International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT241061199

Novel Based Approach towards Diabetic's Classification Using Artificial Intelligence and Internet of Things Environment

Divyashree R, Dr. Sumati Ramakrishna Gowda

Department of Computer Science ,Karnataka State Open University, Mysore, Karnataka, India

ARTICLEINFO

Article History:

Accepted : 20 Nov 2024

Published: 12 Dec 2024

Publication Issue

Volume 10, Issue 6

November-December-2024

Page Number

1646-1651

ABSTRACT

The work is focused on designing effective technique employing AI such as machine learning and Internet of things environment for diabetic classification and management. The model is focused in reliably and energy efficient manner in collecting data using IoT and edge-computing paradigm. The work is focused on designing a novel ML model that can classify diabetics and should address class imbalance issues. The model should also be robust considering different kinds of data/attributes related to diabetics.

Introduction

In today's world, many things contribute to an in appropriate way of life, including irregular eating habits, a lack of nutrition, pollutants, lack of sufficient exercise, unending work, impatience, and elevated stress levels, all of which contribute to worsening of human health. An inactive lifestyle affects up to 40 percent of youngsters, middle-aged people, and career women in various nations. Because of our hectic daily schedules, we barely have time to focus on our health, leading to a variety of health problems. Furthermore, it would be challenging for a doctor to keep a constant eye on the patient. Individual patients may find it challenging to maintain state of their health condition and seek guidance from their doctors. Chronic diseases are defined as those that last a long time and necessitate long-term therapy. Patients with severe conditions are frequently admitted to the hospital for extended periods of time to be monitored on a regular basis. Heart disease, cancer, and diabetes are examples of common chronic diseases [3].

Critical conditions of healthcare issues require continuous monitoring and proper feedback and follow-up from more than one authority such as patients, healthcare practitioners, city planners, to name a few. Internet of Things(IoT) networks are no longer simple sensors that collect information, on the contrary, they are now capable of provide learning

and train data. Therefore, the use of IoT technology in health critical conditions monitoring is important [1].

Machine learning plays an essential role in the detection of numerous diseases. In the last decade, researcher's have looked into a variety of machine learning approaches in the health-care industry. Recent technological advancements have resulted in the successful integration of the Internet of Things and machine learning in a variety of fields. In the clinical setting, the coupling of these two fields i.e., machine learning and IoT can greatly aid in the creation of a receptive and inter connected environment and as a result, healthcare professionals and patients will benefit from a variety of services [2].

Diabetic disease is currently exceedingly serious, as it kills a large number of individuals each year [4]. As a result, in order to live a normal life, the diabetic patient must be monitored on regular basis.. Many organs, including the eyes, neurons, and internal organs, can become defective or worsen as a result of high or low blood sugar levels. Therefore, to prevent the diabetic patient's health from deteriorating, continual and daily monitoring is necessary. Current medical-care software platforms, such as health management systems, patient monitoring systems are insufficient in assisting treatment with necessary information [5]. The objective of this research is to make a patient health monitoring system especially for those who are suffering from chronic diseases like diabetes. It is a novel application designed by using various machine learning models. In this application, the patient's data is extracted by using IoT wearable technology.

Literature Survey

This section studies various existing IoT and AI technology for detecting and managing diabetics, and challenges [6]. Here[7] the goal is to make a smart patient's health monitoring system based on machine learning that helps todetect the presence of a chronic disease in patient early and accurately. For the implementation, the diabetic dataset has been used. To detect the presence of the fatal disease, six different machine learning techniques are used i.e., Support Vector Machine (SVM), Logistic Regression, Artificial Neural Network (ANN),Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM). The performanceof the proposed model is evaluated by using four evaluation metrics i.e.,accuracy, precision, recall, and F1-Score. The RNN outperformed the remaining algorithms in terms of accuracy (81%), precision (75%), and F1-Score (65%).However, the recall (56%) for ANN was higher as compared to SVM and logistic regression, CNN, RNN, and LSTM. With the help of this proposed patient's health monitoring system, doctors will be able to diagnose the presence ofthe disease earlier. In [8] The proposed solution will alert the patients' parents or guardians of their situation when they about to reach critical conditions specially at night by sending alarms and notifications by Short Messages (SMS) along with the patients current location to up to three people.Moreover, a Machine Learning (ML) model is implemented to predict future events where the patient might have serious issues.

In [9] presents a model using a fused machine learning approach for diabetes prediction. The conceptual framework consists of two types of models: Support Vector Machine (SVM) and Artificial Neural Network (ANN) models. These models analyze the dataset to determine whether a diabetes diagnosis is positive or negative. In [10] focused on the diabetes dataset extracted from Kaggle and two unseen real datasets. They developed Synthetic Minority Over-Sampling Technique (SMOTE) algorithm to balance the dataset. Also explored Boruta as the feature selection method. To tune hyper-parameters of different algorithms, we have proposed an improved technique by combining the Grid Search method with the Grey Wolf Optimization algorithm. The Grid Search method requires extensive searching, while

the Grey Wolf Optimization algorithm is easily linked, rapid to seek, and extremely exact.

In [11] A dataset consisting of patient bodily statistics was obtained from a local hospital according to ethical guidelines, such as obtaining the prior consent of both patients and doctors. In [12] First, a novel method based on Local Median-based Gaussian Naive Bayes (LMeGNB) is proposed to compensate for the missing values, combined with the K-means SMOTE method to adjust the positive and negative samples of diabetes to obtain the normalized balanced data. Then, a probability-based multi-stage ensemble is devoted to building ensemble models on the different types of machine learning algorithms. When extreme gradient boosting, random forests, and weighted k nearest neighbors are integrated, the highest classification accuracy of 94.53% is obtained on Pima Indian diabetes dataset. Finally, to evaluate the PE_DIM model, the experiment equally considered two diabetes datasets, RSMH and Tabriz, to demonstrate the generality of the method in diabetes prediction.

In [13] propose diabetes mellitus network framework (DMNet) for type 2 diabetes risk assessment of elderly people. Specifically, we propose tandem long short-term memory to extract the long-term temporal information of different diabetes risk categories. In addition, the tandem mechanism is used to capture the correlation between the diabetes risk factor categories. To balance the label distribution, we adopt the method of synthetic minority over-sampling technique with Tomek links. To form the better feature representations, utilize entity embedding to solve the problem of high-dimensional features. To evaluate the performance of our proposed method, we conduct the experiments on a real-world dataset called Research on Early Life and Aging Trends and Effects.In [14] present the SMART Framework through a case study on the microvascular complications of diabetes using data from the ACCORD clinical trial [15], [16].

In [17], shows model for the long term effects of physical activity on diabetes progression, by exploiting and adapting an existing short-term model of physical activity. A pivotal role in the proposed model is played by interleukin-6 released during physical activity and known to be fundamental in maintaining pancreatic beta cells production and therefore satisfactory insulin secretion [18].In [19] a framework based on synthetic data generation algorithms was developed. Eight medical datasets containing tabular data were used to test this framework. Three different statistical metrics were used to analyze the preservation of synthetic data integrity and six different synthetic data generation sizes were tested.

Objectives of the Proposed Research

The objective of the research work is given below:

- Design a reliable and energy efficient IoT framework for collecting diabetic patient vital information.
- To develop a novel feature selection mechanism and ML-based model for accurate diabetic classification.
- Identify features contributing to diabetics considering imbalanced diabetics dataset.
- Robustness of proposed algorithm will be validated by testing on different publicly available dataset and compare with existing model in terms of accuracy, precision, F-measure etc.

Proposed Methodology

In addressing the problem and meeting research objectives the following methodologies is defined.

1. **Design a reliable and energy efficient IoT framework for collecting diabetic patient vital information.**

In meeting research objectives, the patient and healthy subject are equipped with different sensors and are asked to perform certain task. Then, person vital information like blood pressure, glucose level, etc. are extracted from both patient and healthy subject. A wireless body area networks(WBANs) are

deployed to collect vital information and these information are transmitted to edge-computing server for further analysis. The work will deploy effective transmission mechanism that is energy efficient and reliable.

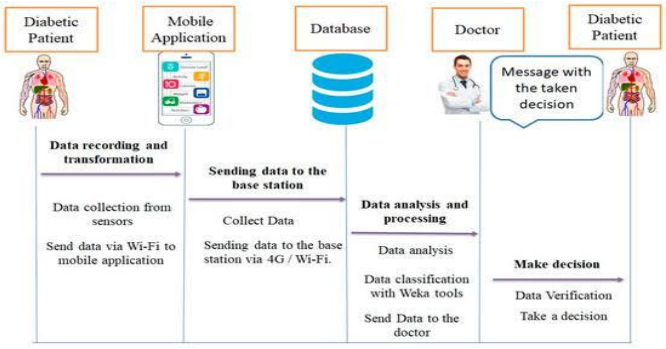

Fig 5.1 Activity Diagram

2. To develop a novel feature selection mechanism and ML-based model for accurate diabetic classification.

In meeting research objectives, the vital information collected from previous objective are labelled by the expert as healthy and diabetics. Later, the labelled data is used for training machine learning model for automatic classification of diabetics. The proposed feature selection optimized ML model performance will be compared with other methodologies. For classification we are using logistic regression refer model. Logistic regression is easier to implement, interpret, and train. It classifies unknown records very quickly. When the dataset is linearly separable, it performs well. Model coefficients can be interpreted as indicators of feature importance. To implement the Logistic Regression using Python, we will use the following steps.

- Data Pre-processing step
- Fitting Logistic Regression to the Training set
- Predicting the test result
- Test accuracy of the result(Creation of Confusion matrix)
- Visualizing the test set result.

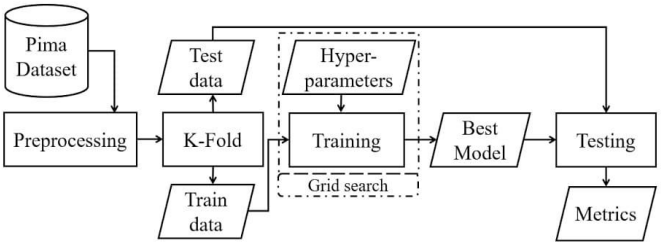

Fig 5.2 Data Flow Diagram

3. Identify features contributing to diabetics considering imbalanced diabetics dataset.

Here a new feature weight optimization mechanism will be designed that selects good feature and then ranks feature according to less error i.e., less miss classification. In meeting a new $K-$ fold cross validation scheme will be modelled for imbalanced classification data.

4. Robustness of proposed algorithm.

It will be validated by testing on different publicly available dataset and compare with existing model in terms of accuracy, precision, F-measure etc.

Here experiment is conducted for different diabetes dataset for validating the robustness of proposed feature selection and machine model. This is to shows one model can be used for solving various similar diabetes dataset like Pima Indian diabetics and OhioT1DM classification problem. The proposed diabetes classification model will be compared with most recent paper/methodologies considering different performance metrics such as a accuracy, precision, F-measure, etc.

Result Analysis

We are evaluating the model using following metrices

Accuracy: Accuracy provides the proportion of correctly classified instances

Accuracy = \frac{True \, Positives + True \, Negatives}{Total}

Precision: Precision focuses on the accuracy of positive predictions

Precision = \frac{True \, Positives }{True\, Positives + False \, Positives}

Recall (Sensitivity or True Positive Rate): Recall measures the proportion of correctly predicted positive instances among all actual positive instances.

Recall = \frac{ True \, Positives}{True\, Positives + False \, Negatives}

F1 Score: F1 score is the harmonic mean of precision and recall

F1 \, Score = 2 * \frac{Precision * Recall}{Precision + Recall}

Expected Outcome Of The Proposed Research

The proposed methodology is expected to achieve very good prediction accuracies and also reduces misclassification. The model can work on different diabetes datasets employing that extract sensory information using IoTs. The proposed model can be adopted for binary and multi-label classification problems.

Conclusion

This proposal studies various data collection and prediction method machine and deep leaning model for diabetics disease employing internet of things and edge-computing server environment. The work identifies the advantage and limitation of various existing work and defined problem definition. The objective has been defined to overcome research problem and limitation. The methodology to meet research objective have been discussed. Finally, the expected outcome has been discussed.

Acknowledgment

I would like to express my deepest gratitude to my research guide, Dr. Sumati Ramakrishna Gowda, Assistant Professor, Department of Computer Science, Karnataka State Open University, Mysore for their invaluable guidance, support, and encouragement throughout the process of preparing this conference paper. Their insightful feedback, expertise, and constant motivation have been instrumental in shaping the direction and quality of this work. I truly appreciate the time and effort they dedicated to reviewing and refining this research, and their mentorship has been a source of inspiration. I am grateful for their belief in my abilities and for providing me with the opportunity to grow as a researcher.

References

[1]. Hasan MK, Alkhalifah A, Islam S, Babiker N, Habib A, Aman AHM, et al.Blockchain technology on smart grid, energy trading, and big data: security issues,challenges, and recommendations. Wireless Commun Mobile Comput. (2022)2022:9065768. doi: 10.1155/2022/9065768.

[2]. Pradhan K, Chawla P. Medical Internet of things using machine learning algorithms for lung cancer detection. J Manage Anal. (2020) 7:591–623.doi: 10.1080/23270012.2020.1811789.

[3]. Garg A, Mago V. Role of machine learning in medical research: a survey.Comput Sci Rev. (2021) 40:100370. doi: 10.1016/j.cosrev.2021.100370.

[4]. Ghazal TM, Hasan MK, Alshurideh MT, Alzoubi HM, Ahmad M, AkbarSS, et al. IoT for smart cities: machine learning approaches in smart healthcare–areview. Future Internet. (2021) 13:218. doi: 10.3390/fi13080218.

[5]. Rghioui A, Lloret J, Sendra S, Oumnad A. A smart architecturefor diabetic patient monitoring using machine learning algorithms. In:Healthcare, Vol. 8. Multidisciplinary Digital Publishing Institute (2020). p. 348.doi: 10.3390/healthcare8030348.

[6]. Farooq MS, Riaz S, Tehseen R, Farooq U, Saleem K. Role of Internet of things in diabetes healthcare: Network infrastructure, taxonomy, challenges, and security model. Digit Health. 2023 Jun 6;9:20552076231179056. doi: 10.1177/20552076231179056.

[7]. Naseem A, Habib R, Naz T, Atif M, Arif M and Allaoua Chelloug S (2022) Novel Internet of Things based approach toward diabetes prediction using deep learning models. Front. Public Health 10:914106. doi: 10.3389/fpubh.2022.914106.

[8]. Alfandi O. An Intelligent IoT Monitoring and Prediction System for Health Critical Conditions. Mobile Netw Appl. 2022;27(3):1299–310. doi: 10.1007/s11036-021-01892-5. Epub 2022 May 18. PMCID: PMC9115747.

[9]. U. Ahmed et al., "Prediction of Diabetes Empowered With Fused Machine Learning," in IEEE Access, vol. 10, pp. 8529-8538, 2022, doi: 10.1109/ACCESS.2022.3142097.

[10]. P. Yadav, S. C. Sharma, R. Mahadeva and S. P. Patole, "Exploring Hyper-Parameters and Feature Selection for Predicting Non-Communicable Chronic Disease Using Stacking Classifier," in IEEE Access, vol. 11, pp. 80030-80055, 2023, doi: 10.1109/ACCESS.2023.3299332.

[11]. S. T. Himi, N. T. Monalisa, M. Whaiduzzaman, A. Barros and M. S. Uddin, "MedAi: A Smartwatch-Based Application Framework for the Prediction of Common Diseases Using Machine Learning," in IEEE Access, vol. 11, pp. 12342-12359, 2023, doi: 10.1109/ACCESS.2023.3236002.

[12]. L. Jia, Z. Wang, S. Lv and Z. Xu, "PE_DIM: An Efficient Probabilistic Ensemble Classification Algorithm for Diabetes Handling Class Imbalance Missing Values," in IEEE Access, vol. 10, pp. 107459-107476, 2022, doi: 10.1109/ACCESS.2022.3212067.

[13]. Z. Yu, W. Luo, R. Tse and G. Pau, "DMNet: A Personalized Risk Assessment Framework for Elderly People With Type 2 Diabetes," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 3, pp. 1558-1568, March 2023, doi: 10.1109/JBHI.2022.3233622.

[14]. B. P. Swan, M. E. Mayorga and J. S. Ivy, "The SMART Framework: Selection of Machine Learning Algorithms With ReplicaTions—A Case Study on the Microvascular Complications of Diabetes," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 2, pp. 809-817, Feb. 2022, doi: 10.1109/JBHI.2021.3094777.

[15]. Theis, W. L. Galanter, A. D. Boyd and H. Darabi, "Improving the In-Hospital Mortality Prediction of Diabetes ICU Patients Using a Process Mining/Deep Learning Architecture," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 1, pp. 388-399, Jan. 2022, doi: 10.1109/JBHI.2021.3092969. (dastaset)

[16]. G. Annuzzi et al., "Impact of Nutritional Factors in Blood Glucose Prediction in Type 1 Diabetes Through Machine Learning," in IEEE Access, vol. 11, pp. 17104-17115, 2023, doi: 10.1109/ACCESS.2023.3244712. (dataset)

[17]. P. F. De Paola, A. Paglialonga, P. Palumbo, K. Keshavjee, F. Dabbene and A. Borri, "The Long-Term Effects of Physical Activity on Blood Glucose Regulation: A Model to Unravel Diabetes Progression," in IEEE Control Systems Letters, vol. 7, pp. 2916-2921, 2023, doi: 10.1109/LCSYS.2023.3290774.

[18]. H. Nemat, H. Khadem, M. R. Eissa, J. Elliott and M. Benaissa, "Blood Glucose Level Prediction: Advanced Deep-Ensemble Learning Approach," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 6, pp. 2758-2769, June 2022, doi: 10.1109/JBHI.2022.3144870.

[19]. A. J. Rodriguez-Almeida et al., "Synthetic Patient Data Generation and Evaluation in Disease Prediction Using Small and Imbalanced Datasets," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 6, pp. 2670-2680, June 2023, doi: 10.1109/JBHI.2022.3196697.