Babylonian Journal of Artificial Intelligence
Vol. 2024, **pp**. 64–72
DOI: https://doi.org/10.58496/BJAI/2024/009 ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BJAI

Research Article

Healthcare Analysis Based on Diabetes Prediction Using a Cuckoo-Based Deep Convolutional Long-Term Memory Algorithm

T. Kavitha[1*] G.Amirthayogam[2] J. Jasmine Hephzipah[3] R.Suganthi[4] Venkata Anjani Kumar G[5] T.Chelladurai[6]

[1] Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu 600062, India .

[2] Department of Information Technology, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu 603103, India.

[3] Department of Electronics and Communication Engineering, R.M.K. Engineering College, Kavaraipettai, Tamil Nadu 601206, India,

[4] Department of Electronics and Communication Engineering, Panimalar Engineering college Poonamallee, Chennai-600123, Tamil Nadu, India.

[5] Department of EEE, Rajiv Gandhi University of Knowledge Technologies(RGUKT) -Andhra Pradesh (Ongole campus), Andhra pradesh-523001, India.

[6] Department of Electronics and Communication and Engineering, PSNA College of Engineering and Technology, Dindigul 624622.Tamil Nadu,India.,

ARTICLE INFO

Article History
Received 17 Mar 2024
Accepted 18 May 2024
Published 11 Jun 2024

Keywords
Internet of Things
Diabetes prediction
healthcare system
CDC-LSTM
Feature selection
FBDT

ABSTRACT

In recent years, the demand for mobile medical applications utilizing the Internet of Things (IoT) for diabetes diagnosis has been progressively increasing. Diabetes is commonly known as a chronic illness that presents a significant danger to individuals, analysing when blood sugar levels surpass the typical range. If diabetes is not promptly treated, hyperosmolar conditions can lead to serious health issues like hyperglycaemia and possibly even death. Since early detection enables lifestyle changes that prevent the disease's progression, it is crucial for diabetes management and health systems. However, diabetes diagnosis has a long computational time and low prediction accuracy. To address this issue, we propose a Cuckoo-Based Deep Convolutional Long-Term Memory (CDC-LSTM) algorithm that increases accuracy by classifying diabetics or non-diabetics. Additionally, we utilize the Standardized Feature Scaler (SFS) method to normalize the variance data by removing the mean of each feature. Moreover, we select the optimal features in the diabetes dataset utilising the Filter-Based Decision Tree (FBDT) technique. Finally, the proposed CDC-LSTM method can be used to distinguish between diabetics and non-diabetics, improving the accuracy of identifying diabetic patients. Additionally, the proposed method can predict diabetes using performance assessments such as precision, recall, and F-measure. Furthermore, the method's accuracy can be improved to 95.18% compared to previous approaches.

1. INTRODUCTION

Today, e-health services are the fastest-growing field of computer-based care and information transfer to enhance healthcare globally. Analyzing electronic medical records to create accurate disease risk prediction models benefits both patient care and improves service delivery by utilizing data-driven healthcare systems. Moreover, medical services stand as a crucial application area within the realm of the Internet of Things (IoT), providing advantages to patients, doctors, hospitals, and insurance companies. EHealth utilizes internet technologies to enhance healthcare services at all levels - from local to global. It can be applied in clinical, educational, preventative, research, and administrative settings, both onsite and remotely within the healthcare sector [1].

Additionally, by integrating smart devices and sensors, IoT technology has completely transformed the medical and healthcare sectors. Biosensors and IoT-based devices are perfect for diagnosing diabetes since they continuously gather

**Corresponding author. Email: kavithaecephd@gmail.com*

precise blood sugar readings. Diabetes is a major, well-recognized chronic disease with a global social impact on the lives of communities [2]. A significant component of the IoT, medical sensors, devices, imaging equipment, and diagnostics are all regarded as smart objects. Healthcare providers anticipate that IoT-enabled diabetes monitoring services will reduce expenses and enhance user satisfaction, mentioning the ability of IoT devices to address remote downtime issues. Furthermore, IoT enhances the monitoring of diabetic patients and facilitates effective resource planning.
Therefore, the monitoring of blood glucose levels is a difficult task that necessitates the development of an appropriate non-invasive sensing and monitoring framework to promote awareness of self-management techniques among individuals with diabetes. Accordingly, it is crucial to regularly monitor blood sugar levels and control them by administering proper doses of insulin. Nevertheless, the consumption of insulin is affected by various factors such as age, calorific intake, and body weight [3]. However, due to the increasing number of patients, the hospitals are struggling to provide adequate treatment. Therefore, the healthcare industry is exploring effective techniques to reduce the burden of patients suffering from chronic diseases and support them through remote patient care.
This paper contributes, to predicting diabetes utilizing the PIMA India Diabetes Database collected from Kaggle. Furthermore, the CDC-LSTM algorithm was proposed to improve accuracy by classifying diabetics or non-diabetics. Additionally, diabetes data can be normalized by removing the average value of each feature using the SFS method. Furthermore, the FBDT technique can be used to select the best features in the diabetic dataset. Finally, the proposed CDC-LSTM method improves the accuracy of analyzing patients with diabetes.

2. LITERATURE SURVEY

In ancient times, accurate prediction and classification of diabetes types were the most important and difficult tasks in the field of medicine to make the correct diagnosis for patients. For this purpose, previous studies have developed different machine learning based detection systems to predict diabetes from a given dataset. To attain the objective, Pivotal Decision Tree (PDT) methodology was deployed. Because of its complexity and time, its cost is comparatively high. For predicting continuous values, this method is not sufficient to use regression [4]. IoT technology has experienced substantial growth in diagnostics in recent years, concentrating on technological aspects that combine epidemiology and clinical information to enhance healthcare. To accurately identify patients with diabetes, a Deep Learning Integrated-Optimization Diabetes Detection (DeO-DiDet) system was created as a result [5]. The novel aimed to create intelligent patient health monitoring systems using Machine Learning (ML) technology for the precise and timely identification of chronic illnesses [6].
In addition, the accuracy of predictions was assessed using the Pima Indian Diabetes Dataset, a key metric commonly utilized in early detection research on diabetes [7]. IoT-enabled sensor devices improve the management of chronic diseases and predict diabetes in patients through a health monitoring system utilizing predictive models [8]. They utilised the Support Vector Machine (SVM) technique to classify the data collected through the LORA process to predict the severity of diabetes [9].

TABLE I. DIABETES PREDICTION BASED ON HEALTH ANALYSIS IN IOT TECHNOLOGY

Author	Year	Technique Used	Limitation	Accuracy
Haq, A.U [10]	2020	Ada Boost and Random Forest	The current diagnostic techniques suffer from drawbacks like extended processing durations and limited prediction precision.	91.9%
S. Kumari [11]	2021	logistic regression, and Naïve Bayes	Diabetes affects many people around the world	79.04%
Ganie SM [12]	2023	XGBoost, AdaBoost, and gradient boost	The inadequate healthcare system has led to a lack of awareness of the symptoms and complications of diabetes.	92.85%
Alfian, G [13]	2018	LSTM	Regular monitoring of blood sugar levels is crucial in reducing diabetes complications.	71%
Umair Muneer Butt [14]	2021	Multilayer Perceptron (MLP)	Due to the inability of patients to use the insulin they produce, diabetes progresses rapidly.	86.8%
S. Gopalakrishnan [15]	2023	Convolutional Neural Network (CNN)	Though, the existing systems have certain limitations. Health care programs should prioritize disease management efforts to reduce hospitalizations and mortality in heart failure patients.	85%

As shown in table 1, the methods obtained from the literature for predicting diabetes can be analyzed and accurately assessed based on the health analysis of IoT technologies and their limitations.

The type of diabetes is determined using Artificial Neural Network (ANN) models to predict the survival rate of diabetic patients effectively [16]. Enhance diabetes treatment applications with better results using technique assistance and patient monitoring. Furthermore, incorporating packaging and reinforcing techniques may improve the precision of diabetes diagnosis [17].

3. PROPOSED METHODOLOGY

In this section, accuracy can be improved by predicting diabetes using the PIMA India Diabetes Database sourced from Kaggle. Accurate prediction of diabetes is critical for early diagnosis and effective treatment. The PIMA India Diabetes Database derived from Kaggle can be used to develop predictive models and provide valuable data to healthcare professionals. Furthermore, standardization of diabetes data is needed to ensure the reliability and performance of predictive models. The SFS method removes the average of each feature, which can standardize the data and improve the performance of the prediction models. The FBDT technique provides a robust method to identify the most relevant features within the decomposed diabetes datasets. By selecting the optimal features, researchers can streamline the analysis process and improve the overall accuracy of predictive models. The CDC-LSTM algorithm may present a proposed approach for improving the accuracy of classifying individuals as diabetic or nondiabetic.

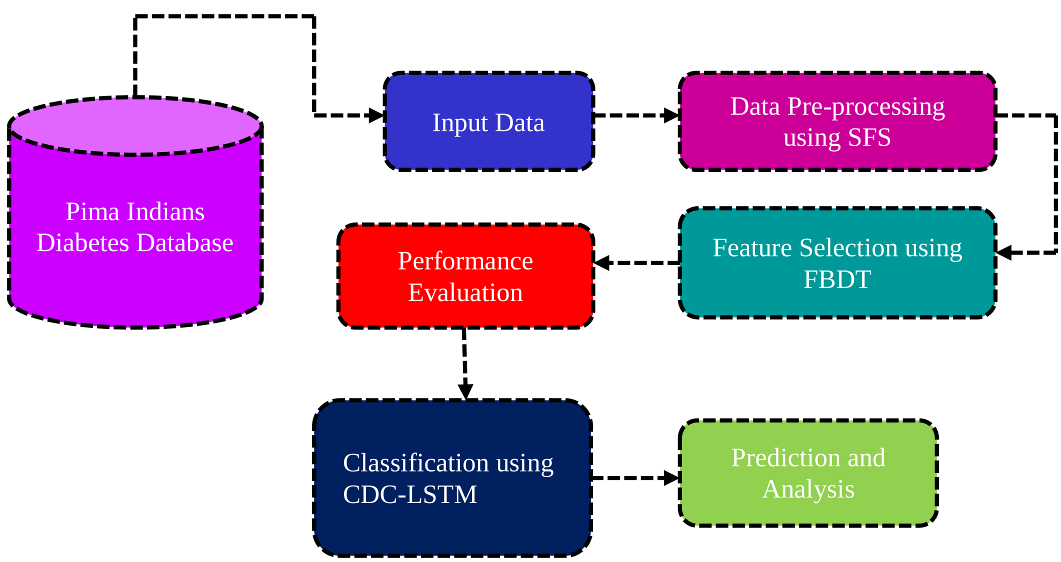

Fig. 1. The Proposed CDC-LSTM Method Architecture Diagram

The objective of the CDC-LSTM technique suggested in figure 1 is to enhance the prediction model's accuracy by determining the most pertinent aspects for the analysis of diabetes cases in the IoT healthcare analysis. This is achieved through the use of the architecture diagram. Additionally, with the CDC-LSTM algorithm, it is possible to predict diabetic patients by distinguishing between diabetics and non-diabetics.

3.1 Dataset Collection

Diabetes can be predicted using the Pima Indians diabetes database collected from Kaggle, and the 769 features identified in these data can be used to classify diabetes. Also, using the diabetes dataset, training-571 data and testing-148 data were used to classify them as diabetic or non-diabetic. In addition, these can be activated at https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database to identify diabetes.

1	Pregnanci	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	6	148	72	35	0	33.6	0.627	50	1
3	1	85	66	29	0	26.6	0.351	31	0
4	8	183	64	0	0	23.3	0.672	32	1
5	1	89	66	23	94	28.1	0.167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	0.201	30	0
8	3	78	50	32	88	31	0.248	26	1
9	10	115	0	0	0	35.3	0.134	29	0
10	2	197	70	45	543	30.5	0.158	53	1
11	8	125	96	0	0	0	0.232	54	1
12	4	110	92	0	0	37.6	0.191	30	0
13	10	168	74	0	0	38	0.537	34	1
14	10	139	80	0	0	27.1	1.441	57	0
15	1	189	60	23	846	30.1	0.398	59	1
16	5	166	72	19	175	25.8	0.587	51	1
17	7	100	0	0	0	30	0.484	32	1
18	0	118	84	47	230	45.8	0.551	31	1
19	7	107	74	0	0	29.6	0.254	31	1
20	1	103	30	38	83	43.3	0.183	33	0
21	1	115	70	30	96	34.6	0.529	32	1
22	3	126	88	41	235	39.3	0.704	27	0
23	8	99	84	0	0	35.4	0.388	50	0

Fig. 2. Diabetes Dataset Feature Collection

Figure 2 shows that diabetes can be classified and analyzed through testing and training using a diabetes dataset. By gathering features from datasets, researchers can identify diseases and develop effective strategies for treatment and prevention.

3.2 Standardized Feature Scalar (SFS)

In this section, the data can be normalized by removing the mean value of each feature based on the SFS method. The SFS technique can be utilized to eliminate outliers in diabetes data, enhancing scaling and efficiency. Additionally, standard deviation creates a distribution standard with a mean of 0 and a variance of 1.

A standard scalar was plotted in Equation 1 to account for missing values in a data set. Let's assume $\widehat{I_x}$ −standard scalar, μ −mean, σ −standard deviation, I-standardized score, and K-normalized value.

$$\widehat{I_k} = \frac{I_k - \hat{I}}{\sigma} \quad (1)$$

$$\mu = \frac{1}{K}\sum_{x=1}^{K}(I_x) \quad (2)$$

$$\sigma = \sqrt{\frac{1}{K}\sum_{x=1}^{K}(I_x - \mu)^2} \quad (3)$$

Calculate the minimum and maximum normalized feature of the standard deviation as shown in equation 4. Let's assume J'_x −scaled value, J-original value, Mi_i −minimum value, Ma_i −maximum value.

$$J'_x = \frac{J - Mi_x}{Ma_x - Mi_x} \quad (4)$$

The minimum and maximum features can be analyzed by normalizing the feature calling in the diabetes dataset.

3.3 Filter Based Decision Tree (FBDT)

In this section, diabetes can be analyzed by employing the FBDT technique for optimal feature selection. Filter-based techniques assess feature relevance based on their correlation with the dependent variable, whereas feature selection methods determine a subset of features through the selection process. Furthermore, the FBDT technique selects the feature with the lowest entropy or highest information gain value. Entropy can be used to quantify the uncertainty in a dataset. Furthermore, entropy can be calculated by selecting the best features in the original dataset. Furthermore, the ratio between the count of elements in a class and the set can be analyzed.

Binary feature selection is calculated using the diabetes dataset as shown in equations 5 and 6. Let's assume I-sample set, r-target label, t-features, I_x −target label, X-output label matrix, J_x −output target classes,

$$\begin{gathered} \mathrm{I=J\{i_1,i_2,\ldots,i_k\}} \\ \mathrm{j=\{j_0,j_1\}} \\ \mathrm{S=\{d_1,d_2,\ldots,d_t\}} \end{gathered} \quad (5)$$

$$D(I,J) = \left\{(i_x, j_x) | I_x \in T^k, J_x \in \{j_0, j_1\}\right\}_{x=1}^{n} \quad (6)$$

As shown in equation 7, calculate the degree of uncertainty entropy in the data set. Let's assume D-root node, S-feature dataset, $S(D)$ −entropy, I-number of elements, $w(i)$ −proportion of number of elements in class,

$$S(D) = \sum_{i \in I} -w(i) \log_2 w(i) \quad (7)$$

A measure of the entropy difference before a feature's root node, divided by the information gain. Equation 8 indicates that the uncertainty in the set is reduced and the information gain is calculated after splitting the root node set into attributes. Let's assumex_o −information gain, D and r-entropy set, r-number of element, U-attribute,

$$x_o(D,U) = S(D) - \sum r \in r_w(r) S(r) = S(D) - S(D|U) \quad (8)$$

The optimal features can be selected by analyzing the ratio of the number of elements in the original subsets created by dividing the features in the diabetes dataset based on entropy. During the process, features with higher information gain are used to select the initial set of root nodes.

3.4 Cuckoo Based Deep Convolutional Long-Short-Term Memory (CDC-LSTM)

In this section, a cuckoo-based deep convolutional long-short-term memory algorithm is proposed for predicting diabetes, which can improve its accuracy by classifying diabetic or non-diabetic. The CDC-LSTM method can utilize large intermediate kernels to efficiently generate features. Early cell output and coding layers can be duplicated and analyzed within the network for predicting diabetes in healthcare. Additionally, the dimensions of the final output match those of the inputs, which are combined and passed to the transformation layer for predicting diabetes. The Cuckoo search technique generates the deep-conv LSTM classifier's weights and biases. Consequently, the cuckoo eggs will be accepted as its own and fed by the host bird. Therefore, they may get more food from the host bird's nest, but if the host bird finds the eggs, they either abandon them or leave the nest to build a new one. The global optimal solution can be analyzed by updating the position using the Levy flight and iterating until the best fitness solution is reached.

Algorithm 1: CDC-LSTM

Input: Optimal feature x_o
Output: Global best positiong_b
Step 1. Initialize the population
Step 2. Select the obtained value as the optimal solution and calculate the fitness value
    For each $Z_y > Z_b$
Step 3. Calculate the new solutions
    End for each
If $N < M_G$
Step 4. Compute the weights between the output layer bias vector and the memory unit. Furthermore, Cuckoo's Levy flight position may create a new solution.

$$\begin{gathered} \gamma \in h\{\gamma^m, \gamma^x, \gamma^y, \gamma^s\} \\ \eta \in \{\eta_k^N, \eta_m^D, \eta_m^N, \eta_y^D, \eta_y^N, \eta_y^o, \eta_x^D, \eta_x^N, \eta_x^o, \eta_s^D, \eta_s^N, \eta_s^o\} \end{gathered} \quad (9)$$

$$O_y^{(n+1)} = O_y^{(n)} + \mu \oplus M(0) \quad (10)$$

Step 5. Estimating the Levy distribution for a large number of steps.

$$L \sim s = N^{-\vartheta} \quad (11)$$

Step 6. Evaluate fitness value

Step 7. A new global optimal solution← g_b
End if
End
As shown in pseudocode 1 of the proposed Cuckoo-based deep convLSTM, diabetes can be predicted by iterating until the optimal fitness solution is reached to generate the global optimal solution. Where Z-fitness value, b-random number, j-nest, $\oplus$ −entry wise multiplication, L-levy flight, η_s^D −input layer weight, η_s^N −weight between memory outputs, η_s^o −cell output, η_x^D −forget gatel, η_y^o −output gate cell, η_m^D −Cell and weight in the middle of the input layer, η_m^N −weight between memory unit, γ^m −bias cel, $O_y^{(N+1)}$ −generate new solution, y-cuckoo.

4. RESULT AND DISCUSSION

In this section, a CDC-LSTM method is proposed to identify the confusion matrix of the performance parameters to predict diabetes by classifying diabetes or non-diabetes using multiple evaluation metrics. The classification performance measures investigated in this study include precision, accuracy, recall and F-measure.

TABLE II. SIMULATION PARAMETER

Simulation	**Value**
Dataset Name	Pima Indians Diabetes Database
Number of Epoch	769
Training	571
Testing	148
Language	Python
Tool	Jupyter

The Pima Indians diabetes database is tested and trained to predict diabetes using simulation parameters. Moreover, diabetes can be evaluated and classified utilising the Jupyter tool using the Python language, as shown in table 2.

TABLE III. CONFUSION MATRIX

Matrices	Formula
Accuracy	$\frac{T_1 + T_0}{T_1 + T_0 + F_1 + F_0}$
Precision(Pre)	$\frac{T_1}{T_1 + F_1}$
Recall(Rec)	$\frac{T_1}{T_1 + F_0}$
F-Measure	$\frac{2 * Pre * Rec}{2 + Pre + REc}$

A suggested method for confusion matrices, which divides true-1 and false-0 into true positives, true negatives, false positives, and false negatives, can improve the performance measurement, as table 3 illustrates.

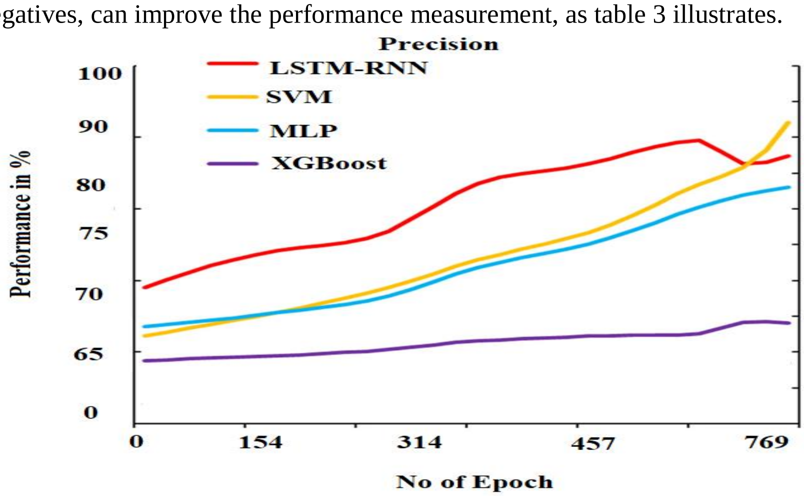

Fig. 3. Analysis of Precision

As described in figure 3, the performance measure can be improved using the proposed method to predict diabetes in precision analysis. Furthermore, using techniques such as SVM, XGBoost and MLP obtained from the literature, their accuracy in predicting diabetes improved to 69%, 74% and 79% by precision analysis. Nevertheless, compared with the previous methods, the proposed CDC-LSTM method can improve the precision analysis by increasing diabetes classification by 81.2%.

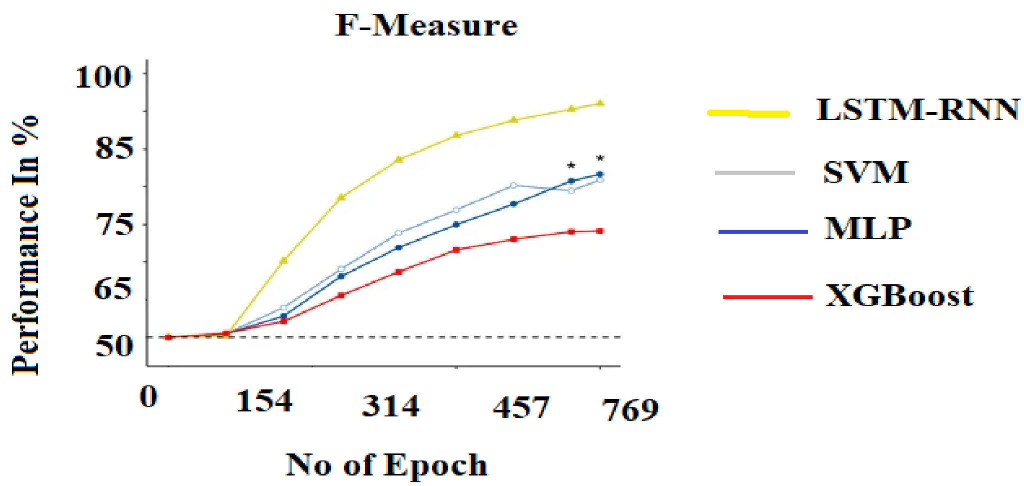

Fig. 4. Analysis of F-Measure

As depicted in figure 4, employing the suggested approach for forecasting diabetes in F-Measurement analysis results in enhanced performance metrics. Furthermore, utilizing literature-based methods like MLP, SVM, and XGBoost, their predictive accuracy for diabetes increased to 73%, 77%, and 80%, respectively, via F-measure analysis. Nonetheless, the proposed CDC-LSTM technique can enhance the F-measure analysis by boosting the diabetes classification by 82.6% compared to the prior methods.

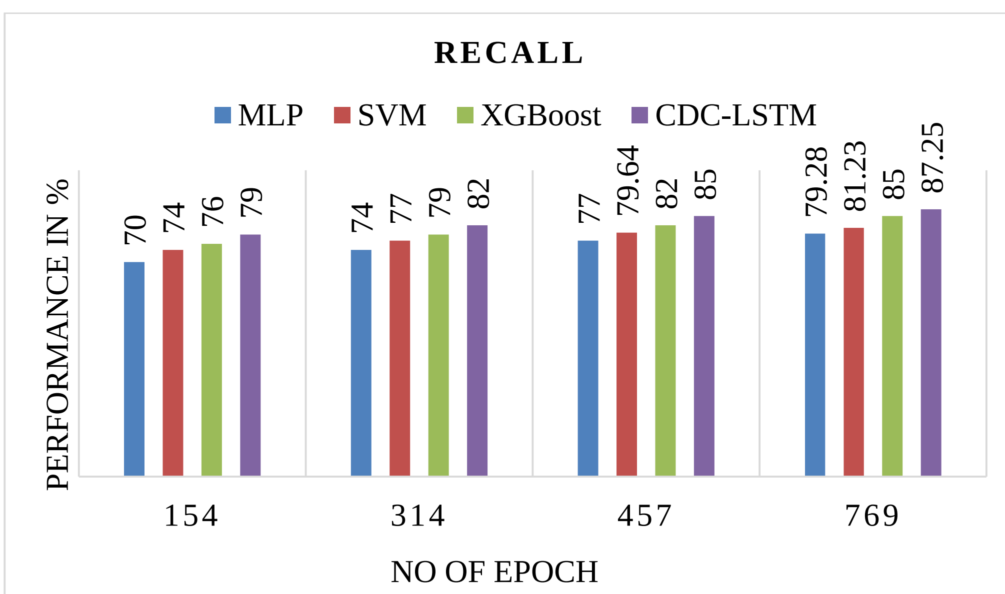

Fig. 5. Analysis of Recall

In figure 5, using the recommended approach for predicting diabetes in recall analysis improves performance metrics. Additionally, applying established methods like XGBoost, MLP, and SVM increases their predictive accuracy for diabetes to 79%, 82%, and 85% respectively through recall analysis. However, the proposed CDC-LSTM technique enhances the recall analysis by improving diabetes classification by 87.2% compared to previous methods.

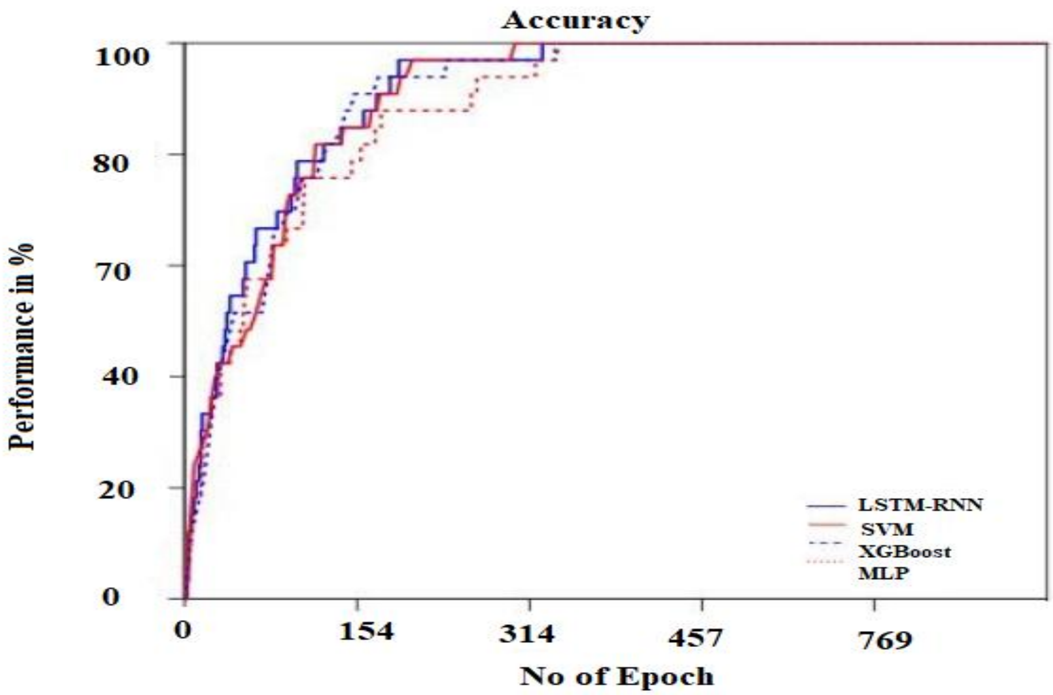

Fig. 6. Analysis of Accuracy

As demonstrated in figure 6, performance metrics are enhanced when the proposed method is applied to accuracy analysis to predict diabetes. Furthermore, with the use of prior techniques like XGBoost, MLP, and SVM, the prediction accuracy of diabetes rose to 82.46%, 87.25%, and 91.36%, respectively. In contrast to previous approaches, the suggested CDC-LSTM technique enhanced diabetes classification by 95.18%, which improved accuracy analysis.

5. CONCLUSION

In this conclusion, they provide an accurate analysis of patient data to better predict and manage diabetes. In addition, the PIMA India Diabetes Database available on Kaggle is a valuable resource for professionals in the field of health analytics. Using the PIMA India diabetes database, the CDC-LSTM algorithm can be proposed as a method to improve the accuracy of classifying diabetics or non-diabetics. By using this methodology, the efficiency and effectiveness of diabetes prediction can be improved. Furthermore, one process that can be achieved by the proposed SFS method is to standardize the diabetes data by removing the average value of each feature. Data standardization allows for accurate analysis and interpretation of information provided in the database. Additionally, the FBDT technique can be used to select the best features of the diabetes dataset to further improve the accuracy and reliability of the predictive model. Finally, the proposed CDC-LSTM method provides a promising approach to improve the diagnostic accuracy of diabetic patients. In addition, classification performance measures include precision, accuracy, recall, and F-measure. Moreover, by employing previous techniques such as XGBoost, MLP and SVM, the diabetes prediction accuracy was enhanced to 82.46%, 87.25% and 91.36%, respectively. The proposed CDC-LSTM technique can improve diabetes classification by 95.18% compared to the traditional methods.

Conflicts Of Interest

The paper states that there are no personal, financial, or professional conflicts of interest.

Funding

No financial contributions or endorsements from institutions or sponsors are mentioned in the author's paper.

Acknowledgements

The author is grateful to the institution for their collaboration and provision of necessary facilities that contributed to the successful completion of this research.

References

[1] M. H. Iqbal, M. F. Faruque, H. Alqahtani, and A. Kalim, "K-Nearest Neighbor Learning based Diabetes Mellitus Prediction and Analysis for eHealth Services," European Union Digital Library, vol. 20, no. 26, 2020.

[2] M. S. Farooq, S. Riaz, R. Tehseen, U. Farooq, K. Saleem, "Role of Internet of things in diabetes healthcare: Network infrastructure, taxonomy, challenges and security model," Digit Health, vol. 9, June 2023, Art. no. 20552076231179056.

[3] P. Valsalan et al., "IoT Based Expert System for Diabetes Diagnosis and Insulin Dosage Calculation," Healthcare, vol. 11, no. 12, 2023. [Online]. Available: https://doi.org/10.3390/healthcare11010012

[4] R. Fayaz et al., "An Intelligent Harris Hawks Optimization (IHHO) based Pivotal Decision Tree (PDT) Machine Learning Model for Diabetes Prediction," International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 4, pp. 415–423, 2022. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/2277

[5] Mr.S. MuthuKumar and Dr.M. Jayakumar, "A Novel Deep Learning Integrated Optimization-based Diabetes Detection (DeO-DiDet) System using IoT," Educational Administration: Theory and Practice, vol. 30, no. 5, pp. 3728-3740, 2024, doi: 10.53555/kuey.v30i5.3525.

[6] A. Naseem, R. H. Tabbasum, and N. Naz, "Novel Internet of Things based approach toward diabetes prediction using deep learning models," Front. Public Health, vol. 10, Sec. Digital Public Health, 24 August 2022. [Online]. Available: https://doi.org/10.3389/fpubh.2022.914106

[7] V. R. Allugunti and K. K. Reddy C., "Prediction of Diabetes Using Internet of Things (IoT) and Decision Trees: SLDPS," January 2021, DOI:10.1007/978-981-15-5679-1_43.

[8] G. Alfian et al., "Utilizing IoT-based sensors and prediction model for health-care monitoring system," Next Gen Tech Driven Personalized Med&Smart Healthcare, pp. 63-80, 2021. [Online]. Available: https://doi.org/10.1016/B978-0-12-822060-3.00009-7.

[9] N. Verma, S. S. Devendra, and D. Prasad, "Machine learning and IoT-based model for patient monitoring and early prediction of diabetes," Comput. Pract. Exper., First published: 02 August 2022, doi: 10.1002/cpe.7219.

[10] A. U. Haq et al., "Intelligent Machine Learning Approach for Effective Recognition of Diabetes in E-Healthcare Using Clinical Data," Sensors, vol. 20, 2649, 2020. [Online]. Available: https://doi.org/10.3390/s20092649

[11] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 40–46, 2021.

[12] P. K. D. Pramanik et al., "An ensemble learning approach for diabetes prediction using boosting techniques," Front. Genet., vol. 14, Art. no. 1252159, Oct. 26, 2023.

[13] G. Alfian et al., "A Personalized Healthcare Monitoring System for Diabetic Patients by Utilizing BLE-Based Sensors and Real-Time Data Processing," Sensors, vol. 18, 2183, 2018. [Online]. Available: https://doi.org/10.3390/s18072183

[14] U. M. Butt et al., "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications," Journal of Healthcare Engineering, vol. 2021, Article ID 9930985, 17 pages, 2021. [Online]. Available: https://doi.org/10.1155/2021/9930985.

[15] G. S. Gopalakrishnan et al., "A Novel Deep Learning-Based Heart Disease Prediction System Using Convolutional Neural Networks (CNN) Algorithm," International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 10s, pp. 516–522, 2023. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/3306

[16] N. Pradhan, G. Rani, V. S. Dhaka, and R. C. Poonia, "Diabetes prediction using artificial neural network," in Deep Learning Techniques for Biomedical and Health Informatics, Springer, Singapore, 2020.

[17] S. Padhy et al., "[Retracted] IoT-Based Hybrid Ensemble Machine Learning Model for Efficient Diabetes Mellitus Prediction," Computational Intelligence and Neuroscience, vol. 2022, Article ID 2389636, 11 pages, 2022. [Online]. Available: https://doi.org/10.1155/2022/2389636.