LEX LOCALIS-JOURNAL OF LOCAL SELF-GOVERNMENT
ISSN:1581-5374 E-ISSN:1855-363X
VOL. 23, NO. S5 (2025)

AI-DRIVEN DIAGNOSIS AND CLASSIFICATION OF DIABETES MELLITUS USING MACHINE LEARNING APPROACHES

Waleed Al Shehri [1] *, Ahmed M. Alghamdi [2]

[1] Department of Computing, College of Engineering and Computing in Al-Lith, Umm Al-Qura University, Makkah, Saudi Arabia
[2] Department of Software Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia

* Corresponding author; Waleed Al Shehri;
washehri@uqu.edu.sa[1]

ABSTRACT: Recent studies have highlighted diabetes as a chronic disease that is spreading worldwide. The World Health Organization (WHO) reports 422 million patients with diabetes worldwide, and this number will increase if diabetes is not adequately controlled. Worldwide, accurate and early diagnosis of diabetes is required. However, there are still deficiencies in the diagnosis, analysis of features, and classification of diabetes types. The proposed model is based on machine learning approaches, where analysis of the features highlights that glucose is the key factor in diagnosing diabetes and insulin is the main feature used to classify diabetes types. The proposed model consists of two major components. The first part discusses model development and training, which involves pre-processing the dataset, extracting features, and training the model on the PIMA Diabetes dataset. This study used four machine learning classifiers for model training: K-Nearest Neighbors, Logistic Regression, Gaussian Naïve Bayes, and Support Vector Machine (SVM). In the second part, the proposed model is evaluated using the PIMA diabetes open-source dataset. After the evolution of the proposed model, the best accuracy was obtained for the training model at 0.9541. The testing achieved an accuracy of 0.9607, which proves that the proposed model performs exceptionally well.

KEYWORDS: Diabetes Diagnosis; Machine Learning; AI-Driven; Mellitus Diagnosis; Glucose Blood Sugar; PIMA.

1 Introduction

Diabetes is a chronic illness affecting many people worldwide. It occurs when the pancreas is unable to produce enough insulin or when the insulin that is produced is not used efficiently to help control blood sugar levels. Blood glucose or blood sugar levels can increase dangerously as a result of diabetes. Glucose is a vital component of human nutrition, serving as the primary source of energy for the body. Insulin, a hormone produced by the pancreas, aids in the transport of glucose from food into cells, where it is converted into energy. Insulin helps regulate glucose and blood sugar levels in the body. Insulin resistance or insufficient insulin production can occasionally cause glucose to remain in the bloodstream and not be adequately absorbed by the cells. Diabetes is a variety of diseases caused by persistently elevated blood glucose levels. Diabetes is harmful because it damages nerves and blood vessels [1].

According to estimates, 8.5% of people aged 18 and over had diabetes in 2014. One and a half million deaths before the age of 70 were caused by diabetes. Diabetes affects 422 million people globally, primarily in low-income and emerging market countries (WHO, 2021). Over the past few years, the number of people with diabetes has increased, and diabetes is currently one of the major causes of death [2].

Type 2 diabetes, sometimes referred to as non-insulin-dependent diabetes, is one of the two primary forms of diabetes. Inefficient use of insulin by the body is the cause of this type of diabetes. Type 2 diabetes affects approximately 95% of patients. Higher body weight and inactivity are causes of type 2 diabetes. Although this type of diabetes is difficult to identify and may not be identified for years after onset, its signs are nearly identical to those of type 1 diabetes. Deficient insulin production causes type 1 diabetes, sometimes referred to as insulin-dependent diabetes, which requires daily insulin injections. Patients with type 1 diabetes typically experience increased thirst, weight loss, polyuria, weariness, and eyesight [3].

Owing to its greater prevalence, type 2 diabetes has a larger impact than type 1 diabetes. The risk of developing type 2 diabetes can be reduced by maintaining a healthy body weight, having a nutritious diet, being physically active, and avoiding tobacco use. Early diagnosis saves time and can provide a prompt test for the presence of type 1 diabetes. Testing is relatively inexpensive, and early diagnosis plays a crucial role in effective treatment. Specifically, early diagnosis can save lives. However, it is not easy to diagnose type 2 diabetes because of its unclear symptoms [3]. There are two significant aspects to diagnosing diabetes: first, determining the presence of diabetes and then classifying the type of diabetes.

The remainder of this paper is structured as follows: Section 2 discusses previously implemented techniques for diabetes diagnosis and classification. Section 3 highlights the proposed methodology and elaborates on the dataset insights along with their detailed parameters. Section 4 presents the statistical results and provides a final discussion of the proposed model. Finally, Section 5 presents the conclusion and suggestions for future work that can enhance the domain with very accurate solutions.

2 Literature Review

Diabetes analysis, diagnosis, and classification are significant challenges that must be addressed to develop optimal, real-time, and efficient solutions. Various techniques have been developed to diagnose and classify type 1 and type 2 diabetes accurately. This is important because type 2 diabetes often does not exhibit symptoms in its early stages; rather, they typically appear later. The following subsections present various techniques for diagnosing and classifying the two types of diabetes.

2.1 Diabetes Diagnosis Using Machine Learning

The ecological impact of e-healthcare and machine learning is an effective technique. They play an essential role in the diagnosis, analysis, and classification of diabetes. Diabetes diagnosis in the e-healthcare domain remains a challenging problem in the research community that needs to be addressed to obtain an accurate diagnosis. Research method for diabetes diagnosis in an e-healthcare environment using machine learning approaches on clinical patient data. Currently implemented approaches are lacking in areas such as accuracy and require considerable computing time [4]. Feature selection was performed using the iterative dichotomized three-technique and a Decision Tree algorithm, and patient data were validated using the k-fold validation approach. The feature choice procedure utilized two additional ensemble algorithms: Random Forest and AdaBoost. Ultimately, the data were classified as either healthy or diabetic using the decision tree technique. The suggested technique was successful because it combined the best features and plasma glucose concentrations. The statistical results show that the proposed research method outperforms the inclusion of machine learning approaches.

Early detection and diagnosis of diabetes are crucial for maximizing lifesaving potential. Various techniques focus on the annual diagnosis of diabetes [5]. They proposed a technique that can diagnose diabetes every year using machine-learning algorithms. To determine the difference between healthy and diabetic participants, machine learning methods such as K-Nearest Neighbors, Support Vector Machine (SVM), Random Forest, Logistic Regression, Naïve Bayes, and Gradient Boosting were employed. The proposed framework was assessed using the PIMA Indian free and open-source datasets. The statistical results show that the proposed research method yields better outcomes, and after refining the entire system, it can be applied in the clinical analysis of patients with diabetes. The comparison results for all these classifiers indicate that this system does not yield the best statistical results and requires improvement in system performance through parameter tuning.

Additionally, this system does not follow the ETL-type process to provide the optimal solution. The maximum accuracy values were obtained using the RF algorithm, but none of the other classifiers yielded better results. Why do these algorithms not perform better on the same dataset? The proposed system did not discuss any pre-processing techniques needed to manage the dataset for better performance. Early diagnosis helps to easily handle serious diseases and avoid type 2 diabetes through early intervention [6].

Classifying type 2 diabetes is a challenging task, and a technique for this task. They used machine learning approaches to the clinical data of patients. The logistic regression algorithm outperformed type 2 prediction on the PIMA open-source dataset. The proposed model utilizes a six-fold classification tree to predict the features most effective in diagnosing diabetes, including BMI, glucose levels, age, and pregnancy status. The statistical evaluation indicates that the proposed model for type 2 prediction outperforms existing models and can be effectively implemented in clinical settings for patient diagnosis [7].

One of the main objectives of diabetes analysis and diagnosis is the identification of diabetic features. These traits were identified using various methods. Researchers have suggested new methods to determine the characteristics of diabetes. The Saudi Arabian dataset and different machine learning techniques for diabetes diagnosis were the primary topics of discussion. They employed a feature-based approach to obtain several useful features. A nearby hospital provided the dataset that was used to assess the model. They suggested gathering the data and using pre-processing methods to prepare the input for the model-building procedure. A total of 16 features were extracted. Subsequently, feature selection techniques were used to minimize the size of the feature vector. To determine whether the data belonged to healthy or diabetic individuals, machine-learning algorithms, including logistic regression, SVM, Decision Trees, Random Forests, and Ensemble Models, were employed. The proposed model outperformed the extracted features, as indicated by the findings from a statistical analysis [8].

Diabetes is a major cause of various human complications. The detection of diabetes complications is a pressing need, and [9] proposed a solution for predicting diabetes complications. Data mining pipelines were embedded with Machine Learning algorithms to obtain insight into the dataset. The proposed system utilized data from one thousand patients to train the model, which was an EU-funded project aimed at delivering a solution for predicting diabetes complications. The purpose of this system was to perform clinical center profiling, specifically for diabetes complications, and to validate the model using state-of-the-art approaches. Pre-processing is applied to remove noise and missing values, and then machine learning algorithms, including Random Forest and Logistic regression, are used for model training. The proposed system was evaluated using the dataset discussed above, achieving an accuracy of 0.838. The major problem was not addressed in the methodology used to predict and analyze early diabetes complications [10].

The number of diabetes patients is increasing day by day due to many factors, such as diet plans, lack of exercise, bad smoking habits, and unhealthy diet. Owing to the large volume of the dataset, an extensive computational and simulation system is required. Big data analytics involves analyzing large volumes of data and performing complex computations. Pre-processing steps were applied to remove noise and analyze the missing values, allowing for their treatment. After pre-processing, machine learning algorithms are used to train on a large volume of data, leveraging analytics of large datasets. This yields better results compared to classical approaches. Still, due to the large volume of training data, the problem of overfitting arises, which does not provide a proper solution for accurately predicting type 2 diabetes [11].

Table 1. State-of-the-art survey for existing approaches

No	Reference	Method	Dataset	Evaluation
1	[4]	e-Health Environment using machine learning Decision tree iterative dichomister 3	Private Dataset	Accuracy
2	[5]	Early diabetes detection using machine learning approaches	PIMA Dataset	Accuracy
3	[6]	Diabetes complications predicted using machine learning approaches	Private Dataset	Accuracy
4	[7]	Big data analysis used for diabetes detection in large volume dataset	Private dataset	Accuracy
5	[8]	Machine learning and ensemble learning approaches used to predict diabetes	Local Hospital dataset	Accuracy
6	[9]	Logistic regression algorithm used for the prediction of diabetes	PIMA Dataset	Accuracy

Table 1 lists existing state-of-the-art approaches, categorizing them by their implemented solutions. After analyzing the various approaches mentioned, the key research gaps still need to be addressed. These research gaps should be addressed in a generalized approach that can be used efficiently.

Paper size: US Letter (8.5″ × 11″ or 21.59 cm × 27.94 cm).

2.1 Research Gap

This literature review reveals some research gaps that can be addressed. Various techniques are discussed in the literature review, and a few key points that can be addressed in the research work are outlined below—namely, accuracy in distinguishing between a patient with symptoms of health or diabetes. Accurate diagnosis using machine learning approaches and real-time solutions is required for hospitals and clinics [12] [13]. Highlight the false-negative type, where a patient already exhibits diabetes symptoms, but the system incorrectly identifies them as a normal patient. Identify the false-positive where a person does not have diabetes symptoms, but the system identifies him as a diabetes patient [14]. The accurate identification of diabetes symptoms in type 2 diabetes remains complex, particularly in distinguishing between type 2 and type 1 [15].

3 Proposed Methodology

Based on the identified research gaps, a model utilizing machine learning techniques is proposed for the diagnosis and classification of diabetes. This model was executed using clinical data alone. The details of how the model was developed, trained, and tested are shown in Figure 1. The proposed methodology was divided into two major parts.

3.1 Material and Design

An open-source dataset [9] was used to train the proposed model. This dataset has some important features that are helpful for this study. The dataset contains information on the diagnosis of diabetes. It originates from the National Institute of Diabetes and is openly available. The data represent female patients aged 21 years and older with PIMA Indian heritage. Data included variables such as disease pedigree, age, skin thickness, glucose level, blood pressure, pregnancy status, and BMI. Table 1 presents these features. Some calculations were performed to analyze the data and strength of each feature in the dataset.

Furthermore, the minimum and maximum values for each feature were identified, and the standard deviation (STD) and mean values were calculated to provide insight into the data. The outcome feature was represented by 0 for healthy patients and 1 for patients with diabetes. All datasets were pre-labeled, and Table 2 provides a description of the datasets and their statistical analysis.

Table 2. Diabetes Dataset Description and Statistical Analysis

S. No	Feature Name	Feature Description	Min-Max	Mean, (±) STD
1	Pregnancies	The total amount of pregnancies	0 - 17	3.703500, (±) 3.306063
2	Glucose	Plasma glucose concentrations	0 - 199	121.182500, (±) 32.068636
3	Blood Pressure	BP (mm Hg)	0 - 122	69.145500, (±)19.188315
4	Skin Thickness	Triceps skinfold thickness(mm)	0 - 110	20.935000, (±) 16.103243
5	Insulin	Serum insulin concentration	0 - 744	80.254000, (±)111.180534
6	BMI	Blood mass index	0 - 80.60	32.193000, (±) 8.149901
7	DFI	Genetic function for diabetes	0.078 - 2.42	0.470930, (±) 0.323553
8	Age	Age in years	21 - 81	33.090500, (±)11.786423
9	Outcome	Healthy = 0 Diabetes = 1	0 - 1	0.342000, (±) 0.474498

3.2 Build and Train the Model

In this phase, the first step was to arrange the diabetes dataset. Data were chosen from open sources or via IoT devices or sensors. This study utilized a large, open-source dataset. Table 2 lists the details of the dataset used in the model. In the next step, the Extract, Transform, Load (ETL) method was used. The data used for training purposes were uploaded and pre-processed to obtain the data in the form required by our model. This pre-processing includes data transformation, data removal, and noise in the data. Techniques were employed to predict the missing values in the data, enabling the model to perform more accurate calculations.

All the data were labeled for training purposes, and the necessary features for classification were extracted. These features can be used in machine-learning algorithms. There is also the option of using deep learning, where feature extraction is automatically performed. However, the focus was on machine-learning approaches, such as K-Nearest Neighbor, Naïve Bayes, Support Vector Machine (SVM), and Logistic Regression methods.

The model was trained using machine learning algorithms, including Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, K-Nearest Neighbor, and Linear Regression. After successful training, the model was validated, and the best model was selected for further testing. The chosen model was advanced to the next phase, where it was utilized for various purposes.

3.3 Deploy and Test the Model

In this phase, the trained model is loaded and used to evaluate its performance. The proposed model was tested under the same conditions, and new unlabeled data were used. Model testing was performed using a trained model with possible test results for diabetes or no diabetes. The proposed model was evaluated to assess the accuracy of the evaluation measures, including accuracy, precision, recall, and F1-measure.

Ultimately, the trained model can be deployed to assist doctors, patients, and hospitals, as well as for specific research purposes.

3.4 Model Evaluation

Accuracy is used as a measure to evaluate the proposed model in combination with all the proposed machine learning algorithms used in the model. Training and testing accuracies were calculated for the given dataset. The equation for accuracy is given by Equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \, x \, 100\% \qquad (1)$$

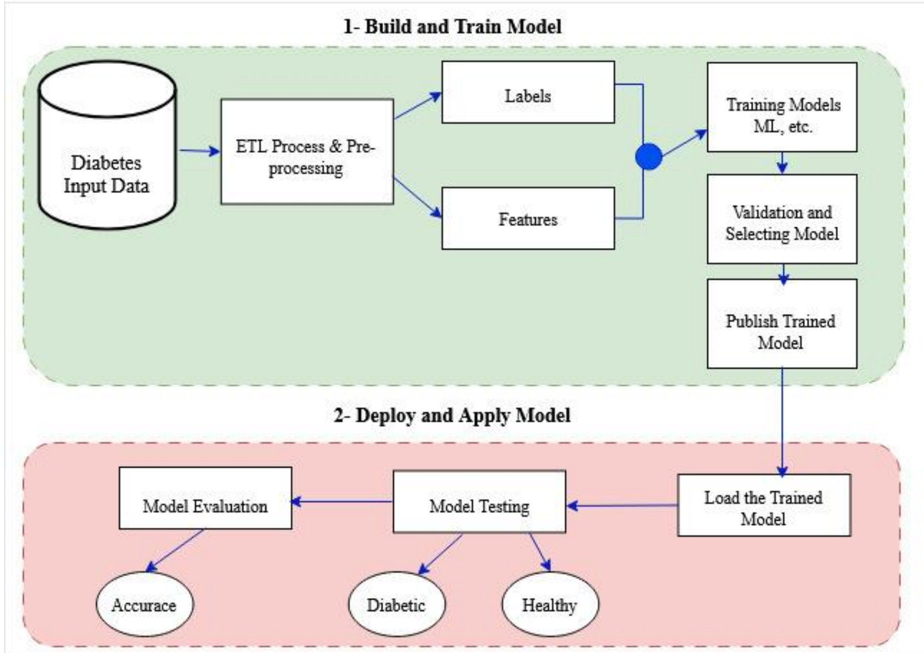

Figure 1. The proposed model methodology

4 Results and Discussion

Diabetes diagnosis and classification of diabetes types are key roles of the proposed model. The model's ability was analyzed to accurately diagnose diabetes and classify type 2 diabetes, a challenging task due to its often hidden symptoms. Existing approaches to these tasks utilized an open-source dataset ("PIMA Indian Dataset"), which was also used to evaluate the proposed model. The Internet of Medical Things (IoMT) and Machine Learning play a significant role in diabetes diagnosis. Various techniques have been applied and utilized, but they have not yielded accurate results in the healthcare domain. This methodology is presented in the following subsections.

4.1 Data Analysis and Interpretation for Features

Each feature was analyzed and presented with a graphical representation of the results. Each factor also has an impact on the model. Correlations were used to determine the statistical relationships between the features or variables. A visualization of this correlation is illustrated using a heat map, as shown in Figure 2.

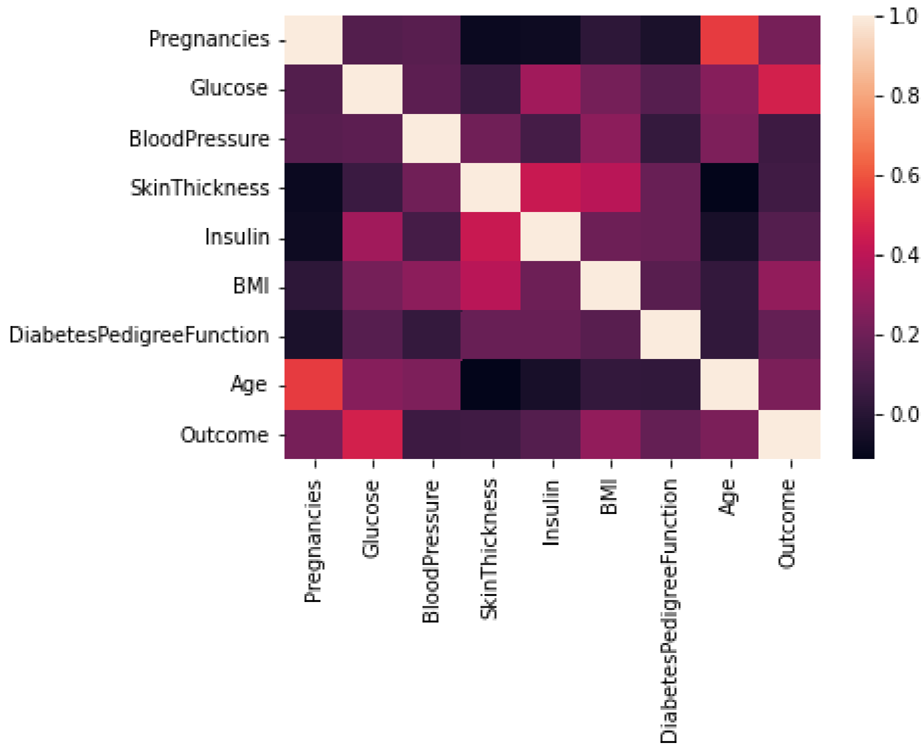

Figure 2. Statistically, relationship analysis between variables using correlation

The heat map analysis showed a correlation between variables, with brighter colors indicating a higher correlation. This indicates that glucose has the highest correlation with the dependent variables in the given data. Insulin had the second-highest correlation among the independent variables; however, it has been medically proven not to be correlated with the outcome. Diabetes types will be predicted using insulin levels. After these experiments, it was concluded that diabetes can be predicted using glucose, and insulin can be used to classify the types of diabetes. Based on the analysis and interpretation, it was concluded that these two variables are very effective [17]-[20].

4.2 Diabetes Diagnosis Results

The proposed model uses four different classifiers: K-Nearest Neighbors, Logistic Regression, Gaussian Naïve Bayes, and Support Vector Machine. To evaluate the proposed model, the PIMA dataset was utilized, along with various evaluation measures, including accuracy, precision, recall, and F1 score. The evaluation reveals that the statistical results for each classifier are excellent; however, Logistic Regression and K-Nearest Neighbors yielded the best results, as shown in Table 3. The maximum values for accuracy, precision, recall, and f-measure were 0.9607, 0.9622, 0.9607, and 0.9612, respectively. This proves that the proposed model performed exceedingly well.

Table 3. Diabetes diagnosis results for the classifiers against different measures

S. No	Classifier	Accuracy	Precision	Recall	F1 Measure
1	K-Nearest Neighbors	0.9607	0.9622	0.9607	0.9612
2	Logistic Regression	0.9607	0.9614	0.9607	0.9604
3	Gaussian Naïve Bayes	0.9411	0.9453	0.9411	0.9423
4	SVM	0.9346	0.9369	0.9346	0.9349

After training the proposed model using the classifiers mentioned above, it was tested with unknown values for glucose and insulin features, as these features have a significant impact on the ability to predict and classify diabetes. Three test types were established for model analysis: normal, pre-diabetic, and diabetic.

In the 1st case, a glucose level of 120 and an insulin level of 30 were used. The system predicted a '1' output, which represents diabetes type 1. The probability shows that it is pre-diabetes, and you have only a 20 % chance of being diabetic.

In the second case, 160 was considered the glucose value, and 30 the insulin value. The proposed model showed result 2'. The probability of prediction indicates that the individual has diabetes, with a 100% chance of being diabetic.

In the 3rd type, it is clear that if the two above types are not present, then the person will be normal and will not exhibit any diabetes symptoms. Therefore, all other values will be for the normal case rather than the values mentioned for pre-diabetes and diabetes.

4.3 Diabetes Type Classification Results

Four distinct classifiers were employed to assess the proposed model, and the model was trained using each classifier to obtain accuracy values of 0.9508, 0.9541, 0.9327, and 0.8934 for Logistic Regression, Gaussian Naïve Bayes, SVM, and K-Nearest Neighbors, respectively. Here, Logistic Regression performed the best, as shown in Table 4. After training, the test was performed on the same classifier, and the accuracy values obtained for the K-Nearest Neighbors, Logistic Regression, Gaussian Naïve Bayes, and SVM were 0.9607, 0.9477, 0.9477, and 0.9346, respectively. Here, for testing purposes, the K-Nearest Neighbors algorithm performed the best, as shown in Table 4.

Table 4. Diabetes type classification results for training and testing the model

S. No	Classifier	Training Accuracy	Test Accuracy
1	K-Nearest Neighbors	0.9508	0.9607
2	Logistic Regression	0.9541	0.9477
3	Gaussian Naïve Bayes	0.9327	0.9477
4	SVM	0.8934	0.9346

After training, some testing was performed on the trained models, and the values for glucose were 180 and for insulin were 300; the results indicated that it was a case of type 2 diabetes. At the same time, the glucose and insulin values are updated to 130 and 30, respectively, and the model indicates that this is a case of type 1 diabetes.

4.4 Comparison with the state-of-the-art

The proposed model outperformed the state-of-the-art techniques under the same circumstances and datasets, as shown in Table 5. Different classifiers were employed to analyze the dataset's features.

Table 5. Comparative analysis with state-of-the-art

S. No	Classifier	Reference [16]	Reference [17]	Proposed Model
1	K-Nearest Neighbors	63.04	73.43	0.9607
2	Logistic Regression	--	77.60	0.9477
3	Gaussian Naïve Bayes	73.48	75.52	0.9477
4	SVM	77.73	65.63	0.9346

As mentioned, the approaches in Table 5, along with their statistical values, utilized K-Nearest Neighbor, Gaussian Naïve Bayes, and Support Vector Machine (SVM), and the values obtained from the computation were not satisfactory [16]. The approach used was based on optimal feature selection; however, the important features could not be utilized to their full potential, which was necessary for the optimal system. In [17], the statistical values are better than those in other studies; however, the proposed system was unable to address diabetes problems properly and failed to train the machine learning algorithms as effectively as the proposed model, as shown in Table 5.

5 Conclusion and Future Work

The proposed model diagnoses diabetes and provides a classification for type 2 or type 1 diabetes. Glucose and insulin were the major components of the analysis. Glucose is used in the diagnosis of diabetes, and insulin is used to classify the types of diabetes. Four classifiers were used: K-Nearest Neighbors, Logistic Regression, Gaussian Naïve Bayes, and SVM, which provided the best results to demonstrate the model's accuracy. In the future, a very large-volume dataset can be used to extract more features. Additional features can help diagnose diabetes more easily and can be applied in clinics or hospitals to reach a large population. Deep learning-based approaches will be applied to large-volume datasets, and these datasets will be managed using blockchain technologies to process their complex structure. IoT devices will be used to collect different types of datasets, and advanced approaches in deep learning models will be utilized to perform real-time diagnoses.

6 Statements

Acknowledgement: Not applicable.
Funding Statement: The authors received no specific funding for this study.
Author Contributions: All authors contributed equally to the research.
Availability of Data and Materials: The data that support the findings of this study are openly available in [Kaggle] at [https://www.kaggle.com/uciml/pima-indians-diabetes-database].

Ethics Approval: Not applicable.
Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

[1]. E. R. F. C. a. others, "Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies," The Lancet, vol. 375, pp. 2215--2222, 2010.
[2]. WHO, "Diabetes," 10 11 2021. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes. [Accessed 6 12 2021].
[3]. J. D. Steinmetz, R. R. A. Bourne, P. S. Briant, S. R. Flaxman, H. R. B. Taylor, J. B. Jonas, A. A. Abdoli, et al.,., "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness about VISION 2020: the right to sight: an analysis for the Global Burden of Disease Study," The Lancet Global Health, vol. 9, no. 2, pp. e144–e160, 2021.
[4]. A. U. Haq, J. P. Li, J. Khan, M. H. Memon, S. Nazir, S. Ahmad, G. A. Khan, and A. Ali, "Intelligent machine learning approach for effective recognition of diabetes in e-healthcare using clinical data," Sensors, vol. 20, no. 9, p. 2649, 2020.
[5]. D. J. Reddy, B. Mounika, S. Sindhu, T. P. R. Reddy, N. S. Reddy, G. J. Sri, K. Swaraja, K. Meenakshi, and P. Kora, "WITHDRAWN: Predictive machine learning model for early detection and analysis of diabetes," 2020.
[6]. A. Dagliati, S. A. S. L. Marini, G. Cogni, M. Teliti, V. Tibollo, P. De Cata, L. Chiovato and R. Bellazzi, "Machine learning methods to predict diabetes complications," Journal of diabetes science and technology, vol. 12, pp. 295--302, 2018.
[7]. A. Mujumdar and V. Vaidehi, "Diabetes prediction using machine learning algorithms," Procedia Computer Science, pp. 292--299, 2019.
[8]. H. F. a. M. H. a. A. H. a. S. M. a. A. A. Ahmad, "Investigating Health-Related Features and Their Impact on the Prediction of Diabetes Using Machine Learning," Applied Sciences, vol. 11, p. 1173, 2021.
[9]. U. M. Learning, "Pima Indians Diabetes Database," [Online]. Available: https://www.kaggle.com/uciml/pima-indians-diabetes-database. [Accessed 6 12 2021].
[10]. M. a. S. R. a. M. A. M. a. I. V. Saravanan, "SMEAD: A secured mobile enabled assisting device for diabetics monitoring," in 2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), IEEE, 2017, pp. 1--6.
[11]. R. F. a. E. A. A. a. N. Mansour, "Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems," IEEE Access, vol. 9, pp. 45137--45146, 2021.
[12]. NIDDk, "What is Diabetes?," [Online]. Available: https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes. [Accessed 6 12 2021].
[13]. S. M. K. Mbengue, O. Diallo, E. H. M. Ndoye, J. J. P. C. Rodrigues, A. Neto, and J. Al-Muhtadi, "Internet of medical things: Remote diagnosis and monitoring

application for diabetics," in 2020 International Wireless Communications and Mobile Computing (IWCMC), IEEE, 2020, pp. 583–588.
[14]. O. a. A. B. a. A. M. a. A. F. AlShorman, "A review of internet of medical things (IoMT)--based remote health monitoring through wearable sensors: A case study for diabetic patients," Indonesian Journal of Electrical Engineering and Computer Science, vol. 20, pp. 414--422, 2020.
[15]. R. a. N. B. a. A. D. G. Basatneh, "Health sensors, smart home devices, and the internet of medical things: an opportunity for dramatic improvement in care for the lower extremity complications of diabetes," Journal of diabetes science and technology, no. 12, pp. 577--586, 2018.
[16]. N. a. G. T. Sneha, "Analysis of diabetes mellitus for early prediction using optimal features selection," Journal of Big data, vol. 6, pp. 1--19, 2019.
[17]. R. Alazaidah, A. Al-Shaikh, M. R. AL-Mousa, H. Khafajah, G. Samara, M. Alzyoud, N. Al-Shanableh, S. Almatarneh, Website Phishing Detection Using Machine Learning Techniques, J. Stat. Appl. Pro. Vol. 13, No. 1 (2024), PP:119-129 doi:10.18576/jsap/130108
[18]. Azhari A. Elhag, Prediction and classification of Tuberculosis using machine learning, J. Stat. Appl. Pro. Vol. 13, No. 3 (May 2024), PP:939-946 doi:10.18576/jsap/130308
[19]. Serkan Onar, Bayram Ali Ersoy, Kostaq Hila, Krisanthi Naka, (T,S)-Intuitionistic Fuzzy Hyperideals of Γ -Hyperrings, Applied Mathematics & Information Sciences, Vol. 15, No. 3 PP: 329-342 (2021), doi:10.18576/amis/150311
[20]. Majed G. Alharbi, Hamiden Abd El- Wahed Khalifa, On Solutions of Fully Fuzzy Linear Fractional Programming Problems Using Close Interval Approximation for Normalized Heptagonal Fuzzy Numbers, Applied Mathematics & Information Sciences, Vol. 15, No. 4 PP: 471-477 (2021) PP: 471-477 doi:10.18576/amis/150409