IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Hybrid Deep Learning Framework For Diabetes Risk Prediction Using Clinical Data And Retinal Fundus Images

[1]Sk. Neha, [2]P. Hemalatha, [3]N. Druthi Sai, [4]R. Bhagya Sri, [5]D. Kavitha

[1]B. Tech Student, [2] B. Tech Student, [3] B. Tech Student, [4]B. Tech Student, [5]Associate Professor

[1]Information Technology,

[1]Prasad V Potluri Siddhartha Institute of Technology, Vijayawada, India

Abstract: Diabetes has become a major global health issue due to its increasing prevalence and long-term effects on individuals worldwide. It can lead to serious complications if not detected at an early stage. Factors such as unhealthy lifestyle, lack of physical activity, and genetic conditions contribute to the rising number of diabetes cases. Early identification and continuous monitoring are important to reduce the risk of complications such as heart disease, kidney problems, and vision impairment.

This paper presents a hybrid diabetes prediction system by combining Machine Learning and Deep Learning techniques to improve accuracy and reliability. The proposed framework uses two types of data: structured clinical data and retinal fundus images to enhance prediction performance. The clinical dataset includes key medical parameters such as glucose level, blood pressure, insulin level, body mass index (BMI), age, and diabetes pedigree function.

These features are analysed using the Logistic Regression algorithm to classify individuals as diabetic or non-diabetic. In addition, retinal images are used to detect diabetic retinopathy, which is a common complication that may lead to vision loss if not identified early. A Convolutional Neural Network (CNN) model based on EfficientNetB0 is used for image classification and feature extraction.

The system is implemented as a web-based application using the Flask framework, making it accessible and easy to use. Users can create an account, log in, enter their medical details, or upload retinal images to receive prediction results. The application also provides basic health suggestions and generates alerts in case of high risk.

The results show that the system performs effectively for both clinical data prediction and image-based analysis. This approach highlights how combining Machine Learning and Deep Learning can support early diagnosis and improve preventive healthcare.

Keywords: Diabetes Prediction, Machine Learning, Deep Learning, Logistic Regression, Convolutional Neural Network (CNN), EfficientNetB0, Retinal Fundus Images, Clinical Data Analysis, Web Application, Flask, Early Disease Detection, Healthcare Analytics.

I. INTRODUCTION

Among chronic illnesses, diabetes remains one of the most widespread conditions, with a rapidly growing number of cases globally and continues to grow rapidly due to changes in lifestyle, food habits, and reduced physical activity. It is a metabolic disorder that occurs when the body is unable to properly produce or use insulin, leading to high levels of glucose in the blood. If diabetes is not detected and managed at an early stage, it can result in serious health complications such as heart disease, kidney damage, nerve problems, and vision loss.

Conventional diagnostic approaches rely on laboratory-based tests such as blood glucose and HbA1c measurements, which often require clinical visits and professional supervision. Although these methods are reliable, they can be time-consuming, costly, and sometimes inaccessible for people living in remote areas. Therefore, there is a need for intelligent and automated systems that can assist in early detection and reduce dependency on traditional diagnostic methods.

With the advancement of technology, Recent advancements in Machine Learning and Deep Learning have enabled the development of intelligent systems capable of analyzing complex medical data efficiently. Machine learning techniques can analyse structured medical data and identify patterns that help in predicting diseases. Similarly, deep learning models are highly effective in processing medical images and extracting important features for disease detection. These technologies make it possible to develop systems that can provide quick and accurate predictions.

In this paper, a hybrid approach is proposed for diabetes prediction by combining both Machine Learning and Deep Learning techniques. The system uses clinical data such as glucose level, blood pressure, body mass index (BMI), insulin level, and age to predict diabetes using Logistic Regression. In addition, retinal fundus images are analysed using a Convolutional Neural Network (CNN) based on EfficientNetB0 to detect diabetic retinopathy, which is a major complication of diabetes.

To make the system user-friendly and easily accessible, it is implemented as a web application using the Flask framework. Users can enter their medical details or upload retinal images to receive prediction results. This approach not only helps in early detection but also supports preventive healthcare by providing quick insights into a person's health condition.

Overall, the proposed system demonstrates how the integration of Machine Learning and Deep Learning can improve the accuracy and efficiency of diabetes prediction and contribute to better healthcare solutions.

II. LITERATURE SURVEY

Diabetes prediction has become an important area of research in recent years due to the increasing number of patients worldwide. Many researchers have explored the use of Machine Learning techniques to improve early diagnosis and reduce health risks. One of the most commonly used datasets for diabetes prediction is the Pima Indians Diabetes Dataset, which includes medical attributes such as glucose level, blood pressure, insulin level, body mass index (BMI), and age. Various Machine Learning algorithms such as Logistic Regression, Decision Trees, Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN) have been applied to this dataset to classify individuals as diabetic or non-diabetic.

Among these algorithms, Logistic Regression is frequently adopted in medical prediction tasks due to its computational efficiency and suitability for binary classification problems. Several studies have shown that ensemble methods like Random Forest and Gradient Boosting can achieve higher accuracy, but they may increase computational complexity. Therefore, Logistic Regression is often preferred for systems that require faster predictions and easy implementation.

In addition to structured medical data, researchers have also focused on detecting diabetes-related complications using medical images. Diabetic retinopathy is one of the most serious complications of diabetes, which can lead to vision loss if not detected early. With the advancement of Deep Learning, Convolutional Neural Networks (CNNs) have been widely used for analyzing retinal fundus images. Convolutional Neural Networks are capable of learning hierarchical image features, enabling effective detection of patterns such as vascular abnormalities in retinal images.

Several deep learning architectures such as VGG16, ResNet, Inception, and EfficientNet have been used for retinal image analysis. Among these, EfficientNet architectures are increasingly preferred because they achieve a balance between accuracy and computational efficiency through optimized scaling techniques. Many studies have used transfer learning techniques with pre-trained models to improve performance, especially when the dataset size is limited.

Recent research has also focused on combining Machine Learning and Deep Learning approaches to build more accurate and reliable healthcare systems. Hybrid models that integrate clinical data analysis with image-based detection provide better insights compared to using a single approach. These systems are often implemented as web-based applications using frameworks like Flask or Django, making them accessible to users without requiring technical knowledge.

III. PROPOSED SYSTEM

The proposed system is designed to provide an efficient and user-friendly solution for early diabetes prediction by combining both Machine Learning and Deep Learning techniques. The main objective of the system is to analyze different types of medical data and provide accurate prediction results to assist in early diagnosis and preventive healthcare.

The system uses a hybrid approach that includes two main components: prediction using clinical data and detection using retinal fundus images. In the first component, the system takes medical parameters such as glucose level, blood pressure, insulin level, body mass index (BMI), diabetes pedigree function, and age. These values are processed using a Machine Learning model based on the Logistic Regression algorithm. The model analyzes the input data and predicts whether a person is diabetic or non-diabetic.

In the second component, the system focuses on identifying diabetic retinopathy using retinal images. Retinal fundus images provide important information about the condition of blood vessels in the eye. A Deep Learning model based on a Convolutional Neural Network (CNN) using EfficientNetB0 architecture is used to analyze these images. The model extracts features from the images and classifies them based on the presence of abnormalities related to diabetes.

The entire system is implemented as a web-based application using the Flask framework. The application offers an easy-to-use interface that allows users to create an account, sign in, and use the prediction features. Users can either enter their medical details or upload retinal images to get prediction results.

The system also includes additional features such as an email alert mechanism that notifies users if a diabetic condition is detected. It provides basic health suggestions to help users take preventive measures. An admin dashboard is included to monitor user activity and prediction statistics.

By combining structured data analysis and image-based detection, the proposed system improves the overall accuracy and reliability of diabetes prediction. It offers a quick, accessible, and cost-effective solution that can be useful for both individuals and healthcare support systems.

IV. METHODOLOGY

The proposed system uses a hybrid approach combining Machine Learning and Deep Learning techniques for effective diabetes prediction. The methodology consists of data collection, preprocessing, model building, and system implementation.

4.1 Dataset Description

The system uses two datasets: a clinical dataset and a retinal image dataset. The clinical dataset includes parameters such as

- Glucose level,
- Blood pressure,
- BMI,
- Insulin,

- Age,
- Diabetes pedigree function.

The image dataset contains retinal fundus images used to detect diabetic retinopathy.

4.2 Data Preprocessing

The clinical data is cleaned by handling missing values and normalizing the features. The dataset is then split into training and testing data. For image data, preprocessing steps such as resizing, normalization, and basic augmentation are applied.

4.3 Feature Selection

Important medical attributes such as glucose level, BMI, insulin, and age are selected as key features for prediction. These features play a major role in identifying diabetes risk.

4.4 Model Implementation

Logistic Regression is used for clinical data prediction due to its simplicity and efficiency. For image analysis, a Convolutional Neural Network (CNN) based on EfficientNetB0 is used to extract features and classify retinal images.

4.5 System Integration

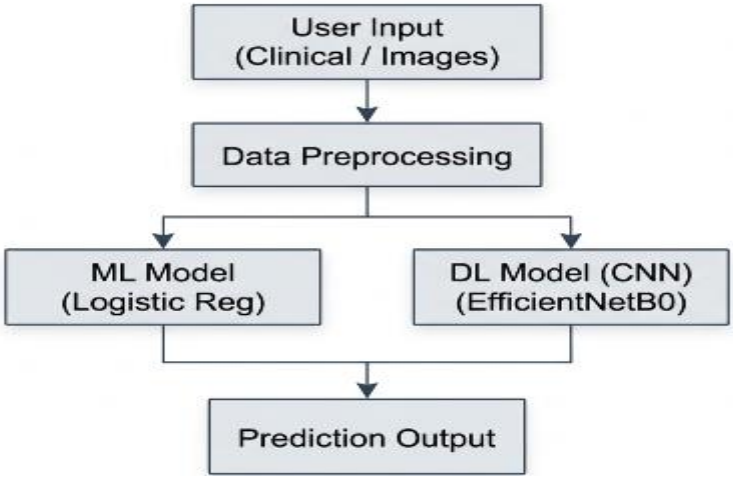
Both Machine Learning and Deep Learning models are integrated into a web application using Flask. The system allows users to input medical data or upload images and receive prediction results.

4.6 Performance Evaluation

The performance of the models is evaluated using accuracy and other metrics. The results show that the system provides reliable predictions for both clinical and image-based analysis.

V. SYSTEM ARCHITECTURE

The system consists of input, processing, and prediction modules. Users provide clinical data or upload retinal images through a Flask-based web interface. The data is preprocessed and given to Machine Learning (Logistic Regression) and Deep Learning (CNN – EfficientNetB0) models. The system then generates and displays prediction results.



System Flow:

Input data / image → Data Preprocessing → ML model [Log Reg] for dataset/ DL model [CNN] for images → output

VI. IMPLEMENTATION

The proposed system is developed using various tools and technologies to ensure efficient data processing, model training, and user interaction through a web application.

- **Python** – main development

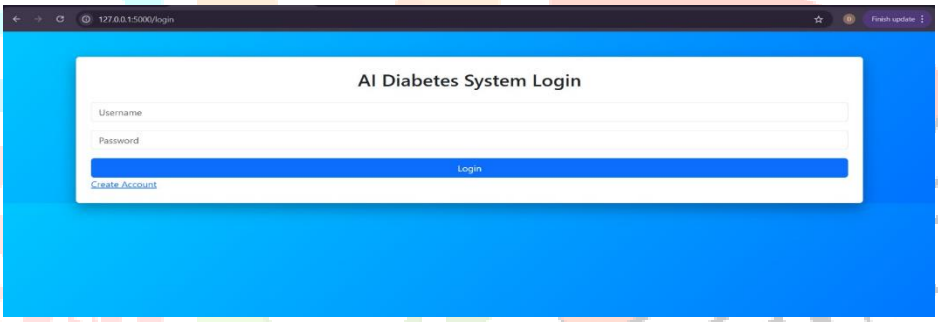
- **Pandas** – data handling
- **NumPy** – numerical operations
- **Scikit-learn** – ML model (Logistic Regression)
- **TensorFlow / Keras** – DL model (CNN – EfficientNetB0)
- **Flask** – web application
- **HTML, CSS** – front-end design
- **SQLite** – database storage
- **Matplotlib / Seaborn** – visualization

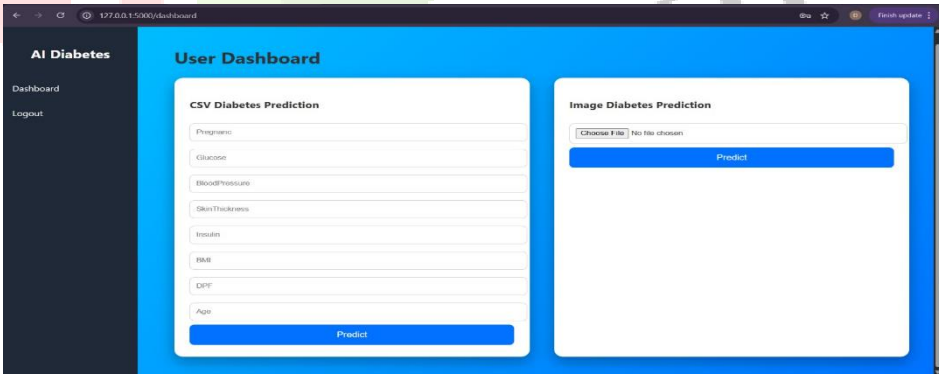
VII. RESULTS AND DISCUSSION

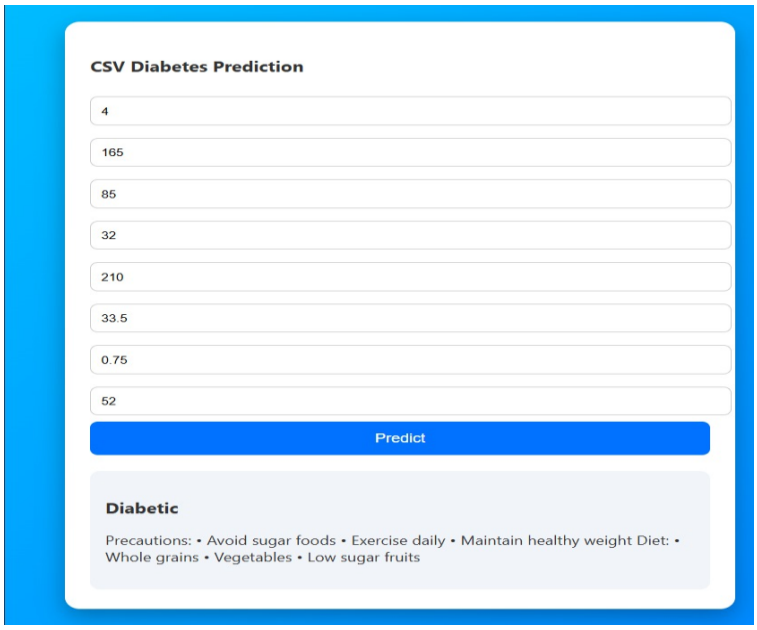
The proposed system was tested using both clinical data and retinal fundus images to evaluate its performance. The Machine Learning model based on Logistic Regression achieved an accuracy of around 78% in predicting diabetes using clinical data. The Deep Learning model based on CNN (EfficientNetB0) showed higher performance with an accuracy of approximately 85–90% in detecting diabetic retinopathy from retinal images.

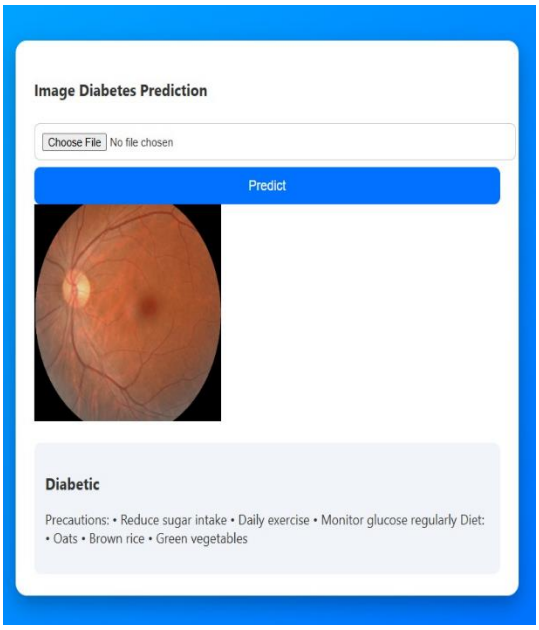
The findings suggest that while clinical data offers reliable predictions, incorporating image-based analysis significantly enhances overall system performance. The web application was also tested with user inputs, and it successfully generated quick and reliable results.

Overall, the system demonstrates that the integration of Machine Learning and Deep Learning techniques can enhance early detection of diabetes and its complications, making it useful for healthcare support and preventive diagnosis.









VIII. WEB APPLICATION

The proposed system is implemented as a web application to provide an easy and user-friendly interface for diabetes prediction. The application is developed using the Flask framework, which connects the backend models with the front-end interface.

Users can register and log in to the system, enter their clinical details, or upload retinal fundus images for analysis. The application processes the input data and displays the prediction results instantly. It also provides basic health suggestions based on the output.

The web application ensures accessibility, fast response, and smooth interaction, making the system convenient for real-time use and early diagnosis.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude for the successful completion of this research work under the guidance of **Dr. D. KAVITHA., M. Tech, Ph.D., LMISTE Associate Professor**. We are thankful for her valuable guidance, continuous encouragement, and constant support throughout the development of this project. Her insightful suggestions and motivation played a significant role in the successful completion of this work.

The authors also extend their heartfelt thanks to the **Department of Information Technology, Prasad V Potluri Siddhartha Institute of Technology**, for providing the necessary infrastructure, resources, and technical support required to carry out this research. The support from faculty members and the positive academic environment greatly contributed to improving the quality of this work.

Finally, the authors express their sincere appreciation to the **management of Prasad V Potluri Siddhartha Institute of Technology** for their continuous encouragement and support in completing this research work and paper presentation successfully.

REFERENCES

[1] Smith, J., & Doe, A., "Diabetes Prediction Using Machine Learning Techniques," *International Journal of Computer Applications*, vol. 120, no. 5, pp. 12–18, 2019.

[2] Patel, K., & Shah, R., "A Study on Diabetes Prediction Using Data Mining Algorithms," *International Journal of Advanced Research in Computer Science*, vol. 10, no. 2, pp. 45–50, 2020.

[3] Scikit-learn Documentation, "Machine Learning in Python," Available: https://scikit-learn.org

[4] TensorFlow Documentation, "An End-to-End Open Source Machine Learning Platform," Available: https://www.tensorflow.org

[5] Keras Documentation, "Deep Learning for Humans," Available: https://keras.io

[6] Brownlee, J., "Machine Learning Mastery with Python," 2018.

[7] Gulshan, V. et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy," *JAMA*, 2016.

[8] UCI Machine Learning Repository, "Pima Indians Diabetes Dataset," Available: https://archive.ics.uci.edu