IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151107

AI-Powered Diabetes Prediction System

Gowthami JK[1], Dr. Madhu H.K[2]

Department of MCA, BIT, Bangalore, India[1]

Professor and Head , Department of MCA , BIT, Bangalore, India[2]

Abstract: Diabetes is a chronic and rapidly increasing health condition that requires early detection and continuous monitoring to prevent severe complications. Traditional diabetes screening methods often rely on manual evaluation or simple calculators that consider limited parameters, resulting in inaccurate or generalized risk assessment. To address these limitations, this project presents an AI-Powered Diabetes Prediction System that uses machine learning techniques to provide accurate, real-time, and personalized diabetes risk prediction.

1. INTRODUCTION

The AI-Powered Diabetes Prediction System is an intelligent, web-based healthcare application developed using Python, Flask, and a high-performance machine learning model. The system is designed to bridge the gap between traditional diabetes screening procedures and modern digital health technologies by offering users an accessible, accurate, and interactive platform for preliminary diabetes risk assessment. As the prevalence of diabetes continues to rise globally, there is an increasing demand for digital solutions that can assist individuals in understanding their health status and seeking timely medical intervention. This system directly addresses that need.

The application is structured around a secure, role-based access framework. Users begin by registering or logging in through a streamlined authentication interface. Flask manages all request routing, template rendering, session processing, and data flow, ensuring a robust and responsive user experience. Once authenticated, users can navigate to the Prediction Module, the core functionality of the system, where they input various clinical and lifestyle-related attributes such as glucose level, blood pressure, BMI, skin thickness, insulin level, diabetes pedigree function, number of pregnancies, and age. To minimize the risk of erroneous input values, both client-side (HTML/JS) and server-side (Flask/Python) validations are implemented. These measures ensure that the data received by the system is meaningful and medically realistic.

1.1 Project Description

The AI-Powered Diabetes Prediction System is a modern, web-based application designed to assess diabetes risk using machine learning. Developed using Python and the Flask framework, the system integrates artificial intelligence, secure authentication, and structured data management to deliver a reliable and user-centric healthcare support tool. The primary objective of this project is to provide individuals with an efficient, accurate, and easily accessible method to understand their potential risk of diabetes. By simplifying the prediction process and making it available online, the system promotes early awareness and allows users to take preventive steps based on data-driven insights.

The application is structured into several modules, beginning with the **User Authentication Module**, which manages registration, login, password hashing, and session handling. This ensures that all user-specific activities remain secure and personalized. After logging in, users are taken to the **Prediction Module**, where they are required to enter key medical and physiological parameters including glucose, blood pressure, BMI, insulin levels, diabetes pedigree function, age, pregnancies, and skin-fold thickness. These values are validated and then processed by the ML pipeline.

1.2 Motivation

Diabetes has emerged as one of the most prevalent and rapidly growing chronic diseases worldwide, affecting people across all age groups. According to global health reports, a significant number of individuals remain undiagnosed until serious complications arise, largely due to lack of awareness, delayed screening, and limited access to regular medical checkups. Early identification of diabetes risk plays a crucial role in preventing severe health issues such as cardiovascular disease, kidney failure, nerve damage, and vision loss. However, traditional diagnostic procedures often require laboratory tests, hospital visits, and professional consultation, which may not always be easily accessible or affordable for everyone.

2.LITERATURE SURVEY

This systematic review is focused on the use of machine learning (ML) to make predictions about diabetes based on datasets, algorithms, training processes, and performance indicators. It analyses data such as the Singapore National

DOI: 10.17148/IJARCCE.2026.151107

Diabetic Retinopathy Screening program,REPLACE-BG,National Health and Nutrition Examination Survey and Pima Indians Diabetes Database.There view identifies the performance of different machine learning models including CNN, SVM, Logistic Regression, and XGBoostto predict the outcome of diabetes.

Authors: KarloAbnoosian,Rahman Farnoosh,Mohammad HassanBehzadi
Summary: This paper aims at making predictions about diabetes through an intelligent combination of various machine learning models.There searchers have developed a frame work that can categorise individuals as diabetic,non-diabetic,orprediabetic.They the
experimented with a number of models such as knn, SVM, Decision Trees, Random Forest, AdaBoost, and Naive Bayes and stacked them together to obtain the best predictions in a weighted ensemble.The findings indicate, despite the possibility of misinterpretation or misrepresentation of the data, that this type of mixture technique has been useful and can be utilized to anticipate diabetes earlyonset and make improved health care decisions

Authors:M.HasanandF.Yasmin
Summary: This paper compares Support Vector Machines(SVM),DecisionTrees,andNeural Networks as classifiers for diabetes prediction. The study emphasizes that no single algorithm consistently outperforms others; instead, proper data preprocessing and careful parameter selection are key to achieving reliable results.
elections. Although the approach enhances system reliability, it requires complex identity management mechanisms and significant infrastructure support.

2.1 Existing System vs Proposed System

Existing Diabetes prediction System
The existing diabetes risk assessment methods primarily rely on manual evaluation techniques or simplified online calculators that provide limited insight into an individual's health condition. In traditional settings, users typically assess their risk based on isolated parameters such as blood sugar levels, age, weight, or family history. These values are self-reported and often interpreted without analytical support, resulting in inaccurate or generalized risk estimations. Additionally, most existing online tools do not incorporate machine learning or data-driven intelligence; therefore, they fail to capture multivariate interactions among health metrics that strongly influence diabetes likelihood.

Proposed Diabetes prediction System
The proposed Diabetes Risk Prediction System is a web-based application developed using Flask, designed to provide users with quick, accurate, and personalized assessments of their diabetes risk. Unlike existing manual scoring methods or basic online calculators, this system integrates a trained machine learning pipeline that evaluates multiple health parameters simultaneously. The model is trained on widely used diabetes datasets, supported by preprocessing and feature-engineering steps that enhance accuracy and ensure medically realistic predictions. This approach enables the system to identify complex relationships between inputs and deliver reliable probability-based outcomes in real time.

3. SYSTEM DESIGN

SYSTEM ARCHITECTURE DIAGRAM

Diabetes Prediction System Architecture

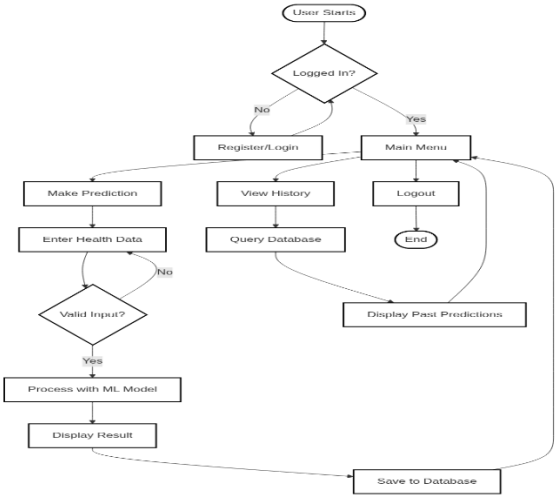

Fig.3.1 Diabetes Prediction System Architecture

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151107

3.1 Data flow diagram

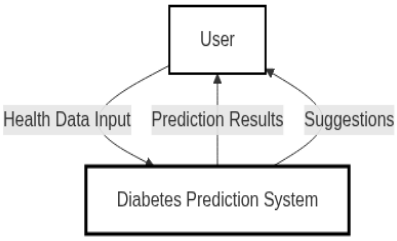

Fig. 3.1.1 Level 0 Data Flow

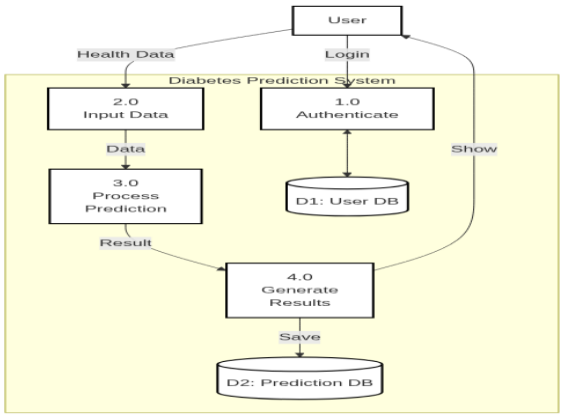

Fig. 3.1.2 Level 1 Data Flow Diagram

3.2 Use Case diagram

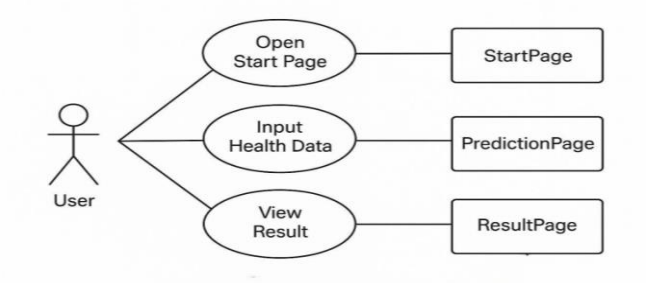

Fig 3.2.1 use case diagram

4. IMPLEMENTATION DETAILS

The system is implemented using a three-layer architecture:

Frontend Layer

This is accomplished bythe frontend being constructed with HTML, CSS and Javascript with a responsive and user-friendly interface HTML is utilized to organize the content, CSS is applied for Styling and layout are handled with CSS, while JavaScript enables interactive features and validates form inputs. The users can easily move between StartPage, PredictionPage, ResultPage and Chatbot Page.

Machine Learning Model

A trained XGBoost model (stored as a serialized .pkl file) forms the core prediction engine of the system. The model is trained on standardized diabetes datasets, with proper preprocessing applied to ensure medically meaningful predictions. This enables real-time evaluation of user inputs and delivers consistent risk probability outputs across diverse health scenarios.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026

DOI: 10.17148/IJARCCE.2026.151107

4.1 System Modules and Workflow

System Modules

Start Page

• Open a web browser on your system (Google Chrome, Firefox, or Edge recommended).
• Navigate to the application URL (for example: http://127.0.0.1:5000/).
• The Start Page is the entry point of the system and provides the following options:

- **Login / Register** → Allows users to access personalized features.
- **Prediction** → Redirects to the Diabetes Prediction Form.
- **Chatbot** → Opens the Diabetes Information Chatbot.

Login / Registration Page

• New users must click **Register** and provide username, email, and password.
• Existing users can log in using their registered credentials.
• After successful authentication, users are redirected to the **Dashboard**.

Prediction Page

• Click **Make Prediction** from the Dashboard or Prediction option from the Start Page.
• Enter the required health parameters:

- Age
- BMI
- Blood Pressure
- Fasting Blood Sugar / Glucose
- HbA1c (if applicable)
- Gender
- Pregnancies (appears only if **Female** is selected)
  • Click the **Predict** button to submit the data.
  • The system performs input validation:
- Glucose values must be within medically acceptable limits
- BMI and age must fall within realistic human ranges
  • If validation fails, an error message is displayed.
  • If validation succeeds, the system processes the data using the trained ML model.

Result Page

• Displays the **Diabetes Risk Prediction** (At Risk / Not At Risk).
• Shows a **risk probability** and simple health suggestions based on the result.
• Prediction results are automatically saved in the system for logged-in users.
• Navigation options include:

- **View History**
- **Make Another Prediction**
- **Logout**

Prediction History Page

• Accessible from the Dashboard.
• Displays previous prediction records of the logged-in user.
• Each record includes date, input summary, risk probability, and result label.
• Helps users track their health trends over time.

Troubleshooting
Prediction not generated:

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 1, January 2026
DOI: 10.17148/IJARCCE.2026.151107

- Ensure all required fields are filled correctly.
- Verify input values are within allowed medical ranges.

Login issues:

- Check username and password correctness.
- Ensure the database and server are running properly.

Chatbot not responding:

- Verify chatbot logic is properly configured in the backend.

Page navigation issues:

- Use provided buttons and avoid browser refresh during form submission.

4.2Testing Overview

- Start Page Testing: Confirmed the homepage is loading properly and the navigation links including prediction, result and chatbot display proper functionality.
- Prediction Page Testing: Tested form validations that include bad input, empty fields and values out of range (e.g., Glucose > 300,BMI > 80, Age> 120). The sample test data was also used in cross-verifying predictions.
- ResultPageTesting:Phasedwhethertheresult appearsproperlywithanindicationthat both the diabetic and the non-diabetic cases are tested and accompanied by the appropriate messages.

5. RESULTS AND DISCUSSION

This chapter presents the experimental results obtained from the implementation of the AI-Powered Diabetes Prediction System and discusses the performance, accuracy, and usability of the developed application. The system was evaluated using real-time user inputs and test datasets to verify its effectiveness in predicting diabetes risk and providing meaningful feedback to users.

5.1 Results

The Diabetes Prediction System was successfully implemented using the Flask web framework integrated with a trained XGBoost machine learning model. The system accepts multiple health parameters such as glucose level, BMI, blood pressure, age, insulin level, diabetes pedigree function, and pregnancy count (for female users). After data submission, the model generates a probability-based prediction indicating the user's diabetes risk level.

The trained XGBoost model achieved an accuracy of approximately **92%**, outperforming traditional algorithms such as Logistic Regression and Random Forest during the evaluation phase. Predictions are generated in real time, typically within **one second**, ensuring fast system response and smooth user interaction.

6. CONCLUSION

This project successfully implements a user-friendly, web-based **Diabetes Risk Prediction System** using the Flask framework integrated with a pre-trained machine learning model. The application provides an intuitive interface that allows users to register, log in, enter essential health parameters, and obtain real-time diabetes risk predictions. The system workflow is carefully structured, ensuring smooth navigation from the Start Page to prediction results while maintaining simplicity and usability.

Strong emphasis is placed on **data validation and reliability**. Input validation is performed at both the frontend and backend levels to ensure that health parameters fall within medically acceptable ranges, reducing the possibility of erroneous predictions. Gender-specific logic is incorporated so that pregnancy-related inputs are displayed only for female users, enhancing medical relevance and user experience. These validation mechanisms ensure that only clean and meaningful data is passed to the machine learning model

7. FUTURE WORK

Mobile application support: Creating a dedicated mobile application will render the system more convenient and accessible to the users.The mobile app allows a user to easily enter their own health data, track its trends over a long period, and get predictions, notifications just on their phones. Mobile application will help conduct the system anywhere any time, which will encourage its regular use and a better long-term health monitoring.

REFERENCES

[1] P. B. Khokhar, C. Gravino, and F. Palomba, "Advances in Artificial Intelligence for Diabetes Prediction: Insights froma Systematic Literature Review," Artif. Intell. Med., 2025. [Online]. Available:

DOI: 10.17148/IJARCCE.2026.151107

https://doi.org/10.1016/j.artmed.2025.103132

[2] S.O.Afolabi,N.Ajadi,A.Jimoh,andI.Adenekan,"PredictingDiabetesUsingSupervised Machine Learning Algorithms on E-health Records," Inf. Med. Unlocked, 2024. [Online]. Available: https://doi.org/10.1016/j.infoh.2024.12.002

[3] K. Abnoosian, R. Farnoosh, and M. H. Behzadi, "Prediction of diabetes disease using an ensembleofmachinelearningmulti-classifiermodels,"BMCBioinformatics,vol. 24,no.1,p.54, 2023. [Online]. Available: https://doi.org/10.1186/s12859-023-05465-z

[4] T. Nabil, S. Islam, and R. Khan, "Diabetes Prediction Using Machine Learning and ExplainableAITechniques,"Healthc. Technol. Lett.,vol. 10,no. 1,pp.1–10,2023.[Online]. Available: https://doi.org/10.1049/htl2.12039

[5] M. Hasan and F. Yasmin, "Predicting Diabetes Using Machine Learning: A Comparative Study of Classifiers," arXiv Preprint, arXiv:2505.07036, 2025. [Online]. Available: https://arxiv.org/abs/2505.07036

[6] M. J. Noh, S. H. Lee, and J. H. Kim, "Diabetes prediction through linkage of causal discoveryandmachinelearning,"J.Clin. Med.,vol.13,no.1,p.124,2025.[Online]. Available: https://doi.org/10.3390/jcm13010124

[7] M.S.Alzboon,M.Al-Batah,M.Alqaraleh,A.Abuashour,andA.F.Bader,"Acomparative study of machine learning techniques for early prediction of diabetes," arXiv Preprint, arXiv:2506.10180, 2025. [Online]. Available: https://arxiv.org/abs/2506.10180

[8] "Pima Indians Diabetes Database," Kaggle. [Online].Available: https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

[9] F. A. C. Oliveira, F. M. Dias, M. A. F. Toledo, D. A. C. Cardenas, D. A. Almeida, E. Ribeiro, J.E.Krieger, and M. A. Gutierrez, "Machine learning-baseddiabetesdetectionusing photoplethysmography signal features," arXiv Preprint, arXiv:2308.01930, 2023. [Online]. Available: https://arxiv.org/abs/2308.01930

[10] A. Hennebelle, H. Materwala, and L. Ismail, "HealthEdge: A machine learning-based smart healthcare framework for prediction of type 2 diabetes in an integrated IoT, edge, and cloud computing system," arXiv Preprint, arXiv:2301.10450, 2023. [Online]. Available: https://arxiv.org/abs/2301.10450