International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.137-146, January 2026.
www.ijersem.com eISSN- 3107-9075

Diabetes Prediction Using Extreme Learning Machine – An Application in Health System

[1]D Janani, [2]G Mahendra, [3]K Lohith Kumar, [4]N Keerthi, [5]G Poojith, [6]B Yugandhar

Department of CSE, Siddartha Institute of Science and Technology, Puttur, India.

[1]sistkcse.janani@gmail.com, [2]mahendramahe689@gmail.com, [3]lohithkumar1266@gmail.com, [4]nellipallikeerthi@gmail.com, [5]poojithg9@gmail.com, [6]bhoopatiyugandhar70@gmail.com

Abstract: Accurate and early prediction of diabetes is critical for effective healthcare management and prevention of long-term complications. Traditional diagnostic approaches rely on manual clinical analysis and laboratory testing, which are often time-consuming and may delay early intervention. Although machine learning techniques have been applied to diabetes prediction, many existing models suffer from limitations such as slow training speed, high computational complexity, sensitivity to noise, and reduced performance on high-dimensional medical datasets. To address these challenges, this paper presents an Extreme Learning Machine (ELM)–based diabetes prediction framework integrated with software engineering principles for reliable health informatics applications. The proposed system utilizes patient health parameters including glucose level, insulin, body mass index (BMI), blood pressure, age, skin thickness, and diabetes pedigree function. Data preprocessing techniques such as normalization, missing value handling, and Principal Component Analysis (PCA) are applied to enhance data quality and reduce dimensionality. The ELM model enables fast and efficient classification of patients into diabetic, non-diabetic, and borderline (pre-diabetic) categories. Experimental evaluation on hospital and Indian diabetes datasets demonstrates improved prediction accuracy and significantly reduced training time compared to traditional machine learning approaches. The system provides a scalable, efficient, and reliable solution for early diabetes detection, supporting timely medical decision-making in modern healthcare environments.

Keywords: Machine Learning, Diabetes Prediction, Extreme Learning Machine, Health Informatics, Principal Component Analysis.

1 Introduction

Diabetes mellitus is one of the most prevalent chronic metabolic disorders and has become a major global public health concern due to its increasing incidence and severe long-term complications. It occurs when the body is unable to regulate blood glucose levels effectively because of insufficient insulin production or improper insulin utilization. If not diagnosed and managed at an early stage, diabetes can lead to serious health complications such as cardiovascular diseases, kidney failure, nerve damage, and vision impairment, thereby increasing morbidity and mortality rates [1]. Early detection of diabetes plays a critical role in preventing disease progression and reducing healthcare burden. However, traditional diagnostic methods primarily rely on manual clinical evaluation and laboratory testing, which are often time-consuming and may delay early medical intervention. Meanwhile, hospitals and healthcare institutions generate large volumes of patient health data, including glucose levels, insulin concentration, body mass index (BMI), blood pressure, age, and hereditary indicators, which remain largely underutilized for predictive analysis and early diagnosis [2].

With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), intelligent systems have been developed to analyze medical datasets and support disease prediction. Several machine learning techniques such as Logistic Regression, Support Vector Machines, Random Forest, K-Nearest Neighbors, and Neural Networks have been applied to diabetes prediction. Although these approaches have demonstrated promising results, many suffer from limitations including high computational complexity, slow training speed, sensitivity to noise, and reduced performance when dealing with high-dimensional medical data [3]. To overcome these limitations, Extreme Learning Machine (ELM) has emerged as an efficient alternative for medical data classification. ELM is a single hidden layer feedforward neural network in which input weights and biases are randomly assigned, and output weights are computed analytically without iterative backpropagation. This learning mechanism significantly reduces training time while maintaining strong generalization performance, making ELM suitable for real-time healthcare applications. Motivated by these advantages, this paper proposes a diabetes prediction system using Extreme Learning Machine integrated with software engineering principles to improve prediction accuracy, system reliability, and decision-making efficiency in health informatics applications [4].

2 LITERATURE REVIEW

Several studies have explored the application of machine learning techniques for early diabetes prediction using clinical and demographic data. M. T. Islam et al. employed Logistic Regression to predict diabetes using features such as glucose level, BMI, and age. Although the model achieved reasonable accuracy, it assumed linear relationships between variables, limiting its ability to model complex, non-linear medical patterns commonly found in real-world healthcare datasets [1]. T. Fatima et al. presented a comprehensive review of machine learning and data mining approaches applied to diabetes research. Their study highlighted the effectiveness of algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests. However, the authors noted that these models often require extensive parameter tuning and suffer from high computational complexity when applied to large-scale healthcare data [2].

R. Asif et al. proposed a diabetes prediction system using K-Nearest Neighbors (KNN) and Naïve Bayes classifiers. While their approach demonstrated moderate predictive performance, it was highly sensitive to feature scaling, noisy data, and class imbalance, which reduced reliability in real clinical environments [3]. T. Ding et al. investigated the use of Artificial Neural Networks (ANN) for diabetes prediction to capture complex feature interactions. Although ANN models improved accuracy compared to traditional classifiers, they required iterative backpropagation, resulting in slow training time and high computational cost, limiting their suitability for real-time healthcare systems [4]. Dimensionality reduction techniques have also been studied to improve diabetes prediction performance. G. Prabhakar et al. demonstrated that Principal Component Analysis (PCA) effectively reduces feature redundancy and noise while preserving essential information. However, PCA alone does not guarantee improved classification performance unless combined with efficient learning algorithms [5].

To overcome limitations of backpropagation-based neural networks, Kishi et al. introduced the Extreme Learning Machine (ELM), a single hidden layer feedforward neural network with randomly initialized input weights and analytically computed output weights. Their work showed that ELM significantly reduces training time while maintaining strong generalization capability [6]. Chandramouli et al. applied ELM to medical diagnosis tasks, including diabetes prediction, and reported faster convergence and improved accuracy compared to conventional machine learning models. However, the study focused mainly on binary classification and did not address multi-class prediction scenarios such as identifying pre-diabetic or borderline cases [7]. N. Gupta et al. explored ensemble learning techniques for diabetes prediction by combining multiple classifiers. While ensemble models improved predictive accuracy, they increased system complexity and computational overhead, making deployment difficult in resource-constrained healthcare environments [8].

Recent research has emphasized integrating machine learning with healthcare informatics systems. A. Taha and S. J. Malebary proposed a clinical decision support system using machine learning for disease prediction. Although the system improved diagnostic support, it lacked structured software engineering practices, affecting scalability, maintainability, and long-term system reliability [9]. Ghorbani et al. highlighted the importance of incorporating software engineering methodologies into intelligent systems to ensure robustness, validation, testing, and maintainability. Despite this, limited studies have integrated structured software engineering frameworks with Extreme Learning Machine models for diabetes prediction, indicating a significant research gap [10]-[12].

3 PROBLEM STATEMENT

Early and accurate prediction of diabetes is a critical requirement in modern healthcare systems due to the increasing prevalence of the disease and its severe long-term complications. However, existing diagnostic approaches primarily rely on manual clinical assessment and laboratory testing, which are often time-consuming, resource-intensive, and may delay early medical intervention. Such methods also place a significant burden on healthcare professionals, especially in environments with limited medical resources.

Several machine learning–based diabetes prediction systems have been proposed; however, many existing solutions suffer from limitations such as high computational complexity, slow training speed, sensitivity to noisy and high-dimensional medical data, and reduced generalization performance. Traditional models often fail to effectively capture complex non-linear relationships among clinical attributes, leading to inconsistent and unreliable prediction outcomes. Additionally, most existing systems focus on binary classification, providing only diabetic or non-diabetic results, and do not adequately address borderline or pre-diabetic conditions that require early preventive care [13].

Many current predictive systems lack proper integration with software engineering practices, resulting in poor scalability, limited reliability, and difficulty in deployment within real-world healthcare environments. The absence of a fast, accurate, and scalable diabetes prediction framework that can efficiently utilize hospital data and support early medical decision-making highlights a significant research gap. This motivates the need for an intelligent, efficient, and reliable machine learning–based diabetes prediction system capable of delivering accurate multi-class classification while ensuring practical applicability in healthcare settings [14].

4 Proposed System

The proposed system implements a Diabetes Prediction Framework based on the SEMLHI (Software Engineering with Machine Learning for Health Informatics) methodology, integrated with the Extreme Learning Machine (ELM) for accurate and efficient disease prediction. The system is designed to analyze real-world healthcare datasets and provide reliable diabetes prediction along with performance evaluation. The overall workflow of the proposed system is illustrated in the block diagram shown in Fig. 1. The architecture emphasizes the seamless integration of software engineering principles, machine learning models, and health data analytics, ensuring scalability, reliability, and improved prediction accuracy.

4.1. Input Data Layer

As depicted in Fig. 1, the system accepts Real-World Datasets, including Indian Diabetes Data, as the primary input. These datasets contain patient health parameters such as glucose level, insulin, body mass index (BMI), blood pressure, age, skin thickness, and diabetes pedigree function. The use of real hospital and regional datasets enhances the practical applicability and reliability of the prediction system.

4.2. SEMLHI Framework

The collected datasets are processed through the SEMLHI Framework, which acts as the core integration layer of the system. SEMLHI ensures that machine learning development follows structured software engineering practices, improving system quality and maintainability. Within the SEMLHI framework, the input data is routed through four major components:

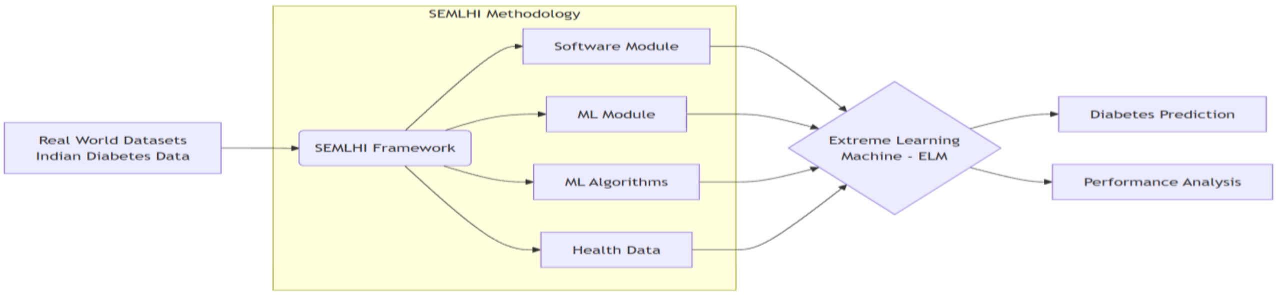

Fig. 1. Block Diagram of the Proposed Diabetes Prediction Using Extreme Learning machine

4.3. Software Module

The Software Module, shown in the block diagram, is responsible for handling system-level operations such as data validation, workflow management, verification, and testing. This module ensures that the data flow between different components is reliable and error-free. It also supports modular development, enabling easier system updates, testing, and reuse of components in future healthcare applications.

4.4. Machine Learning Module

The ML Module manages the overall machine learning pipeline, including data preprocessing, feature normalization, and preparation of datasets for training and testing. This module ensures that the health data is converted into a machine-readable format suitable for efficient learning. It bridges the gap between raw health data and machine learning algorithms.

4.5. Machine Learning Algorithms Module

As illustrated in Fig. 1, the ML Algorithms Module supports multiple machine learning techniques for comparative analysis. Algorithms such as Logistic Regression, KNN, Naïve Bayes, Random Forest, Linear SVC, and Extreme Learning Machine are applied to the processed data. This enables performance comparison and validation, ensuring that the most efficient algorithm is selected for diabetes prediction.

4.6. Health Data Processing

The Health Data component focuses on preprocessing medical attributes by handling missing values, removing noise, and reducing dimensionality. Techniques such as Principal Component Analysis (PCA) are applied to eliminate redundant features and enhance learning efficiency. Clean and optimized health data significantly improves model accuracy and reduces computational complexity.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.137-146, January 2026.
www.ijersem.com eISSN- 3107-9075

4.7. Extreme Learning Machine (ELM)

At the center of the proposed system is the Extreme Learning Machine (ELM), which receives inputs from the Software Module, ML Module, ML Algorithms Module, and Health Data processing unit. ELM is a single hidden layer feedforward neural network where input weights and biases are randomly assigned, and output weights are calculated analytically. This architecture eliminates iterative backpropagation, resulting in faster training speed and strong generalization performance.

4.8. Output Layer: Prediction and Performance Analysis

As shown in Fig. 1, the output of the ELM model is directed to two major components:

- **Diabetes Prediction**: The system classifies patients into diabetic, non-diabetic, or borderline (pre-diabetic) categories, supporting early diagnosis and preventive healthcare.
- **Performance Analysis**: The system evaluates model performance using metrics such as accuracy, efficiency, and prediction reliability. This analysis validates the effectiveness of ELM compared to traditional machine learning models.

The proposed system effectively integrates real-world healthcare data, structured software engineering practices, and fast machine learning algorithms under the SEMLHI methodology. By utilizing Extreme Learning Machine, the system achieves faster training, improved prediction accuracy, and reliable performance evaluation. The modular design ensures scalability, robustness, and adaptability, making the system suitable for real-time healthcare and clinical decision support applications.

5 System Architecture

The system architecture of the proposed Diabetes Prediction System using Extreme Learning Machine (ELM) is designed according to the SEMLHI (Software Engineering with Machine Learning for Health Informatics) methodology. The architecture integrates real-world healthcare datasets, software engineering components, and machine learning modules to achieve accurate diabetes prediction and performance evaluation. The complete architecture is shown in Fig. 2. The architecture follows a layered and modular approach, ensuring proper coordination between data handling, machine learning processes, and prediction output. Each module shown in the architecture performs a specific role in the overall prediction workflow.

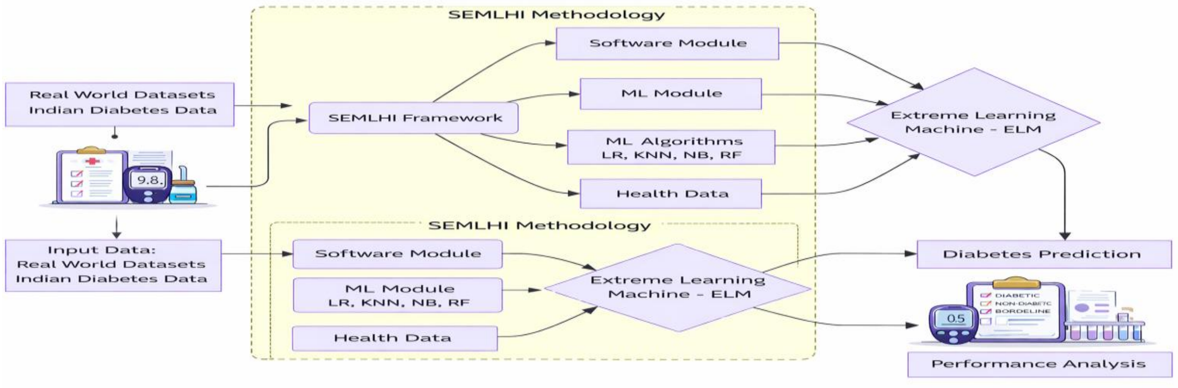

Fig. 2. System Architecture

5.1. Input Data Layer

As shown in Fig.2, the system begins with Real World Datasets, specifically Indian Diabetes Data. This dataset contains patient health attributes such as:

- Glucose level
- Insulin
- Body Mass Index (BMI)
- Blood pressure
- Skin thickness
- Age
- Diabetes pedigree function

These datasets act as the primary input to the system.

5.2. SEMLHI Framework

The input datasets are passed to the SEMLHI Framework, which forms the core integration layer of the architecture. The SEMLHI framework ensures that machine learning development follows structured software engineering practices such as verification, validation, testing, and workflow management. The SEMLHI framework distributes data to four major internal components:

- Software Module
- Machine Learning (ML) Module
- ML Algorithms Module
- Health Data Module

5.3. Software Module

The Software Module, shown inside the SEMLHI methodology block, is responsible for:

- Managing system workflow
- Coordinating data flow between modules
- Performing verification and validation
- Ensuring system reliability and correctness

This module ensures smooth interaction between machine learning processes and the overall system logic.

6 ALGORITHMS USED

The proposed Diabetes Prediction System employs Extreme Learning Machine (ELM) as the core prediction algorithm, supported by data preprocessing techniques and Principal Component Analysis (PCA) for dimensionality reduction. These algorithms are selected to achieve fast training, improved accuracy, and reduced computational complexity when working with medical datasets.

6.1. Data Preprocessing Algorithm

Data preprocessing is the initial and essential step in the proposed system. Medical datasets often contain missing values, noise, and attributes with different value ranges, which can negatively affect prediction accuracy.
The preprocessing algorithm performs the following operations:

- Removal of missing and null values
- Elimination of noisy or inconsistent data
- Normalization of numerical features
- Preparation of clean and structured health data

Normalization ensures all features contribute equally during learning and prevents bias toward attributes with larger numeric ranges. After preprocessing, the dataset becomes clean, noise-free, complete, and model-ready, as mentioned in the PPT.

6.2. Principal Component Analysis (PCA) Algorithm

Principal Component Analysis (PCA) is used to reduce the dimensionality of the healthcare dataset while preserving important information. PCA removes redundant and correlated features, improving learning efficiency and reducing computational cost. The PCA algorithm works as follows:

1. Compute the mean of each feature
2. Calculate the covariance matrix of the dataset
3. Compute eigenvalues and eigenvectors
4. Select principal components with maximum variance
5. Transform original data into reduced feature space

The covariance matrix is computed as:

$$C = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T$$

where
X_i represents a data sample and
μ represents the mean of the dataset.

6.3. Extreme Learning Machine (ELM) Algorithm

The Extreme Learning Machine (ELM) is the core algorithm used for diabetes prediction. ELM belongs to the family of Single Layer Feedforward Neural Networks (SLFNs) and is designed for fast learning and efficient classification. Unlike traditional neural networks, ELM does not use iterative weight updates or backpropagation. Instead, input weights and biases are randomly assigned, and output weights are computed analytically. Working of ELM is given below.

1. Randomly assign input weights and hidden layer biases
2. Compute the hidden layer output matrix
3. Calculate output weights using a mathematical solution

The hidden layer output is calculated as:

$$H = g(XW + b)$$

where
$X$is the input feature matrix,
$W$is the randomly assigned weight matrix,
$b$is the bias vector,
$g(\cdot)$is the activation function.

The output weights are computed as:

$$\beta = H^{\dagger}T$$

where
$H^{\dagger}$is the Moore–Penrose pseudoinverse of the hidden layer matrix,
$T$is the target output matrix.
Because ELM avoids iterative training, it provides very fast learning speed, making it suitable for large medical datasets such as diabetes records.

7 Experimental Setup

The experimental setup of the proposed Diabetes Prediction System using Extreme Learning Machine (ELM) is designed to evaluate the effectiveness of the SEMLHI-based framework in predicting diabetes accurately and efficiently. The experiments were conducted using real-world healthcare datasets, appropriate hardware and software environments, and standard evaluation procedures to validate prediction performance and reliability.

7.1. Hardware Requirements

The experimental evaluation of the proposed system was carried out on a computing platform with the following hardware specifications, as mentioned in the research requirements:

- Processor: Intel Core i3 or higher
- RAM: Minimum 4 GB
- Hard Disk: Minimum 250 GB
- Operating System: Windows / Linux

These specifications are sufficient to handle medical datasets and support machine learning model training and testing.

7.2. Software Environment

The software environment used for the development and experimentation of the system includes:

- Operating System: Windows / Linux
- Programming Language: Python (Version 3.7)
- Development Environment: Python IDLE / Anaconda / Jupyter Notebook / Google Colab
- Machine Learning Libraries: NumPy, Pandas, Scikit-learn
- Visualization Tools: Matplotlib

7.3. Dataset Description

The proposed system was evaluated using real-world hospital datasets and Indian diabetes datasets, as specified in the PPT. The dataset consists of patient health records containing the following attributes:

- Glucose level

- Insulin
- Body Mass Index (BMI)
- Blood pressure
- Skin thickness
- Age
- Diabetes pedigree function

Each record is labeled to indicate diabetic or non-diabetic conditions, and borderline cases are identified during prediction. The dataset was preprocessed to handle missing values and noise before being used for model training and testing.

7.4. Data Preprocessing and Feature Reduction

Before training the machine learning model, the dataset was subjected to preprocessing operations including:

- Removal of missing and null values
- Noise elimination
- Feature normalization

After preprocessing, Principal Component Analysis (PCA) was applied to reduce dimensionality and eliminate redundant attributes. This step improves learning efficiency and reduces computational complexity while preserving essential medical information.

7.5. Model Training and Configuration

The Extreme Learning Machine (ELM) was used as the core prediction model in the proposed system. During training:

- Input weights and biases were randomly assigned
- Hidden layer outputs were computed
- Output weights were calculated analytically using a mathematical solution

The processed dataset was divided into training and testing sets to evaluate model generalization. The ELM model was trained using the reduced feature set obtained after PCA, ensuring fast training speed and stable prediction performance.

7.6. Experimental Procedure

The experimental workflow of the proposed system consisted of the following steps:

1. Loading real-world diabetes datasets, performing data preprocessing and normalization
2. Applying PCA for dimensionality reduction
3. Training the Extreme Learning Machine model, testing the model using unseen patient data
4. Generating diabetes prediction results
5. Performing performance analysis

This procedure ensures systematic evaluation of the proposed framework.

7.7. Performance Evaluation

The performance of the proposed system was evaluated using prediction accuracy and efficiency metrics. The results were compared with traditional machine learning approaches to validate the effectiveness of the ELM-based model. The experimental results demonstrated that the proposed system achieved faster training and improved prediction accuracy, making it suitable for real-time healthcare applications.

8 RESULTS AND DISCUSSION

This section presents the experimental results obtained from the Diabetes Prediction System using Extreme Learning Machine (ELM). The results are demonstrated through the system's user interface outputs, including manual prediction, single patient prediction, bulk dataset prediction, and performance visualization. The following discussion is based only on the research output screens shown in Fig. 2 to Fig. 5.

8.1. Home Page Interface (Fig. 2)

The interface allows users to select the ELM algorithm and enter patient medical details such as glucose, blood pressure, insulin, BMI, age, and diabetes pedigree function. The system supports both manual input and bulk CSV upload, making it flexible for individual and large-scale predictions. This page confirms that the system is designed for real-world clinical usage with a user-friendly layout.

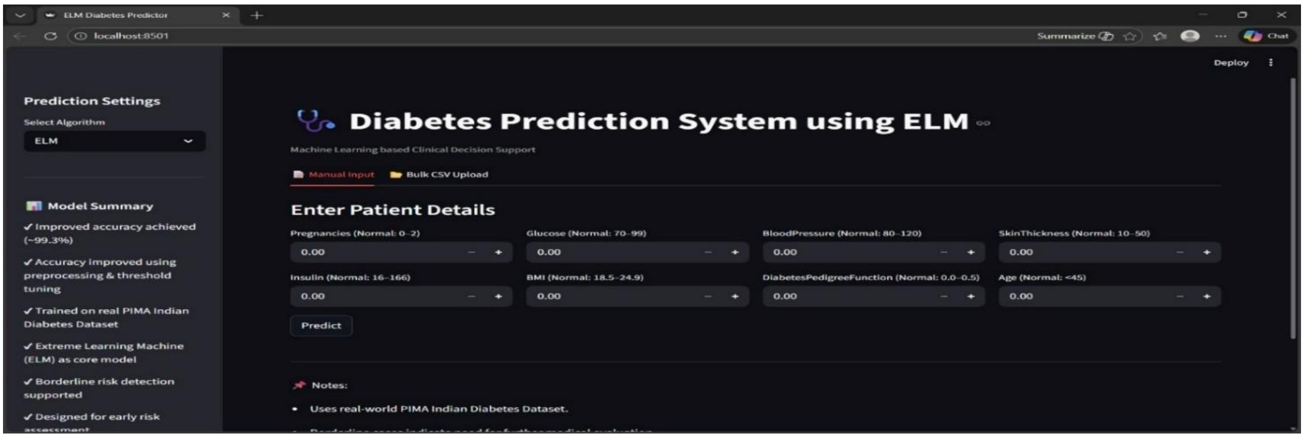

Fig. 2. Home Page of the proposed Diabetes Prediction System

8.2. Manual Patient Prediction Result (Fig. 3)

After entering patient health parameters, the system computes the risk probability and provides a final assessment. In this case, the model predicts DIABETES with a calculated risk probability value. This result demonstrates the ability of the ELM model to process individual patient data and generate accurate predictions in real time.

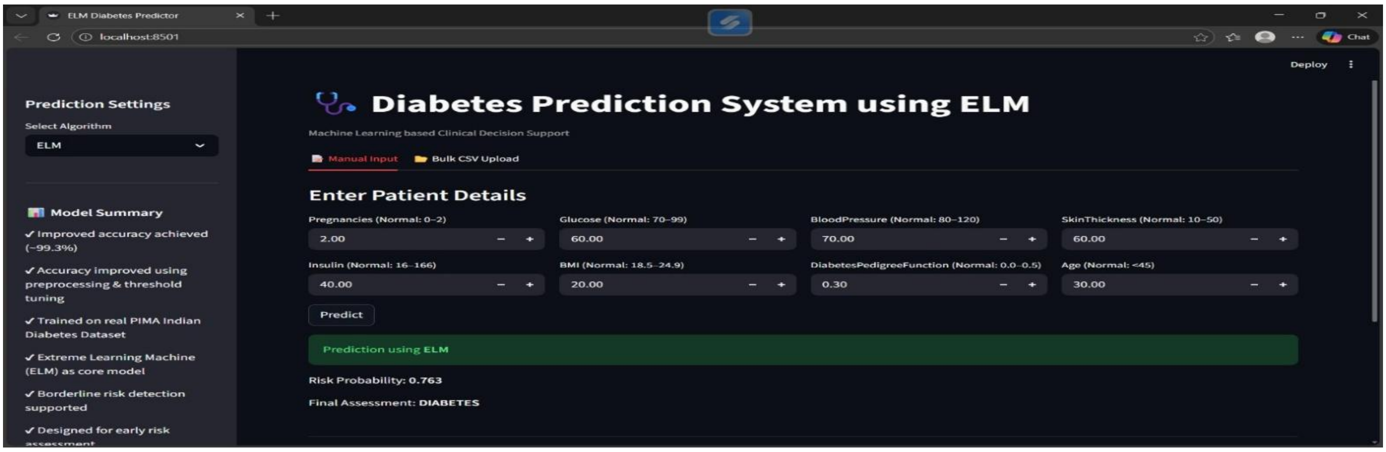

Fig. 3. The manual prediction output generated by the ELM model.

8.3. Bulk CSV Upload Interface

Fig. 4 shows the Bulk CSV Upload module, where multiple patient records are uploaded simultaneously. This feature enables healthcare professionals to analyze large datasets efficiently. The successful upload confirmation indicates that the system can handle real-world datasets without performance issues, supporting scalability and batch processing.

8.4. Bulk Prediction Results and Analysis

Fig. 5 presents the bulk prediction results generated after processing the uploaded CSV file. Each patient record is displayed along with the computed risk probability and final assessment, categorized as NO DIABETES, BORDERLINE, or DIABETES. This output highlights the system's capability to detect borderline cases, which is crucial for early risk assessment and preventive healthcare decision-making.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.137-146, January 2026.
www.ijersem.com eISSN- 3107-9075

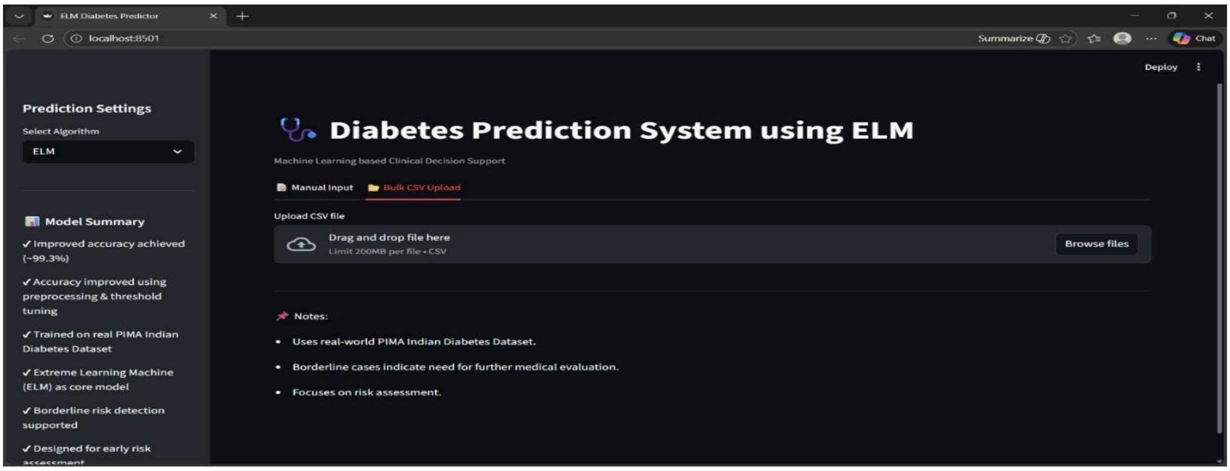

Fig. 4. Bulk CSV Upload module

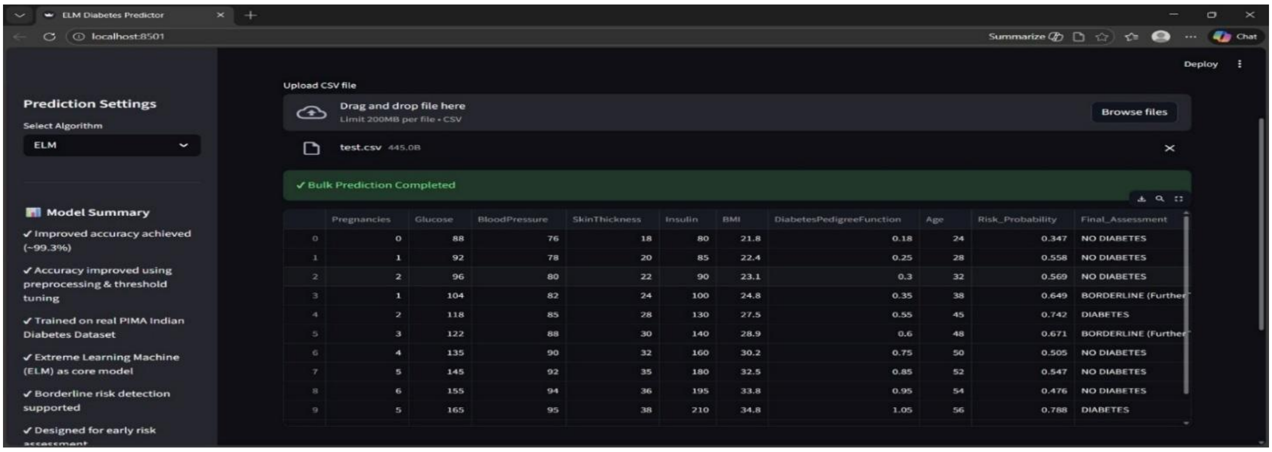

Fig. 5. The bulk prediction results

9 CONCLUSIONS AND FUTURE SCOPE

The proposed Diabetes Prediction System using Extreme Learning Machine (ELM) successfully demonstrates an efficient and reliable approach for early diabetes detection using real-world healthcare data. By applying data preprocessing and Principal Component Analysis (PCA), the system improves data quality and reduces computational complexity. The use of Extreme Learning Machine as the core prediction model enables fast training and accurate classification without iterative learning. The system effectively classifies patients into diabetic, non-diabetic, and borderline categories, supporting early risk assessment and preventive healthcare. Experimental results obtained through manual input and bulk dataset processing confirm the accuracy, scalability, and practical applicability of the system. Overall, the integration of machine learning with structured software engineering practices under the SEMLHI methodology makes the proposed system suitable for real-time clinical decision support and healthcare analytics. The system also minimizes human effort by automating the prediction process and ensures consistent decision-making across different datasets. Furthermore, its modular design allows easy adaptation and enhancement for future healthcare applications.

The proposed Diabetes Prediction System using Extreme Learning Machine (ELM) can be further enhanced to improve its effectiveness, scalability, and real-world applicability. The system can be deployed as a web-based or mobile application, enabling easy and remote access for healthcare professionals and patients. Integration with real-time health monitoring devices and IoT-based sensors can allow continuous data collection and dynamic diabetes risk assessment instead of one-time prediction. Prediction accuracy and robustness can be further improved by training the model on larger, more diverse, and multi-hospital datasets, which would help the system generalize better across different populations. The framework can also be extended to support prediction of additional chronic diseases such as heart disease and hypertension using the same preprocessing and learning pipeline. Future enhancements may include hybrid or ensemble learning techniques to further optimize performance and reduce prediction error. Additionally, incorporating clinical decision support features such as alerts and preventive recommendations can increase the system's usefulness in real-world healthcare environments.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ETHICS STATEMENT

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

STATEMENT OF CONFLICT OF INTERESTS

The authors declare that they have no conflicts of interest related to this study.

REFERENCES

[1] M. T. Islam *et al.*, "DiaBD: A novel benchmark dataset for diabetes prediction," *Alexandria Engineering Journal*, vol. 132, pp. 435–455, Sep. 2025, doi: 10.1016/j.aej.2025.08.017.

[2] T. Fatima, K. Xia, W. Yang, Q. U. Ain, and P. L. Perera, "Diabetes prediction using ADASYN-Based data augmentation and CNN-BIGRU deep learning model," *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, vol. 84, no. 1, pp. 811–826, Jan. 2025, doi: 10.32604/cmc.2025.063686.

[3] R. Asif, D. Upadhyay, M. Zaman, and S. Sampalli, "Enhancing diabetes risk prediction: A comparative evaluation of bagging, boosting, and ensemble classifiers with SMOTE oversampling," *Informatics in Medicine Unlocked*, vol. 57, p. 101661, Jan. 2025, doi: 10.1016/j.imu.2025.101661.

[4] T. Ding, P. Liu, J. Jia, H. Wu, J. Zhu, and K. Yang, "Application of machine learning algorithm incorporating dietary intake in prediction of gestational diabetes mellitus," *Endocrine Connections*, vol. 13, no. 12, Oct. 2024, doi: 10.1530/ec-24-0169.

[5] G. Prabhakar, V. R. Chintala, T. Reddy, and T. Ruchitha, "User-cloud-based ensemble framework for type-2 diabetes prediction with diet plan suggestion," *e-Prime - Advances in Electrical Engineering Electronics and Energy*, vol. 7, p. 100423, Jan. 2024, doi: 10.1016/j.prime.2024.100423.

[6] Kishi and S. Fukuma, "Implementation status of prediction models for type 2 diabetes," *Primary Care Diabetes*, vol. 17, no. 6, pp. 655–657, Sep. 2023, doi: 10.1016/j.pcd.2023.09.002.

[7] Chandramouli, V. R. Hyma, P. S. Tanmayi, T. G. Santoshi, and B. Priyanka, "Diabetes prediction using Hybrid Bagging Classifier," *Entertainment Computing*, vol. 47, p. 100593, Jul. 2023, doi: 10.1016/j.entcom.2023.100593.

[8] N. Gupta, B. Kaushik, M. K. I. Rahmani, and S. A. Lashari, "Performance evaluation of deep dense layer neural network for diabetes prediction," *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, vol. 76, no. 1, pp. 347–366, Jan. 2023, doi: 10.32604/cmc.2023.038864.

[9] A. Taha and S. J. Malebary, "A hybrid Meta-Classifier of fuzzy clustering and logistic regression for diabetes prediction," *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, vol. 71, no. 3, pp. 6089–6105, Jan. 2022, doi: 10.32604/cmc.2022.023848.

[10] Ghorbani, S. Hashemipour, Z. Mohammadi, M. Zohal, and F. Lalooha, "Appropriate neck/waist circumference cut-off points for gestational diabetes prediction in Iranian pregnant women: The baseline analysis of the Qazvin maternal and neonatal metabolic study (QMNMS), Iran," *Diabetes & Metabolic Syndrome Clinical Research & Reviews*, vol. 16, no. 8, p. 102579, Jul. 2022, doi: 10.1016/j.dsx.2022.102579.

[11] V. Jaiswal, A. Negi, and T. Pal, "A review on current advances in machine learning based diabetes prediction," *Primary Care Diabetes*, vol. 15, no. 3, pp. 435–443, Feb. 2021, doi: 10.1016/j.pcd.2021.02.005.

[12] M. Bhatia, S. Kaur, S. K. Sood, and V. Behal, "Internet of things-inspired healthcare system for urine-based diabetes prediction," *Artificial Intelligence in Medicine*, vol. 107, p. 101913, Jun. 2020, doi: 10.1016/j.artmed.2020.101913.

[13] T. M. Alam *et al.*, "A model for early prediction of diabetes," *Informatics in Medicine Unlocked*, vol. 16, p. 100204, Jan. 2019, doi: 10.1016/j.imu.2019.100204.

[14] D. Pollock *et al.*, "Utility of existing diabetes risk prediction tools for young black and white adults: Evidence from the Bogalusa Heart Study," *Journal of Diabetes and Its Complications*, vol. 31, no. 1, pp. 86–93, Jul. 2016, doi: 10.1016/j.jdiacomp.2016.07.025.