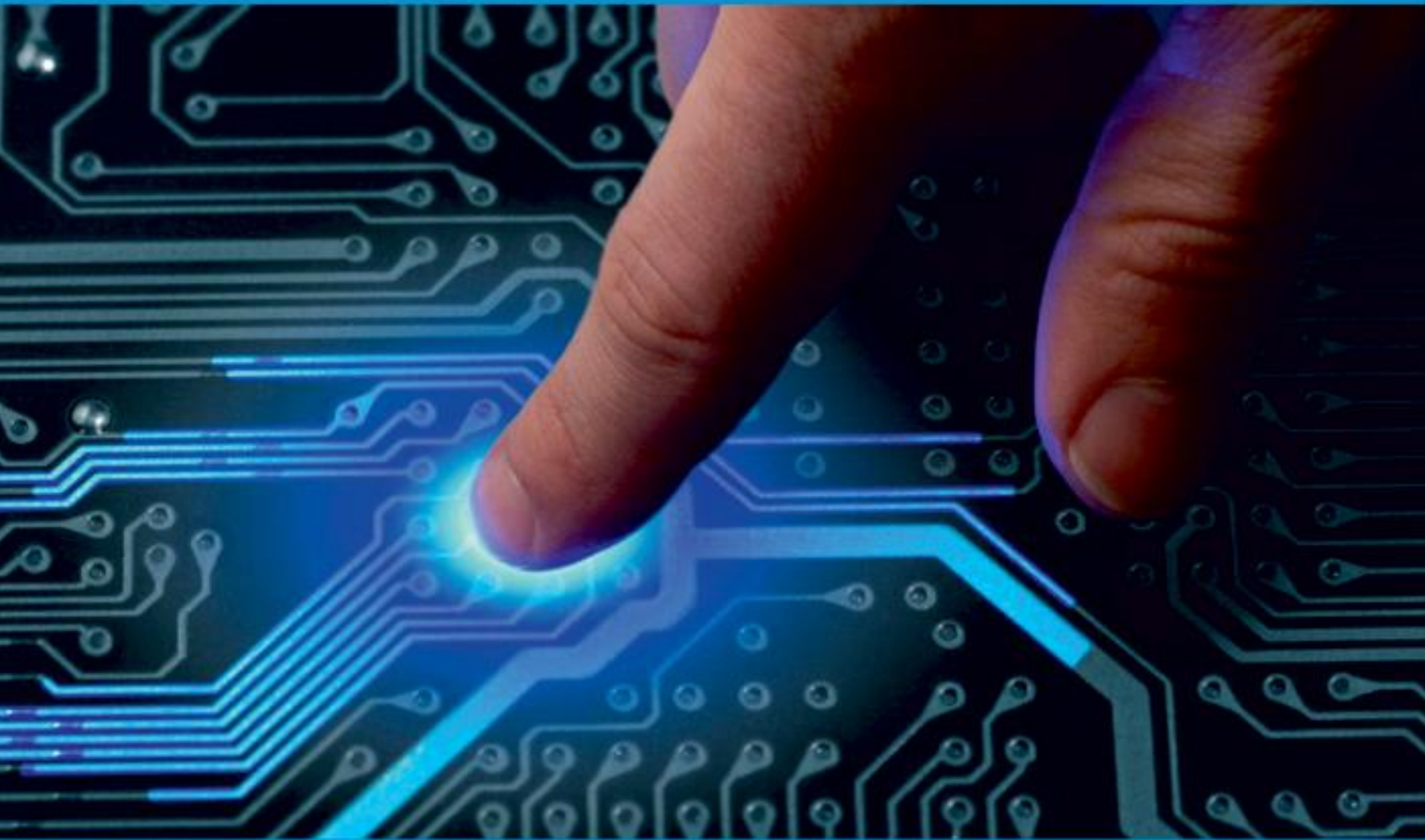
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

Impact Factor: 8.379

Identifying Diabetics Using Diabdetect Software

Devika G Pillai, Joyal B Mathew, Shon Jacob Roy, Snehamol S V,Shahida M , Smita C Thomas

U.G. Student, Department of Computer Engineering, Mount Zion College of Engineering, Kadamanitta, Kerala, India

U.G. Student, Department of Computer Engineering, Mount Zion College of Engineering, Kadamanitta, Kerala, India

U.G. Student, Department of Computer Engineering, Mount Zion College of Engineering, Kadamanitta, Kerala, India

U.G. Student, Department of Computer Engineering, Mount Zion College of Engineering, Kadamanitta, Kerala, India

Assistant Professor, Department of Computer Engineering, Mount Zion College of Engineering, Kadamanitta, Kerala, India

Associate Professor, Department of Computer Engineering, MountZion College of Engineering, Kadamanitta, Kerala, India

ABSTRACT: Detecting diabetes early is crucial in the medical field to prevent its adverse effects. With the prevalence of sugar and fat in modern diets, the risk of diabetes has escalated globally. Leveraging machine learning (ML) algorithms for disease detection has become indispensable. This article introduces a novel approach—a fusion of Support Vector Machine (SVM) and Artificial Neural Network (ANN) models—for diabetes prediction. The framework entails training and testing data sets in a 70:30 ratio. The outputs from SVM and ANN models serve as input membership functions for a fuzzy logic model, ultimately determining the diabetes diagnosis. Fused models are stored in a cloud storage system for future utilization. By analyzing real-time patient medical records, the fused ML model accurately predicts diabetes status with a remarkable 94.87% accuracy, surpassing previous methods.

KEYWORDS: Diabetic prediction, fuzzy system, fused machine learning model, diabetic symptoms, disease prediction.

I. INTRODUCTION

Diabetes stands as one of the most prevalent chronic metabolic disorders worldwide, encompassing Type-1 and Type-2 variants. Type-1 diabetes arises from immune system attacks on pancreatic Beta cells, leading to insufficient or absent insulin release. Conversely, Type-2 diabetes emerges from cellular resistance to insulin or inadequate insulin production by pancreatic cells, resulting in elevated blood glucose levels and compromised carbohydrate, fat, and protein metabolism. Symptoms of diabetes span from polyuria and polydipsia to obesity and delayed healing, with significant health complications contributing to millions of annual deaths globally. The exponential rise in diabetes-related mortality, witnessed over the years, underscores the urgent need for intelligent machine learning (ML) diagnostic systems capable of early disease detection. Effective ML frameworks rely on meticulously curated datasets, such as those sourced from the University of California Irvine (UCI) Machine Learning repository and compiled by institutions like the Sylhet Hospital in Bangladesh. These datasets furnish the necessary features for robust model training and validation, enhancing prediction accuracy. Preventive strategies against malnutrition and obesity, common diabetes precursors, emphasize lifestyle modifications and a balanced diet. Routine medical examinations, coupled with laboratory tests, facilitate timely disease diagnosis. For Type-2 diabetes patients, insulin therapy remains indispensable, underlining the critical importance of early detection and intervention. Harnessing Artificial Intelligence (AI), particularly in medical diagnosis, offers promising avenues for disease detection and prevention. This research introduces a sophisticated framework leveraging machine learning fusion for the early identification of diabetic patients, representing a significant stride towards proactive healthcare management.

II. RELATED WORK

Presenting the Fused Model for Diabetes Prediction (FMDP), this innovative approach comprises two primary phases: the Training Layer and the Testing Layer. The Training Layer encompasses distinct stages: Data Acquisition, Preprocessing, Classification, Performance Evaluation, and Machine-Learning Fusion. Leveraging datasets from the renowned UCI Machine Learning Repository, the Data Acquisition phase ensures the availability of feature-rich datasets essential for diabetes prediction. Subsequently, data undergoes meticulous cleaning, normalization, and partitioning into training and test sets during Preprocessing. Notably, Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) are employed for prediction, chosen for their efficacy following preliminary experimentation. Performance Evaluation involves assessing various accuracy metrics like accuracy, specificity, sensitivity, precision, and F1 score. If model performance falls short, iterative retraining is conducted until learning requirements are met. Once achieved, SVM and ANN outputs serve as inputs for machine-learning fusion, where fuzzy rules are applied to produce final predictions. The resultant fused model is stored in the cloud for future accessibility. The Testing Layer, representing the second phase, involves acquiring datasets from medical databases and retrieving preprocessed training models from the cloud. Utilizing the fused model, predictions regarding diabetes diagnosis positivity or negativity are made. Prediction accuracy is evaluated by comparing the expected output with the actual output. For model training, the preprocessed dataset is split into training and test data with a 70:30 ratio, utilizing Bayesian regularization with a 5% validation set and the remaining 90% for training. This systematic approach ensures robust model training and validation, paving the way for accurate diabetes prediction.

III. METHODOLOGY

In response to the persistent quest for more accurate diabetes prediction models, this study introduces a pioneering method: a diabetes decision support system leveraging machine learning and decision-level fusion. By synergizing two prevalent machine learning approaches via fuzzy logic, the proposed model achieves a remarkable accuracy of 94.87%, outperforming current systems. This heightened precision offers potential in averting fatalities and curbing the mortality rates linked with diabetes, emphasizing the significance of early detection and preventive strategies.

IV. EXPERIMENTAL ISSUES

The proposed Fused Model for Diabetes Prediction (FMDP) underwent rigorous evaluation in two main phases: the Training Layer and the Testing Layer.

In the Training Layer, a series of steps were followed, commencing with data acquisition from the UCI Machine Learning Repository. Subsequently, the dataset underwent preprocessing, involving cleaning, normalization, and partitioning into training and test datasets. Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) were then trained using the preprocessed data for diabetes prediction, chosen for their efficacy following initial experimentation. Performance evaluation of these trained models ensued, utilizing a range of accuracy measures such as accuracy, specificity, sensitivity, precision, and F1 score. Any model failing to meet performance criteria underwent retraining until satisfactory results were achieved. Upon meeting learning requirements, SVMs and ANNs outputs were fused using fuzzy rules to yield the final prediction, with the fused model stored in the cloud for future use.

Moving to the Testing Layer, the performance of the trained model was assessed using an independent dataset sourced from a medical database. The preprocessed training model stored in the cloud was retrieved, and the fused model was applied to predict diabetes diagnoses. Prediction accuracy was determined by comparing predicted outcomes with actual outcomes.

During the training of the ANN model, the preprocessed training dataset was divided into training, testing, and validation sets in a 70:30 ratio, respectively. Bayesian regularization was employed during training, allocating 5% of the data to testing and validation subsets, with the remaining 90% used for actual training.

Overall, experimental results underscore the efficacy of the proposed FMDP framework in accurately predicting diabetes diagnoses, with the fused model demonstrating high prediction accuracy when tested on independent datasets.

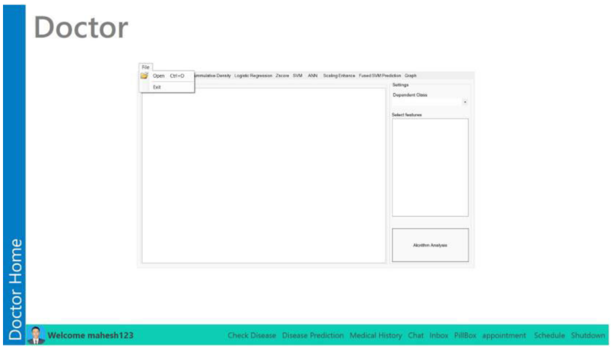
International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205240 |

Fig 1 and 2 shows how to select a training data set



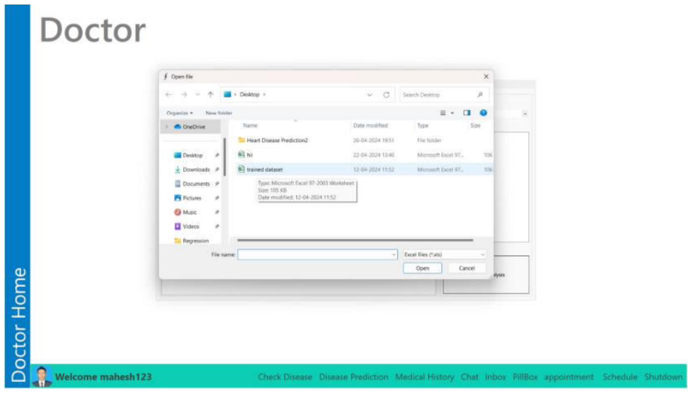


Fig 3 and 4 shows select trained dataset

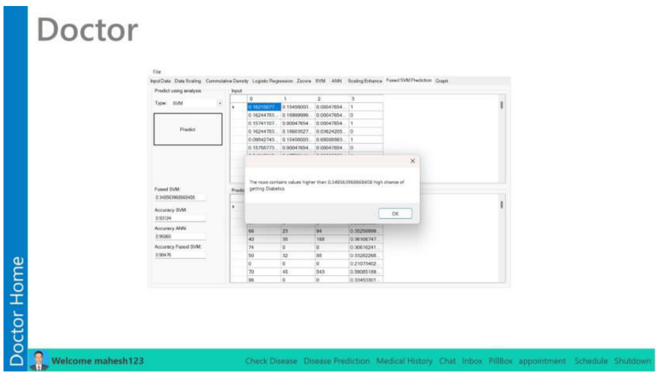


Fig 5 and 6 shows prediction of diabetic

V. CONCLUSION

Machine learning methods and techniques are beginning to play an ever-increasing role in the domain of healthcare for the analysis of medical data to support the diagnosis and inform on the optimum treatment of critical health conditions. However, the existence and availability of sufficiently large datasets for the training and testing of machine learning models remain a barrier to achieving better-performing algorithms. Thus, given the prevailing restrictions in the scope and quality of available data, the paper reports on the development and evaluation of the performance of an approach based on the fusion of two machine learning methods for the prediction of diabetes. Though different models had been used for the prediction of diabetes, the accuracy of the proposed models in disease prediction has always been the main concern of researchers. Therefore, a new model is required in order to achieve higher prediction accuracy in diabetes prediction. This research proposed a machine learning based diabetes decision support system by using decision level fusion. Two widely used machine learning techniques are integrated in the proposed model by using the fuzzy logic. The proposed fuzzy decision system has achieved the accuracy of 96.87, which is higher than the other existing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205240 |

systems In the future, other machine learning classifiers such as Random Forest, Decision Tree can also be consider at machine learning fusion layer .

REFERENCES

[1] F. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra, Computer Vision and Machine Intelligence in Medical Image Analysis. London, U.K.: Springer, 2019.

[2] World Health Organization (WHO). (2020). WHO Reveals Leading Causes of Death and Disability Worldwide: 2000–2019. Accessed: Oct. 22, 2021. [Online]. Available: https://www.who.int/news/item/09-12-2020-whoreveals-leading-causes-of-death-anddisability-worldwide-2000-2019

[3] A. Frank and A. Asuncion. (2010). UCI Machine Learning Repository. Accessed: Oct. 22, 2021. [Online]. Available: http://archive.ics.uci.edu/ml

[4] G. Pradhan, R. Pradhan, and B. Khandelwal, ‘‘A study on various machine learning algorithms used for prediction of diabetes mellitus,’’ in Soft Computing Techniques and Applications (Advances in Intelligent Systems and Computing), vol. 1248. London, U.K.: Springer, 2021, pp. 553–561, doi: 10.1007/978-981-15-7394-1_50.

[5] S. Kumari, D. Kumar, and M. Mittal, ''An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier,'' Int. J. Cogn. Comput. Eng., vol. 2, pp. 40–46, Jun. 2021, doi: 10.1016/j.ijcce.2021.01.001.

[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, doi: 10.23919/IConAC.2018.8748992.

[7] S. K. Dey, A. Hossain, and M. M. Rahman, ''Implementation of a web application to predict diabetes disease: An approach using machine learning algorithm,'' in Proc. 21st Int. Conf. Comput. Inf. Technol. (ICCIT), Dec. 2018, pp. 21–23, doi: 10.1109/ICCITECHN.2018.8631968.