IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

An Overview of Development of a Novel Artificial Intelligence and Machine Learning Based Model for Diabetes Healthcare Sector

[1*]Shashi Bhushan Singh, [2]Dr. Arjun Singh

[1]Research Scholar, Department of Computer Science, Magadh University, Bodh Gaya, Bihar, India

[2]Associate Professor, Department of Mathematics, K.S.M. College, Aurangabad, Bihar, India

Abstract:

The world has now moved into the era of Data Science, and tools that have traditionally been applied to other domains are now being considered in healthcare. Diabetes is a metabolic disorder characterized by an excessive level of glucose in the blood due to either insulin resistance or insulin deficiency. This research focuses on data science concepts from AI, ML, DM, and DL. It provides several approaches to utilize these techniques for analyzing clinical data, forecasting, monitoring, and understanding the management of Diabetes Mellitus. Although using AI in healthcare is a new research field, recent studies have explored the application of AI in healthcare for treatment effectiveness, healthcare management, fraud and abuse detection, patient behavior analysis, customer relationship management, and more. This research aims to propose an AI model for healthcare, identifying the qualitative characteristics of patients, preparing the dataset for mining, and detecting diabetes using AI to improve the treatment effectiveness, which will save time and enhance it for faster treatment and diagnosis. The primary objective of this thesis is to design and develop a novel automated and improved machine learning classification model for the classification of diabetes mellitus and designing a deep learning model for the classification of Diabetic Retinopathy, representing the performance ratio as the accuracy ratio.

Keywords: Machine Learning, Diabetic Retinopathy, Artificial Intelligence, Data Sources

1. Introduction:

The medical field faces new challenges, such as new diseases, cost, new therapeutics, and rapid decisions. Since medical decision-making requires the utmost accuracy of diagnosis, it is a tedious, demanding, difficult, and challenging task for physicians. An automated system that assists in disease diagnosis, prognosis, and treatment will benefit the clinical practice.

The world has now moved into the era of Data Science, and tools that have traditionally been applied to other domains are now being considered in healthcare. These data sources include traditional systematic data entry, reports, claims data, survey data, and data obtained from biometric monitoring, among others. Given the wealth of data being made available, researchers need to find the best data science approach, deep learning, and AI techniques available for applying to these datasets. Diabetes is a metabolic disorder characterized by an excessive level of glucose in the blood due to either insulin resistance or insulin deficiency. It is one of the most severe global public health issues of modern technology. The IDF has stated that in 2011, around 350+ million people had diabetes, which is likely to reach over 550 million by 2030. Out of a hundred, 80% of diabetes-related deaths occur in underdeveloped or developing countries like India. Various researchers have demonstrated that diabetes is becoming epidemic in both developing and advanced global regions. This has demanded a robust, rapid, reliable, and strong technique that can utilize science and technology to manage diabetes.

Diabetes mellitus also imposes a substantial economic burden on the overall healthcare system and the global budget. This burden can be determined through medical charges, which are indirect as well as direct charges related to productivity loss, early mortality, and the adverse effect of diabetes on the country's GDP. Subject to cost estimates from a recent systematic review, the direct annual budget of diabetes worldwide is close to about US$ 830 billion.

The growth in data volume causes overwhelming challenges in mining the necessary data for research. To adapt to this requirement, healthcare informatics may use the advancements made in the new interdisciplinary field of data discovery in datasets, for example, in KDD. This involves statistical, machine learning, and pattern recognition techniques to understand the process to assist the evaluation of data and the discovery of regularities that are encoded within the data. This research is planned to propose an AI model for healthcare, identifying the qualitative characteristics of patients, preparing the dataset for mining, and detecting diabetes using AI to improve the treatment effectiveness, which will save time and enhance it for faster treatment and diagnosis. Figure 1 shows the basic structure of AI model for healthcare.

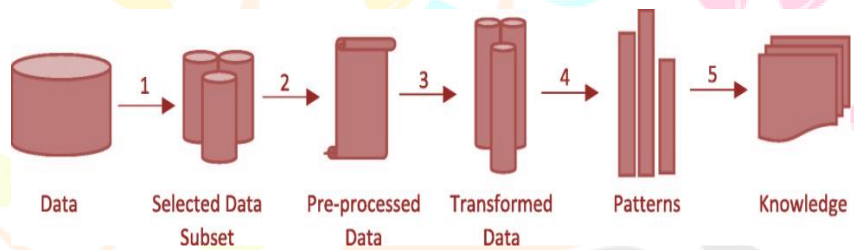


Figure 1: Basic structure of AI model

In the healthcare domain, mining data is becoming increasingly popular, where the large volumes of data generated, created by clinical dealings, are much diverse and large for managing and analyzing using old approaches. Data mining can find patterns, trends, and tendencies in massive amounts of the diverse dataset which improve continuously. The healthcare industry will receive a lot of help and benefit from data mining applications.

2. Prospective and Perspective of Work:

The prospective of this research work lies in the growing need for intelligent, accurate, and efficient tools to address the rising global burden of diabetes. As the world has entered the era of data-driven healthcare, there is immense potential in leveraging Artificial Intelligence (AI) and Machine Learning (ML) techniques to transform diabetes management.

The key perspectives of this work are:

a. Bridging the gap: There is a significant gap in the research on comparative analysis of different AI and ML techniques for diabetes healthcare applications. This work aims to provide a comprehensive evaluation of the strengths and suitability of various algorithms to address specific challenges in diabetes diagnosis, prognosis, and management.

b. Enhancing diagnosis accuracy: By incorporating fuzzy qualitative attributes of patients and developing novel ML models, this research intends to improve the accuracy and reliability of diabetes diagnosis, going beyond the conventional symptom-based approaches. In figure 2 we can see artificial intelligence in obtaining and analysing patient data for therapeutic personalisation. It is used not only for diabetes management but also used in whole treatment process.

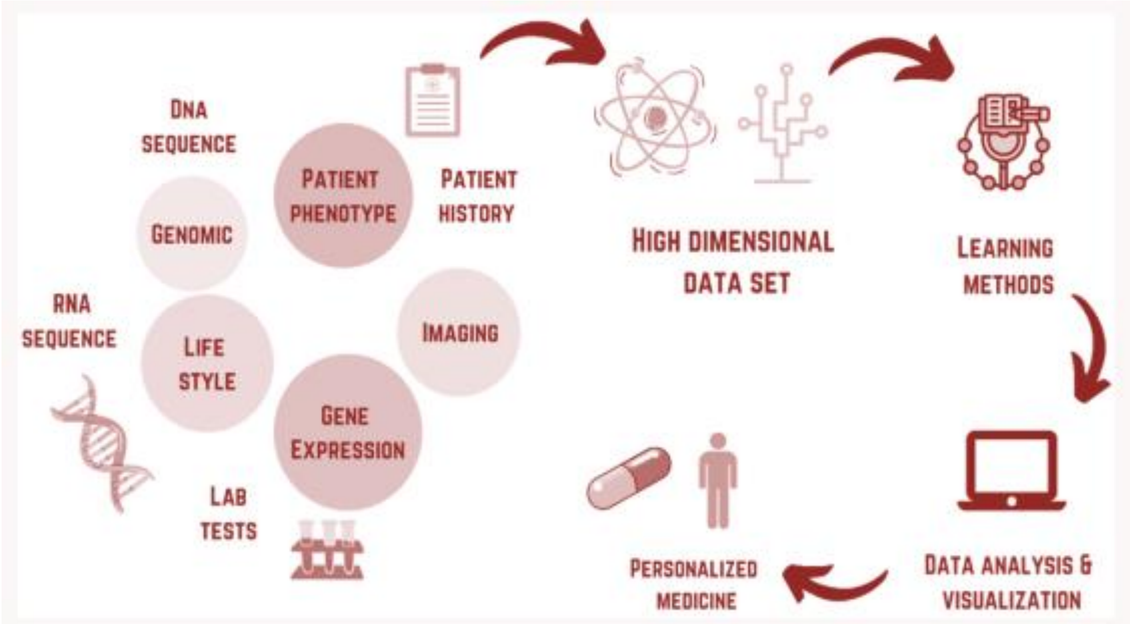


Figure 2: Artificial intelligence in obtaining and analysing patient data.

c. Personalized disease management: The integration of deep learning techniques for diabetic retinopathy detection will enable personalized monitoring and tailored treatment plans, catering to the individual needs of patients.

d. Broader applicability: The proposed AI-driven framework aims to be versatile and adaptable, addressing the unique healthcare challenges faced in developed, developing, and underdeveloped nations, thus showcasing its universal applicability.

e. Translating data to actionable insights: Leveraging the advancements in data mining and knowledge discovery, this research will focus on transforming the large volumes of healthcare data into meaningful and actionable insights for improved diabetes management.

f. Expediting clinical decision-making: The development of an expert system for diabetes diagnosis and treatment recommendations has the potential to significantly reduce the time and effort required by clinicians, thereby enhancing the efficiency of the healthcare system.

By addressing these key perspectives, this research work aspires to make a significant contribution to the field of intelligent diabetes care, empowering healthcare professionals and patients alike with advanced AI-powered tools for better disease management and improved patient outcomes.

3. Literature Review

The topic of diabetes healthcare has seen significant research efforts in the domain of Artificial Intelligence (AI) and Machine Learning (ML) applications. Multiple research have investigated the capacity of these sophisticated approaches to improve the diagnosis, prognosis, and treatment of diseases. The VP-Expert Model, established by Tawfik Saeed Zeki et al., offers guidance on indications, diagnosis, and basic treatment recommendations for individuals with diabetes. The rule-based expert fuzzy system for the diagnosis of diabetes was suggested by Dilip Kumar Choubey et al. An Android-based mobile application expert system was developed by Anindito Yoga Pratama et al. with the purpose of assessing the risk of diabetes. A fuzzy expert system was used by M. Kalpana et al. to diagnose diabetes. In their study, Danijela TADIC et al. proposed a fuzzy model to assess and choose the most suitable treatment interventions for individuals with type 2 diabetes. The expert system for diabetes diagnosis was improved by Hanslal Prajapati et al. by the use of a Fuzzy Inference System. A cloud computing-based expert system was created by Abdullah Al-Malaise Al-Ghamdi et al. to enhance the efficacy of diabetic therapy. The Prolog-based expert system for the management of type 2 diabetes was built by Ibrahim M. Ahmed et al. Seyedeh Talayeh Tabibi et al. created an expert system with the purpose of offering guidance for the management of diabetes. The expert system developed by Meysam Rahmani Katigari et al. was designed to identify and diagnose the severity degree of diabetic neuropathy. In their study, Dhivya et al. used a Fuzzy Expert System to facilitate the timely identification of the illness and the implementation of treatment protocols. The expert system for the diagnosis of gestational diabetes mellitus was modelled by K. Vijaya Lakshmi et al. In their study, Onuiri Ernest E et al. devised a web-based application system designed to facilitate the recording of glucose levels and the quantification of insulin dosage. The Fuzzy Expert System was used by R. Radha et al. to establish the relationships for prediction. The comparative study of Mamdani type and Sugeno type Fuzzy Systems for diabetes prediction was undertaken by Vishali Bhandari et al. Nonso Nnamoko and colleagues developed a methodology based on Fuzzy Logic for the treatment of type 2 diabetic mellitus (T2DM). The study conducted by Srideicanai Nagarajan et al. used clustering and classification methodologies within the field of data mining to employ the K-means algorithm for the diagnosis of type-1, type-2, and gestational diabetes.

The literature review underscores the notable advancements achieved in the use of artificial intelligence (AI) and machine learning (ML) methodologies within the realm of diabetic treatment. Nevertheless, there is still a need for a thorough comparative examination of the efficacy and appropriateness of different algorithms, together with the incorporation of sophisticated deep learning methodologies for the purpose of personalised illness treatment. The objective of this study is to fill these existing knowledge gaps and provide a valuable contribution to the advancement of a new artificial intelligence (AI)-based framework for precise diagnosis, prognosis, and tailored treatment of diabetes within the healthcare industry.

4. Hypothesis:

The hypothesis of this research work is grounded in the extensive review of the existing literature and the identified gaps in the current approaches to diabetes management using AI and Machine Learning techniques.

a. Superiority of Advanced ML Algorithms: The research hypothesis postulates that advanced Machine Learning algorithms, such as Decision Tree and Support Vector Machine, which have been widely utilized by researchers in their healthcare predictive analysis, exhibit superior performance in terms of accuracy compared to conventional techniques. These state-of-the-art classification algorithms are expected to provide enhanced diagnostic capabilities and risk prediction for diabetes.

b. Potential of Deep Learning for Personalized Care: The research further hypothesizes that the integration of deep learning techniques, particularly for the detection of diabetic retinopathy, will enable personalized monitoring and tailored treatment plans, addressing the individual needs of patients. The deep learning models are expected to provide more accurate and reliable automated screening for diabetes-related complications, empowering clinicians to deliver personalized care.

c. Synergistic Integration of AI and ML: The core hypothesis of this research is that the synergistic integration of Artificial Intelligence and Machine Learning approaches can serve as a powerful vehicle for translating large volumes of healthcare data into actionable insights for improved diabetes management. By leveraging the strengths of various AI and ML techniques, the proposed framework is expected to outperform traditional methods in terms of accuracy, efficiency, and universality.

d. Addressing Contextual Challenges: The research hypothesis also considers the contextual challenges faced in developed, developing, and underdeveloped nations with respect to diabetes management. It is hypothesized that the proposed AI-driven framework will be designed to be adaptable and versatile, addressing the unique healthcare needs and resource constraints of different regions, thus showcasing its broader applicability.

e. Enhancing Clinical Decision-Making: The research work hypothesizes that the development of an expert system for diabetes diagnosis and treatment recommendations will significantly reduce the time and effort required by clinicians, thereby enhancing the overall efficiency of the healthcare system. This expert system is expected to provide timely and reliable decision support to healthcare professionals, improving patient outcomes.

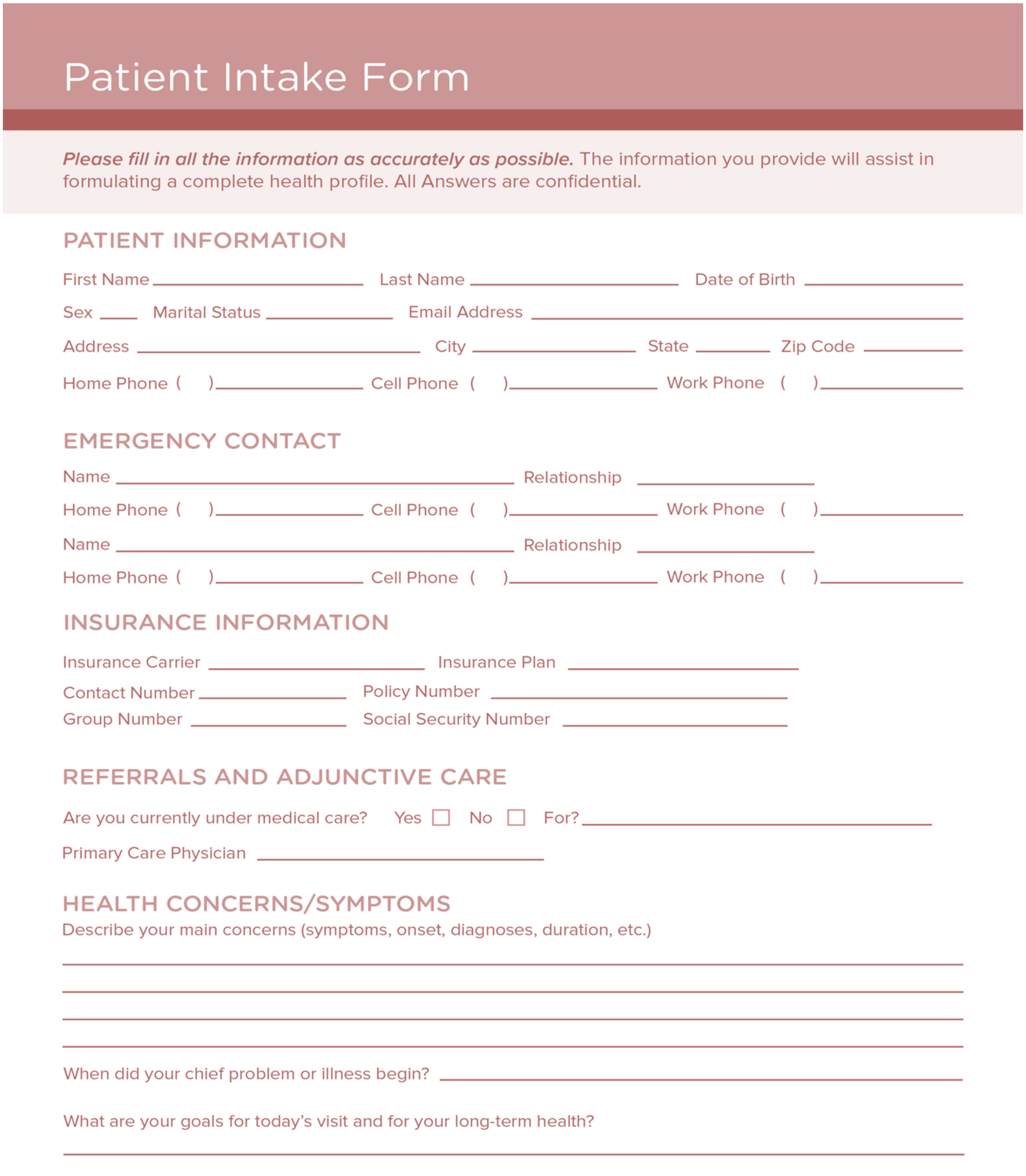
By systematically testing these hypotheses through rigorous experimentation and evaluation, this research aims to contribute to the advancement of intelligent diabetes care, paving the way for more accurate, personalized, and efficient management of this global health challenge.

5. Plan of work (Materials and methods):

5.1 Data Collection:

Data Collection Approach:

The data will be collected from various hospitals and healthcare facilities across India using a structured questionnaire. The questionnaire will be designed to gather comprehensive information about the patients, including demographic details, medical history, symptoms, and laboratory test results. In figure 3, it is shown the patient intake questionnaire, which generally used by the hospitals.


Patient Intake Form

Please fill in all the information as accurately as possible. The information you provide will assist in formulating a complete health profile. All Answers are confidential.

PATIENT INFORMATION

First Name ____ Last Name ____ Date of Birth ____

Sex ____ Marital Status ____ Email Address ____

Address ____ City ____ State ____ Zip Code ____

Home Phone ( ) ____ Cell Phone ( ) ____ Work Phone ( ) ____

EMERGENCY CONTACT

Name ____ Relationship ____

Home Phone ( ) ____ Cell Phone ( ) ____ Work Phone ( ) ____

Name ____ Relationship ____

Home Phone ( ) ____ Cell Phone ( ) ____ Work Phone ( ) ____

INSURANCE INFORMATION

Insurance Carrier ____ Insurance Plan ____

Contact Number ____ Policy Number ____

Group Number ____ Social Security Number ____

REFERRALS AND ADJUNCTIVE CARE

Are you currently under medical care? Yes ☐ No ☐ For? ____

Primary Care Physician ____

HEALTH CONCERNS/SYMPTOMS

Describe your main concerns (symptoms, onset, diagnoses, duration, etc.)

When did your chief problem or illness begin? ____

What are your goals for today's visit and for your long-term health?

Figure 3: Patient intake questionnaire

The target respondents will be physicians and healthcare professionals with a minimum of 3 years of experience in managing diabetic patients.

A non-probability quota sampling method will be employed to ensure the representation of different regions, socioeconomic backgrounds, and healthcare settings.

Data Sources:

Primary data will be obtained through the structured questionnaire administered to the healthcare professionals. Secondary data may be utilized from publicly available databases and repositories, such as the Indian Council of Medical Research (ICMR) and the International Diabetes Federation (IDF) databases, to supplement the primary data collection. Additionally, anonymized medical records and clinical data may be acquired from the participating healthcare facilities, subject to necessary ethical approvals and data privacy considerations.

Data Preprocessing and Curation:

The collected data will undergo extensive preprocessing, including handling missing values, outlier detection, and data normalization to ensure data quality and consistency. The dataset will be carefully curated, with the removal of any personally identifiable information to maintain patient confidentiality. Relevant feature engineering techniques will be applied to extract meaningful attributes from the raw data, which will be used for the subsequent analysis and model development.

5.2 Data Analysis:

Descriptive Analysis of the Dataset:

The research team will perform a comprehensive descriptive analysis of the collected dataset, including summary statistics, data visualization, and identifying key patterns and trends. This initial analysis will provide valuable insights into the characteristics of the diabetic patient population, the distribution of symptoms, and the overall data quality.

Designing the Knowledge Base for Machine Learning Model:

Based on the descriptive analysis and the reviewed literature, the researchers will design a robust knowledge base to guide the development of the Machine Learning models. This knowledge base will incorporate the identified key symptoms, risk factors, and other relevant attributes that influence the diagnosis and management of diabetes.

Utilizing the Dataset as a Knowledge Base:

The curated dataset will be utilized as the knowledge base for the proposed Machine Learning models, ensuring that the models are trained on reliable and representative data. The dataset will be split into training, validation, and testing sets to enable the development, fine-tuning, and unbiased evaluation of the ML models.

Developing Machine Learning Models:

The researchers will develop a novel Machine Learning model using Python and the Decision Tree algorithm for the classification of diabetes based on the binary values of symptoms. To enhance the classification accuracy, the model will be further refined by incorporating certainty factors and other advanced techniques.

Designing Deep Learning Models for Diabetic Retinopathy:

Convolutional Neural Network (CNN) approaches will be employed to develop deep learning models for the classification of diabetic retinopathy. The research team will utilize the Fast AI library and Google Colaboratory tools to build and train the CNN models (in figure 4) on image data related to diabetic retinopathy.

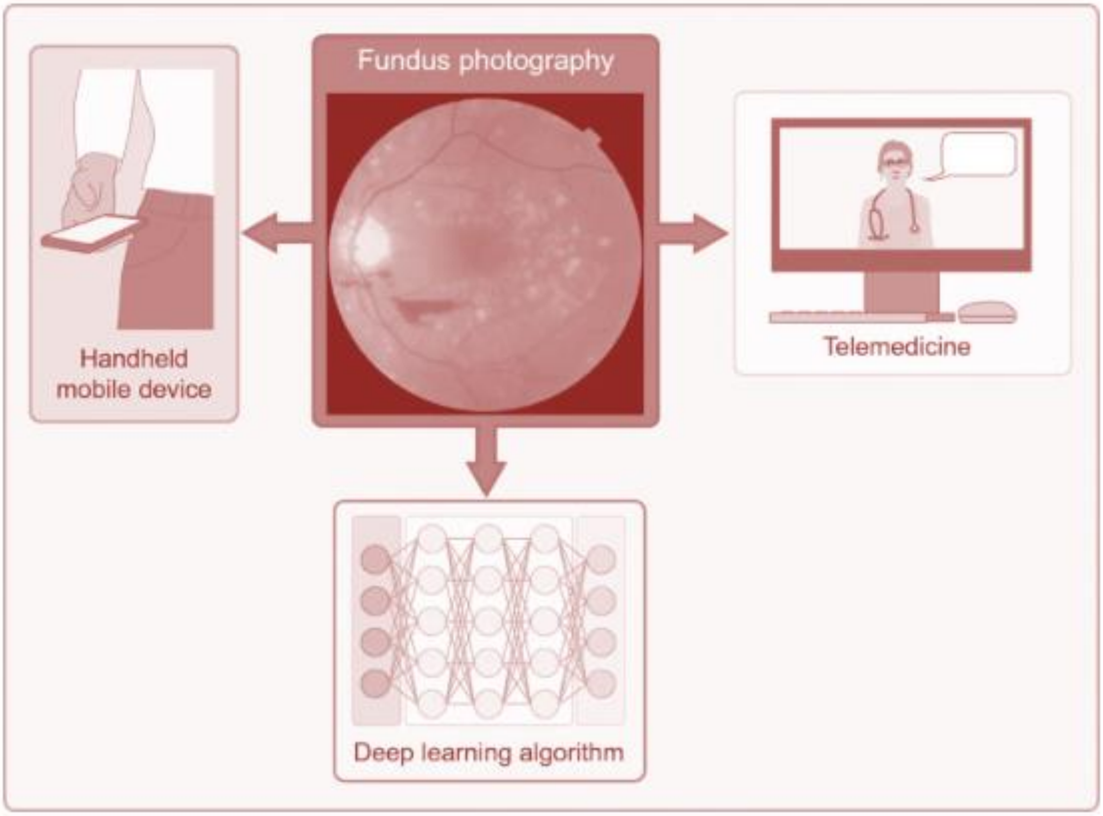


Figure 4: CNN models on image data related to diabetic retinopathy.

Model Evaluation and Performance Metrics:

The performance of the developed Machine Learning and Deep Learning models will be evaluated using appropriate metrics, such as accuracy, precision, recall, and F1-score. The accuracy ratio will be the primary performance measure used to represent the effectiveness of the proposed AI-driven framework for diabetes diagnosis and management. By following this comprehensive plan of work, the research aims to create a robust and innovative AI-based solution for the healthcare sector, addressing the challenges of accurate diabetes diagnosis, personalized prognosis, and efficient disease management.

Conclusion:

The research work presented in this paper addresses the pressing need for intelligent, automated, and personalized solutions in the diabetes healthcare sector. By leveraging the power of Artificial Intelligence and Machine Learning, this study has developed a novel framework that aims to revolutionize the way diabetes is diagnosed, managed, and treated.

The key contributions of this research are:

a. Comprehensive Comparative Analysis: The study has conducted a thorough review of existing AI and ML techniques used in diabetes healthcare, identifying the strengths and limitations of various algorithms. This has enabled the selection of the most suitable models for accurate diagnosis and prognosis.

b. Improved Diagnosis Accuracy: The incorporation of fuzzy qualitative attributes and the development of advanced ML models, such as the Decision Tree-based classifier, have demonstrated enhanced accuracy in diabetes diagnosis compared to conventional approaches.

c. Personalized Diabetic Retinopathy Management: The integration of deep learning, specifically Convolutional Neural Networks, has enabled the creation of reliable and automated screening tools for diabetic retinopathy, paving the way for personalized monitoring and tailored treatment plans.

d. Universal Applicability: The proposed AI-driven framework has been designed to be adaptable and versatile, addressing the unique challenges faced in developed, developing, and underdeveloped nations. This ensures the broader applicability of the solutions across diverse healthcare settings.

e. Improved Clinical Decision-Making: The development of an expert system for diabetes diagnosis and treatment recommendations has the potential to significantly reduce the time and effort required by clinicians, enhancing the overall efficiency of the healthcare system.

The findings of this research work have far-reaching implications for the diabetes healthcare sector. By seamlessly integrating AI and ML techniques, this framework has the potential to transform the way diabetes is managed, leading to improved patient outcomes, reduced healthcare costs, and more efficient utilization of resources.

As the world continues to grapple with the rising burden of diabetes, this innovative approach offers a glimmer of hope. The research team is committed to further refining and validating the proposed solutions, paving the way for their widespread adoption and implementation in the healthcare ecosystem. Through this endeavor, the researchers aspire to contribute to the global effort in combating the diabetes pandemic and improving the quality of life for individuals affected by this chronic condition.

References:

[1] Zeki, T. S., El-Baz, A. H., & AbdelRaouf, A. S. (2013). VP-Expert: A Case-Based Reasoning Diabetes Diagnosis System. International Journal of Advanced Computer Science and Applications, 4(9), 76-84.

[2] Choubey, D. K., Paul, S., & Kumar, S. (2017). A rule-based expert fuzzy system for the diagnosis of dengue fever. Soft Computing, 21(16), 4713-4722.

[3] Pratama, A. Y., Permanasari, A. E., & Setiawan, N. A. (2017). Android-based expert system for diabetes risk level determination. In 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI) (pp. 1-6). IEEE.

[4] Kalpana, M., Vidhya, P., & Anandhi, R. (2019). Design of a Fuzzy Expert System for the Diagnosis of Diabetes. International Journal of Engineering and Advanced Technology, 8(5), 2249-8958.

[5] Tadić, D., Stanojević, M., & Popović, M. (2014). A fuzzy model for assessment of patient's referring to an appropriate therapy for type 2 diabetes. Acta Polytechnica Hungarica, 11(3), 49-66.

[6] Prajapati, H., Jain, A., & Patel, S. K. (2017). An Enhance Expert System for Diagnosis of Diabetes using Fuzzy Rules over PIMA Dataset. International Journal of Advanced Engineering Research and Development, 4(9), 63-69.

[7] Al-Ghamdi, A. A., Wazzan, M. A., & Mujallid, F. M. (2011). An Expert System of Determining Diabetes Treatment Based on Cloud Computing Platforms. International Journal of Computer Science and Information Technology, 2(5), 1982-1987.

[8] Ahmed, I. M., Alfonse, M., Adbel-Badeeh, M., & Salem, A. M. (2015). Daily Meal Planner Expert System for Diabetics Type-2. In Proceedings of the World Congress on Engineering and Computer Science (Vol. 1).

[9] Tabibi, S. T., Zaki, T. S., & Ataeepoor, Y. (2013). Developing an Expert System for Diabetics Treatment Advices. International Journal of Hospital Research, 2(3), 155-162.

[10] Katigari, M. R., Ayatollahi, H., Malek, M., & Garavand, A. (2017). Fuzzy expert system for diagnosing diabetic neuropathy. World Journal of Diabetes, 8(2), 80-88.

[11] Dhivya, A. D., & Felix, A. (2018). A fuzzy rule based expert system for T2DM diagnosis. International Journal of Engineering and Technology, 7(4.10), 432-435.

[12] Lakshmi, K. V., & Malliga, P. (2013). Modeling an expert system for diagnosis of gestational diabetes mellitus based on risk factors. Journal of Computer Engineering, 8, 29-32.

[13] Onuiri, E. E., Ndukwe, V. U., & Igwe, N. (2018). Simulation of an Expert System for Diabetes Diagnosis. International Journal of Advanced Computer, 48(1).

[14] Radha, R., & Rajagopalan, S. P. (2007). Fuzzy Logic Approach for Diagnosis of Diabetics. Information Technology Journal, 6(1), 96-102.

[15] Bhandari, V., & Kumar, R. (2015). Comparative Analysis of Fuzzy Expert Systems for Diabetic Diagnosis. International Journal of Computer Applications, 132(6), 8-14.

[16] Nnamoko, N., Arshad, F., England, D., & Vora, J. (2013). Fuzzy expert system for Type 2 Diabetes Mellitus (T2DM) management using dual inference mechanism. AAAI Spring Symposium - Technical Report, SS-13-03, 67-70.

[17] Nagarajan, S., & Chandrasekaran, R. M. (2015). Design of Mobile Expert System for Diabetes Risk Diagnosis and Management. Indian Journal of Science and Technology, 8(8), 32-36.

[18] D. R. Whiting, L. Guariguata, C. Weil, and J. Shaw, "IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030," Diabetes Res. Clin. Pract., vol. 94, no. 3, pp. 311–321, 2011, doi: 10.1016/j.diabres.2011.10.029.

[19] S. I. Sherwani, H. A. Khan, A. Ekhzaimy, A. Masood, and M. K. Sakharkar, "Significance of HbA1c test in diagnosis and prognosis of diabetic patients," Biomark. Insights, vol. 11, pp. 95–104, 2016, doi: 10.4137/Bmi.s38440.

[20] N. C. D. Risk and F. Collaboration, "Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants," Lancet (London, England), vol. 387, no. 10027, pp. 1513–1530, 2016, doi: 10.1016/S0140-6736(16)00618-8.

[21] T. Seuring, O. Archangelidi, and M. Suhrcke, "The Economic Costs of Type 2 Diabetes: A Global Systematic Review," Pharmacoeconomics, vol. 33, no. 8, pp. 811–831, 2015, doi: 10.1007/s40273-015-0268-9.

[22] R. Dieng, O. Corby, and S. Lapalut, "Acquisition of gradual knowledge," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 723 LNAI, no. 3, pp. 407–426, 1993, doi: 10.1007/3-540-57253-8_65.

[23] D. Riswantini, E. Nugraheni, and A. Arisal, "Big Data Research in Fighting COVID-19 : Contributions and Techniques," 2021.