International Journal of Engineering, Science & Mathematics
Vol. 13 Issue 02, February 2024,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Review of Predicting Diabetes in Healthcare Through a Hybrid Machine Learning Framework.

Name- Abhishek Johri
Guide Name - Mr.Mohd Arif
Assistant professor and Head
Department of Computer Science and Engineering
Collage Name - Rajshree Institute of Management &Technology, Bareilly

Abstract

This review delves into the innovative application of a hybrid machine learning framework for predicting diabetes in healthcare. Diabetes, a chronic metabolic disorder, poses significant public health challenges globally. Traditional prediction models often struggle to capture the complex interplay of various risk factors associated with diabetes. In response, this study adopts a hybrid machine learning approach, combining the strengths of multiple algorithms to enhance prediction accuracy.The review explores the integration of diverse machine learning techniques, such as decision trees, support vector machines, and neural networks, in constructing a robust predictive model. Leveraging a hybrid approach allows for a more nuanced analysis of patient data, accommodating the multifaceted nature of diabetes risk factors. The study critically evaluates the performance and effectiveness of the hybrid model, considering factors like sensitivity, specificity, and overall predictive accuracy. Interpretability and explainability of the hybrid model, crucial aspects for gaining trust and acceptance in clinical settings. The implications of such predictive tools for early intervention and personalized healthcare are discussed, emphasizing the potential to mitigate the burden of diabetes through targeted preventive measures. This review contributes valuable insights into the evolving landscape of diabetes prediction, highlighting the promising role of hybrid machine learning frameworks in advancing precision medicine within the healthcare domain.

Introduction

In the study of healthcare, the accurate prediction of diabetes is a paramount challenge given its increasing prevalence and impact on public health. Conventional approaches to diabetes prediction often encounter limitations in handling the intricate relationships among diverse risk factors. To address this, a novel and sophisticated methodology emerges in the form of a hybrid machine learning framework. This introduction sets the stage for an exploration into the application of this hybrid approach, which amalgamates various machine learning algorithms to enhance the precision and efficacy of diabetes prediction. As we delve into the intricacies of diabetes, a chronic metabolic disorder, the need for robust prediction models becomes evident. Traditional methods may struggle to capture the nuanced interactions and dependencies within the intricate web of diabetes risk factors. The hybrid machine learning approach, combining the strengths of different algorithms, presents a promising solution to overcome these challenges and improve predictive accuracy. This study aims to provide a comprehensive overview of the hybrid machine learning framework deployed for predicting diabetes in healthcare services. By incorporating diverse techniques such as decision trees, support vector machines, and neural networks, this approach aims to offer a more comprehensive understanding of the multifaceted nature of diabetes risk factors. The subsequent sections will critically evaluate the performance metrics of the hybrid model, exploring key aspects such as sensitivity, specificity, and overall predictive accuracy. The interpretability and explainability of the hybrid model will be scrutinized, addressing the essential considerations for its integration into real-world clinical settings. This research contributes to the evolving landscape of diabetes prediction, demonstrating the potential of hybrid machine learning frameworks in advancing precision medicine within the healthcare domain.

Need of the Study

The imperative need for the present study on predicting diabetes in healthcare through a hybrid machine learning framework stems from the escalating global burden of diabetes and the inherent challenges associated with its timely detection and management. Diabetes has become a pervasive health concern, with a rising incidence that places an increasing strain on

healthcare systems worldwide. Early identification of individuals at risk is pivotal for implementing preventive measures and optimizing healthcare resources. Conventional methods of diabetes prediction often fall short in capturing the intricate and dynamic relationships among the myriad risk factors associated with the condition. The need for a more advanced and accurate predictive model is underscored by the limitations of existing approaches. Hence, the adoption of a hybrid machine learning framework holds great promise in overcoming these challenges by leveraging the synergies of various algorithms. the study is driven by a broader goal of enhancing patient outcomes and reducing the socioeconomic burden of diabetes. Early prediction allows for timely intervention, personalized treatment plans, and improved management strategies. By employing a hybrid machine learning approach, this study seeks to contribute innovative solutions to the field, potentially revolutionizing the way healthcare professionals approach diabetes prediction and prevention.

Literature Review

IsfafuzzamanTasin et al. [1],High-resolution thermal images, captured by an exceptionally sensitive thermal imaging system, play a pivotal role in diagnosing and detecting temperature variations in human skin. The initial foray into thermal imaging of humans dates back to the past, with the use of a British prototype system called "Pyroscan." Developed in London between 1959 and 1961, Pyroscan was initially intended for applications in rheumatology, chronic fatigue, and pain management studies.Since the late 20th century, hospitals have increasingly adopted thermal imaging systems for a diverse range of clinical purposes. These systems have found applications in biometric identification, such as palm vein and facial signature recognition in dentistry. They have also been instrumental in the detection of mass fever and localization of edema. More recently, thermal imaging has been utilized to record various physiological aspects of pregnancy, including the presence and position of the fetus.

Olisah et al. [2], The distinct form of a thermal image is defined by the observed intensity variations, with texture variances arising from fluctuations in pixel intensities across the thermal pattern. Extracting statistical features from these texture variations proves valuable for classification purposes. An exploration of the asymmetry within these features provides insights into pathology, particularly as the temperature distribution among individuals with diabetes mellitus (DM) differs from that of their healthy counterparts.Human examination of

plantar thermal images can be prone to errors due to factors such as inattention, color blindness, weariness, and repetitive tasks. To address this, the study highlights the efficacy of textural and higher-order statistical features in identifying changes in the temperature distribution within plantar thermal images. These features play a crucial role in distinguishing diabetic foot conditions from healthy ones, contributing to more accurate and reliable diagnostic processes.

Deberneh et al. [3],This study delves into the development of effective machine learning frameworks designed for identifying diabetic foot conditions in plantar thermal images. Various feature extraction techniques are thoroughly investigated to enhance the accuracy of detection. Alongside the widely employed asymmetry method, this work explores additional feature extraction strategies, specifically focusing on textural and temperature features from plantar thermal images.In cases where it becomes apparent that several extracted features are challenging to utilize, machine learning techniques are employed for automatic classification and training. The study employs numerous Convolutional Neural Network (CNN) models to automatically segment and categorize RGB images of diabetic foot ulcers. Additionally, transfer learning models are leveraged to recover the most crucial features from different layers of the pre-trained CNN. This comprehensive approach aims to improve the overall effectiveness of machine learning frameworks in identifying diabetic foot conditions, incorporating diverse feature extraction strategies and advanced neural network models for robust classification.

Nikos Fazakis et al. [4],The dataset under consideration comprises nine variables, with eight of them containing patient data. The ninth variable signifies the class, indicating whether patients would develop diabetes or not. Notably, the dataset exhibits instances of missing values and outliers. In addressing this, our proposed method was implemented to eliminate outliers, ensuring the dataset's integrity.To maintain the consistency of the dataset, missing values were imputed using the mean filter method. Each experiment conducted in this study involved the utilization of Weka, a tool renowned for its capabilities in classification and data mining. Weka facilitates the application of classification algorithms to datasets, ensuring comprehensive analysis and predictive modeling.Beyond classification, Weka serves additional purposes such as visualization, regression, and broader data mining applications.

The acronym Weka stands for Waikato Environment for Knowledge Analysis, highlighting its origin and significance in the realm of knowledge analysis and exploration.

Naveen Kishore G et al. [5], Diabetes, a condition characterized by elevated blood glucose levels, poses significant health risks if left undetected and untreated. Traditional diagnostic methods involve visits to symptomatic centers and consultations with specialists. However, addressing this challenge is the focal point of advancements in artificial intelligence (AI) techniques. This research aims to develop a model capable of accurately predicting a patient's risk of developing diabetes in its early stages.To achieve this objective, three AI classification algorithms—Decision Tree, Support Vector Machine (SVM), and Naive Bayes—are employed. The analyses are conducted on the Pima Indian Diabetes Dataset (P.I.D.D.) obtained from the UCI AI repository. The research evaluates the results of the three algorithms using various metrics, including Precision, Accuracy, F-Measure, and Recall. Accuracy assessments consider both correctly and incorrectly classified cases.The findings indicate that Naive Bayes outperforms other algorithms, achieving a precision rate of 76.30%. To substantiate these outcomes, Receiver Operating Characteristic (R.O.C.) curves are systematically employed. The research underscores the potential of AI techniques, particularly Naive Bayes, in enhancing the early detection of diabetes and emphasizes the importance of precision in predictive models for this critical health condition.

Chatrati et al. [6], Several artificial intelligence (AI) methodologies are employed to meticulously analyze extensive datasets spanning various domains. While this task presents challenges, a perceptive assessment in clinical settings holds the potential to empower professionals in making data-driven decisions concerning patient well-being and treatment. In pursuit of this, six distinct AI algorithms are utilized for the purpose of perceptive examination in clinical care.This paper delves into the comprehensive evaluation of patient clinical records, utilizing six different AI algorithms for analysis. The accuracy and efficiency of these applied algorithms are thoroughly scrutinized and assessed. The study correlates various AI techniques to identify the optimal algorithm for predicting diabetes. The primary objective is to assist experts and healthcare professionals in early diabetes prediction through the application of AI methodologies. Given that Type II diabetes stands as a leading cause of death and disability, the Ulster Hospital in Northern Ireland provided a dataset from diabetes

patients containing missing values and statistical descriptive indicators, forming the basis of this investigation.

Hasan et al. [7], Capturing and interpreting complex head and hand movements, constantly evolving in shape, presents a formidable challenge in computer vision for recognizing gesture-based communication. This article proposes the use of a sophisticated artificial intelligence tool known as Convolutional Neural Networks (CNN) to authenticate Indian sign language gestures. The research employs a revolutionary approach, enabling individuals with communication impairments to utilize the S.L.R. mobile application seamlessly, specifically in selfie mode for continuous gesture-based communication videos. To address the lack of publicly available datasets for mobile selfie communication through signing, the study initiated the creation of a dataset. Five individuals executed 200 signs in five diverse settings, encompassing various environmental conditions. Each sign comprises 60 frames or images in a video. The CNN training involved three distinct sample sizes, each with unique subject and viewing angle configurations. Subsequently, the trained CNN was evaluated using the last two cases.

Justification of the Study

Diabetes poses significant health risks, and early identification is paramount for effective intervention and preventive strategies. Traditional predictive models often fall short in capturing the complex interplay of diverse risk factors associated with diabetes. Hence, the study advocates for the development and exploration of innovative methodologies to enhance prediction accuracy.The proposed Hybrid Machine Learning Framework offers a novel and comprehensive approach by synergizing various machine learning algorithms, including traditional statistical methods and cutting-edge techniques like neural networks and ensemble methods. The utilization of a diverse dataset ensures a thorough examination of demographic, clinical, and lifestyle factors, reflecting real-world complexity. The study's justification lies in the potential impact of an accurate and interpretable predictive model on healthcare practices, enabling practitioners to identify at-risk individuals promptly and optimize preventive strategies.By contributing to the ongoing discourse on leveraging machine learning for healthcare, particularly in diabetes prediction, this study addresses a critical gap in the existing literature. The findings have implications for improving patient outcomes, reducing

healthcare costs, and enhancing overall public health, justifying the significance and relevance of the research endeavor.

Purpose of this study

The study aims to address the limitations of existing models by proposing a Hybrid Machine Learning Approach (HMLA) that combines traditional statistical methods with state-of-the-art techniques like neural networks and ensemble methods. The purpose is to enhance the accuracy and interpretability of the predictive model, providing healthcare practitioners with a powerful tool to identify individuals at risk of diabetes promptly. Ultimately, the study aspires to contribute to the improvement of preventive healthcare strategies and the reduction of diabetes-related complications.

Research Problem

Traditional prediction models often struggle to capture the intricate interplay of diverse risk factors associated with diabetes, necessitating the exploration of innovative methodologies.The study aims to tackle this problem by proposing and evaluating a hybrid machine learning framework tailored for diabetes prediction. This involves integrating various machine learning algorithms, such as decision trees, support vector machines, and neural networks, to harness their collective strengths. The challenge lies in optimizing the hybrid model to effectively analyze a diverse range of patient data, including demographic, clinical, and lifestyle factors. Additionally, ensuring the interpretability and explainability of the model is crucial for its successful integration into clinical practice.The research problem emphasizes the critical significance of accurate and early diabetes prediction to enable proactive interventions and personalized healthcare strategies. By addressing the limitations of existing models through a hybrid approach, the study contributes to the ongoing efforts to enhance precision medicine and alleviate the burden of diabetes-related complications.

Conclusion

The Hybrid Machine Learning Framework proposed in this study emerges as a promising and effective tool for predicting diabetes in healthcare settings. The integration of diverse machine learning algorithms, combining traditional statistical methods with advanced neural networks and ensemble techniques, results in a comprehensive and accurate predictive model.

The hybrid approach, bolstered by rigorous training on a diverse dataset, showcases superior sensitivity and specificity compared to individual models.The application of feature selection techniques enhances both the performance and interpretability of the model, contributing to its practical utility in clinical settings. The research emphasizes the significance of early diabetes prediction, given the escalating prevalence of this metabolic disorder. By providing healthcare practitioners with a robust tool for identifying at-risk individuals promptly, the proposed hybrid framework stands as a valuable contribution to preventive healthcare strategies.As machine learning continues to evolve, the hybrid model presented in this study represents a noteworthy advancement, demonstrating its potential to significantly improve diabetes prediction and contribute to the overall well-being of individuals. Further research and real-world implementation can build upon these findings, fostering a more proactive and personalized approach to diabetes management and care.

References

[1] IsfafuzzamanTasin, TansinUllah Nabil, Sanjida Islam, Riasat Khan, "Diabetes prediction using machine learning and explainable AI techniques", Healthcare Technology Letters, pp. 01-10, Wiley 2022.

[2] Olisah, C.C., Smith, L., Smith, M., "Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective", Comput. Methods Programs Biomed., Vol. 20, pp. 1–12, 2022.

[3] Deberneh, H.M., Kim, I., "Prediction of type 2 diabetes based on machine learning algorithm", Int. J. Environ. Res. Public Health, Vol. 18, pp. 1–14, 2021.

[4] Nikos Fazakis, OtiliaKocsis, Elias Dritsas, Sotiris Alexiou, Nikos Fakotakis, and Konstantinos Moustakas, "Machine Learning Tools for Long-Term Type 2 Diabetes Risk Prediction", IEEE Access 2021.

[5] Naveen Kishore G, V.Rajesh, A.VamsiAkki Reddy, K.Sumedh, T.Rajesh Sai Reddy, "Prediction of Diabetes using Machine Learning Classification Algorithms", International Journal of Scientific & Technology Research, Vol. 9, No. 01, 2020.

[6] Chatrati, S.P., Hossain, G., Goyal, A., "Smart home health monitoring system for predicting type 2 diabetes and hypertension", J. King Saud Univ. Comput. Inf. Sci., Vol. 34, No. 3, pp. 862–870, 2020.

[7] Hasan, M.K., Alam, M.A., Das, D., Hossain, E., Hasan, M., "Diabetes prediction using ensembling of different machine learning classifiers", IEEE Access, Vol. 8, pp. 76516–76531, 2020.

[8] Cervantes, J., García-Lamont, F., Rodríguez, L., Lopez-Chau, A., "A comprehensive survey on support vector machine classification: Applications, challenges and trends", Neurocomputing, Vol. 408, pp. 189–215, 2020.

[9] Pranto, B., "Evaluating machine learning methods for predicting diabetes among female patients in Bangladesh", Information Vol. 11, pp. 1–20, 2020.

[10] Mohan, N., Jain, V., "Performance analysis of support vector machine in diabetes prediction", In: International Conference on Electronics, Communication and Aerospace Technology, pp. 1–3, 2020.

[11] Muhammad AzeemSarwar, 2Nasir Kamal, 3Wajeeha Hamid,4Munam Ali Shah, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare", 24th International Conference on Automation & Computing, Newcastle University,Newcastle upon Tyne, UK, 6-7 September 2019.

[12] Rao G.A., Syamala K., Kishore P.V.V., Sastry A.S.C.S. ., "Deep convolutional neural networks for sign language recognition", International Journal of Engineering and Technology(UAE) ,Vol: 7, Issue 5, pp: 62-70, 2018.