# Diabetic Report Analyzer Review Paper

*Dnyaneshwar Sonawane[1], Vaishnavi Punalkar[2], Raj Pandharpatte[3], Prerna Bodake[4], Prof. Ravishankar Bhaganagare[5]*

[1,2,3,4]Student, Computer Engineering, SKN Sinhgad Institute of Technology and Science Kusgaon Pune, Maharastra, India
[5]Lecturer, Computer Engineering, SKN Sinhgad Institute of Technology and Science Kusgaon Pune, Maharastra, India

**Abstract—**

Diabetes is a type of metabolic disorder with a high level of blood glucose. Due to the high blood sugar, the risk of heart-related diseases like heart attack and stroke gets increased. The number of diabetic patients worldwide has increased, and it is considered to be a major life-threatening disease worldwide. The diabetic disease cannot be cured but it can be controlled and managed by timely detecting it. Artificial Intelligence (AI) with Machine Learning (ML) empowers automatic early diabetes detection which is found to be much better than a manual method of diagnosis. At present, there are many research papers available on diabetes detection using ML techniques.

*Index Terms— Classification, Diabetes, Machine Learning, Prediction Algorithm.*

## I. Introduction

Major growth in biotechnology and the good throughput computation is contributing to rapid and economic data creation. One of the most important applications is in diabetes disorder auguring. Diabetes mellitus (DM) is one kind of human threatening disease-causing lot of other health problems [1]. Diabetes is among the most widespread life-threatening disease. In 2012, death caused due to diabetes is about 1.5 million, and the mortality rate of 2.2 million due to heart disease, and kidney problems. In 2017, about 8.8% of the world population was affected by diabetic Mellitus. In India, nearly 80 million people are with high blood glucose and India is in the second for having the highest count of diabetic patients in the world. According to the National Diabetes Statistics Report 2020, about 34.2 million population in the United States are affected by high blood glucose. Only 26.9 million population have detected diabetes and the remaining 7.3 million were not aware of this diabetic condition. A diabetes diagnosis can either be done by the manual method through the physician or by a device. Both the diagnosis method has their advantages and disadvantages. The main benefit of the manual method of diagnosis is there is no need for any help from an automatic device. But in most cases, the symptoms of diabetes are very low and it is difficult to find out even by a medical professional expert at the earlier stage. The manual diagnosis is an uncomfortable and painful invasive method, sometimes it is infective. Due to the advancement in AI and ML, automated diagnosis is a more possible efficient method that assists the manual diagnosis. Many researchers have done diabetes-related research but there are only a few published review articles(Chaki et al., 2020; Jaiswal et al., 2021. As a result, the effort has been taken to consider and examine ML and AI-based diabetes report analyzer.

## II. Diabetes disease

Once the food is taken, the body will convert the food into sugar/glucose and move it into the blood. Insulin is produced by the pancreas which is the hormone that helps to move sugar to the cells from the bloodstream. If the body is unable to produce enough insulin that condition is commonly named high blood sugar. People with high blood sugar face a high risk of getting many secondary disorders such as heart problems and nerve-related diseases. The main reason for diabetes is not identified but the researchers believe that gene and the living lifestyle is the main reason for diabetes. Detecting diabetes at an earlier stage and taking treatment for it can reduce the harmful complications and reduces the risk of other health issues. The diabetic disease cannot be cured but it can be controlled and managed by timely detection.

### A. Prediabetes

Prediabetes arises when the glucose in the blood level gets increase, in the prediabetes the symptoms of this stage are very low so doctors are not able to find the prediabetes stage very fast. The prediabetes stage increases the rate of heart related risks and also it gives to rise to Type-2 stage.

*B. Type-I Diabetes*

The Type1diabetes arises usually in infancy. The condition at which the body generates no insulin/very little insulin is referred to as Type 1 diabetes. Insulin injections can be used by the patients to control type 1 diabetes. Symptoms of this type of DM are unusual weight loss, unusual hunger, and thirst, abnormal urination, kidney, and eyes related disorders. Symptoms of type 1 diabetes will increase the higher risk of stroke and heart-related disease (Katsarou et al., 2017; Pranto et al., 2020).

*C. Type-II Diabetes*

Type 2 diabetes (T2D) arises when the body doesn't react to insulin and usually occurs in adults. Symptoms of type 2 diabetes are weight gain and a high rise in blood pressure.T2D increases the probability of getting heart-related disorders and stroke (Jaiswal et al., 2021; Pranto et al., 2020).

## III. Methods of Study Selection

Significant studies were selected based on the IEEE protocol. Journal search was performed on Web of Science, Scopus, PubMed databases from the year 1988 to 2021. This strategic selection of studies was conducted using the keywords like "machine learning", "prediction", "detection", and "Artificial Intelligence".

*A. Selection Method*

The work of study selection was done based on IEEE. Finally, some articles were short listed by team. The papers were searched by the team authors considering the guidelines. During the inceptive stage examination of research papers was based on the name and the abstract of the studies and only the studies that matches the article selection conditions were chosen manually. At future stage of the search, the full subject matter of the paper is considered and only the studies that matches the article selection conditions were selected. Extraction of data includes the name of the author, year of publication, ML algorithm, and the model with the best performance.

*2. Knowledge about Machine Learning*

- Machine Learning is a subfield of Artificial Intelligence (AI). The main goal to the machine learning technology is to understand the structure of the data and fit that into specific models which then later can be able to understand the data and used by humans for various application throughout life.

- In general, there are two types of Machine Learning: (i) deductive machine learning and (ii) inductive machine learning. The deductive ML method learns from the previous data or existing data and the inductive ML method learns by taking examples.

- Machine Learning has three main types of learnings; they are; Supervised learning, unsupervised learning and reinforcement learning. Many algorithms are used to predict and classify diabetes in research articles. The most commonly used algorithms for the prediction of diabetes are: -

- Support Vector Machine (SVM): Support Vector Machine SVM is a supervised ML method used for binary classification purposes. SVM transforms the input data to a required data using a set of functions like linear function, nonlinear function, and polynomial function. They are mostly used for medical applications. SVM lower the error in the empirical classification and enlarge the geometrical margin and hence it is named as maximum margin classifier(Brereton & Lloyd, 2010; Singh & Chaturvedi, 2012).

- K-Nearest Neighbor (KNN) Algorithm :- The KNN algorithm can be used in the research of supervised learning and unsupervised ML problems. In the present situation, KNN is the most needed ML algorithm. KNN algorithm stores the dataset at the raining phase and when the algorithm gets the new data, then it classifies that data into category which is more similar to the new data. KNN algorithm can be used for regression as well as for classification.

*C. Machine Learning-based Diabetes Prediction Techniques*

There are very popular ML techniques for predicting different health related issues in medical applications. Many scientists have tried to develop diabetes prediction systems using various machine learning techniques. Some of the research in diabetes detection is mentioned in this content. At the start, the ML algorithm based on neural network using the PIMA dataset was used for the estimation of diabetes (US Department of Health and Human Services, 2020). Currently, many diabetes prediction models using neural networks had been developed (Yang & Wang, 2020; Zhou et al., 2020). Omar and Eman(S.Soliman&AboElhamd, 2014) introduced a technique for type 2 diabetes. For prediction a hybrid method of LS-SVM (Least Squares SVM) and MPSO (Modified Particle Swarm Optimization) was used.

Sridar and Shanthi (M.E. & M.E., 2014) introduced a diabetes prediction system by implementing Apriori and a backpropagation algorithm. In this system, the medical data records are taken from PIMA. In this study, Backpropagation achieved an accuracy of 83.5%, Apriori algorithm achieved an accuracy of 71.2% and by combining Apriori and backpropagation algorithm an accuracy of 91.2% were achieved.

Daghestani and Alshammari (Daghistani&Alshammari, 2020) had complete study on RF and logistic regression ML algorithm-based diabetes detection. From the Saudi Arabia region; the data for this model is been considered. Many performance criteria were calculated; and for validation 10 fold cross-validation method had been used.

## IV. Motivation And Aim

The Hyperglycemia (high blood sugar) condition is major health issue which causes various serious ramification. According to World Health Organization, 1.4 million populations gets affected by diabetes every year, and half of the population remains undetected. Time to time prediction is very important thing that helps to take on time treatment so that future consequences can be avoided. Consequently by considering the condition of diabetes, several diabetes detections in reference to the ML algorithms are developed.

Application of new ML techniques is required but without the understanding of the existing problems and the limits in previous diabetes detection, the advancement is impossible. In the latest research, the new ML algorithms are giving performance than the existing applications. Survey of diabetes prediction report that, the prediction is made using Machine Learning -based neural network methods.

The main aim of our application is to develop a system that predicts the diabetes with better accuracy. In this system the ML algorithms are used to classify diabetes patients. The classification accuracy is achieved by the classifying diabetes patients. The aim of introducing a diabetes prediction system is to shift from high precision to high reliability for real time applications(system).
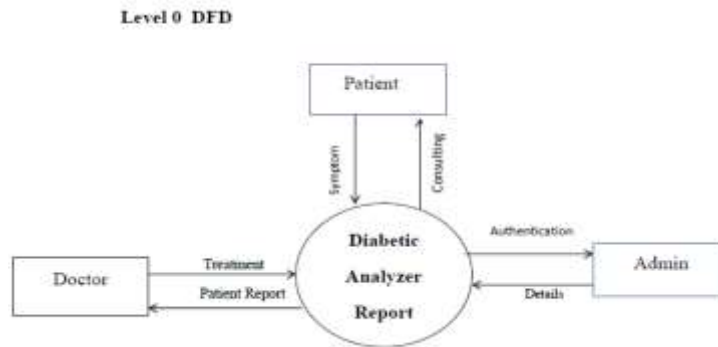
### A. Figures and Tables

### 1) Use case diagram of Diabetes Report Analyzer:

The Use case diagram is used to identify the interaction between system and actors. Our system have 3 actors such as – Admin, Doctor and Patient. These three interact with their own use-cases. Below mention is use-case diagram:-
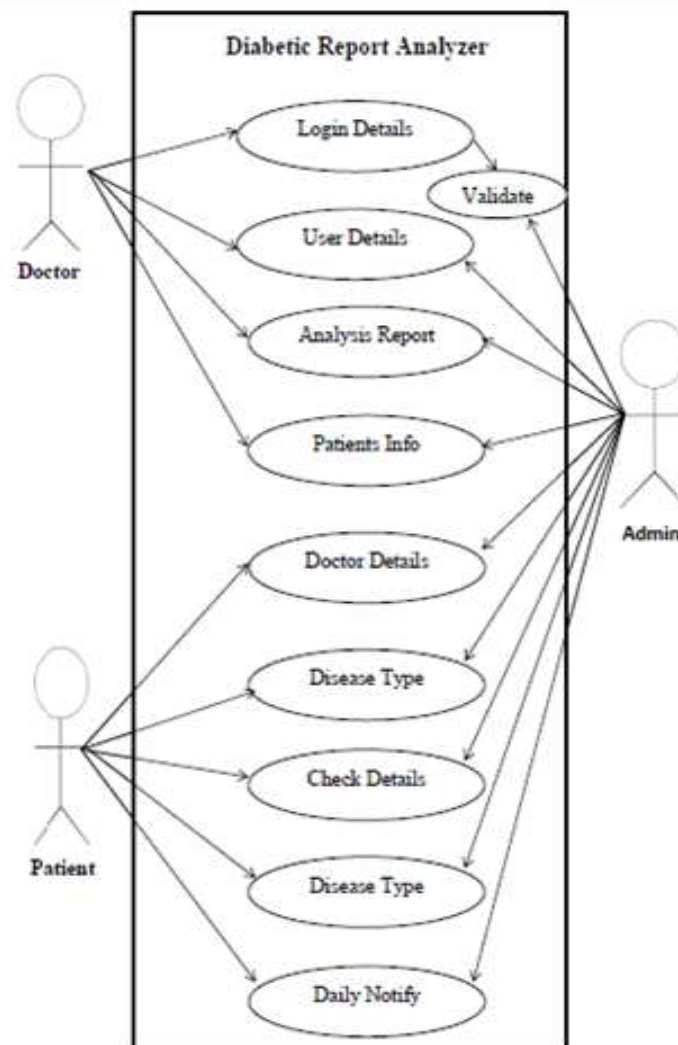
*2) Data Flow Diagram (Level 0)*



*Acknowledgment*

## References

1. G. Swapna, K. P. Soman, and R. Vijayakumar, "Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals," Procedia Computer Science, vol. 132, no. Iccids, pp. 1253–1262, 2018.

2. P. Saeediet al., "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition," Diabetes Research and Clinical Practice, vol. 157, p. 107843, 2019.

3. S. Larabi-Marie-Sainte, L. Aburahmah, R. Almohaini, and T. Saba, "Current techniques for diabetes prediction: Review and case study," Applied Sciences (Switzerland), vol. 9, no. 21, 2019.

4. I. Qureshi, J. Ma, and Q. Abbas, "Recent development on detection methods for the diagnosis of diabetic retinopathy," Symmetry, vol. 11, no. 6, pp. 1–34, 2019.

5. S. S. G. and M. K., "Diagnosis of diabetes diseases using an optimized fuzzy rule set by grey wolf optimization," Pattern Recognition Letters, vol. 125, pp. 432–438, 2019.

6. S. Natarajan, A. Jain, R. Krishnan, A. Rogye, and S. Sivaprasad, "Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening with an Offline Artificial Intelligence System on a Smartphone," JAMA Ophthalmology, vol. 137, no. 10, pp. 1182–1188, 2019.

7. K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," Clinical Epidemiology and Global Health, vol. 7, no. 4, pp. 530–535, 2019.

8. M. Jahangir, H. Afzal, M. Ahmed, K. Khurshid, and M. Faisal, "Abbas (2019) Auto-MeDiSine: An Auto-Turnable Medical Decision Support Engine Using an Automated Class Outlier Detection MEthod and Auto AMLP. Neural Computing and Applications. ISSN 1433–3058 Downloaded from : https://e-space.mmu.ac.uk/623523/ Versio," 2019.

9. N. Yuvaraj and K. R. SriPreethaa, "Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster," Cluster Computing, vol. 22, 2019.

10. N. Sneha and T. Gangil, "Analysis of diabetes mellitus for early prediction using optimal features selection," Journal of Big Data, vol. 6, no. 1, 2019.

11. H. Zhou, R. Myrzashova, and R. Zheng, "Diabetes prediction model based on an enhanced deep neural network," Eurasip Journal on Wireless Communications and Networking, vol. 2020.

12. 1. J. Chaki, S. Thillai Ganesh, S. K. Cidham, and S. AnandaTheertan, "Machine learning and artificial intelligence-based Diabetes Mellitus detection and self-management: A systematic review," Journal of King Saud University - Computer and Information Sciences, no. XXXX, 2020.

13. V. Jaiswal, A. Negi, and T. Pal, "A review on current advances in machine learning-based diabetes prediction," Primary Care Diabetes, vol. 15, no. 3, pp. 435–443, 2021.

14. B. Pranto, S. M. Mehnaz, E. B. Mahid, I. M. Sadman, A. Rahman, and S. Momen, "Evaluating machine learning methods for predicting diabetes among female patients in Bangladesh," Information (Switzerland), vol. 11, no. 8, 2020.

15. B. A, V. v S, S. D, and U. S, "Diagnosis of Diabetic Retinopathy Using Machine Learning," Innovations in Information and Communication Technology Series, vol. 6956, no. July, pp. 477–481, 2020.

16. T. Daghestani and R. Alshammari, "Comparison of statistical logistic regression and random forest machine learning techniques in predicting diabetes," Journal of Advances in Information Technology, vol. 11, no. 2, pp. 78–83, 2020

17. G. Bansal and M. Singla, "Ensembling of non-linear SVM models with partial least square for diabetes prediction," Lecture Notes in Electrical Engineering, vol. 569, pp. 731–739, 2020.

18. R. D. Howsalya Devi, A. Bai, and N. Nagarajan, "A novel hybrid approach for diagnosing diabetes mellitus using farthest first and support vector machine algorithms," Obesity Medicine, vol. 17, p. 100152, 2020.

19. H. Naz and S. Ahuja, "Deep learning approach for diabetes prediction using PIMA Indian dataset," Journal of Diabetes and Metabolic Disorders, vol. 19, no. 1, pp. 391–403, 2020.

20. M. Rahman, D. Islam, R. J. Mukti, and I. Saha, "A deep learning approach based on convolutional LSTM for detecting diabetes, "Computational Biology and Chemistry, vol. 88, p. 107329, 2020.

21. J. M. M. Rumbold, M. O'Kane, N. Philip, and B. K. Pierscionek, "Big Data and diabetes: the applications of Big Data for diabetes care now and in the future," Diabetic Medicine, vol. 37, no. 2, pp. 187–193, 2020.

22. H. Naz and S. Ahuja, "SMOTE-SMO-based expert system for type II diabetes detection using PIMA dataset," International Journal of Diabetes in Developing Countries, no. 0123456789, 2021.

23. 70. R. Patra and B. khuntia, "Analysis and Prediction Of Pima Indian Diabetes Dataset Using SDKNN Classifier Technique," IOP Conference Series: Materials Science and Engineering, vol. 1070, p. 012059, 2021.