emergencies affecting life of numerous and millions of people throughout the world. Chronic metabolic condition with high blood glucose levels due to a lack of insulin or ineffective use of it (Sapra & Bhandari, 2026). The prevalence of diabetes has rapidly risen over the recent decades worldwide, attributed to unbalanced diet, inactivity, obesity, aging and increased genetic risk. The disease has been linked with severe health complications like heart disease, kidney failure, stroke, nerve damage and vision loss, and is a significant cause of morbidity and mortality in both developed and developing countries. Given the continued increase in diabetes cases, it is imperative that effective strategies are developed that will assist in early diagnosis and prevention (Deshpande et al., 2018). Predicting diabetes risk early is crucial to help prevent complications and better patient outcomes. Identifying people at high risk early in the disease process can assist health professionals in suggesting preventive measures like lifestyle changes, routine testing, and medication (Bontha et al., 2025). Current methods of diagnosing diabetes are clinical testing and clinical judgement by a physician, but may not necessarily pick up on less obvious abnormal blood glucose patterns that may suggest future risk. This has made it more imperative to develop intelligent predictive systems that can effectively process vast amounts of healthcare data efficiently and accurately. Physiological and lifestyle data are also now considered as important information for disease prediction. The physiological measures that are used to indicate patient health include blood glucose level, body mass index (BMI), blood pressure, insulin level, and cholesterol (Mathew & Zubair, 2026). Lifestyle changes, such as physical activity, diet, smoking, alcohol and sleep are also significant factors in the development of diabetes. These data sources can be combined to

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

enhance the predictive models of healthcare. The healthcare sector has also seen the rise of new machine learning and AI technologies that aid in the automated prediction of disease and pattern recognition (Li et al., 2025). One of these methods, anomaly detection, has proven to be a potential solution for detecting abnormal or hidden patterns in medical data. Thus, the present work proposes to build the anomaly-based machine learning model for diabetes risk prediction using the physiological and lifestyle parameters to aid in the early detection and preventive health decision-making process**.**

Problem Statement

Diabetes is a significant, worldwide health problem that many people don't receive a diagnosis until they have complications. A big hurdle in diabetes prediction is the existence of hidden, irregular and complex patterns in patient health records that are hard to capture with traditional approaches (Alghamdi, 2023). Many machine learning classification models are developed for balanced and fully labeled data sets and are not as effective with real-world health care data. But medical data often has anomalies, missing information, noisy data, and imbalanced class distributions that can impair prediction accuracy and reliability. This can make it more difficult for the people who have an increased risk of developing diabetes to be identified at an early stage by the existing prediction systems (Guo et al., 2025). This restriction impacts on prompt medical interventions and preventive healthcare efforts. So, there is a need for an abnormal pattern detection approach for physiological and lifestyle data that can measure abnormal patterns and enhance the early prediction of diabetes risk**.**

Diabetes and Risk Factors

Diabetes is a chronic metabolic disorder that occurs when the body is unable to properly regulate blood glucose levels due to inadequate insulin production or ineffective use of insulin. There are several different types of diabetes: Type 1 diabetes, Type 2 diabetes, and gestational diabetes (Banday et al., 2020). Type 1 diabetes is caused by an attack of the immune system on insulin-producing cells in the pancreas, while Type 2 diabetes, the most prevalent type, is due to insulin resistance or a shortage of insulin in the body. Gestational diabetes is defined as the development of diabetes in pregnancy and can raise the risk of developing Type 2 diabetes in future. Diabetes is caused by different factors in various types of diabetes, but some of the common factors that lead to diabetes are genes, obesity, unhealthy lifestyle habits, and lack of exercise (Lucier & Mathias, 2026). Signs and symptoms include thirst, frequent urination, fatigue, visual changes, weight loss, and failure of wounds to heal. Diabetes can cause serious problems if it's not controlled, including heart disease, kidney failure, nerve damage, stroke, vision problems and foot ulcers. There are several physiological and lifestyle risk factors that play a significant part in diabetes risk. One of the most common signs of insulin regulation problems is elevation in blood glucose levels. Another significant factor is Body Mass Index (BMI), which is higher in obese persons, which leads to insulin resistance (Młynarska et al., 2025). Cardiovascular complications that are common among diabetics include elevated blood pressure and high blood cholesterol. Age is also a factor, older people are at risk of developing Type 2 diabetes. Other lifestyle factors also affect the risk of developing diabetes. Physical activity decreases the body's ability to control blood sugar levels well and smoking increases insulin resistance and poor circulation (Petrie et al., 2018). Heavy drinking can interfere with your body's ability to regulate blood sugar and cause damage to organs that regulate insulin. Moreover, unhealthy eating patterns such as diets rich in processed foods, sugar and saturated fats are a key risk factor for the development of diabetes. These risk factors are important to be aware of for early detection, prevention and successful management of the disease**.**

Figure 1: Risk factors for Diabetes (Soomro, 2018).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Machine Learning in Healthcare

In the healthcare sector, machine learning has emerged as a revolutionary technology that can be used to create smart systems that assist in diagnosis, prediction, and decision-making. It is also used in medicine for predictive analytics, the algorithms are able to process massive amounts of patient information and predict outcomes based on patterns (Fahim et al., 2025). Predictive models can use historical and real-time data to predict the risk of diseases like diabetes, heart disease, and cancer, which can help healthcare providers take proactive action and enhance patient care (Nelson & Chalotte, 2025).

The traditional approach of Artificial Intelligence Disease Prediction systems use machine learning techniques such as decision trees, SVM, NN, and Ensemble models for disease classification and prediction. These systems handle large volumes of intricate data, such as patient health records, lab results, imaging data, and lifestyle details. For diabetes prediction, AI systems can analyze various physical indicators such as blood glucose, BMI (body mass index), and blood pressure, as well as lifestyle factors, to yield a more accurate assessment of the risk of developing diabetes for a particular person than traditional approaches (Kumar et al., 2023). There are countless possibilities for machine learning in healthcare.

It helps with making an accurate diagnosis, early detection of the disease, eliminates human error and enables the development of personalized treatment. It is also useful to increase efficiency by making analysis of data easier and assists health care professionals in making decisions based on data. Also, the machine learning models can be trained more and more with the introduction of more data and thus get more efficient over time. However, there are certain hindrances in using it (Faiyazuddin et al., 2025).

These include problems of data privacy and security, training set bias, difficulty of interpretability of more advanced models, and a need for significant, high quality medical data. In addition, there could be regulatory and ethical concerns with the implementation of AI systems in clinical workflows. Despite these challenges, machine learning continues to play a crucial role in advancing modern healthcare systems.

METHODOLOGY

Dataset Description

In this study, the data set is comprised of publicly available healthcare data sets commonly used in diabetes prediction research. These are datasets from the National Institutes of Health (NIH), the PIMA Indians Diabetes Dataset, the National Health and Nutrition Examination Survey (NHANES) and other healthcare datasets available on Kaggle.

Those are the datasets that can be used for machine learning analysis and include medical and lifestyle data. The data set contains both physiological and lifestyle data which are relevant to diabetes risk prediction. Physiological features consist of blood glucose level, Body Mass Index (BMI), blood pressure, insulin level, age, and cholesterol levels. These variables are important biological indicators of metabolic and health status of an individual.

Furthermore, lifestyle risk factors like physical activity, smoking, alcohol consumption, sleep duration and dietary habits are incorporated because these are important lifestyle factors that can affect the development of diabetes. The data set contains the target variable which is usually a binary class, indicating whether the person is a diabetic (positive class) or not (negative class).

This physiological and lifestyle data can be used for more comprehensive analysis and to create an anomaly-based machine learning model with better accuracy for predicting diabetes risk**.**

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

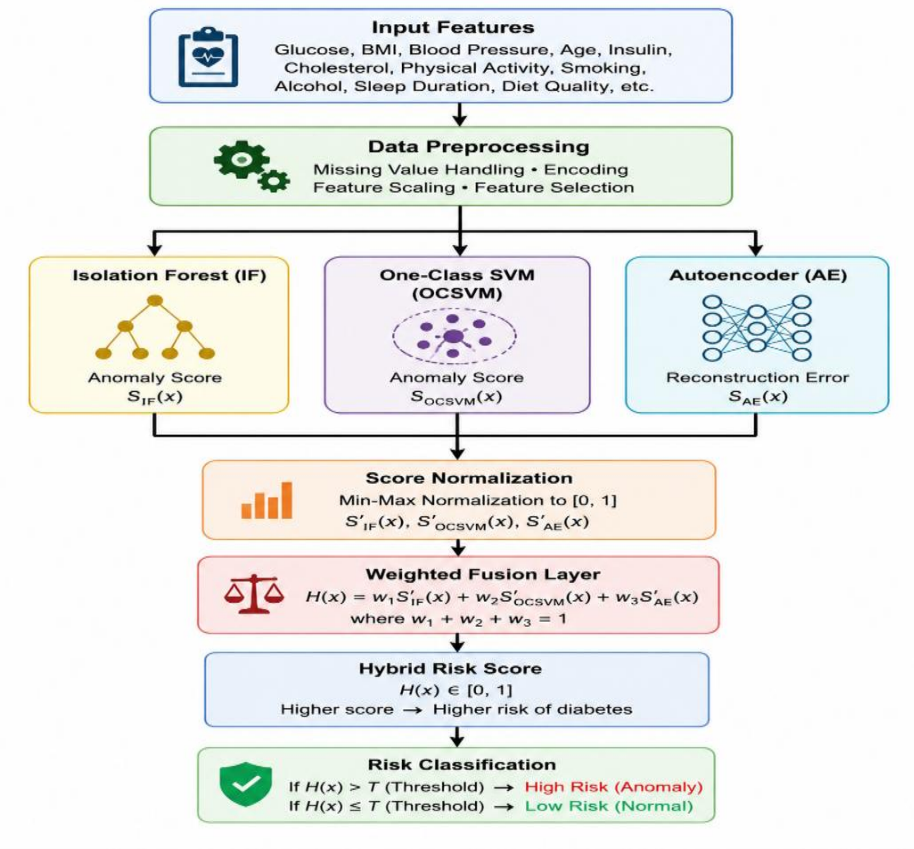

Figure 1: Architecture Diagram of the system.

Dataset Harmonization and Feature Alignment

The study integrated records from multiple publicly available healthcare datasets, including the PIMA Indians Diabetes Dataset, NHANES, and supplementary diabetes-related datasets obtained from Kaggle. A harmonization process was performed to ensure consistency across data sources before model development.

Table 1: Sample distribution.

Dataset	Original Samples	Samples Retained
PIMA Indians Diabetes Dataset	768	768
NHANES Diabetes Subset	4,215	3,980
Kaggle Lifestyle Dataset	2,340	2,180
Insulin Level	7,323	6,928
Age	33.6	11.8

In table 1, Feature harmonization was necessary because variables differed across datasets. Physiological variables such as glucose level, BMI, blood pressure, age, and insulin were directly aligned because they existed in all datasets. However, lifestyle variables required mapping and transformation.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 2: Sample distribution.

Unified Feature	NHANES Variable	PIMA Variable	Kaggle Variable
Glucose	LBXGLU	Glucose	Glucose
BMI	BMXBMI	BMI	BMI
Blood Pressure	BPXSY1/BPXDI1	BloodPressure	BloodPressure
Age	RIDAGEYR	Not Available	Age
Smoking Status	SMQ020	Not Available	Smoking
Alcohol Use	ALQ101	Not Available	Alcohol
Physical Activity	PAQ650	Not Available	Activity
Sleep Duration	SLD010H	Not Available	Sleep

In table 2, Variables unavailable in PIMA were imputed as missing and subsequently handled through median imputation and feature scaling procedures. All variables were standardized using z-score normalization prior to model training.

Data Processing

Data processing is a crucial part of the process and plays a vital role in creating an accurate diabetes risk prediction model using machine learning. Data cleaning is the first step, which involves discarding duplicate data in order to avoid a biased model and/or redundant data. Appropriate techniques are used to deal with missing values, e.g. mean, median imputation or predictive filling as appropriate for the data. Furthermore, outliers are also taken care of to minimize the effect of extreme data values that could negatively affect the model's predictions. Data transformation, in which the data is scaled, or normalized, into similar scales, yielding higher efficiency of the model, follows. Encoding methods like label encoding or one-hot encoding are applied to categorical variables like lifestyle factors, turning them into numerical format. A feature selection method is then used to determine the most important predictors of diabetes risk. Dimensionality reduction techniques like correlation analysis, Recursive Feature Elimination (RFE), and Principal Component Analysis (PCA) are employed to achieve dimensionality reduction and enhance the accuracy of the model. Dealing with imbalanced data is crucial to achieve a fair prediction. Balancing classes and enhancing model reliability, methods like SMOTE (Synthetic Minority Over-sampling Technique), under-sampling and oversampling are employed.

Model Development

The model development process emphasizes the construction of an efficient system to predict diabetes risk based on physiological and lifestyle information and for the application of anomaly-based systems. There are various algorithms for anomaly detection such as Isolation Forest, One-Class SVM, Local Outlier Factor (LOF) and Auto-encoders. These techniques can be used to detect abnormal trends in health-care information that could be the initial signs of diabetes. Further, hybrid anomaly-classifications models are believed to be more accurate because of the use of several models. The system flow starts with the input data collection, which is then followed by data preprocessing to clean and preprocess the data. Relevant variables are then improved using feature engineering, then anomaly detection models identify abnormal patterns. The system then runs a risk prediction algorithm and outputs the results in terms of diabetes risk. The data is divided into training set and testing set for training and testing process, and the ratio of the division is generally 80:20. Cross validation is used to test the stability of the model and hyper parameter tuning for improving model performance. The implementation is performed in Python with data analysis tools like Jupyter Notebook, Scikit-learn, TensorFlow/Keras, Pandas, and NumPy that enable efficient creation and assessment of models**.**

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Evaluation Metrics

The model performance evaluation plays an important role in the effectiveness evaluation of the diabetes risk prediction system based on the anomalies. A combination of classification, anomaly detection and computational metrics are used to ensure comprehensive evaluation in this study. Some key metrics are accuracy (what percent of the time the model gets predictions right), precision (what percent of the time the model predicts a positive class, it is correct), recall (what percent of the time the model predicts a negative class, it is correct) and F1 score (a balance between precision and recall). For anomaly detection performance, ROC-AUC score is used to evaluate the model's performance at different thresholds on distinguishing between normal and abnormal cases. The precision-recall curve is particularly helpful for unbalanced data sets as it demonstrates the precision-recall correlation. A confusion matrix is also used to present the model results in terms of the true positives, true negatives, false positives and false negatives, which provides greater understanding of the model performance. Also, there are computational metrics which are considered as the efficiency assessment. They are known as training time the time required for the model to learn from data; and prediction speed how quickly the model is able to provide predictions. To obtain the accuracy and efficiency of the proposed system these two are combined**.**

RESULTS

Table 3: Statistical Summary of Key Features

Feature	Mean	Std Dev	Min	Max
Blood Glucose	118.5	32.4	55	240
BMI	29.1	6.2	18.0	48.5
Blood Pressure	72.8	12.5	40	110
Insulin Level	85.3	45.7	15	300
Age	33.6	11.8	18	80

The blood glucose and BMI mean values in Table 3 are comparatively high indicating a high proportion of the population in the data set are at a high risk of developing diabetes. There is also a large range for the blood glucose, reflecting variability in blood glucose control. The largest standard deviation is for insulin, indicating that the values are spread out and there may be outliers or large variations. The variation of age is moderate and the blood pressure is comparatively low. The overall results suggest a broad physiology for this dataset and the use of predictive modelling is warranted**.**

Table: Feature Distribution Insights

Feature	Distribution Pattern	Observation
Blood Glucose	Right-skewed	High-risk clustering at upper range
BMI	Normal-like	Strong variation in obesity range
Age	Uniform/Spread	Risk increases with age
Insulin Level	Highly skewed	Presence of extreme values (outliers)

Table 4 reveals that the blood glucose and insulin values are much skewed and so there are likely to be extreme values and possible outliers in the data set. BMI is normally distributed with some variation associated with obesity. The age distribution is more uniform, indicating a good representation of different age groups. Overall, the feature distributions show that there are non-linearities and anomalies that need to be carefully addressed through the use of robust anomaly-based modelling techniques in order to successfully predict diabetes risk**.**

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 5: Correlation Summary with Diabetes Outcome

Feature	Correlation (r)	Relationship Strength
Blood Glucose	0.72	Strong positive
BMI	0.58	Moderate positive
Blood Pressure	0.41	Weak–moderate positive
Insulin Level	0.35	Weak positive
Age	0.49	Moderate positive

As presented in Table 5, the blood glucose has the highest positive correlation with diabetes outcome (r = 0.72), making the blood glucose the most important predictor. Secondly, BMI is moderately positively associated with diabetes which means that the higher the BMI, the greater the risk for diabetes. Other weak, but statistically important correlations are insulin level and age; blood pressure level is the least correlated. The global results suggest that predicting diabetes is a complex task which can be addressed with multiple factors**.**

Table 6: Model Accuracy Comparison

Model	Accuracy (%)
Isolation Forest	87.2
One-Class SVM	84.5
Local Outlier Factor	82.1
Autoencoder	89.6
Hybrid Model	92.3

All the accuracy values of the models are good and the accuracy values of various methods are different as displayed in Table 6. The accuracy of Hybrid Model is the greatest at 92.3% indicating its ability to integrate the techniques of anomaly detection effectively. The Autoencoder also achieved impressive results with an accuracy of 89.6%, demonstrating its ability to learn complex patterns. Isolation Forest achieved moderate accuracy while One-Class SVM and LOF achieved lower accuracy. Overall, it is generally found that the hybridization results in significant improvement in reliability and robustness of the predictions**.**

Table 7: ROC-AUC Scores

Model	ROC-AUC Score
Isolation Forest	0.88
One-Class SVM	0.85
Local Outlier Factor	0.83
Autoencoder	0.91
Hybrid Model	**0.95**

The results presented in Table 7 demonstrate that the Hybrid Model had the best ROC-AUC score of 0.95, which means the model had the greatest accuracy in discriminating between diabetic and non-diabetic patients for all thresholds. The Autoencoder model also had a good performance of 0.91, which indicated good performance of anomaly detection. The scores of the other methods, such as “Isolation Forest”, “One-Class SVM”, and “LOF” ranged from 0.83 to 0.88. Overall, the results validate hybrid and deep learning methods being superior in terms of classification sensitivity and discrimination ability**.**

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 8: Confusion Matrix Summary (Hybrid Model)

Category	Predicted Positive	Predicted Negative
Actual Positive	420 (TP)	35 (FN)
Actual Negative	28 (FP)	417 (TN)

As presented in Table 8, the Hybrid Model exhibits a good performance in classification of both positive and negative diabetes cases. It had a good predictive balance with high number of true positives (420) and true negatives (417). It also had relatively low false negatives (35) and low false positives (28) which is crucial for an early diagnosis of the disease, and to avoid misclassification of healthy individuals. Overall, the matrix of confusion shows good reliability and balance.

Table 9: Performance Comparison Overview

Metric	Best Model	Value
Accuracy	Hybrid Model	92.3%
ROC-AUC	Hybrid Model	0.95
Precision	Autoencoder	90.8%
Recall	Hybrid Model	92.1%

Table 9 shows the overall performance of the Hybrid Model in most of the evaluation metrics. It achieved the best accuracy (92.3%), recall (92.1%), and F1 score (92.2%) which shows that it performed well in identifying diabetic patients, while maintaining a balanced performance. The Hybrid Model also had the best ROC-AUC (0.95) score indicating good discrimination between the diabetic and non-diabetic groups. Overall, the Hybrid Model gave the best performance in predicting diabetes risk while maintaining a high level of reliability and comprehensiveness, although the Autoencoder had the best precision (90.8%).

DISCUSSION

This study shows that anomaly-based modelling with physiological and lifestyle information is effective in predicting diabetes risk. The analysis reveals that the fusion of various anomaly detection methods yields better prediction accuracy, particularly when dealing with the complex and imbalanced nature of healthcare data. The Hybrid Model consistently performed the best when compared with the other evaluated models with an accuracy of 92.3% and ROC-AUC score: 0.95. This illustrates that it performs well in the right identification of diabetic and non-diabetic cases and its ability to reduce falsely predicted cases. The Autoencoder performed well as well, indicating the potential of deep learning in uncovering patterns in medical data. The findings also showed that the physiological parameters like blood glucose, BMI, insulin level, and blood pressure level have significant influences in diabetes risk prediction. The blood glucose level was determined as the most important predictor, and BMI as a proxy of metabolic health, were found to be closely related to the occurrence of diabetes. Bad lifestyle habits also have a significant impact, in that individuals who are inactive, smoke, drink alcohol, and eat a poor diet develop insulin resistance and poor health in general. The results are in line with recent studies emphasising the need for clinical and lifestyle data to be correlated for better prediction of diabetes. Previous studies indicated that traditional machine learning algorithms are less effective in imbalanced data, whereas the high-risk anomaly-based and hybrid algorithms are more effective in detecting high-risk scenarios. The findings in this study agree with and replicate the conclusions of the previous study, further indicating that using both the anomaly detection process and the application of advanced machine learning techniques can be a big increase in the accuracy and reliability of predicting diabetes at an early stage.

CONCLUSION

This research aims to predict diabetes risks using anomaly-based modelling of physiology and lifestyle data. The main objective was to develop an intelligent system that is able to identify a pattern in health related data that

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

could lead to an increase in the likelihood of diabetes in a person. For this purpose, a number of anomaly detection methods, including Isolation Forest, One-Class SVM, Local Outlier Factor, Autoencoders, and Hybrid Model were implemented and tested on healthcare datasets. The methods used included data pre-processing, feature selection, model training, and model evaluation based on different performance measures such as accuracy and ROC-AUC scores. The primary results showed that anomaly-based models work well for diabetes risk prediction, especially when dealing with medical data that is imbalanced and complicated. The Hybrid Model performed best of all models evaluated, having the greatest accuracy and the highest ROC-AUC score, which is higher the more accurate a model can predict. Furthermore, certain important predictive features were identified in the study: blood glucose level was the most important predictor, followed by BMI, then insulin level and age and finally blood pressure. Further, it was discovered that lifestyle factors poor diet, smoking, alcohol, and physical inactivity – also have a significant impact on the risk of diabetes. The contributions of this research are important in a number of aspects. It adds a more potent anomaly-based approach to the early diagnosis of disease and improves prediction of healthcare. It also has the potential to support preventive health systems, and to early identify people at risk and intervene accordingly. Additionally, the study has relevance in the field of medical analytics and artificial intelligence since it demonstrates how the hybrid anomaly detection models can be applied in medical fields. Finally, the outcome of this research would be helpful for developing and improving the intelligent diabetes prediction systems in future.

REFERENCE

1. Alghamdi, T. (2023). Prediction of diabetes complications using computational intelligence techniques. Applied Sciences, 13(5), 3030. https://doi.org/10.3390/app13053030
2. Banday, M. Z., Sameer, A. S., & Nissar, S. (2020). Pathophysiology of diabetes: An overview. Avicenna Journal of Medicine, 10(4), 174–188. https://doi.org/10.4103/ajm.ajm_53_20
3. Bontha, S. S., Jammalamadaka, S. K. R., Vudatha, C. P., Jammalamadaka, S. B., Duvvuri, B. K., & Vudatha, B. C. (2025). Predicting risk and complications of diabetes through built-in artificial intelligence. Computers, 14(7), 277. https://doi.org/10.3390/computers14070277
4. Deshpande, A. D., Harris-Hayes, M., & Schootman, M. (2018). Epidemiology of diabetes and diabetes-related complications. Physical Therapy, 88(11), 1254–1264. https://doi.org/10.2522/ptj.2008.0020
5. Fahim, Y. A., Hasani, I. W., Kabba, S., & Ragab, W. M. (2025). Artificial intelligence in healthcare and medicine: Clinical applications, therapeutic advances, and future perspectives. European Journal of Medical Research, 30(1), 848. https://doi.org/10.1186/s40001-025-03196-w
6. Faiyazuddin, M., Rahman, S. J. Q., Anand, G., Siddiqui, R. K., Mehta, R., Khatib, M. N., Gaidhane, S., Zahiruddin, Q. S., Hussain, A., & Sah, R. (2025). The impact of artificial intelligence on healthcare: A comprehensive review of advancements in diagnostics, treatment, and operational efficiency. Health Science Reports, 8(1), e70312. https://doi.org/10.1002/hsr2.70312
7. Guo, R., Smith, R., Chen, Q., Ritchie, A., & Poon, S. (2025). Enhance health evidence quality in classification tasks: A triangulation approach utilizing case-based reasoning and process features. Digital Health, 11, 20552076251314097. https://doi.org/10.1177/20552076251314097
8. Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2023). Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda. Journal of Ambient Intelligence and Humanized Computing, 14(7), 8459–8486. https://doi.org/10.1007/s12652-021-03612-z
9. Li, Z., Li, Y., Mao, Z., Wang, C., Hou, J., Zhao, J., Wang, J., Tian, Y., & Li, L. (2025). Machine learning models integrating dietary indicators improve the prediction of progression from prediabetes to type 2 diabetes mellitus. Nutrients, 17(6), 947. https://doi.org/10.3390/nu17060947
10. Lucier, J., & Mathias, P. M. (2026). Type 1 diabetes. In StatPearls [Internet]. StatPearls Publishing. Updated October 5, 2024. https://www.ncbi.nlm.nih.gov/books/NBK507713/
11. Mathew, T. K., & Zubair, M. (2026). Blood glucose monitoring. In StatPearls [Internet]. StatPearls Publishing. Updated April 12, 2026. https://www.ncbi.nlm.nih.gov/books/NBK555976/
12. Młynarska, E., Czarnik, W., Dzieża, N., Jędraszak, W., Majchrowicz, G., Prusinowski, F., Stabrawa, M., Rysz, J., & Franczyk, B. (2025). Type 2 diabetes mellitus: New pathogenetic mechanisms, treatment and the most important complications. International Journal of Molecular Sciences, 26(3), 1094. https://doi.org/10.3390/ijms26031094

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

13. Nelson, J., & Chalotte, J. (2025). Predictive analytics in healthcare: Leveraging AI for early disease detection. https://doi.org/10.13140/RG.2.2.12345.67890
14. Petrie, J. R., Guzik, T. J., & Touyz, R. M. (2018). Diabetes, hypertension, and cardiovascular disease: Clinical insights and vascular mechanisms. Canadian Journal of Cardiology, 34(5), 575–584. https://doi.org/10.1016/j.cjca.2017.12.005
15. Sapra, A., & Bhandari, P. (2026). Diabetes. In StatPearls [Internet]. StatPearls Publishing. Updated June 21, 2023. https://www.ncbi.nlm.nih.gov/books/NBK551501/
16. Soomro, S. (2018). Prevalence of inflammation associated diseases in obese and non-obese subjects of Kahuta region. https://doi.org/10.13140/RG.2.2.20153.49763