IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Review Of Diabetes Classification And Prediction Based On Artificial Intelligence Techniques

[1]Prateeksha Singh, [2]Dr. Shikha Agrawal, [3]Dr. Sanjay Silakari,
Department of Computer Science & Engineering
University Institute of Technology, RGPV, Bhopal, India

Abstract : It is crucial to the profession of medicine to be able to anticipate illnesses early on in order to avoid them. Diabetes is widely recognized as one of the most lethal illnesses on a global scale. Because of the current lives we lead, sugar and fat are frequently included in our dietary habits; as a result, the risk of developing diabetes has grown. The metabolic condition known as diabetes mellitus (DM) is characterized by elevated levels of glucose in the blood. The pathology may show itself in a variety of disorders, the most prominent of which is neuropathy, which is brought on by diabetes illness. It is essential to have a thorough understanding of the disease's symptoms in order to make an accurate prognosis. In the field of illness diagnosis, machine learning (ML) techniques are now quite useful. The purpose of this study is to provide an overview of machine learning approaches for the categorization and prediction of diabetes.

IndexTerms - AI, Machine Learning, Diabetes Classification, Disease**.**

I. INTRODUCTION

Diabetes, a chronic metabolic disorder characterized by high blood sugar levels, affects millions of people worldwide and poses significant challenges to public health systems. Early detection and accurate classification of diabetes can help in effective management and personalized treatment of the disease. Artificial intelligence (AI) techniques have emerged as powerful tools in the field of healthcare, including diabetes classification and prediction. Leveraging machine learning algorithms, AI has the potential to improve the accuracy, efficiency, and accessibility of diabetes diagnosis and prognosis.

Diabetes classification involves categorizing individuals into different classes based on their diabetes status, such as normal, pre-diabetes, or different types of diabetes (type 1 or type 2). On the other hand, diabetes prediction aims to estimate the risk of developing diabetes in individuals who are currently non-diabetic. By analyzing various factors such as clinical data, genetic information, lifestyle patterns, and environmental factors, AI techniques can uncover complex relationships and patterns that may contribute to diabetes onset and progression.

The application of AI techniques in diabetes classification and prediction offers several potential benefits. Firstly, it can enhance the accuracy and speed of diagnosis, enabling early detection and intervention. Early identification of individuals at high risk for diabetes can lead to lifestyle modifications, prevention programs, or targeted interventions to delay or prevent the onset of the disease. Secondly, AI can assist healthcare professionals in making informed decisions by providing additional insights and predictions based on comprehensive data analysis. This can aid in personalized treatment plans and optimizing patient care. Lastly, AI techniques have the potential to streamline and automate certain aspects of diabetes management, reducing the burden on healthcare systems and improving patient outcomes.

However, the integration of AI techniques in diabetes classification and prediction also presents challenges. These challenges include the availability and quality of data, feature selection and extraction, class imbalance, interpretability of complex models, ethical considerations, generalization across diverse

populations, and integration with clinical workflows. Overcoming these challenges requires interdisciplinary collaboration, robust methodologies, and addressing ethical concerns to ensure the reliability, fairness, and effectiveness of AI-based approaches.

AI techniques have shown promise in improving diabetes classification and prediction, enabling early diagnosis and personalized treatment. By leveraging machine learning algorithms, these techniques can analyze diverse data sources and uncover patterns that may contribute to the onset and progression of diabetes. While there are challenges to be addressed, the potential benefits of AI in diabetes management make it a compelling area of research. Continued advancements in AI techniques, data availability, interpretability, and ethical frameworks will contribute to the development of accurate, efficient, and reliable AI models for diabetes classification and prediction.

II. LITERATURE REVIEW

N. Bhaskar et al.,[1] presented an automated medical system that can determine if a person has type 2 diabetes based on their exhaled air. Because it includes many of the same gases that are dissolved in the blood, human breath may be utilized as a diagnostic sample for the purpose of diagnosing a wide variety of disorders. Analysis of exhaled breath stands out among the many non-invasive methods of detection due to the fact that it produces more accurate forecasts and offers a number of benefits. The presence of type 2 diabetes may be determined by analyzing the level of acetone in exhaled breath, as this study demonstrates. In order to identify the illness, a brand new sensing module made up of a number of sensors has been put into operation. This module is responsible for monitoring the acetone concentration. “

G. Annuzzi et al.,[2] Diabetes type 1, often known as T1D, is an autoimmune condition that affects millions of individuals all over the globe. Patients with type 1 diabetes have a significant challenge when it comes to regulating their postprandial glucose response (PGR), which is accomplished by administering the appropriate amount of insulin bolus injections before meals. The Artificial Pancreas (AP), which combines self-administered insulin administration with continuous monitoring of blood glucose levels, is an intriguing potential treatment. However, in order to properly administer boluses, modern APs need a number of pieces of information, such as an estimate of the amount of carbohydrates consumed during the day. This is mostly attributable to the lack of understanding about the factors that determine PGR.

S. T. Himi et al.,[3] One of the technologies that is now expanding at the most rapid rate and is among the most powerful is health information technology. Because going to the doctor and getting pathological tests done may take a lot of time and money, most of the time this technology is utilized for anticipating disease and getting drugs promptly. This has encouraged a large number of researchers to make a contribution by creating new illness prediction systems or upgrading ones that already exist. This work presents a smartwatch-based prediction system called 'MedAi' for multiple diseases that use machine learning algorithms, including ischemic heart disease, hypertension, respiratory disease, hyperthyroidism, hypothyroidism, stroke, myocardial infarction, kidney failure, gallstones, diabetes, and dyslipidemia. It is composed of three primary modules: a prototype wristwatch called 'Sense O'Clock' that is fitted with eleven sensors to gather body statistics; a machine learning model that will evaluate the data and make a forecast; and a mobile application that will show the outcome of the prediction.

P. Hu et al.,[4] presented a radiomics pipeline to estimate the postoperative risk of patients having distal pancreatectomy of developing new-onset diabetes. To be more specific, we extract multiscale image features using 3D wavelet transformation, and we integrate information on patients' characteristics, body composition, and pancreatic volume as clinical features. Following this, we provide a technique for the selection and fusion of image and clinical characteristics called the multi-view subspace clustering guided feature selection method (MSCUFS). In the end, a prediction model is built using a traditional machine learning classifier. An known distal pancreatectomy cohort was used for the experiments, and the findings revealed that the SVM model with integrated imaging and EMR information provided strong discriminating, with an AUC value of 0.824.

T. Zhu et al.,[5] The availability of large amounts of data from continuous glucose monitoring (CGM), in conjunction with the most recent developments in deep learning techniques, has paved the way for a new paradigm of algorithm design for personalized blood glucose (BG) prediction in type 1 diabetes (T1D) with superior performance. This paradigm was made possible by the availability of large amounts of data from continuous glucose monitoring (CGM). However, the broad deployment of deep learning algorithms in real

clinical settings is hampered by a number of obstacles. These obstacles include uncertain prediction confidence and a lack of training data for newly diagnosed individuals with type 1 diabetes. In order to address these clinical issues, we present a revolutionary deep learning architecture called the Fast-adaptive and Confident Neural Network (FCNN). In specifically, an attention-based recurrent neural network is used in order to train representations from CGM input and then forward a weighted sum of hidden states to an evidentiary output layer.

M. Dodek et al.,[6] presented, a unique nonparametric identification approach for estimating impulse responses of the generic two-input single-output linear system is presented, with the intention of applying it to the individualization of an empirical model of type 1 diabetes. The approach that has been suggested relies on correlation functions and a developed version of the Wiener-Hopf equation for systems that have two inputs. Additionally, the program takes into consideration the stochastic features of the output measurements. In the end, this way of handling the deconvolution issue may be considered as an alternative to methods of prediction error that are often utilized. The generalized least squares approach was used in order to estimate the impulse response coefficients. This was done in order to represent nonuniform variances and nonzero covariances of the stochastic estimate of the cross-correlation functions, which ultimately led to the generation of the minimal variance estimator. ”

R. Marzouk et al.,[7] Diabetes, in all of its forms, results in unacceptably high personal, social, and economic consequences for nations of all socioeconomic levels. This research attempted to construct a web-based customized diabetes monitoring system in addition to an analytical prediction model that was based on machine learning strategies. The latter was done so that type-2 diabetes could be predicted. via constantly monitoring the patient's vital data, a patient's medical history may be compiled and made available for analysis using machine learning methods. This is accomplished via the monitoring of the patient's vital data. The use of a patient's QR card as part of a diabetic monitoring system that connects patients, physicians, and other medical professionals to the Internet of Things is one option being considered. As a result, individuals are able to offer up-to-date information on their health state (for example, data pertaining to insulin), and they are able to visit various healthcare facilities. The suggested technology would assist medical professionals in making choices based on patient data and will improve care for patients.

Z. Ye et al.,[8] The amount of glucose in the blood is an essential health indication. It is of the utmost importance to have glucose detection and monitoring techniques and equipment that are not intrusive and are simple to use, particularly for individuals who have diabetes. The testing result demonstrates that both the E-Nose system and the suggested analytical models are capable of identifying blood glucose levels with an accuracy of 90.4% and a very modest average error of 0.69 mmol/L in blood glucose concentration. This research suggests that the E-Nose system, which is equipped with machine learning, is a technique that is both effective and accurate in terms of diagnosing diseases in a non-invasive and cost-effective manner.

“M. Ismail et al.,[9] The idea of a tumor's field effect suggests that cancer is a systemic illness whose effects extend well beyond the boundaries of the visible tumor. For example, in glioblastoma (GBM), an aggressive form of brain cancer, a rise in intracranial pressure caused by the weight of the tumor often results in a herniation of the brain and poor patient outcomes. Our research is predicated on the theory that highly aggressive tumors have a propensity to grow uncontrollably, which results in pronounced biomechanical tissue deformations in the normal parenchyma. These deformations, when combined with local morphological differences in the tumor confines as seen on MRI scans, will comprehensively capture the tumor field effect. To be more specific, we provide an integrated descriptor that is based on MRI called radiomic-Deformation and Textural Heterogeneity (r-DepTH). This descriptor includes measurements of the minor perturbations in tissue deformations that are caused by the mass effect across the normal parenchyma that is around the affected area. ”

V. Felizardo et al.,[10] presented a broad range of issues for public health officials working on a variety of fronts. Some of these difficulties may be overcome with the use of computerized systems, which may pave the way not only to offer medical professionals with information into the situations of their patients everywhere and at any time, but also to enable diabetes patients to take charge of their own health management. These many approaches to illness management occur in a wide variety of forms and configurations; however, the ones that incorporate expert systems that contain specialized knowledge and make use of predictive models, feature engineering, and reasoning are now showing the greatest promise. The current state of the art in reasoning and prediction models relating to blood glucose level or

hypoglycaemia occurrences is presented in this article. The most important things that have been learned are that predictive models have space for development, namely in terms of increasing their accuracy and their capacity to foresee future occurrences over a longer period of time.

III. CHALLENGES

Diabetes classification and prediction using artificial intelligence techniques have gained significant attention in recent years. While these methods have shown promise, they also face several challenges. Here are some of the key challenges in diabetes classification and prediction based on artificial intelligence techniques:

- Data quality and availability: One of the primary challenges in diabetes classification and prediction is the quality and availability of data. Accurate and comprehensive data are crucial for training robust machine learning models. However, obtaining large, well-annotated, and diverse datasets with reliable labels can be challenging. Limited data can result in overfitting or biased models that may not generalize well to different populations.

- Feature selection and extraction: Identifying relevant features from the available data is crucial for building effective classification and prediction models. However, in diabetes, there are numerous potential features, including clinical, genetic, lifestyle, and environmental factors. Selecting the most informative and relevant features while avoiding noise and irrelevant variables is a complex task. Feature extraction techniques, such as dimensionality reduction, can help address this challenge.

- Class imbalance: Class imbalance refers to an unequal distribution of samples across different diabetes classes (e.g., normal, pre-diabetes, and diabetes). In real-world datasets, the number of instances belonging to the minority class (e.g., diabetes) is often much smaller than the majority class (e.g., normal). Class imbalance can lead to biased models that favor the majority class and have lower predictive accuracy for the minority class. Specialized techniques like oversampling, undersampling, or synthetic sample generation can be used to mitigate this issue.

- Interpretable models: While complex artificial intelligence models like deep learning can achieve high accuracy in diabetes classification and prediction, they often lack interpretability. Interpretable models are important in healthcare applications to gain trust from healthcare professionals and patients. Developing models that strike a balance between accuracy and interpretability is a challenge in this domain.

- Ethical and legal considerations: Using artificial intelligence techniques for diabetes classification and prediction raises ethical and legal considerations. Privacy, data security, informed consent, and bias are critical issues that need to be addressed. Ensuring fairness and transparency in the models and decision-making process is crucial to avoid discrimination and ensure patient trust.

- Generalization to diverse populations: Diabetes affects individuals from diverse populations, including different ethnicities, age groups, and geographical regions. It is important to develop classification and prediction models that can generalize well across these diverse populations. Models trained on one population may not perform well on another due to variations in genetic, lifestyle, and environmental factors.

- Integration with clinical workflows: For artificial intelligence techniques to be effectively used in clinical practice, they need to be seamlessly integrated into existing clinical workflows. This requires bridging the gap between research and practical implementation. The development of user-friendly interfaces and tools that can be easily used by healthcare professionals is a challenge that needs to be addressed.

These challenges require collaboration between researchers, clinicians, and data scientists. Further advancements in data collection, feature engineering, model development, interpretability, and ethical frameworks will contribute to the successful integration of artificial intelligence techniques in diabetes classification and prediction.

IV. CONCLUSION

Diabetes mellitus is a condition that lasts for a long time and poses a significant threat to people's health all over the globe. The World Diabetes Federation estimates that there are presently 246 million individuals in the globe who suffer from diabetes, but that figure is predicted to more than double by the year 2025, reaching 380 million people. In addition, diabetes-related complications are responsible for 3.8 million fatalities worldwide each year. It has been shown that early detection of those at risk may avoid or postpone the onset of 80 percent of the problems that are associated with type 2 diabetes. Techniques from the field of machine learning may be used to investigate and make predictions about the diabetes categorization. The purpose of this study is to provide an overview of machine learning approaches for the categorization and prediction of diabetes. In the near future, we will build an effective machine learning or deep learning approach to create a more accurate prediction model for the diagnosis of diabetic disorders.

REFERENCES

1. N. Bhaskar, V. Bairagi, E. Boonchieng and M. V. Munot, "Automated Detection of Diabetes From Exhaled Human Breath Using Deep Hybrid Architecture," in IEEE Access, vol. 11, pp. 51712-51722, 2023, doi: 10.1109/ACCESS.2023.3278278.
2. 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.
3. 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.
4. P. Hu et al., "Prediction of New-Onset Diabetes After Pancreatectomy With Subspace Clustering Based Multi-View Feature Selection," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 3, pp. 1588-1599, March 2023, doi: 10.1109/JBHI.2022.3233402.
5. T. Zhu, K. Li, P. Herrero and P. Georgiou, "Personalized Blood Glucose Prediction for Type 1 Diabetes Using Evidential Deep Learning and Meta-Learning," in IEEE Transactions on Biomedical Engineering, vol. 70, no. 1, pp. 193-204, Jan. 2023, doi: 10.1109/TBME.2022.3187703.
6. M. Dodek, E. Miklovičová and M. Tárník, "Correlation Method for Identification of a Nonparametric Model of Type 1 Diabetes," in IEEE Access, vol. 10, pp. 106369-106385, 2022, doi: 10.1109/ACCESS.2022.3212435.
7. R. Marzouk, A. S. Alluhaidan and S. A. El_Rahman, "An Analytical Predictive Models and Secure Web-Based Personalized Diabetes Monitoring System," in IEEE Access, vol. 10, pp. 105657-105673, 2022, doi: 10.1109/ACCESS.2022.3211264.
8. Z. Ye, J. Wang, H. Hua, X. Zhou and Q. Li, "Precise Detection and Quantitative Prediction of Blood Glucose Level With an Electronic Nose System," in IEEE Sensors Journal, vol. 22, no. 13, pp. 12452-12459, 1 July1, 2022, doi: 10.1109/JSEN.2022.3178996.
9. M. Ismail et al., "Radiomic Deformation and Textural Heterogeneity (R-DepTH) Descriptor to Characterize Tumor Field Effect: Application to Survival Prediction in Glioblastoma," in IEEE Transactions on Medical Imaging, vol. 41, no. 7, pp. 1764-1777, July 2022, doi: 10.1109/TMI.2022.3148780.
10. V. Felizardo, D. Machado, N. M. Garcia, N. Pombo and P. Brandão, "Hypoglycaemia Prediction Models With Auto Explanation," in IEEE Access, vol. 10, pp. 57930-57941, 2022, doi: 10.1109/ACCESS.2021.3117340.