Next-Gen Public Healthcare: Modeling an Advanced Framework for Diabetes Management

Kamlesh Kumar Yadav[1], Dhablia Dharmesh Kirit[2]

[1]*Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India*
[2]*Research Scholar, Department of CS & IT, Kalinga University, Raipur, India*

KEYWORDS

Diabetes management, crow search-driven dynamic random forest (CS-DRF), Diagnosis, healthcare, PID dataset, Machine learning (ML), Public health

ABSTRACT

Advancements in biotechnology and public health technologies have significantly increased public health data, aiding early disease detection and prevention, particularly in diabetes, which can lead to serious public health issues. In this study, we propose a novel crow search-driven dynamic random forest for (CS-DRF) for diabetes management reorganization. The goal of this integrated strategy is to enhance diabetes treatment results and diagnosis. The Pima Indian Diabetes (PID) dataset is gathered from open-source Kaggle website for diabetes management. We implement our recommended evaluation technique using Python software. The findings showed that the CS-DRF performed more effectively than the others regarding F1-score-93.25%, accuracy-95.46%, specificity-92.38 %, sensitivity-92.22%, and precision-93.28%. The study's validation results demonstrate how advanced the structure model based on machine learning is in detecting diabetes management.

1. Introduction

Intense glucose management dramatically lowers long-term diabetic complications, as the diabetic control and complications trial (DCCT) plays a major role in public health [1]. Frequent self-monitoring of blood glucose (SMBG) and routine HbA1c readings were part of public health in this aggressive therapy [3]. Modern developments in glucose monitoring technologies, including continuous glucose monitoring (CGM), enable customized treatment and offer substitute glucose control measures like time in range (TIR) [2]. Because TIR is applied for public health to improve everyday life and there is growing evidence that it can predict long-term diabetes problems and pregnancy outcomes, it has become a favored metric among patient's health [14]. However, interpreting TIR can be difficult for researchers and clinicians used to HbA1c and conventional blood glucose readings. The significance of TIR, the connection between TIR and hyperglycemia, and the necessity of taking time below range (TBR) interpretation guidelines for medical professionals are included along with the present roadblocks to TIR's wider use in diabetes control [4]. Personalized medicine, which customizes treatment medications is changing diabetes care by using a patient's biology, lifestyle decisions, and medical information as a foundation [5]. This enables patients to make knowledgeable decisions about their food, exercise routine, and medication adherence by allowing real-time data utilization and personalized feedback [15]. This comprehensive strategy aims to improve diabetes treatment outcomes and diagnostics.

The rest of the paper is divided into **p**hrases. The second phrase organizes the related works, phrase 3 explains the approach, phrase 4 outlines the experimental findings, and phrase 5 describes the conclusion.

2. Related work

Nomura et al. [7] provided artificial intelligence/ machine learning (AI/ML)-based diabetes public health prediction models and medical devices [6]. When sudden diabetes public health prediction utilized ML techniques, its effectiveness was currently on a level with traditional risk stratification models that depend on statistical methodologies. It helps to significantly increase the diabetic illness prediction models' accuracy. The public health development and distribution of an ML-based software solution by Sowah et al. [8] enhanced diabetes public health care. The tensor flow neural network model for food classification was constructed to utilize the suggested framework, which takes into consideration the problem of managing diabetes patients' public health [17]. The system in place would address the issue of diabetics' activity of public health management, dietary advice, and medication notifications. Clinical practice guidelines were created by Navaneethan et al. [9].

They investigated the association between blood glucose in Pasquel et al. [16], recommended for managing persons having diabetes. They suggested unaltered modified odds ratios were determined by the use of vicariate and multivariable logistic regression modeling, respectively. Various ML techniques in Soni and Varma [11] were employed to forecast diabetes in a patient or the human body at an earlier stage with a higher degree of accuracy for applying the machine learning methods [10]. The project effort produced a model that is correct or more accurate, proving that the model can predict diabetes with any degree of precision. Agrawal et al. demonstrated non-invasive glucose levels for public health measurement with ML models Agrawal et al [12]. Machine learning techniques were employed for the prediction of public health-related issues. The device seemed like the perfect non-invasive option for continuous glucose monitoring.

2. Methodology

Dataset

The data is referred from open source Kaggle website: https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database. This sample was initially obtained from the nationalized organization of kidney diseases, diabetes and digestive disorder. The objective of the sample challenge is to use certain diagnostic parameters that are part of the collection to diagnose and perhaps rule out diabetes in patients. The information set attempts to forecast a person's diagnosis as diabetic using the specific diagnostic criteria that were included in the collection. More specifically, all of the patients at this public health facility are twenty-one years old.

Classification for diabetes management using Crow Search-driven Dynamic Random Forest (CS-DRF)

In diabetes management recognition, we presented a unique Crow Search-driven Dynamic Random Forest (CS-DRF) in this study. Selecting features is optimized by the CS algorithm, while classification accuracy and flexibility are improved by the dynamic random forest.

Dynamic Random Forest

The irrigation area's statistics are used to approximate the $N\ (0,1)$ normal distribution to get the desired value. The following is the standardized formula (1):

$$\tilde{z} = \frac{z-\bar{z}}{\sqrt{\sigma}} \qquad (1)$$

Where the area's standard deviation is denoted by $\sqrt{\sigma}$ and its mean value is represented by $\tilde{z}$. To create a training set and test set utilizing the diabetes data and the DRF model created during this inquiry.

The many decision trees of the random forest are its most significant characteristic. To optimize the prediction model trees, more accurate tree model parameters must be provided by many magnitude judgments. By taking into consideration both MAE and MSE, enhance the evaluation index model. The arithmetic mean value [AMM] impure function model, or average MSE-MAE model, is used to achieve this, as shown in Equation 2.

$$INDEX = Average\{MSE + MSE\} \qquad (2)$$

The function with the purpose is min $INDEX$. By choosing the proper segmentation factors and points, the improved random forest model can be optimized.

Crow search optimization

Crow search optimization improves the accuracy of predictions and flexibility and quickly finds the best treatment method that improves diabetes control. The crow's ability to hide surplus food and recover it when needed is the foundation of the CSA, or crow search algorithm. When the owner leaves, crows might follow other birds to where their food is hidden and take it. Because they have experience stealing from others, the best way to keep other crows from stealing their food. Assume a d-dimensional environment contains M crows. In the search space, the vector $w^{j,iter}, where j = 1,2,\ldots,M, iter =$

$1,2,\dots.iter_{max} and iter_{max}$ is the maximum iteration number, indicates the position of the j^{th} crow at time/iteration $iter$. One workable solution to the problem is to find the crow's position. Every crow has a memory that helps it retain the location of its favorite hiding spot thus far. The position and memory of each crow are randomly initialized in the traditional crow search algorithm (CSA), and each position's fitness value is then calculated.

Additionally, each time a new position is generated for the j^{th} crow, crow i is chosen at random, causing crow j to follow crow i to find its concealed location. Ultimately, the following is how the new role is created (3):

$$w^{j,iter+1} = \begin{cases} w^{j,iter} + q \times EK^{j,iter} \times \left(n^{i,iter} - w^{j,iter}\right), if rand > AP^{j,iter} \\ brandom position otherwise \end{cases} \quad (3)$$

Where $n^{i,iter}$, and $AP^{j\ iter}$ are the concealed location of the i^{th} crow and awareness probability of the i^{th} crow at iteration $iter$, respectively, and q is an arbitrary value between 0 and 1.

$EK^{j,iter}$ is the air travel distance of the j^{th} crow at iteration $iter$. i^{th} The crow's memory is then refreshed, and the fitness value of the new location is determined.

$$n^{i,iter+1} = \begin{cases} w^{i,iter+1}, if e\left(w^{i,iter+1}\right) is better then e\left(n^{i,iter+1}\right) \\ n^{i,iter}, \qquad otherwise \end{cases} \quad (4)$$

Where the hidden location of the i^{th} crow at each iteration $iter$ is denoted by $n^{i,iter}$, and $e(\cdot)$ represents the fitness value. All crows go through this process again until the termination condition is satisfied.

The awareness probability (AP) in the traditional CSA regulates the trade-off between intensification and diversity; a small AP compels CSA to enhance its local search ability, while a big AP encourages CSA to comb over the search space.

A novel method for managing diabetes called CSDRF makes use of the cleverness of crow search algorithms to maximize the dynamic feature selection process in random forest models. By constantly altering the feature set by the search algorithm's exploration and exploitation capabilities, this strategy increases prediction accuracy. In diabetes care, CSDRF effectively manages sizable and intricate datasets, enhancing the forecasting of blood glucose levels and patient results.

3. Results and discussion

An Intel i7 core Windows 10 laptop with 8GB RAM and Tensor Flow/Keras was modeled with Python 3.10.1 software and the scikit-learn method. The suggested approach, CS-DRF, is contrasted with other approaches, including (DT), (KNN), and (RF) [13]. The metrics F1 score, precision, specificity, sensitivity, and accuracy are employed. The overall outcome values are displayed in Table 1.

Table 1 overall outcome values of precision, sensitivity, F1 score, accuracy, and specificity

Methods	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-score (%)	Specificity (%)
Random Forest (RF) [13]	82.26	80.45	83.47	82.26	84.07
Decision tree (DT) [13]	81.02	77.98	83.02	81.1	84.05
K-nearest neighbor (KNN) [13]	80.55	79.5	81.2	80.54	81.59
CS-DRF [Proposed]	95.46	92.22	93.28	93.25	92.38

Accuracy and precision: When it comes to managing diabetes, precision denotes how consistently blood glucose readings from the system correspond to actual blood glucose levels, while accuracy refers to how closely these measurements match the system readings when repeated under the same circumstances. Good decision-making regarding therapy is facilitated by high precision, which guarantees accurate measurements. A high degree of precision indicates dependability in the readings, giving confidence while tracking trends. To maximize the efficiency and dependability of diabetes management systems, both metrics are essential.

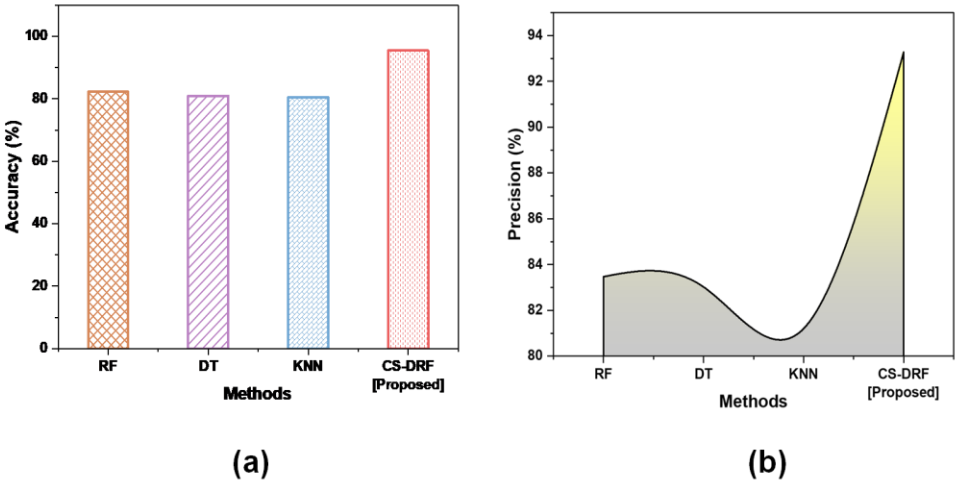

Figure 1 Accuracy and Precision Performance

Figure 1 and Table 1 display the accuracy and precision performance. The accuracy of the proposed (CS-DRF-95.46%) outperforms the existing systems such as RF-82.26%, DT-81.02%, and KNN - 80.55%, respectively. The precision for the proposed method (CS-DRF-93.28%) outperforms the existing systems such as RF-83.47%, DT-83.02%, and KNN-81.2%, respectively. Our suggested method is more effective than an existing method for diabetes management for public health.

Sensitivity, specificity, and F1 score: Sensitivity (recall) indicates the percentage of real diabetics that are successfully diagnosed, which is important for early detection and intervention. Specificity determines the percentage of non-diabetes who are correctly identified, minimizing false positives and unneeded interventions. To maintain equilibrium between accuracy and recall, the F1-score offers a single metric for the overall effectiveness of the model that accounts for false negatives as well as false positives. The performance of the F1-score, specificity, and sensitivity are shown in Figure 2.

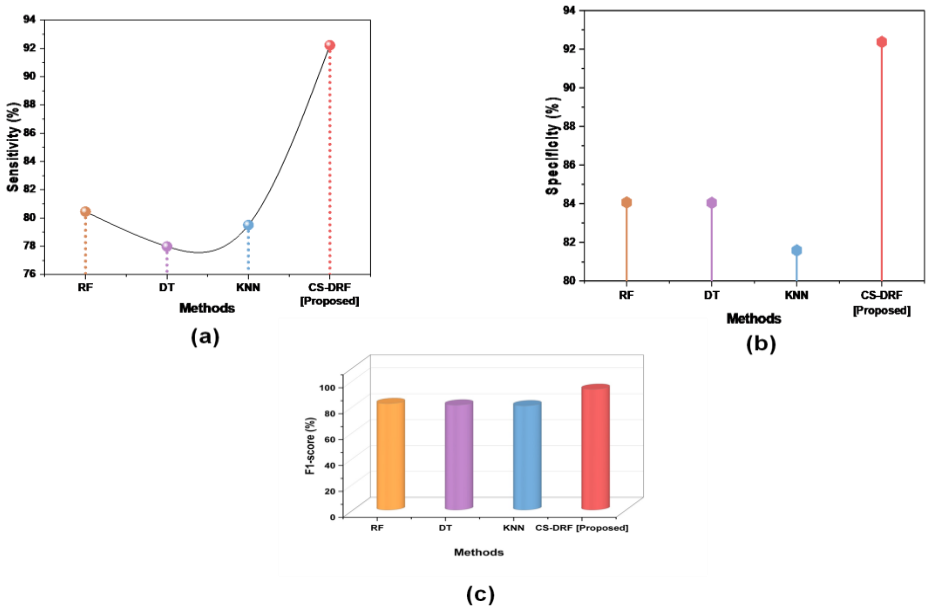

Figure 2 Sensitivity, F1 score, and specificity performance

The sensitivity for the proposed (CS-DRF-92.22%) outperforms the existing systems such as RF-80.45%, DT-77.98%, and KNN-79.5%, respectively. The specificity for the proposed (CS-DRF-

92.38%) outperforms the existing systems such as RF-84.07%, DT-84.05%, and KNN-81.59%, respectively. The F1-score for the proposed (CS-DRF-93.25%) outperforms the existing systems such as RF-82.26%, DT-81.1%, and KNN-80.54%, respectively. Our recommended method is better than the other existing methods for diabetes management for public health.

4. Conclusion and future scope

The study provides a novel CS-DRF for diabetes management for public health reorganization for this paper. Improving diabetes treatment outcomes and diagnostics is the aim of this comprehensive approach. The method was validated using the PID database. The outcome demonstrates that the CS-DRF has outperformed the other in terms of F1-score-93.25%, accuracy-95.46%, specificity-92.38 %, sensitivity-92.22%, and precision-93.28%. Our suggested method for diabetes management for public health outperforms the existing method. The management of diabetes presents several obstacles, including the requirements for ongoing real-time monitoring to maintain accurate forecasts, concerns about data privacy, and significant individual diversity in treatment responses. In the future, diabetes management could be greatly enhanced by developments in AI and machine learning wearable technology integration and customized treatment programs, which would allow for more accurate and patient-specific care.

Reference

[1] C.P. Hurst, N. Rakkapao, and K. Hay, "Impact of diabetes self-management, diabetes management self-efficacy, and diabetes knowledge on glycemic control in people with Type 2 Diabetes (T2D): A multi-center study in Thailand, "*PLoS One*, *15*(12), p.e0244692, 2020. DOI:https://doi.org/10.1371/journal.pone.0244692.

[2] A. Advani, "Positioning time in range in diabetes management," *Diabetologia*, *63*(2), pp.242-252, 2020. DOI:https://doi.org/10.1007/s00125-019-05027-0.

[3] S. Neelima, Manoj Govindaraj, Dr.K. Subramani, Ahmed ALkhayyat, & Dr. Chippy Mohan. (2024). Factors Influencing Data Utilization and Performance of Health Management Information Systems: A Case Study. Indian Journal of Information Sources and Services, 14(2), 146–152. https://doi.org/10.51983/ijiss-2024.14.2.21

[4] B.S. Senevirathne, M.A., Jayasinghe, D. Pavalakumar, and C.G. Siriwardhana, "Ceylon cinnamon: a versatile ingredient for futuristic diabetes management," *Journal of future foods*, *2*(2), pp.125-142. 2022. DOI:https://doi.org/10.1016/j.jfutfo.2022.03.010.

[5] F.N.U. Sugandh, M. Chandio, F.N.U. Raveena, L. Kumar, F.N.U. Karishma, S. Khuwaja, U.A. Memon, K. Bai, M. Kashif, G. Varrassi, and M. Khatri, "Advances in the management of diabetes mellitus: a focus on personalized medicine," *Cureus*, *15*(8). 2023. DOI:https://doi.org/10.7759%2Fcureus.43697.

[6] Alamer, L., Alqahtani, I. M., & Shadadi, E. (2023). Intelligent Health Risk and Disease Prediction Using Optimized Naive Bayes Classifier. Journal of Internet Services and Information Security, 13(1), 01-10.

[7] A. Nomura, M. Noguchi, M. Kometani, K. Furukawa, and T. Yoneda, "Artificial intelligence in current diabetes management and prediction," *Current Diabetes Reports*, *21*(12), p.61. 2021. DOI:https://doi.org/10.1007/s11892-021-01423-2

[8] R.A. Sowah, A.A. Bampoe-Addo, S.K. Armoo, F.K. Saalia, F. Gatsi, and B. Sarkodie-Mensah, "Design and development of diabetes management system using machine learning, "*International journal of telemedicine and applications*, *2020*(1), p.8870141. 2020. DOI:https://doi.org/10.1155/2020/8870141.

[9] S.D. Navaneethan, S. Zoungas, M.L. Caramori, J.C. Chan, H.J.Heerspink, C. Hurst, A. Liew, E.D. Michos, W.A. Olowu, T. Sadusky, and N. Tandon, "Diabetes management in chronic kidney disease: synopsis of the 2020 KDIGO clinical practice guideline," *Annals of internal medicine*, *174*(3), pp.385-394. 2021. DOI:https://doi.org/10.7326/M20-5938.

[10] Sofiene, M., Souhaila, B., & Souhir, C. (2024). Machine Learning for Early Diabetes Detection and Diagnosis. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 15(1), 216-230.

[11] M. Soni, and S."Varma, "Diabetes prediction using machine learning techniques," *International Journal of Engineering Research & Technology (IJERT)*, *9*(09), pp.2278-0181. 2020. DOI:

[12] H. Agrawal, P. Jain, and A.M. Joshi, "Machine learning models for non-invasive glucose measurement: towards diabetes management in smart healthcare," *Health and Technology*, *12*(5), pp.955-970. 2022. DOI:https://doi.org/10.1007/s12553-022-00690-7.

[13] V. Chang, M.A. Ganatra, K. Hall, L. Golightly, and Q.A. Xu, "An assessment of machine learning models and algorithms

for early prediction and diagnosis of diabetes using health indicators," *Healthcare Analytics*, *2*, p.100118. 2022. DOI:https://doi.org/10.1016/j.health.2022.100118.

[14] S. Alam, M.K. Hasan,S.Neaz, N. Hussain, M.F. Hossain, and T. Rahman, "Diabetes Mellitus: insights from epidemiology, biochemistry, risk factors, diagnosis, complications, and comprehensive management," *Diabetology*, *2*(2), pp.36-50. 2021. DOI:https://doi.org/10.3390/diabetology2020004.

[15] B.T. Seboka, T.M. Yilma, and A.Y. Birhanu, "Factors influencing healthcare providers' attitude and willingness to use information technology in diabetes management," *BMC Medical Informatics and Decision Making*, *21*, pp.1-10. 2021. DOI:https://doi.org/10.1186/s12911-021-01398-w

[16] F.J. Pasquel, M.C. Lansang, K. Dhatariya, and G.E. Umpierrez, "Management of diabetes and hyperglycemia in the hospital," *The Lancet Diabetes & endocrinology*, *9*(3), pp.174-188. 2021. DOI:https://doi.org/10.1016/S2213-8587(20)30381-8.

[17] Bobir, A.O., Askariy, M., Otabek, Y.Y., Nodir, R.K., Rakhima, A., Zukhra, Z.Y., Sherzod, A.A. (2024). Utilizing Deep Learning and the Internet of Things to Monitor the Health of Aquatic Ecosystems to Conserve Biodiversity. *Natural and Engineering Sciences*, *9*(1), 72-83.