A Review on Applications of Machine Learning Techniques in Diabetes Handling

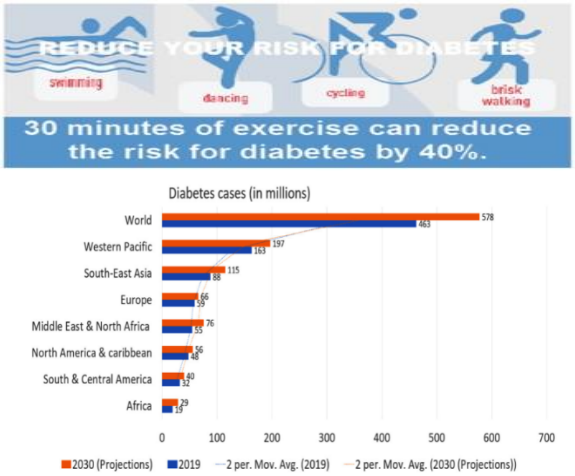
[1]Samridhi Puri, [2]Satinder Kaur [3]Satveer Kour [4] Kumari Sarita
[1,2,3]*Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar*
[4]*Deparment of Computer Science, Guru Nanak Dev University, Amritsar*

Abstract - Diabetes mellitus is a chronic metabolic disorder affecting millions of people worldwide. Early diagnosis and effective management of diabetes are crucial to prevent complications and improve patient outcomes. In recent years, the integration of machine learning (ML) techniques has shown promising results in aiding diabetes diagnosis and management. Various Machine learning techniques can be used like Support Vector Machine (SVM), K Nearest Neighbour (KNN), Random Forest, Decision Tree in detection and treatment of diabetes mellitus. This paper provides an overview of different machine learning applications used in diabetes handling. In addition, it also discusses how to use ML in this discipline. It is concluded that ML techniques is a good approach for controlling diabetes mellitus at an early stage.

Keywords – Machine learning, Diabetes mellitus, Support Vector Machine, Random Forest

INTRODUCTION

Diabetes is a chronic disease and commonly stated by health professionals or doctors as diabetes mellitus (DM), which describes a set of metabolic diseases in which the person has blood sugar, either insulin production inefficient, or because of the body cell do not return correctly to insulin, or by both reasons. This will increase concentration levels of glucose in the blood. Most cases of diabetes can be broadly classified in two categories, type 1 and type 2, although some cases are difficult to classify. Many complications occur if diabetes remain untreated. Therefore, it is not only a disease but also a creator of different kinds of diseases like heart attack, blindness, kidney diseases, etc. Diabetes has become one of the major causes of national disease and death in most countries. According to the International Diabetes Federation report, this figure is expected to rise to more than 642 million in 2040, so early screening and diagnosis of diabetes patients have great significance in detecting and treating diabetes on time. (Firdous et.al,2022) The analysis of diabetes data is a challenging issue because most of the medical data are nonlinear, abnormal, correlation structured, and complex in nature. Applying machine learning methods in diabetes mellitus research is a key approach to utilizing large volumes of available diabetes-related data for extracting knowledge. We have also discussed Machine learning cycle for implementation in diabetes mellitus at an early stage.



Present scenario and projections of diabetes cases in different parts of the world (International Diabetes Federation [IDF] (Saria et.al,2022)

HOW TO USE ML TECHNIQUES IN DIABETES

Various researchers have used different ML techniques like Support Vector Machine, Random Forest, Decision Tree, K Nearest Neighbour. (Ahmed et al.,2021); (Saru et al.,2019)

1. Data Collection and Preprocessing: Various researchers have used different ML techniques like Support Vector Machine, Random Forest,

Decision Tree, K Nearest Neighbour. ML models require a diverse and representative dataset to train effectively. Relevant features might include age, gender, BMI, family history, blood pressure, glucose levels, and other biomarkers. Data preprocessing involves cleaning, normalizing, and handling missing values to ensure the quality of input data. (Cardozo et al. ,2022)

2.Feature Selection and Engineering:

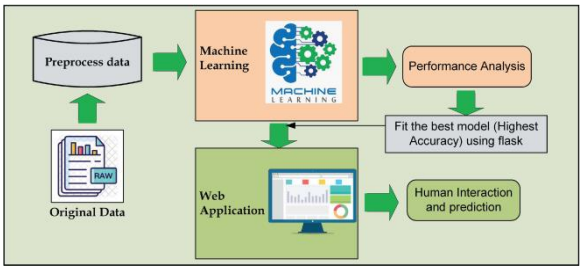


Figure 1 (Ahmed et.al,2021)

Selecting the right features can significantly impact the performance of a predictive model. Feature engineering involves transforming or creating new features that can improve prediction accuracy. For instance, deriving a feature like BMI from height and weight can provide valuable insights. (Nadeem et.al,2021); (Kaur et.al,2022)

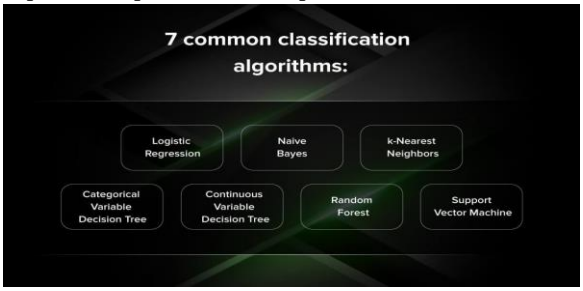
3.Model Selection: There are various ML algorithms suitable for diabetes prediction, including:

Logistic Regression: A simple algorithm that models the probability of an event occurring based on input features.

Random Forest: An ensemble method that combines multiple decision trees for improved accuracy. (Nadeem et.al,2021)

Support Vector Machines (SVM): These find a hyperplane that best separates data points of different classes.

Neural Networks: Deep learning models that can capture complex relationships in the data.



4.Model Training and Evaluation: The selected model is trained using a portion of the dataset and evaluated using another portion. Common evaluation metrics include accuracy, precision, recall, F1-score, and area under the Receiver Operating Characteristic (ROC-AUC) curve. (Shafi et.al,2021)

5.Hyperparameter Tuning: ML models have parameters that need to be fine-tuned for optimal performance. Techniques like grid search or random search are used to find the best combination of hyperparameters.

6.Validation and Testing: Once the model is trained and tuned, it needs to be validated using a separate dataset to ensure it generalizes well to unseen data. Cross-validation techniques help in assessing the model's stability. (Firdous et.al,2022)

7.Deployment and Monitoring: A well-performing model can be deployed in clinical settings to assist healthcare professionals in making informed decisions. Continuous monitoring is essential to ensure the model's predictions remain accurate over time as new data becomes available.

8.Ethical Considerations: Patient privacy, data security, and fairness are critical aspects when deploying ML models in healthcare. Ensuring that the model does not inadvertently perpetuate biases is important.

In conclusion, ML techniques have the potential to revolutionize diabetes prediction by analysing patient data and providing early insights for better management and care. However, it's essential to approach these techniques with caution, considering ethical concerns and continuously validating their performance in real-world scenarios.

USES OF ML TECHNIQUES IN DIABETES

Machine learning (ML) techniques have found various applications in the field of diabetes management and research. These techniques leverage the power of data analysis and pattern recognition to improve prediction, diagnosis, treatment, and overall understanding of diabetes. Here are some applications of ML in diabetes. (Saria et.al,2022)

1.Early Diagnosis and Risk Prediction: ML algorithms can analyse patient data, including medical history, genetic information, and lifestyle factors, to predict the risk of developing diabetes or its complications. This

can aid in early intervention and personalized preventive measures.

2.Glucose Monitoring and Insulin Dosage: ML models can process continuous glucose monitoring (CGM) data to predict blood glucose levels and optimize insulin dosages for individuals with diabetes. This helps in maintaining stable blood sugar levels and avoiding dangerous fluctuations.

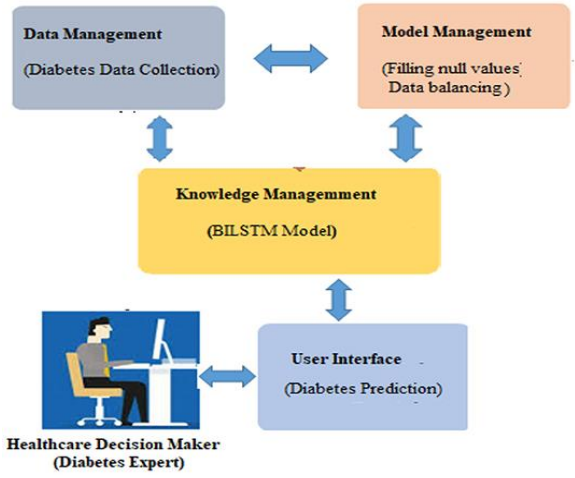
3.Hypoglycemia and Hyperglycaemia Detection: ML algorithms can detect patterns that lead to hypoglycaemic (low blood sugar) or hyperglycaemic (high blood sugar) episodes. Early detection can prompt timely interventions to prevent severe complications.

4.Diabetes Subtype Classification: ML techniques can analyse genetic and clinical data to categorize different diabetes subtypes, such as type 1, type 2, and gestational diabetes. This aids in targeted treatment and management strategies. (Ahmed et.al,2021)

5.Risk Stratification: ML techniques can assess a patient's risk of developing diabetes or its complications based on factors such as genetics, medical history, lifestyle, and biomarkers. This information can guide preventive measures and early interventions.

6. Complication Risk Assessment: ML techniques can assess a patient's risk of developing diabetes-related complications by analysing various factors such as medical history, lab results, and lifestyle data. This can aid in early intervention and preventive measures.

7. Dietary Management: ML algorithms can analyse dietary habits and glucose level patterns to provide personalized dietary recommendations. This can help individuals with diabetes make healthier food choices and manage their blood sugar levels more effectively.



CONCLUSION

It is concluded that Machine learning techniques are very much useful in diabetes detection at an early stage. As discussed above various Machine learning techniques can be used in applications of diabetes handling. Applying machine learning methods in diabetes mellitus research is a key approach to utilizing large volumes of available diabetes-related data for extracting knowledge.

REFERENCE

[1] Ahmed, N., Ahammed, R., Islam, M. M., Uddin, M. A., Akhter, A., Talukder, M. A., & Paul, B. K. (2021). Machine learning based diabetes prediction and development of smart web application. *International Journal of Cognitive Computing in Engineering*, *2*, 229-241.

[2] Cardozo, G., Pintarelli, G. B., Andreis, G. R., Lopes, A. C. W., & Marques, J. L. B. (2022). Use of Machine Learning and Routine Laboratory Tests for Diabetes Mellitus Screening. *BioMed research international*, *2022*.

[3] Das, D., Biswas, S. K., & Bandyopadhyay, S. (2022). A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning. *Multimedia Tools and Applications*, *81*(18), 25613-25655.

[4] Firdous, S., Wagai, G. A., & Sharma, K. (2022). A survey on diabetes risk prediction using machine learning approaches. *Journal of Family Medicine and Primary Care*, *11*(11), 6929.

[5] Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology & metabolic syndrome*, *13*(1), 1-22.

[6] Gouri Saria, M. K., Jee, G., Harshvardhan, G. M., Singh, V., Singh, P. K., & Workneh, T. C. (2022). Data science appositeness in diabetes mellitus diagnosis for healthcare systems of developing nations. *IET Communications*, *16*(5), 532-547.

[7] Kaur, H., & Kumari, V. (2022). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied computing and informatics*, *18*(1/2), 90-100.

[8] Khan, F. A., Zeb, K., Al-Rakha mi, M., Dahab, A., & Bukhari, S. A. C. (2021). Detection and

prediction of diabetes using data mining: a comprehensive review. *IEEE Access*, *9*, 43711-43735.

[9] Nadeem, M. W., Goh, H. G., Ponnusamy, V., Ando Novic, I., Khan, M. A., & Hussain, M. (2021, October). A fusion-based machine learning approach for the prediction of the onset of diabetes. In *Healthcare* (Vol. 9, No. 10, p. 1393). MDPI.

[10] Saru, S., & Subashree, S. (2019). Analysis and prediction of diabetes using machine learning. *International journal of emerging technology and innovative engineering*, *5*(4).

[11] Shafi, S., & Ansari, G. A. (2021, May). Early prediction of diabetes disease & classification of algorithms using machine learning approach. In *Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021)*.

AUTHORS

First Author- Satinder Kaur, PhD, Guru Nanak Dev University Amritsar

Second Author- Satveer Kour, PhD, Guru Nanak Dev University Amritsar

Third Author- Samridhi Puri, Pursuing MTech, Guru Nanak Dev University Amritsar