International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 220-223, June-2023

https://doi.org/10.47001/IRJIET/2023.706033

Diabetes Disease Prediction Powered by Fused Machine Learning

[1]Dr. Kavyashree N, [2]Ruba Naaz S, [3]Kruthika V R

[1]Assistant Professor, Department of MCA, Sri Siddhartha Institute of Technology, Tumkur, Karnataka, India

[2,3]Student, Department of MCA, Sri Siddhartha Institute of Technology, Tumkur, Karnataka, India

***Abstract* - To avoid diseases, early disease prediction is crucial in the medical industry. One of the most hazardous diseases in the world is diabetes. Our dietary habits in modern lifestyles are often high in sugar and fat, which has raised the risk of diabetes. Understanding the disease's symptoms is crucial for making predictions about it. Machine-learning (ML) techniques are useful right now for identifying diseases. To anticipate and analyse in this case, we employed the support vector machine and the k-nearest neighbour technique. These models examine the data set to evaluate whether a positive or negative diabetes diagnosis has been made.**

Keywords: Machine-learning, k-nearest neighbour technique, support vector.

I. INTRODUCTION

A collection of metabolic illnesses known as DIABETES are characterised by hyperglycaemia, or increased blood glucose levels. Nearly 500 million people are thought to have diabetes, and both type 1 diabetes (T1D) and type 2 diabetes are becoming more common. T1D is hypothesised to be triggered by the autoimmune death of pancreatic beta-cells, which causes a complete lack of insulin. To keep blood glucose (BG) levels in a medically suitable range, people with T1D must practise lifetime self-management. Failure to do so increases the danger of both hypoglycaemia, which can result in coma or, in the worst circumstances, death, and hyperglycaemia, which can cause microvascular and macrovascular problems.

The goal of decision support systems in diabetes care is to attenuate these negative glycaemic events and lessen the costs on persons with diabetes. In this perspective, precise BG prediction is a helpful tool for improving these systems. In instance, BG prediction allows for proactive interventions like glucose warnings in continuous glucose monitoring (CGM) and the predicted low-glucose basal insulin suspension in already-available CGM-augmented pumps. With a predetermined sample frequency (for example, every five minutes), CGM enables real-time monitoring of blood glucose levels and has been proven to be an effective tool for managing T1D. CGM can be used as a closed-loop artificial pancreas or in conjunction with insulin pumps as sensor-augmented therapy.

Connecting CGM to smartphone apps is becoming more popular in order to display retrospective BG trajectories and enable users to log daily activities (such as meals, insulin dosages, and physical activity) that affect blood sugar levels. Our project's goal is to analyse and determine whether a patient has diabetes or not. In order to better categorise diabetes, we have suggested a diabetes prediction model that takes into account a few additional characteristics in addition to the usual ones like glucose, body mass index, age, insulin, etc.

Our project's main goal is to analyse and forecast if a patient has diabetes or not by looking at their blood sugar levels. In order to better categorise diabetes, we have suggested a diabetes prediction model that takes into account a few additional characteristics in addition to the usual ones like glucose, body mass index, age, insulin, etc.

- To execute the suggested machine learning algorithm and accurately estimate the project's condition. The remaining structure of the paper of the is structured as follows, the Section.2 is discussed about the literature Survey. The existing method is presented in Section 3. The Section.4 represents the proposed methodology Section.5 presents the Implementation of the work Section.6. Represents the summary of the paper.

II. LITERATURE SURVEY

Pereverzyev Sergei [1] Using clinical data, we examine the issue of 30-min prediction of blood glucose levels determined by continuous glucose monitoring systems. The machine does not need to re-calibrate on the new patients in the data set, unlike most studies of this kind that deal with one patient at a time. Instead, we use a portion of the patients in the data set as training data and test on the remaining patients. In this case, we show how deep learning can outperform shallow networks. One innovative idea is to show how domain knowledge may be used to build a compact deep representation.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 6, pp 220-223, June-2023

https://doi.org/10.47001/IRJIET/2023.706033

- It might lower costs.
- Time requirements are minimal.

Too Many False Negatives is a drawback.

- Low probability of failure in the run.

Alessandro Aliberti's [2] approach: The blood glycaemic value of a diabetic patient can be measured with continuous glucose monitoring systems (CGMSs) at a high sampling rate, yielding a significant amount of data. Machine learning approaches can efficiently use these data to predict future glycaemic concentration levels, enabling the earlier detection of harmful hyperglycaemic or hypoglycaemic situations and improved optimisation of the diabetes treatment. Long short-term memory (LSTM) networks and non-linear autoregressive (NAR) neural networks were used to accomplish this goal. Three literature-based approaches—feed-forward neural networks (FNNs), autoregressive (AR) models, and recurrent neural networks (RNN)—were experimentally evaluated with these answers.

Advantage: The capacity to anticipate changes.

It can be frightening and expensive, which are drawbacks.

Kezhi Li and Pau Herrero's [3] approach: For patients with type 1 diabetes, a convolutional neural network (CNN) model is provided to predict their future blood glucose levels. The model is a modified version of the recently proposed Wave Net model, which is particularly beneficial in the processing of acoustic signals. The model is mostly constructed by casual dilated CNN layers and makes use of quick Wave Net methods by converting the task into a classification issue. The four input fields—glucose levels, insulin events, carbohydrate intake, and time index—come from the OhioT1DM dataset. Data and targets for the glucose change in 30 minutes are supplied into the network.

Meng Tong, Wenbo wang Methodology: Time-varying,[4] nonlinear, and non-stationary characteristics characterise the blood glucose concentration time series in diabetics. A short-term blood glucose prediction model (VMD-IPSO-LSTM) combining variational modal decomposition (VDM) and improved Particle swarm optimisation optimising Long short-term memory network (IPSO-LSTM) was developed to increase the accuracy of blood glucose prediction. In order to reduce the non-stationarity of the blood glucose time series, the intrinsic modal functions (IMF) of the blood glucose components were first extracted from the time series of blood glucose concentration of the patients using the VMD approach.

Eliasson, Bjorn [5] Personalised precision health and disease management are made possible by low-cost sensors that continually monitor blood glucose levels at intervals of a few minutes, mobile platforms, and machine learning (ML) solutions. diverse sensor technologies, analysis jobs, and people require diverse ML solutions. This brings up the scale challenge for developing such tailored ML solutions. We outline a method for forecasting diabetics' blood glucose levels up to an hour in advance. The method relies only on the patient's past glucose levels and is built on recurrent neural networks that were trained end-to-end. Users can better understand the expected levels because the model outputs the prediction together with an assessment of its certainty.

- Rapid deployment, low cost, high sensing fidelity, fault tolerance, and adaptability.

A drawback is that nodes are prone to failure.

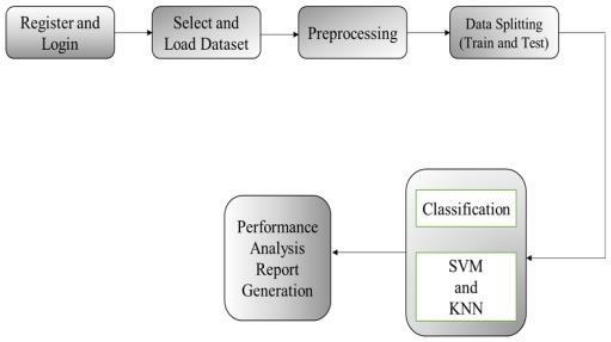

Figure 1: Proposed Architecture

III. EXISTING MODEL

Since they often concentrated on pre-processing methods, data balancing, and various forms of supervised and semi-supervised learning models, earlier research models were less accurate than contemporary ML algorithms for the detection of diseases. In order to combine the accuracy of various machine learning algorithms with high illness detection accuracy, a novel technique with decision level fusion must be developed. In order to do this, a fused machine learning model combining ANN and SVM, two supervised machine learning techniques, is proposed.

- When compared to the proposal, the loss value is relatively considerable.
- Time usage is considerable.
- Theoretical boundaries.

IV. PROPOSED METHODOLOGY

Diabetes evaluation and diabetes prediction have always been significant components of the medical industry. The next

ISSN (online): 2581-3048
Volume 7, Issue 6, pp 220-223, June-2023
https://doi.org/10.47001/IRJIET/2023.706033

stage is to implement data pre-processing. In order to prevent incorrect predictions, we must deal with the missing values in this stage. Then, after dividing the dataset into test and train, we must implement clustering for comparable data grouping. Ratio-based data splitting is used. The majority of the data will be present in train. A reduced portion of the data will be present during the test. The model is assessed during the training phase, and predictions are made during the testing phase. The purpose of feature selection is to choose the classification algorithm's best features.

- It works well with numerous datasets. When compared to the current system, the experimental outcome is superior, and the time commitment is little.
- Deliver precise prediction outcomes.

V. IMPLEMENTATION

Dataset

- The input data was gathered from the online website's UCI repository.
- The test data set in this work includes both the test dataset and the train dataset.
- This step involved reading our dataset that was gathered.

Login/User Registration

- Before logging in, users should register. Only new users are required to complete this registration process.
- User registration information such as branch, username, password, and confirm password.

Data Preprocessing

Data pre-processing is the process of eliminating unnecessary data from the dataset. The dataset is transformed using pre-processing data transformation techniques into a structure appropriate for machine learning. Missing values and Nan values are replaced by 0 during this process to remove missing data.

Splitting of Data

Data are necessary for machine learning in order for learning to occur. In addition to the data needed for training, test data are also necessary to assess how well the algorithm performs, although in this case, the training and testing dataset are separate. We must separate training and testing in our process into x_train, y_train, x_test, and y_test.

Classification

- The K-NN algorithm saves all of the information that is available and categorises new data based on similarity. This means that utilising the K-NN method, fresh data can be quickly and accurately sorted into a suitable category.
- Support Vector Machine, often known as SVM, is one of the most well-liked algorithms for supervised learning and is applied to both classification and regression issues.

VI. CONCLUSIONS

We should look at more effective teaching methods that can accommodate analysis and prediction of diabetic occurrence. Our machine learning method produces results with great accuracy. Accurate prediction status has been attained for accuracy, precision, recall, and specificity. The challenge, however, is how to effectively exercise glucose-level control by properly predicting short-term future glucose levels. Although a variety of models have been used to forecast diabetes, researchers have always been more concerned with the suggested models' ability to accurately predict diseases. Therefore, a new model is necessary to increase diabetes prediction accuracy.

REFERENCES

[1] Computer Vision and Machine Intelligence in Medical Image Analysis by F. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra. 2019; London, UK: Springer.
[2] The World Health Organisation (WHO). Leading Causes of Death and Disability Worldwide: 2000-2019, according to WHO. 22 October 2021, accessed. [Online].
[3] Asuncion and Frank, A. (2010). Machine Learning Repository at UC Irvine.
[4] G. Pradhan, R. Pradhan, and B. Khandelwal, "A Study on Various Machine Learning Algorithms Used for Prediction of Diabetes Mellitus," in Advances in Intelligent Systems and Computing, vol. 1248, Soft Computing Techniques and Applications. London.
[5] 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 Engineering, Volume 2, June 2021, Pages 40–46.
[6] M. A. Sarwar, N. Kamal, W. Hamid, and M. A. Shah., "Prediction of diabetes using machine learning algorithms in healthcare," in Proc. 24th Int. Conf. Autom. Comput. (ICAC), Sep. 2018, pp. 6-7.
[7] S. K. Dey, A. Hossain, and M. M. Rahman., Implementation of a web application to forecast diabetic disease: A method employing machine learning algorithms, in Proceedings of the 21st

ISSN (online): 2581-3048
Volume 7, Issue 6, pp 220-223, June-2023
https://doi.org/10.47001/IRJIET/2023.706033

International Conference on Computer and Information Technology (ICCIT).

[8] "Diabetes disease prediction using machine learning on big data of healthcare," by A. Mir and S. N. Dhage, 4th International Conference on Computer Communication and Control Automation (ICCUBEA), August 2018.

[9] S. Subashree and S. Saru. Diabetes Analysis and Prediction Using Machine Learning. 22 October 2022, accessed. [Online].

[10] P. Sonar and K. JayaMalini, "Diabetes prediction using different machine learning approaches," in Proc. 3rd Int. Conf. Comput. Methodologies Commun. (ICCMC), Mar. 2019, pp. 367-371.

Citation of this Article:

Dr. Kavyashree N, Ruba Naaz S, Kruthika V R, "Diabetes Disease Prediction Powered by Fused Machine Learning" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 6, pp 220-223, June 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.706033

****

